

上海交通大学博士学位论文

DarkSHINE 实验模拟框架设计与正电子束流模式下的暗光子产生分析

姓　　名：朱栩量

学　　号：021426910016

导　　师：李数教授

院　　系：李政道研究所

学 科/专 业：粒子物理

申 请 学 位：理学博士

2026 年 11 月 13 日

A Dissertation Submitted to

Shanghai Jiao Tong University for the Degree of Doctor

DESIGN OF DARKSHINE SIMULATION FRAMEWORK AND PROSPECTIVE STUDY OF POSITRON-BEAM MODE

Author: Xuliang Zhu

Supervisor: Prof. Shu Li

Tsung-Dao Lee Institute

Shanghai Jiao Tong University

Shanghai, P.R. China

November 13^{th}, 2026

上海交通大学

学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全知晓本声明的法律后果由本人承担。

学位论文作者签名：

日期：　　年　　月　　日

上海交通大学

学位论文使用授权书

本人同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。

本学位论文属于：

□**公开论文**

□**内部论文**，保密□1 年/□2 年/□3 年，过保密期后适用本授权书。

□**秘密论文**，保密____年（不超过 10 年），过保密期后适用本授权书。

□**机密论文**，保密____年（不超过 20 年），过保密期后适用本授权书。

（请在以上方框内选择打“✓”）

学位论文作者签名：　　指导教师签名：

日期：　　年　　月　　日　　日期：　　年　　月　　日

摘 要

天文和宇宙学观测表明暗物质约占宇宙物质总量的百分之 85%，但其粒子属性以及与标准模型粒子的非引力相互作用仍未被确认。暗光子模型通过动力学混合将标准模型与隐扇区连接，为兆电子伏至吉电子伏质量区间的轻暗物质提供了具有明确实验特征的理论框架。DarkSHINE 是拟依托上海硬 X 射线自由电子激光装置建设的固定靶实验，利用高重复频率的单电子束流以及对末态粒子的近全覆盖测量，通过缺失能量和缺失动量寻找不可见衰变的暗光子。

本论文介绍 DarkSHINE 软件系统框架的设计与实现，以及其在探测器优化和物理研究中的应用。该系统以 Geant4 和 ROOT 为核心，将可配置的探测器几何、粒子产生与输运、探测器数字化、重建及物理分析连接起来，并通过稀有过程截面偏置和事例筛选支持专用背景样本的产生。该框架为探测器几何扫描与性能评估提供了模拟基础。在 8 GeV 电子束流模拟中，所选七层标记径迹探测器的重建效率为 99.81%，动量分辨率为 0.87%；在 ECAL 单电子模拟中，数字化后相对能量分辨率的随机项系数小于 2%。

基于该软件系统，论文进一步研究了 DarkSHINE 使用 8 GeV 正电子束流搜索不可见暗光子的可行性。除与电子束流相同的暗韧致辐射过程外，正电子还可与靶原子电子发生非共振和共振湮灭，从而扩展暗光子的产生机制。针对三类信号的不同末态特征，论文构建了三个正交信号区，并利用完整探测器模拟评估信号效率和主要稀有背景。研究表明，湮灭过程能够在特定质量区间显著增强正电子模式的探索能力，尤其是在 8 GeV 束流对应的共振质量附近。在十年累计 3×10^{14} 个正电子在靶的基准曝光量下，采用本文的背景模型，90 MeV 暗光子的联合预期 90% 置信水平上限达到 $\epsilon^2 = 5.33 \times 10^{-13}$。该结果是在拟合背景模型与 HCAL 单元阈值假设下，对设计阶段物理灵敏度的前瞻性估计。论文最后讨论在多束流能量下利用湮灭共振扫描轻暗物质参数空间的研究方向。

关键词： 暗物质，暗光子，固定靶实验，蒙特卡洛模拟，正电子束流，缺失能量

Abstract

Dark matter constitutes approximately 85% of the matter in the Universe, yet its particle nature and non-gravitational interactions remain unknown. Kinetic mixing with a dark photon connects the Standard Model to a hidden sector and enables searches for MeV–GeV dark matter. DarkSHINE is a proposed fixed-target experiment at the Shanghai High Repetition Rate XFEL and Extreme Light Facility. It will use a high-repetition-rate single-electron beam and a nearly hermetic detector to search for invisibly decaying dark photons through missing energy and momentum.

This thesis presents the design and implementation of the DarkSHINE software framework for detector optimisation and physics studies. Based on Geant4 and ROOT, it connects particle generation and transport, configurable geometry, digitisation, reconstruction, and analysis. Rare-process biasing and event filters provide dedicated background samples. In 8 GeV electron-beam simulation, the seven-layer tagging tracker achieves 99.81% reconstruction efficiency and 0.87% momentum resolution. The isolated-electron ECAL study gives a stochastic coefficient below 2% for the digitised relative energy resolution.

Using this framework, the thesis investigates an 8 GeV positron-beam search for invisibly decaying dark photons. Alongside dark bremsstrahlung, positrons enable non-resonant and resonant annihilation with atomic electrons. Three orthogonal signal regions use the different final-state signatures; full detector simulation evaluates their signal efficiencies and principal rare backgrounds. Annihilation improves the reach near the resonant mass accessible with an 8 GeV beam. At a mediator mass of 90 MeV, the combined expected 90% confidence-level limit reaches $\epsilon^2 = 5.33 \times 10^{-13}$ for the ten-year exposure of 3×10^{14} positrons on target. This is a prospective sensitivity estimate under the fitted background model and assumed HCAL single-cell thresholds. The thesis also discusses scanning light-dark-matter parameter space with several positron-beam energies.

Key words: dark matter, dark photon, fixed-target experiment, Monte Carlo simulation, positron beam, missing energy

Contents

Chapter 1 Introduction 1
1.1 Scientific Motivation 1
1.2 The DarkSHINE Approach 2
1.2.1 From an incident particle to a candidate event 2
1.2.2 The role of a common simulation chain 3
1.3 Motivation for a Positron-Beam Mode 3
1.4 Scope and Contributions of This Thesis 4
1.5 Thesis Organisation 5

Chapter 2 Dark Matter and the Vector Portal 7
2.1 Evidence and Properties of Dark Matter 7
2.1.1 Historical and Astrophysical Evidence 7
2.1.2 Small-Scale Structure 8
2.1.3 Mass, Temperature, and Formation Constraints 9
2.2 Dark sectors and portal interactions 13
2.2.1 Kinetic mixing and mass eigenstates 14
2.2.2 Visible and invisible mediator decays 15
2.2.3 Thermal target for light dark matter 17
2.2.4 Fixed-target kinematic scales 18
2.2.5 From production rate to selected yield 20
2.3 Dark-Matter Production and Abundance 21
2.3.1 Thermal Freeze-out 21
2.4 Experimental Searches for Dark Matter 24
2.4.1 Direct Detection 24
2.4.2 Searches for ultralight fields 25
2.4.3 Accelerator Production 26
2.4.4 Dark-Photon Production in Lepton Fixed-Target Experiments 26
2.4.5 Measurement strategies and their interpretation 28

Chapter 3 The DarkSHINE Experiment and Detector System 31
3.1 Experimental Setup and Beamline 31
3.2 Magnetic Field System 32
3.3 Tracking System 34
3.3.1 Requirements and Sensor Technology 34
3.3.2 Tagging Tracker 36
3.3.3 Recoil Tracker 39
3.3.4 Sensor Prototype Performance 39
3.4 Tungsten Target 39
3.5 Electromagnetic Calorimeter 49
3.5.1 Reference Design 49
3.5.2 Expected Performance 53
3.6 Hadronic Calorimeter 57
3.6.1 Role and Physics Requirements 57
3.6.2 Reference Conceptual Design 64
3.6.3 Expected Performance 65
3.7 Subsystem Prototyping and Instrumentation Validation 66
3.7.1 AC-LGAD sensor characterisation 66
3.7.2 LYSO crystal and SiPM response 70
3.7.3 ECAL readout chain 72
3.7.4 ECAL beam-test methodology 76
3.7.5 Radiation environment and calibration stability 78
3.7.6 HCAL active unit and optical readout 82
3.8 Mechanical Integration 84

Chapter 4 Event Simulation and Reconstruction in DarkSHINE 91
4.1 The DarkSHINE Simulation Software Framework 91
4.1.1 Workflow 92
4.1.2 Data structure: `AnaEvent` and `DEvent` 93
4.1.3 `DSimu` configuration overview 96
4.1.4 `DSimu` software architecture 98
4.1.5 Processing an event through the framework 98

4.1.6 Staged transport and filter decisions 101
4.1.7 Simulation performance 102
4.2 Hit Digitisation 103
4.2.1 Optical-photon simulation and lookup table 104
4.2.2 Electronics response and calibration 104
4.2.3 Waveform formation and the energy estimator 108
4.2.4 Parameterised response and stored optical information 108
4.3 Track Reconstruction 109
4.3.1 Candidate construction and the circle pre-fit 110
4.3.2 State propagation and measurement updates 112
4.3.3 ACTS integration in the reconstruction framework 112
4.4 ECAL Cluster Reconstruction 115
4.5 Signal Generation and Generator Validation 116
4.5.1 External event-generator chain 116
4.5.2 Transport-native signal process 122
4.6 Rare Standard Model Processes 122
4.6.1 Photonuclear and electronuclear interactions 124
4.6.2 Muon-pair conversion 124
4.6.3 Biasing weights and effective statistics 125
4.7 Detailed Calorimeter Feature Reconstruction 130
4.7.1 Energy-ordering and spatial moments 130
4.7.2 Topological clusters and track-assisted reconstruction 131
4.8 Software Quality and Reproducible Production 135
4.9 Chapter Summary 136

Chapter 5 Positron-Beam-Mode Analysis 139
5.1 Physics scope and simulated samples 139
5.1.1 Signal generation and visible final states 140
5.1.2 Standard Model production and sample normalisation 141
5.2 From final-state topology to signal regions 142
5.2.1 Tracking and calorimeter discrimination 143
5.2.2 Logical separation of the measured categories 144

5.2.3 Efficiency and migration between categories 145
5.2.4 Background rejection through the selection sequence 146
5.3 Luminosity and selected signal yields 149
5.3.1 Target luminosity and the production normalisation 149
5.3.2 The response matrix and its normalisation 160
5.3.3 Combining channels and changing the reference coupling 162
5.3.4 Example calculations with the benchmark yield table 163
5.3.5 Background exposure and the role of effective POT 163
5.4 Background prediction beyond the simulated statistics 164
5.4.1 Dedicated rare-process estimates 164
5.4.2 Cross-check with the inclusive sample 168
5.5 Expected sensitivity and its physical interpretation 169
5.5.1 Statistical construction used for the projection 169
5.5.2 Relative importance of the production mechanisms 171
5.5.3 Comparison with other searches and thermal targets 174
5.6 Summary 178

Chapter 6 Summary and Outlook 179
6.1 Simulation and reconstruction 179
6.2 Positron-beam search 179
6.3 Assumptions of the projection 180
6.4 Further work 181

Bibliography 183

Appendix A Exotic Higgs decays at the CEPC 191
A.1 Physics question and scalar benchmark 191
A.1.1 Vacuum conditions and physical masses 191
A.1.2 Exotic width and visible scalar decays 192
A.1.3 Thermal interpretation of the collider target 194
A.2 Collision sample and detector response 194
A.2.1 Production kinematics and the mass grid 194
A.2.2 Generation, transport, and reconstruction 195

A.3 From reconstructed particles to scalar candidates 196
A.3.1 Lepton tag and recoil selection 196
A.3.2 Four-jet assignment and heavy flavour 198
A.4 Classification across the scalar-mass grid 200
A.4.1 Information carried by the inputs 200
A.4.2 Boosting, model complexity, and the final score 204
A.5 Uncertainties and statistical treatment 204
A.5.1 Rate variations and detector response 204
A.5.2 Nuisance parameters and sensitivity 206
A.5.3 Profile tests and expected upper bounds 208
A.6 Projected reach and its interpretation 208
A.6.1 Yield normalisation and decay assumptions 208
A.6.2 Mass dependence of the published result 209
A.6.3 Sensitivity scaling and limits on model parameters 211

List of Research Achievements 213

List of Figures

Figure 2.1 Dark matter mass scales, candidates, and representative experiments 12
Figure 3.1 SHINE and DarkSHINE extraction kickers..... 32
Figure 3.2 Reference DarkSHINE detector layout..... 33
Figure 3.3 Tracker assembly and square Helmholtz coils..... 34
Figure 3.4 Longitudinal magnetic-field profile and spatial map 35
Figure 3.5 Tracking-system layout and AC-LGAD sensor..... 36
Figure 3.6 Tagging-tracker layer-number scan at 8 GeV 37
Figure 3.7 Recoil-tracker geometry scans at 30 MeV 40
Figure 3.8 Recoil momentum resolution and algorithm comparison..... 42
Figure 3.9 Target-material dependence of dark-photon production 45
Figure 3.10 Rare-background production in candidate target materials 47
Figure 3.11 ECAL energy and recoil angle for signal and background..... 50
Figure 3.12 Reference ECAL crystal array 52
Figure 3.13 Isolated-electron ECAL response, resolution, and linearity 53
Figure 3.14 ECAL energy-cut efficiency versus dark-photon mass..... 56
Figure 3.15 ECAL observables for dark photons and rare backgrounds..... 57
Figure 3.16 Neutron energies after the ECAL energy selection..... 63
Figure 3.17 Reference uniform-absorber HCAL geometry..... 64
Figure 3.18 Alternating HCAL scintillator and absorber layers 65
Figure 3.19 Neutron-veto inefficiency versus HCAL depth 66
Figure 3.20 HCAL sensitive-layer response to incident hadrons 67
Figure 3.21 ECAL–HCAL energy correlation in inclusive events..... 68
Figure 3.22 Electrical characterisation of AC-LGAD prototypes 69
Figure 3.23 AC-LGAD laser-scan setup and sensor module 69
Figure 3.24 AC-LGAD time resolution along two laser scan lines..... 70
Figure 3.25 LYSO–SiPM component and module prototypes 71
Figure 3.26 Measured and simulated intrinsic LYSO spectrum..... 73
Figure 3.27 ECAL preamplifier, ADC, and FPGA prototypes..... 75

Figure 3.28 ECAL waveform and frequency-domain readout tests..... 75
Figure 3.29 Crystal beam-test and SiPM laser-test arrangements..... 77
Figure 3.30 Crystal-channel response at four electron-beam energies 78
Figure 3.31 Spatial distributions of ECAL radiation exposure..... 82
Figure 3.32 Cosmic-ray setup for HCAL strip measurements 84
Figure 3.33 HCAL photoelectron yield versus fibre configuration 85
Figure 3.34 HCAL readout board and digitisation electronics 87
Figure 3.35 Mechanical supports for the DarkSHINE detector 88
Figure 3.36 Tracker and target support assembly 88
Figure 3.37 ECAL support and maintenance rails 89
Figure 3.38 Independent support frame for the HCAL..... 89
Figure 4.1 DarkSHINE Software event-processing workflow..... 94
Figure 4.2 DarkSHINE event-data organisation..... 95
Figure 4.3 Current simulation detector geometry 99
Figure 4.4 Simulation initialisation and event loop..... 100
Figure 4.5 Optical-photon detection efficiency and LUT validation 105
Figure 4.6 Optical-photon transit times and LUT validation 106
Figure 4.7 ECAL pulse model and digitised waveforms 107
Figure 4.8 Calibration of the pre-digitised ECAL response 107
Figure 4.9 Greedy track-finding workflow 110
Figure 4.10 Comparison of track-fitting momentum spectra 113
Figure 4.11 Dark-bremsstrahlung cross section 117
Figure 4.12 Generator-level dark-bremsstrahlung kinematics 118
Figure 4.13 Comparison of dark-photon generators 123
Figure 4.14 CLAS recoil-proton validation configuration 125
Figure 4.15 Recoil-proton process and momentum comparisons..... 126
Figure 4.16 Muon-pair-conversion validation observables 128
Figure 4.17 ECAL energy distributions for rare-process samples 131
Figure 4.18 ECAL shower-shape distributions 133
Figure 4.19 ECAL second-moment distributions 136
Figure 5.1 Positron dark-photon production cross sections..... 140

Figure 5.2 Tagging-to-recoil momentum difference 143
Figure 5.3 ECAL and HCAL energy correlations 144
Figure 5.4 Total ECAL energy for signal and background 146
Figure 5.5 Largest ECAL cell energy for annihilation signals 148
Figure 5.6 Signal efficiencies in the recoil-positron category SR1 150
Figure 5.7 Signal efficiencies in the photon category SR2 153
Figure 5.8 Signal efficiencies in the low-activity category SR3 156
Figure 5.9 Inclusive and dedicated background survival comparison 165
Figure 5.10 Dedicated-background sideband fits for SR1 166
Figure 5.11 Photonuclear-background sideband fits for SR2 168
Figure 5.12 Photonuclear-background survival surfaces for SR3 170
Figure 5.13 Inclusive-background sideband fit for SR1 172
Figure 5.14 Inclusive-background sideband fit for SR2 173
Figure 5.15 Inclusive-background survival surface for SR3 174
Figure 5.16 Expected sensitivity by production mechanism 175
Figure 5.17 Expected positron sensitivity and search comparisons 176
Figure 5.18 Expected sensitivity and thermal-relic targets 177

List of Tables

Table 2.1 Observables and detector information used by accelerator search strategies. … 29
Table 3.1 Connection between the missing-energy measurement and the functions of the DarkSHINE detector subsystems. …… 31
Table 3.2 Key geometric parameters of the tracking detectors. …… 37
Table 3.3 Physical parameters of candidate target materials[46]. …… 44
Table 3.4 Irreducible neutrino-induced missing-energy background yields for different target materials (3×10^{14} EOT, $E_{\mathrm{beam}} = 10$ GeV)…… 49
Table 3.5 Final-state particle composition from EN and PN background processes. …… 63
Table 4.1 Simulation global configuration variables …… 96
Table 4.2 ROOT output-manager options…… 96
Table 4.3 Event-collection output options …… 97
Table 4.4 Cross-section biasing options…… 97
Table 4.5 Production-filter options…… 98
Table 4.6 Current simulation detector parameters …… 99
Table 4.7 Roles of sampling and selection operations…… 102
Table 4.8 Simulation computing benchmarks …… 103
Table 4.9 Calorimeter electronics-response parameters…… 104
Table 5.1 Signal mediator-mass samples…… 141
Table 5.2 Background production and equivalent exposure …… 142
Table 5.3 Signal-region selection requirements …… 142
Table 5.4 Background cutflow for SR1…… 159
Table 5.5 Background cutflow for SR2…… 160
Table 5.6 Background cutflow for SR3…… 161
Table 5.7 Produced and selected benchmark signal yields …… 172

Chapter 1 Introduction

1.1 Scientific Motivation

Evidence from galactic rotation curves, gravitational lensing, the cosmic microwave background, and large-scale structure requires a non-luminous matter component beyond the Standard Model. Although its gravitational effects are well established, no particle dark-matter candidate has yet been identified. This motivates experiments that test interactions between a dark sector and Standard Model particles over a broad range of masses and couplings.

Light dark matter in the MeV–GeV range is particularly well matched to accelerator experiments. A widely used benchmark is the vector portal, in which a massive dark photon A' couples to the electromagnetic current through kinetic mixing. If the dark photon decays predominantly to lighter dark-sector states, it is invisible to ordinary detectors. Its production must then be inferred from an imbalance in the measured energy and momentum of the visible particles.

An accelerator search tests whether a weakly coupled particle can be produced in a controlled laboratory interaction and escape the apparatus. The beam defines the incident energy and direction, while the target defines the available interaction material. A prediction can be expressed in terms of a production cross section and the probability that the resulting visible particles satisfy the event selection. The small couplings of interest make the beam exposure and rejection of ordinary interactions part of the same sensitivity problem.

A deficit of visible energy is not unique to a new particle. Standard Model particles can also leave the instrumented acceptance or deposit too little energy to be identified. For a missing-energy search, the relevant background is the small subset of ordinary events that survives all measurements and vetoes. Predicting that rate requires the simulation to follow the hard interaction and its secondary particles through the detector response to the final selection, including the rare tails of electromagnetic showers.

The vector portal permits this problem to be studied with a defined set of parameters. The mediator mass controls the available kinematics; kinetic mixing controls its production rate; and the dark-sector masses and coupling determine the invisible-decay assumption. The same laboratory result can then be compared with a thermal-relic target after fixing the additional

dark-sector parameters.

1.2 The DarkSHINE Approach

DarkSHINE is a proposed fixed-target experiment using the high-repetition-rate single-electron beam available from the SHINE facility. A tagging tracker measures the incident particle, a thin tungsten target provides the interaction material, and a recoil tracker together with electromagnetic and hadronic calorimeters measures or vetoes the visible final state. This arrangement is designed to identify events with one low-momentum recoil particle and substantial missing energy while rejecting rare Standard Model processes that produce neutral hadrons or muons.

The small signal cross section and the required background rejection make simulation central to the experiment. Detector optimisation requires a configurable geometry and realistic response model, while the study of rare backgrounds requires biased event generation without distorting physical weights. The resulting samples must be processed through a common digitisation, reconstruction, and analysis chain so that detector choices can be propagated consistently to the final sensitivity.

1.2.1 From an incident particle to a candidate event

The tagging tracker establishes the charged particle entering the target and provides the momentum reference for the downstream measurement. The target is the location of the intended hard interaction, but it also introduces energy loss and multiple scattering. The recoil tracker measures charged particles after that interaction. Its acceptance and reconstruction affect both the measured momentum and the event's track multiplicity. The calorimeters then measure the visible energy or reject activity associated with particles that could account for the apparent deficit.

For dark bremsstrahlung, this sequence tests whether a well-measured incident particle is followed by a low-momentum recoil and sufficiently little calorimeter activity. A neutral particle produced in a rare Standard Model interaction can carry energy through the ECAL and must be tested by the HCAL. A photon converting to muons produces another possible route to a low electromagnetic deposit. Particles produced in the target and particles produced inside the ECAL encounter different amounts of instrumented material. Their production

locations and ancestry consequently enter the background classification.

The measurements also have distinct failure modes. A charged particle outside the tracker acceptance changes the reconstructed topology, while a poorly contained shower changes the visible-energy sum.

1.2.2 The role of a common simulation chain

The detector description supplies geometry, materials, and fields to particle transport. Transport supplies energy deposits and secondary particles. Digitisation and reconstruction convert those quantities into the observables available to the analysis.

The DarkSHINE software exchanges event collections through a common data model. Monte Carlo truth can identify why an event enters a selected category, while the reconstructed collections provide the quantities on which the selection acts. The common event record allows a simulated interaction to be followed from transport to final selection.

Rare-process production requires further bookkeeping. Enhancing a process in simulation increases the frequency with which its events are sampled. Physical weights and the represented beam exposure must accompany the sample to recover the predicted rate. Production filters remove events outside the generated region and must be distinguished from the final analysis cuts. The framework records these definitions and carries them through reconstruction.

Full optical transport is used to study light collection, while a lookup table reduces the cost of repeated response calculations.

1.3 Motivation for a Positron-Beam Mode

The baseline electron-beam programme produces dark photons primarily through dark bremsstrahlung. Replacing electrons with positrons retains this channel and adds annihilation on atomic electrons. Non-resonant annihilation produces a visible photon in association with the dark photon, whereas resonant annihilation can transfer nearly the full centre-of-mass energy to an invisible dark photon. These complementary final states motivate separate signal regions and can improve the reach around the annihilation resonance. They also introduce positron-specific Standard Model backgrounds. Evaluating the benefit of positron running therefore requires the production cross sections together with the detector response

and selection.

At an incident positron energy of 8 GeV, annihilation on a stationary electron has a centre-of-mass energy near 90 MeV. The nuclear bremsstrahlung process has a different target and kinematic scale. Adding annihilation therefore changes the mass dependence of the available signal rate as well as the visible final state. The benchmark studied here uses the same incident particle count and reference target as the earlier electron-mode projection, allowing the effect of the additional production mechanisms to be examined at a matched exposure.

The three production mechanisms suggest three observable categories: a recoil positron with missing energy, a photon-like deposit with missing energy, and an event with little visible downstream activity. Transport and reconstruction can move an event between these categories: a recoil can fail to produce a reconstructed track, and a photon can escape acceptance or fall below the energy threshold. Nine channel-to-region efficiencies describe the contribution of each production mechanism to each mutually exclusive category.

The comparison evaluates how much of the additional annihilation rate survives the detector selection and how these selected events combine with dark bremsstrahlung over the mediator-mass range. A larger production cross section can be offset by a smaller efficiency, and a contribution reconstructed in another category still carries signal information. The selected yields, with migration retained, provide the inputs to the sensitivity calculation.

Positron running also changes the Standard Model processes that must be simulated. Ordinary annihilation contributes photons, which can interact or escape along with particles from the electromagnetic shower. The projected advantage of the beam mode depends on the background treatment and veto performance.

1.4 Scope and Contributions of This Thesis

The author is one of the original designers of the DarkSHINE software framework. The work connects detector-simulation development to a prospective positron-beam analysis. The author led the following contributions:

1. the development of the `DSimu` simulation framework and the `DEvent` event model, providing configurable transport and a shared representation of event information, together with interfaces supporting detector-geometry optimisation and performance analysis;

2. the development of rare-process biasing and staged filtering, which concentrate simulation resources on the backgrounds of interest;
3. the integration of ACTS into the DarkSHINE reconstruction software; and
4. a positron-beam search using three signal regions for dark bremsstrahlung and the two annihilation mechanisms, with studies of signal efficiency, background extrapolation, and projected sensitivity.

The `DAna` framework, digitisation, other track-reconstruction algorithms, calorimeter reconstruction, and generator-validation studies are collaboration results. The detector hardware results are also collaboration contributions.

1.5 Thesis Organisation

Chapter 2 introduces the dark-matter and dark-photon framework and reviews the relevant search strategies. Chapter 3 describes the DarkSHINE experiment and detector subsystems. The simulation, digitisation, and reconstruction software is presented in Chapter 4. Chapter 5 develops the positron-beam-mode analysis and its projected sensitivity. The principal results, limitations, and future work are summarised in the final chapter.

Chapter 2 Dark Matter and the Vector Portal

The gravitational evidence for dark matter constrains its abundance and its role in structure formation. An accelerator search requires additional assumptions about interactions with ordinary particles. The vector portal supplies a benchmark in which those assumptions can be stated through a mediator mass and couplings. Its production channels determine the visible particles that a fixed-target experiment measures, while its dark-sector parameters connect the measurement to an abundance calculation. Throughout the kinematic expressions below, natural units with $\hbar = c = 1$ are used unless a conversion factor is displayed explicitly.

2.1 Evidence and Properties of Dark Matter

2.1.1 Historical and Astrophysical Evidence

In 1933, Zwicky estimated the mass of the Coma cluster in two different ways[1]. The first method was to count galaxies and add up their luminous masses, using a mass-to-light ratio calibrated from local stellar systems, $\frac{M_{\text{tot}}}{L_{\text{tot}}} \sim 3\frac{M_\odot}{L_\odot}$ (modern value $\frac{M_{\text{tot}}}{L_{\text{tot}}} \sim 160\frac{M_\odot}{L_\odot}$[2]). He assumed the cluster contained about 800 galaxies, each with a typical mass of about 10^9 solar masses. From this, he obtained a total luminous mass $8 \times 10^{11} M_\odot \sim 1.6 \times 10^{45}$ g of the Coma system (modern value is $\sim 10^{13} M_\odot$). The second method infers the cluster mass from its gravitational effects. Assuming dynamical equilibrium, the virial theorem relates the kinetic energy of its members to their gravitational potential energy. From Doppler shifts of the radiation coming from eight galaxies, he obtained a typical velocity dispersion of $\sigma \approx 1019 \pm 30\,\text{km/s}$ (modern $\sigma \approx 1082 \pm 30\,\text{km/s}$[3]). The estimated radius of the Coma cluster was $\sim 10^6$ light years. This gives a mass $M = \frac{5}{3}\frac{\langle v^2 \rangle R}{G} \sim 3 \times 10^{47}$ g (modern value $\sim 10^{48}$ g). The average mass density in the Coma cluster is larger than the density inferred from the luminous matter alone by 2 orders of magnitude.

Galactic rotation curves use Doppler velocities to trace the gravitational potential, as in Zwicky's second method. If most of the mass is concentrated in a central region of total mass M, orbital velocities outside that region should fall as $v \propto 1/\sqrt{r}$. Models based on the stellar disk and gas therefore predict a central peak followed by a decline at larger radii. In 1985,

van Albada et al. studied the rotation curve of the spiral galaxy NGC 3198[4]. Typically for $v_{\max}(\text{disk}) \simeq 140\,\text{km/s}$ at $r \simeq 5\,\text{kpc}$, we would expect $v \simeq 70\,\text{km/s}$ at $r \simeq 30\,\text{kpc}$. However the observed circular velocity is about $150\,\text{km/s}$ and stays almost constant out to $r \simeq 30\,\text{kpc}$. If we want the rotation curve to be flat, the enclosed mass within radius r must grow as $M(r) \propto r$. This suggests a "dark halo" reaches beyond the stellar distribution. Since the 1970s, there have been many studies of rotation curves in different systems[5-8]. These studies generally find rotation curves that are close to flat at large radii.

The modern statement based on improved observations and improved simulations is that every galaxy and our universe is embedded in an extended halo of dark matter, with densities approximately described by NFW (Navarro-Frenk-White) or Einasto profile.

2.1.2 Small-Scale Structure

While collisionless cold dark matter (CDM) simulations successfully describe large-scale structures, they systematically over-predict the density of dark matter on small scales ($\leq 10\,\text{kpc}$)[9]. Two classic issues that arise from comparing these dark matter-only simulations to observations of dwarf galaxies are known as the "cusp-core" and "too big to fail" problems.

The "cusp-core" problem refers to the distribution of dark matter at the center of galaxies. CDM-only simulations predict that the dark matter density should continue to rise steeply at small distances, forming a density "cusp" ($\rho \propto 1/r$). However, observations of small dwarf galaxies, such as those in the THINGS[10] and LITTLE THINGS[11] surveys, suggest that the density profile actually flattens out, forming a "core" of about 0.1 to 1 kpc in size.

The "too big to fail" problem is a related issue regarding the masses of dwarf galaxies. Simulations predict the existence of massive, dense dark matter subhalos that should naturally attract enough gas to form stars, making them "too big to fail" at star formation. Yet, observations show fewer of these massive and dense satellites than expected[12-13].

If we assume dark matter is purely collisionless, without self-interaction, these discrepancies might be resolved by properly including baryonic matter in the simulations[14]. Because ordinary matter interacts via electromagnetism, it can dissipate energy and dominate the gravitational potential at the centers of halos. Specifically, if a galaxy has a "bursty" star formation history, supernova feedback can trigger large outflows of gas. This rapid transport

of baryonic mass perturbs the gravitational potential, causing the dark matter to slosh around and breaking up the steep cusp into a flattened core.

Dark-matter self-interactions could also alter the small-scale structure[15-17]. If dark matter particles can scatter off each other, the energy transfer would reduce the central density and produce the observed cores. To significantly change the collisionless picture, the scattering rate must be roughly once per dynamical time (about a billion years in a dwarf galaxy). This requires a self-interaction cross-section over mass ratio of $\sigma/m \sim 0.1 - 1\,\mathrm{cm}^2/\mathrm{g}$. For a dark matter mass of $\sim 1\,\mathrm{GeV}$, this interaction is comparable to the QCD scale. However, the cross-section cannot be arbitrarily large, as values much larger than this are constrained by observations of non-equilibrium colliding systems like the Bullet Cluster[18].

2.1.3 Mass, Temperature, and Formation Constraints

The expanding Universe is described by the Einstein equations of general relativity. On large scales, it can be approximated as spatially homogeneous and isotropic, with matter and energy represented by perfect fluids. Its geometry is described by the Friedmann-Lemaître-Robertson-Walker (FLRW) metric,

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + a(t)^2 \left(\frac{\mathrm{d}r^2}{1 - kr^2} + r^2\,\mathrm{d}\Omega^2 \right). \tag{2.1}$$

The evolution of the cosmic scale factor $a(t)$ is controlled by the Friedmann equation,

$$H \equiv \frac{1}{a}\frac{da}{dt}, \quad H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \tag{2.2}$$

where ρ is the energy density and the second term vanishes in a spatially flat universe ($k = 0$). A perfect fluid is characterised by an equation-of-state parameter w defined by $p = w\rho$. Energy conservation then gives $\rho \propto a^{-3(1+w)}$ for constant w.

For matter, $w = 0,\ p = 0,\ \rho \propto a^{-3}$, which dilutes with volume as space expands.

For radiation, $w = 1/3,\ \rho = 3p,\ \rho \propto a^{-4}$, with dilution from the expanding volume and an additional energy loss as wavelengths stretch.

For dark energy, $w = -1,\ p = -\rho,\ \rho \propto a^0$, current constraint is $w = -1.01 \pm 0.04$[19].

Define $1 + z \equiv \frac{a_0}{a(t)}$, where a_0 is the present-day scale factor and z is the redshift.

The specific energy densities of these perfect fluids are measured and constrained by the Planck 2018 cosmological parameters[20].

The early universe was radiation dominated. The total radiation energy density can be expressed as

$$\rho_{\text{radiation}} = \rho_\gamma + N_{\text{eff}}\rho_\nu, \quad N_{\text{eff}} \equiv 3 + \Delta N_\nu, \tag{2.3}$$

where ρ_ν is the energy density of a single species of neutrino; N_{eff} is the number of "effective neutrino species", which counts contributions to the radiation energy density, including any additional light degrees of freedom. In the Standard Model calculation, $(N_{\text{eff}})_{\text{SM}} = 3.046$.

Big Bang nucleosynthesis (BBN) places an upper limit on additional relativistic degrees of freedom at temperatures of order 1 MeV[21]. The recombination of free electrons and protons into neutral hydrogen provides a later constraint through the cosmic microwave background, for example $N_{\text{eff}} = 3.15 \pm 0.23$ in the cited analysis[22]. These measurements constrain light species that were in thermal contact with the Standard Model. In particular, sub-MeV dark matter must generally have a lower temperature or a sufficiently weak coupling to avoid modifying BBN and recombination. Energy injected by late annihilation or decay is constrained independently[23].

The CMB is a snapshot of the cosmos at redshift $z \sim 1000$. For $T \geq 1$ eV, the Universe is almost fully ionized. Then the ionization level drops rapidly at $z \sim 1000$, $T \sim 0.4$ eV, photons scatter poorly off neutral atoms, and the Universe becomes transparent. Photons in this bath then free-stream to the present day. Prior to the release of photons, the Universe is a plasma of ions, electrons, photons, neutrinos, and maybe dark matter. Small perturbations in density and temperature oscillate over time, as gravity drives density up, and radiation pressure drives density down. The resulting temperature fluctuations are imprinted on the CMB. Temperature and polarization anisotropies are measured in detail by Planck[20], ACT[24], SPT[25], etc. The observed CMB requires a component that gravitates without experiencing radiation pressure, or has equivalent effects, and is roughly 5 times more abundant than ordinary matter. This dark-matter component is hypothesised to be the same one inferred around galaxies and clusters. Dark matter must therefore be present by $z \sim 1000$, $t \sim$ few 100 000 years long, before stars, galaxies, etc. This eliminates many late-forming compact objects as the dominant candidate. Primordial black holes formed in the early Universe remain possible in restricted mass windows, but existing constraints make it difficult for them to constitute all dark matter[26-28]. Modern CMB measurements also test the "darkness" of dark matter

through small interactions between DM + baryons or DM + radiation[29].

After the CMB photons decouple, these density fluctuations continue to grow under gravity to form the structures we see today. The early dark-matter velocity distribution defines three limiting cases for structure formation:

Cold Dark Matter (CDM): The dark matter is very non-relativistic. Small clumps form first and then accrete to form larger structures, creating the "cosmic web". This is currently favored by observations.

Hot Dark Matter (HDM): The dark matter is relativistic. Its large free-streaming length erases small-scale structure, meaning large structures would form first and then fragment. This would predict that clusters are older than galaxies, which is observationally excluded (setting a limit of $m \geq 0.67$ eV for thermal relics)[30].

Warm Dark Matter (WDM): Dark matter has an intermediate velocity, erasing structures below a specific critical scale. A cutoff in the matter power spectrum constrains this case. Observations of the Lyman-α forest give $m_{\mathrm{DM}} > 5.3$ keV at 2σ[31]. Dark matter must therefore have sufficiently low velocities to preserve the observed small-scale structure.

By bounding the size of the smallest dark matter structures, the Lyman-α forest also sets a fundamental lower bound on the dark matter wavelength. Since structure formation is suppressed on scales smaller than this wavelength, the dark matter wave cannot be larger than the small kiloparsec halos we observe. This sets a model-independent lower bound of $m_{\mathrm{DM}} \gtrsim 2 - 3 \times 10^{-21}$ eV[32].

For fermions, Pauli exclusion limits the number of particles that can occupy a small dark-matter halo. The Tremaine-Gunn bound constrains their phase-space density more strongly than the wavelength limit:

$$\frac{n}{p^3} \sim \frac{\rho}{m_{\mathrm{DM}} \left(m_{\mathrm{DM}} v\right)^3} \leq 2$$

Using typical dwarf-galaxy densities and velocities gives:

$$m_{\mathrm{DM}} \gtrsim \left(\frac{\rho}{v^3}\right)^{1/4} \sim \text{few hundred eV}$$

Therefore, particle dark matter below the ~ 100 eV scale must be purely bosonic.

At the low-mass end, these bounds favour light, cold, bosonic fields such as the QCD axion. At MeV masses and above, dark matter can be fermionic or bosonic, and thermal or

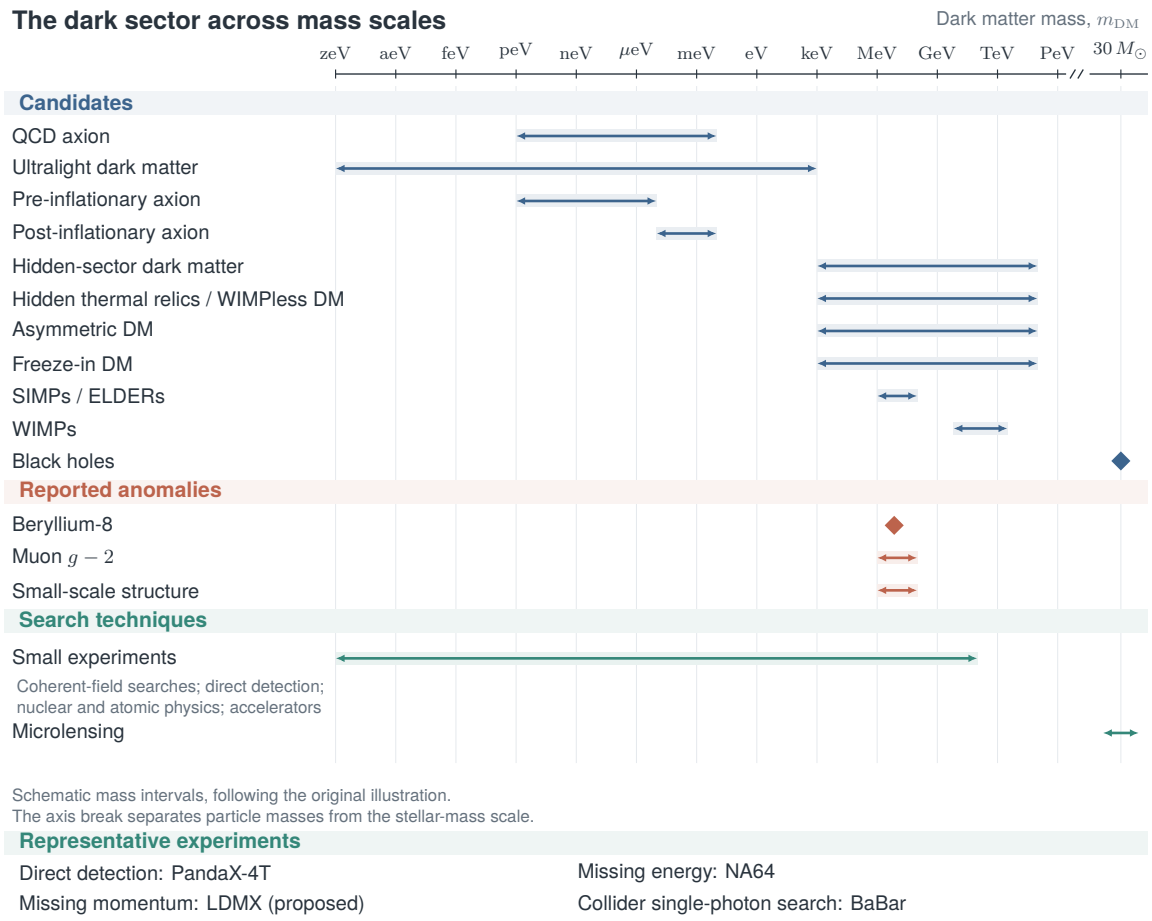


Figure 2.1 Summary of cosmological constraints on the dark matter mass scale, dark sector candidates, anomalies, and search techniques. For $m_{\rm DM} < 10^{-21}$ eV, the wavelength is too large to fit inside small halos. The 10^{-21} eV – keV range includes bosonic, non-thermal, and cold dark matter. The keV – MeV range is tightly constrained or ruled out by Big Bang nucleosynthesis (BBN) for thermal relics. In the MeV – 100 TeV scale, dark matter can be bosonic or fermionic, and thermal or non-thermal. The region above 100 TeV includes macroscopic dark matter and primordial black holes (PBHs). The named experiments illustrate direct-detection and accelerator techniques[33-36].

non-thermal; WIMPs are one example.

The different observations constrain complementary properties. Orbital motions determine the gravitational mass enclosed within a region, subject to assumptions about the tracers and dynamics. Cosmological measurements constrain how a matter component affects the expansion and the growth of perturbations. Structure on small scales restricts its early velocity distribution and, for a specified particle model, its free-streaming or wavelength scale.

Connecting these gravitational measurements to a laboratory rate requires an interaction model, including couplings and mediator branching fractions. A laboratory limit then constrains its specified production and detection processes. The vector portal supplies that model for the fixed-target search considered here.

2.2 Dark sectors and portal interactions

A dark-matter theory can contain a visible sector, with Standard Model (SM) fields, and a dark sector, with the stable particle or particles responsible for the relic density. Interactions between these sectors must respect both sets of gauge symmetries. This allows a dark particle to interact weakly with ordinary matter even though it carries no SM gauge charge. At energies below the scale of additional heavy states, the interaction Lagrangian may be written schematically as

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \mathcal{L}_{\mathrm{dark}} + \sum_i \frac{c_i}{\Lambda^{d_i-4}} O^{(i)}_{\mathrm{SM}} O^{(i)}_{\mathrm{dark}}, \tag{2.4}$$

where d_i is the operator dimension, c_i a dimensionless coefficient, and Λ the scale at which the effective description is generated. Interactions with $d_i > 4$ become rapidly weaker at low energy. Renormalisable operators are therefore particularly important because very small couplings can remain observable without requiring a low ultraviolet cutoff.

The Higgs, neutrino, and vector portals allow renormalisable interactions between Standard Model fields and a neutral hidden sector[37]. A scalar singlet can couple to the Higgs bilinear $H^\dagger H$; a neutral fermion can couple through the lepton–Higgs operator LH; and a new Abelian vector field can mix kinetically with hypercharge. These interactions have operator dimension at most four after the corresponding dark-sector fields are included.

The axion portal is also commonly considered, but its coupling to gauge fields usually enters through dimension-five operators. For example, $aF_{\mu\nu}\widetilde{F}^{\mu\nu}$ contains a pseudoscalar field

of dimension one and two field strengths of dimension two, so its coefficient has inverse-mass dimension. This interaction belongs to the higher-dimensional terms in the effective Lagrangian above. DarkSHINE targets the vector portal because an electrically neutral spin-one mediator can be produced by a lepton beam and leave a missing-energy signature if it decays invisibly.

2.2.1 Kinetic mixing and mass eigenstates

Before electroweak symmetry breaking, a new $U(1)_D$ gauge field A'_μ may mix with the hypercharge field B_μ through

$$\mathcal{L}_{\text{mix}} = -\frac{\epsilon_Y}{2} B_{\mu\nu} F'^{\mu\nu}, \qquad F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu. \tag{2.5}$$

The operator has dimension four and is gauge invariant. Even when its coefficient vanishes at tree level, particles carrying both hypercharge and dark charge can generate it radiatively; a small value is thus technically natural[38-39]. After diagonalising the kinetic terms and rotating to the electroweak mass eigenstates, the leading low-energy interaction is

$$\mathcal{L}_{\text{int}} \supset \epsilon e A'_\mu J^\mu_{\text{EM}} + g_D A'_\mu J^\mu_D, \tag{2.6}$$

where J^μ_{EM} and J^μ_D are the electromagnetic and dark currents, e and g_D are their gauge couplings, and $\epsilon \simeq \epsilon_Y \cos\theta_W$ for a mediator well below the Z mass. Equation (2.6) shows why the production rate in an electron or positron experiment is proportional to ϵ^2.

Below the electroweak scale, denote the photon and dark-vector fields before canonical normalisation by $\widehat{A}_\mu$ and $\widehat{X}_\mu$. Their kinetic Lagrangian has the form

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}\widehat{F}_{\mu\nu}\widehat{F}^{\mu\nu} - \frac{1}{4}\widehat{X}_{\mu\nu}\widehat{X}^{\mu\nu} - \frac{\kappa}{2}\widehat{F}_{\mu\nu}\widehat{X}^{\mu\nu}. \tag{2.7}$$

Here κ denotes the photon-component mixing before field normalisation; its magnitude is $|\epsilon_Y \cos\theta_W|$ at leading order in small mixing. The change of variables

$$\widehat{A}_\mu = A_\mu - \kappa \widehat{X}_\mu, \qquad X_\mu = \sqrt{1-\kappa^2}\,\widehat{X}_\mu \tag{2.8}$$

removes the mixed kinetic term. Substitution leaves

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}X_{\mu\nu}X^{\mu\nu}. \tag{2.9}$$

A coupling of $\widehat{A}_\mu$ to the electromagnetic current consequently also contains a coupling of X_μ to the same current, with magnitude $|\kappa|e/\sqrt{1-\kappa^2}$. The overall sign depends on the field and charge conventions. The production probabilities considered here depend on its square; Eq. (2.6) fixes the interaction convention.

This change of basis gives the massive vector a weak coupling to the visible current while retaining its dark-current coupling. The electromagnetic current fixes the relative visible couplings to charged particles. At the sub-GeV scales considered here, the low-energy description uses the physical mediator mass and the small effective mixing ϵ.

For one mediator-emission vertex, the production amplitude contains one factor of ϵe. Squaring the amplitude gives a rate proportional to ϵ^2, with the remaining dependence supplied by the ordinary electromagnetic vertices, phase space, and target form factors. A smaller mixing reduces the number of produced events without changing the on-shell final-state masses. This permits a reference signal sample to determine the detector response at a fixed mediator mass, followed by a change in its overall normalisation. This rescaling also requires fixed decay assumptions.

The dark photon must also acquire a mass. A dark Higgs field can break $U(1)_D$ spontaneously in close analogy with the SM Higgs mechanism. Alternatively, a Stueckelberg mass may be introduced without a physical dark Higgs at the scale of the experiment. For the missing-energy phenomenology studied here, the production and decay kinematics depend primarily on $m_{A'}$, ϵ, the dark coupling $\alpha_D = g_D^2/(4\pi)$, and the masses of kinematically accessible dark states. Details of the mass-generation mechanism enter only if additional dark-sector particles are light enough to be produced.

2.2.2 Visible and invisible mediator decays

For $m_{A'} > 2m_f$, the partial width to a charged SM fermion f is

$$\Gamma(A' \to f\bar{f}) = \frac{N_c}{3}\alpha\epsilon^2 m_{A'}\left(1+\frac{2m_f^2}{m_{A'}^2}\right)\sqrt{1-\frac{4m_f^2}{m_{A'}^2}}, \tag{2.10}$$

where N_c is the colour factor. Hadronic final states require the measured $e^+e^- \to$ hadrons spectral function once the mediator mass passes the hadronic threshold. When no lighter dark state is available, the dark photon decays visibly and can be sought as a narrow invariant-mass resonance or a displaced vertex[40].

The invisible-production benchmark assumes a stable dark particle χ with $m_{A'} > 2m_\chi$. For a Dirac fermion,

$$\Gamma(A' \to \chi\bar{\chi}) = \frac{1}{3}\alpha_D m_{A'} \left(1 + \frac{2m_\chi^2}{m_{A'}^2}\right) \sqrt{1 - \frac{4m_\chi^2}{m_{A'}^2}}. \tag{2.11}$$

When $\alpha_D \gg \alpha\epsilon^2$, the invisible branching fraction is close to unity. The mediator lifetime is then normally short compared with detector scales, but its decay products traverse the apparatus without depositing measurable energy. The experimental signature is consequently inferred from four-momentum imbalance rather than from reconstructing the mediator decay.

The square-root factors in Eqs. (2.10) and (2.11) describe the closing of two-body phase space. Defining $\beta_\chi = \sqrt{1 - 4m_\chi^2/m_{A'}^2}$, the invisible width tends to zero as $\beta_\chi \to 0$, even when the dark coupling is appreciable. Well above threshold, the Dirac-fermion expression tends to $\Gamma_{\rm inv} \simeq \alpha_D m_{A'}/3$. The condition $m_{A'} > 2m_\chi$ specifies whether the channel is open, while the mass ratio determines its phase-space suppression. The relative coupling strengths alone do not fix the branching fraction arbitrarily close to threshold.

With the set of open channels fixed, the total width and invisible branching fraction are

$$\Gamma_{\rm tot} = \Gamma_{\rm inv} + \Gamma_{\rm vis}, \qquad {\rm Br}_{\rm inv} = \frac{\Gamma_{\rm inv}}{\Gamma_{\rm inv} + \Gamma_{\rm vis}}. \tag{2.12}$$

Here $\Gamma_{\rm vis}$ is the sum over allowed visible final states. When $\Gamma_{\rm vis} \ll \Gamma_{\rm inv}$, its first correction to unit invisible branching fraction is ${\rm Br}_{\rm inv} \simeq 1 - \Gamma_{\rm vis}/\Gamma_{\rm inv}$. For charged-lepton and Dirac-dark-matter channels sufficiently far above their thresholds, their width ratio is controlled by $\alpha\epsilon^2/\alpha_D$, apart from the number of available visible channels. This explains the coupling hierarchy assumed in the invisible benchmark.

The width also fixes the proper lifetime, $\tau = 1/\Gamma_{\rm tot}$ in natural units. A mediator of momentum magnitude $p_{A'}$ has mean laboratory decay length

$$\ell_{A'} = \frac{p_{A'}}{m_{A'}} \frac{\hbar c}{\Gamma_{\rm tot}}. \tag{2.13}$$

A visible-decay experiment must also account for the probability of a decay inside its fiducial volume. In an invisible-decay search, the response depends on the visible particles accompanying production and the absence of a detectable interaction by the invisible daughters.

Visible and invisible searches also differ in their coupling dependence. Visible-decay limits depend mainly on the production coupling ϵ, whereas an invisible signal requires both

production and decay into the dark sector. In the on-shell regime and for nearly unit invisible branching fraction, the yield retains the simple ϵ^2 scaling. Off shell, the rate contains the product $\epsilon^2\alpha_D$ and depends more strongly on m_χ. The on-shell benchmark permits accelerator experiments to be compared at fixed decay assumptions. The dimensionless parameter y connects these production limits to thermal freeze-out.

On-shell factorisation expresses the selected invisible rate as production times branching fraction times detector efficiency. Keeping the same efficiency while changing the branching fraction multiplies that rate by $\mathrm{Br}_{\mathrm{inv}}$. This statement assumes that the change does not alter the production spectrum or introduce additional visible particles. A simultaneous change in the masses, width, or final-state content requires revisiting those assumptions.

2.2.3 Thermal target for light dark matter

For $m_{A'} > m_\chi$ and away from the s-channel pole, annihilation of light dark matter into SM fermions has the parametric form

$$\langle\sigma v\rangle_{\chi\bar{\chi}\to f\bar{f}} \propto \frac{\epsilon^2\alpha\alpha_D m_\chi^2}{m_{A'}^4} = \frac{\alpha}{m_\chi^2}\underbrace{\epsilon^2\alpha_D\left(\frac{m_\chi}{m_{A'}}\right)^4}_{y}. \tag{2.14}$$

The variable

$$y \equiv \epsilon^2\alpha_D\left(\frac{m_\chi}{m_{A'}}\right)^4 \tag{2.15}$$

absorbs the leading mediator-mass and coupling dependence. For a specified spin and mass ratio, the observed relic abundance maps to a curve in the (m_χ, y) plane[41-42]. This is the origin of the thermal-target curves used in Chapter 5. Their exact positions depend on whether annihilation is s- or p-wave, whether the dark state is self-conjugate, and whether coannihilation or an inelastic mass splitting is present.

The origin of the mediator-mass factor can be understood from its propagator. If the momentum transfer through the mediator is small compared with $m_{A'}$, the interaction between the two currents has an effective coefficient proportional to $\epsilon e g_D/m_{A'}^2$. Its square contributes $\epsilon^2\alpha\alpha_D/m_{A'}^4$, and the remaining mass dimension in the annihilation cross section is supplied by the dark-matter kinematics. Near the annihilation pole the full propagator must be retained, and this simple scaling is insufficient.

For fixed $r = m_{A'}/m_\chi$ and α_D, the coordinate transformation is

$$m_\chi = \frac{m_{A'}}{r}, \qquad y = \frac{\alpha_D}{r^4}\epsilon^2. \tag{2.16}$$

An experimental bound on mediator production can be expressed in the $(m_{A'}, \epsilon^2)$ plane under its decay assumptions. Its placement relative to a thermal target additionally requires the mass ratio, dark coupling, and dark-matter model.

For an accelerator search, the thermal target becomes harder to reach at lower dark-matter mass because it moves to smaller couplings. Intense beams and strong background rejection are needed to compensate. Accelerator production does not depend on the present Galactic velocity distribution and remains useful when direct-detection recoil energies fall below threshold. Its event rate is expressed through laboratory quantities: integrated luminosity, target thickness, production cross section, acceptance, and efficiency.

2.2.4 Fixed-target kinematic scales

In a fixed-target collision of a beam lepton with energy E_b and an atomic electron approximately at rest, the centre-of-mass energy is

$$s = (p_b + p_e)^2 = 2m_e^2 + 2m_e E_b \simeq 2m_e E_b. \tag{2.17}$$

For $E_b = 8$ GeV, $\sqrt{s} \simeq 90.4$ MeV. This centre-of-mass energy follows from the stationary target and lies in the sub-GeV mass range of interest. Positron annihilation can produce an on-shell mediator when $m_{A'}^2 \simeq s$. Neglecting the atomic-electron momentum gives

$$E_b^{\mathrm{res}} \simeq \frac{m_{A'}^2}{2m_e}, \qquad m_{A'}^{\mathrm{res}} \simeq \sqrt{2m_e E_b}. \tag{2.18}$$

Atomic motion, energy loss in the target, and the mediator width broaden this ideal relation. Nevertheless, it explains both the sensitivity enhancement near 90 MeV for an 8 GeV positron beam and the possibility of scanning mediator mass by changing the beam energy.

For the invisible final state $e^+e^- \to A'^* \to \chi\bar{\chi}$, the dark photon is an internal propagator. The finite-width cross section includes off-shell exchange away from $s = m_{A'}^2$, with the final-state threshold $s \geq 4m_\chi^2$. The mediator pole mass and the invariant mass of the dark-matter pair are separate quantities. An undecayed A' produced as an external particle is described by an on-shell final state.

Taking the beam direction along z, the initial four-vectors are $p_b = (E_b, 0, 0, p_b^z)$ and $p_e = (m_e, 0, 0, 0)$, with $(p_b^z)^2 = E_b^2 - m_e^2$. Their sum gives

$$(E_b + m_e)^2 - (p_b^z)^2 = 2m_e(E_b + m_e), \tag{2.19}$$

which yields Eq. (2.17). Here E_b is the total beam energy; the kinetic energy differs from it by m_e. The centre-of-mass frame moves with

$$\beta_{\rm cm} = \frac{p_b^z}{E_b + m_e}, \qquad \gamma_{\rm cm} = \frac{E_b + m_e}{\sqrt{s}}. \tag{2.20}$$

The large boost explains how a final state with a centre-of-mass mass of order 100 MeV can carry several GeV of laboratory energy.

The exact stationary-electron resonance relation is $E_b^{\rm res} = m_{A'}^2/(2m_e) - m_e$. Its high-energy form appears in Eq. (2.18). For the radiative two-body process, $s \geq m_{A'}^2$ is required because the photon energy must be non-negative. At the endpoint the photon energy tends to zero. A distribution of incident energies or atomic-electron momenta replaces the single value of s by a distribution.

For dark bremsstrahlung, the relevant hard scale is not limited by Eq. (2.17) because the target nucleus absorbs recoil momentum. In the improved Weizsäcker–Williams picture, the nuclear Coulomb field supplies an effective flux of virtual photons and the process is approximated by a hard lepton–photon scattering multiplied by an effective photon flux[43]. The mediator is preferentially emitted in the forward direction and often carries a large fraction $x = E_{A'}/E_b$ of the beam energy. The typical emission angle scales roughly as

$$\theta_{A'} \sim \frac{m_{A'}}{E_b}, \tag{2.21}$$

up to terms involving the lepton mass and $1 - x$. The visible recoil lepton then has reduced momentum, providing the basic tagging variable in a missing-momentum experiment.

The non-resonant process $e^+e^- \to \gamma A'$ is a two-body final state at parton level. In its centre-of-mass frame, the photon energy is

$$E_\gamma^* = \frac{s - m_{A'}^2}{2\sqrt{s}}, \tag{2.22}$$

and boosts into a forward laboratory photon accompanied by missing momentum. This topology differs from both dark bremsstrahlung, which retains a recoil positron, and resonant

annihilation, which can leave no energetic visible final-state particle. The three production mechanisms therefore motivate the orthogonal signal regions developed in Chapter 5.

To obtain Eq. (2.22), take the two final particles back to back in their centre-of-mass frame. With photon momentum magnitude k^*, energy conservation reads

$$\sqrt{s} = k^* + \sqrt{(k^*)^2 + m_{A'}^2}. \tag{2.23}$$

Solving for $k^* = E_\gamma^*$ gives Eq. (2.22). If the photon makes an angle θ_γ^* with the beam direction, its laboratory energy and longitudinal momentum are

$$E_\gamma = \gamma_{\rm cm} E_\gamma^* (1 + \beta_{\rm cm} \cos\theta_\gamma^*), \tag{2.24}$$

$$p_{\gamma,z} = \gamma_{\rm cm} E_\gamma^* (\cos\theta_\gamma^* + \beta_{\rm cm}), \qquad p_{\gamma,T} = E_\gamma^* \sin\theta_\gamma^*. \tag{2.25}$$

The transverse momentum is unchanged by the longitudinal boost. The laboratory photon angle follows from $p_{\gamma,T}/p_{\gamma,z}$, with the sign of the longitudinal momentum retained when identifying its direction. Forward centre-of-mass emission receives the largest energy boost, while backward emission has lower laboratory energy. The photon spectrum therefore depends on the production angular distribution as well as the mass-dependent value of E_γ^*.

These relations also give a kinematic missing-mass estimator when the incident four-momentum and the photon energy and direction are measured. For an invisible remainder $q = p_b + p_e - k_\gamma$,

$$m_{\rm miss}^2 = q^2 = s - 2E_\gamma\big(E_b + m_e - p_b^z \cos\theta_\gamma\big). \tag{2.26}$$

An ideal $\gamma A'$ event has $m_{\rm miss}^2 = m_{A'}^2$. Energy and angle resolutions, knowledge of the beam, and the target-electron approximation all enter this reconstruction.

2.2.5 From production rate to selected yield

For a thin target, the probability of producing a dark photon through channel c is approximately $n_c\sigma_c$, where n_c is the appropriate areal density. The selected yield is

$$N_c = N_{\rm beam}\, n_c\, \sigma_c\, {\rm Br}(A' \to {\rm inv})\, \mathcal{A}_c \varepsilon_c. \tag{2.27}$$

Here $\mathcal{A}_c$ denotes geometric and kinematic acceptance, while ε_c contains reconstruction and selection efficiency within that acceptance. For bremsstrahlung n_c is the areal density of nuclei; for annihilation it is the electron areal density. In Eq. (2.27), theory supplies the cross

section and event kinematics. Detector simulation and analysis then determine the acceptance and efficiency.

The thin-target approximation expresses the normalisation. Transport simulation accounts for the energy loss and multiple scattering before and after a hard interaction; photons can convert or interact photonuclearly; and secondary particles may escape through inactive material. These effects correlate production position with acceptance and background rejection. DarkSHINE therefore evaluates $\mathcal{A}_c\varepsilon_c$ using the full detector model described in Chapters 3 and 4.

For material of density ρ, physical thickness t, molar mass M, and atomic number Z, the two relevant areal densities are

$$n_N = \frac{\rho t N_A}{M}, \qquad n_e = Z n_N. \tag{2.28}$$

The product $\rho t/M$ has units of moles per area, so multiplication by Avogadro's constant gives the number of atoms per area. A cross section quoted per atom must be multiplied by n_N; one quoted per electron must be multiplied by n_e. The integrated luminosity $N_{\text{beam}} n_c$ then has inverse-area units. Its product with a cross section is an expected count before the branching fraction and detector response are applied. A change of scattering-centre convention without the corresponding change of luminosity would alter the predicted yield by a factor of Z.

Acceptance and efficiency are conditional probabilities. Their product can be estimated directly as the fraction of generated events that satisfy the full detector-level selection. Separating them is useful when comparing geometry and reconstruction effects, but requires a common definition of the intermediate accepted sample. In a search with several exclusive categories, each production mechanism has a selected fraction in every category. The observed category contributions must then be summed over production mechanisms, as implemented by the response matrix in Chapter 5.

2.3 Dark-Matter Production and Abundance

2.3.1 Thermal Freeze-out

For a particle species χ in thermal contact with the Standard Model, the number density evolves according to the Boltzmann equation

$$\dot{n}_\chi + 3Hn_\chi = -\langle\sigma v\rangle \left(n_\chi^2 - n_{\chi,\text{eq}}^2\right), \tag{2.29}$$

where H is the Hubble expansion rate and $\langle\sigma v\rangle$ is the thermally averaged annihilation cross section. At early times the interaction rate $\Gamma_\chi = n_{\chi,\mathrm{eq}}\langle\sigma v\rangle$ exceeds H, maintaining chemical equilibrium. Freeze-out occurs near $\Gamma_\chi \simeq H$; afterwards the comoving abundance is approximately conserved and $n_\chi \propto a^{-3}$.

If decoupling occurs while χ is relativistic, its equilibrium density scales as T^3 and its late abundance resembles that of a hot relic. Structure-formation and cosmological measurements restrict such a component to a small fraction of the dark matter. The scenario relevant to conventional weakly interacting massive particles and many light-dark-matter models is non-relativistic freeze-out. In this case,

$$n_\chi \simeq \begin{cases} (m_\chi T)^{3/2}\exp(-m_\chi/T), & T > T_f, \\ n_{\chi,\mathrm{eq}}(T_f)(a_f/a)^3, & T < T_f, \end{cases} \tag{2.30}$$

up to factors that depend on the internal degrees of freedom and the entropy density. A larger annihilation cross section keeps the species in equilibrium longer and therefore produces a smaller relic abundance. This inverse relation defines the thermal target tested by accelerator searches.

The two terms on the left-hand side of the number-density equation describe the time variation of the density in an expanding volume. Multiplying by a^3 gives

$$\frac{\mathrm{d}(n_\chi a^3)}{\mathrm{d}t} = -a^3\langle\sigma v\rangle(n_\chi^2 - n_{\chi,\mathrm{eq}}^2). \tag{2.31}$$

Without annihilation or production, $n_\chi a^3$ is constant. When the density exceeds its equilibrium value, the right-hand side is negative and annihilation reduces the excess. In equilibrium, annihilation and the reverse production reaction balance. The departure from equilibrium becomes appreciable when reactions can no longer follow the changing equilibrium density as the Universe cools. The comparison of Γ_χ with H expresses this competition of time scales.

For a non-relativistic species with negligible chemical potential, the Maxwell–Boltzmann expression is

$$n_{\chi,\mathrm{eq}} = g_\chi\left(\frac{m_\chi T}{2\pi}\right)^{3/2} e^{-m_\chi/T}, \tag{2.32}$$

where g_χ counts the internal states included in the number-density convention. Its exponential decrease makes the time of decoupling important for the remaining abundance. A

quantitative solution also retains the temperature dependence of the annihilation cross section and the thermodynamic degrees of freedom.

It is useful to divide the number density by the entropy density $s_{\rm ent}$, introducing the yield $Y = n_\chi / s_{\rm ent}$. For conserved comoving entropy, $\dot{s}_{\rm ent} + 3Hs_{\rm ent} = 0$, and the density equation becomes

$$\dot{Y} = -s_{\rm ent}\langle\sigma v\rangle(Y^2 - Y_{\rm eq}^2). \tag{2.33}$$

If the entropy degrees of freedom are approximately constant over the interval considered, $T \propto a^{-1}$. The inverse temperature $x = m_\chi / T$ then satisfies $\dot{x} = Hx$, giving

$$\frac{{\rm d}Y}{{\rm d}x} = -\frac{s_{\rm ent}\langle\sigma v\rangle}{Hx}(Y^2 - Y_{\rm eq}^2). \tag{2.34}$$

A change in entropy degrees of freedom modifies the relation between time and x; it must be included when solving across such a transition.

After $Y_{\rm eq}$ becomes negligible, integration gives

$$\frac{1}{Y(\infty)} = \frac{1}{Y(x_f)} + \int_{x_f}^{\infty} \frac{s_{\rm ent}\langle\sigma v\rangle}{Hx}\,{\rm d}x, \tag{2.35}$$

under the same entropy approximation. Increasing the annihilation rate increases the integral and decreases the final yield. The present energy density follows from $\rho_{\chi,0} = m_\chi s_{{\rm ent},0} Y(\infty)$ once the particle is non-relativistic. This relation connects a microscopic interaction rate to an observable cosmological abundance. The value of the integral depends on the velocity dependence of annihilation and the expansion history. A thermal target thus specifies a cosmological history as well as a particle interaction.

Whether n_χ counts one particle species, the sum of particles and antiparticles, or a self-conjugate state changes the counting factors multiplying the collision term. The equation above adopts the normalisation already used for $\langle\sigma v\rangle$ in this section. Numerical thermal curves for different dark-matter models include their own state counting and annihilation channels.

For vector-mediated annihilation, Eq. (2.14) expresses the leading coupling dependence through y, while the integral also depends on phase space, spin, and velocity. Fixing a mass ratio and a dark coupling allows an accelerator production limit to be mapped to this plane; comparing that limit with a relic target then tests the corresponding thermal model[42].

2.4 Experimental Searches for Dark Matter

2.4.1 Direct Detection

Traditional dark matter direct detection experiments search for Weakly Interacting Massive Particles (WIMPs) with masses above 10 GeV. These experiments measure the nuclear recoil energy deposited when dark matter scatters off target nuclei. PandaX-4T has searched for dark-matter–nucleon interactions through a dark mediator[33]. For sub-GeV dark matter, the maximum nuclear recoil energy drops below the detector thresholds, requiring different experimental strategies.

The mass dependence follows from non-relativistic two-body kinematics. For a dark-matter particle of speed v scattering elastically from a stationary target of mass m_T, define the reduced mass $\mu_{\chi T} = m_\chi m_T/(m_\chi + m_T)$. The largest recoil energy is

$$E_R^{\max} = \frac{2\mu_{\chi T}^2 v^2}{m_T}. \tag{2.36}$$

For $m_\chi \ll m_T$, the reduced mass tends to m_χ, so the maximum energy falls quadratically with dark-matter mass. Lowering the detector threshold or changing the target addresses this kinematic suppression. For m_χ much larger than m_T, the reduced mass instead tends to m_T. Bound-electron excitations require the electronic structure of the material in addition to this free-target picture.

If the dark matter belongs to a hidden sector, it is neutral under Standard Model forces but interacts through a new mediator. A standard benchmark is the vector portal, where the interaction between the dark sector and the Standard Model is mediated by a dark photon A'. The Lagrangian includes a kinetic mixing term:

$$\mathcal{L}_V \supset -\frac{\epsilon}{2\cos\theta_W} B_{\mu\nu} F'^{\mu\nu}$$

where ϵ is the kinetic mixing parameter, $B_{\mu\nu}$ is the hypercharge field strength, and $F'^{\mu\nu}$ is the dark photon field strength. Theoretical models motivated by dark matter production mechanisms provide specific cross section targets. These models include elastic scalar dark matter, asymmetric fermions, Elastically Decoupling Relics (ELDER), Strongly Interacting Massive Particles (SIMP), and Majorana fermions.

For a mediator lighter than 1 keV, the cross section for dark matter scattering is enhanced at low momentum transfers q. The direct detection cross section is parameterized as:

$$\sigma_{DD} \sim 4\pi\alpha_D\epsilon^2\alpha\frac{\mu_{\chi,e}^2}{q^4}$$

where α_D is the dark fine structure constant, α is the electromagnetic fine structure constant, and $\mu_{\chi,e}$ is the reduced mass of the dark matter and electron system.

To probe this parameter space, experiments measure electron recoils instead of nuclear recoils. For dark matter masses between 1 MeV and 1 GeV, scattering off an electron extracts a larger fraction of the dark matter kinetic energy. Experiments use semiconductors like silicon or germanium. Dark matter collisions excite electrons from the valence band to the conduction band, producing measurable ionization signals. The threshold for this process is around 1 eV. For dark matter masses down to 1 keV, superconducting aluminum targets are proposed. The binding energy of Cooper pairs is on the order of 1 meV, allowing sensitivity to much smaller energy depositions.

For nuclear recoils, using light target nuclei improves kinematic matching. Superfluid helium targets can detect multi-phonon or roton excitations. The energy threshold for these excitations is around 1 meV, enabling sensitivity to dark matter masses down to 1 MeV. Bubble chambers using fluorine-based liquids search for spin-dependent couplings between dark matter and protons.

2.4.2 Searches for ultralight fields

Laboratory searches also probe ultralight bosonic dark matter with mass ranging from 10^{-22} eV to 1 keV. A primary candidate in this mass range is the QCD axion, which arises from the Peccei-Quinn mechanism to solve the strong CP problem. The Peccei-Quinn symmetry is broken at a high energy scale f_a, and the axion obtains a mass:

$$m_a \approx \frac{\Lambda_{QCD}^2}{f_a}$$

Due to the small mass, the number density of these particles in the galactic halo is high. The dark matter acts as a classical coherent oscillating field rather than individual particles. The field oscillates at a frequency determined by its mass, $\omega = m_a/\hbar$.

Experiments search for continuous-wave signals from these macroscopic fields. Haloscope experiments use microwave cavity resonators placed inside strong magnetic fields. The axion field interacts with the magnetic field to produce microwave photons. The cavity

enhances the signal when its resonant frequency matches the axion mass.

For lower axion masses, the corresponding electromagnetic wavelength exceeds the practical size of a cavity. Experiments use lumped element LC circuits and toroidal magnets. The oscillating dark matter field acts as an effective alternating current, inducing a magnetic flux that is read out by sensitive magnetometers like SQUIDs.

Other search methods target different couplings of the ultralight dark matter. Nuclear Magnetic Resonance techniques search for the axion-gluon coupling. The dark matter field induces an oscillating electric dipole moment in nucleons. When an external magnetic field matches the nuclear spin Larmor frequency to the axion Compton frequency, the transverse magnetization is amplified and measured. Atom interferometers and torsion balances measure the time-varying differential accelerations or spin precessions caused by scalar or vector ultralight dark matter.

2.4.3 Accelerator Production

Accelerator experiments produce dark matter relativistically from collisions of Standard Model particles. This approach does not depend on the local galactic halo velocity distribution and is not limited by kinematic thresholds that affect direct detection. It is suited for testing the thermal relic target for sub-GeV dark matter.

In the thermal freeze-out scenario, dark matter depletes its abundance in the early universe by annihilating into Standard Model particles. For scalar dark matter annihilating into leptons through a dark photon mediator, the process depends on a dimensionless variable y:

$$y \equiv \varepsilon^2 \alpha_D \left(\frac{m_\chi}{m_{A'}}\right)^4, \tag{2.37}$$

where $\alpha_D = g_D^2/(4\pi)$ is the dark-sector coupling and ε is the kinetic-mixing parameter. For fixed mass and spin assumptions, the observed relic density selects a target band in the (y, m_χ) plane that can be tested by accelerator experiments.

2.4.4 Dark-Photon Production in Lepton Fixed-Target Experiments

After diagonalising the kinetic terms, the dark photon couples to the electromagnetic current with strength εe. A minimal vector-portal Lagrangian may be written as

$$\mathcal{L} \supset -\frac{1}{4} F'_{\mu\nu} F'^{\mu\nu} + \frac{1}{2} m_{A'}^2 A'_\mu A'^\mu - \frac{\varepsilon}{2} F'_{\mu\nu} F^{\mu\nu} + g_D A'_\mu J_D^\mu. \tag{2.38}$$

When $m_{A'} > 2m_\chi$ and the dark-sector coupling dominates over the visible couplings, A' decays predominantly to $\chi\bar{\chi}$ and appears as missing energy.

Three production mechanisms are relevant to this thesis. Dark bremsstrahlung,

$$e^\pm N \to e^\pm N A', \tag{2.39}$$

is available to both electron and positron beams[43]. A positron beam additionally permits non-resonant annihilation,

$$e^+e^- \to \gamma A', \tag{2.40}$$

and resonant annihilation, $e^+e^- \to A'$, on atomic electrons. For an electron approximately at rest, the resonant condition is

$$m_{A'} \simeq \sqrt{2m_e E_{e^+}}, \tag{2.41}$$

which corresponds to about 90 MeV for an 8 GeV positron beam. The distinct visible final states of these processes motivate the three orthogonal signal regions developed in Chapter 5.

Missing mass experiments collide a high-intensity electron or positron beam with a fixed target. Dark photons are produced via processes like $e^+e^- \to \gamma A'$. The mass of the dark photon is reconstructed entirely from the kinematics of the visible outgoing photon. This method requires precise initial state knowledge and does not rely on the dark matter interacting in the detector.

Missing momentum and missing energy experiments measure an electron scattering off a nucleus. The electron radiates a dark photon that decays into dark matter. The detector tracks the scattered electron and measures the total energy deposited. A signal event is characterized by a single low-energy scattered electron and substantial missing energy carried away by the dark matter. This approach requires strict detector hermeticity to veto Standard Model background events.

Beam dump experiments direct a high-intensity proton or electron beam into a dense block of material. The collisions produce dark photons, either directly or through the decay of secondary mesons like neutral pions. The dark photons decay into dark matter particles, which travel through thick shielding that absorbs all Standard Model radiation except neutrinos. The dark matter is detected when it scatters off electrons or nucleons in a downstream detector.

2.4.5 Measurement strategies and their interpretation

The choice of measured quantity determines which parts of an interaction must be resolved. A missing-energy measurement compares a known incident energy with the visible energy recorded in the detector. It is sensitive to invisible production only if ordinary energy leakage and unobserved Standard Model particles are sufficiently suppressed. A missing-momentum measurement additionally resolves the incident and recoil trajectories. It can use their directions and momentum magnitudes to test whether a low-energy recoil is compatible with the expected signal kinematics. Both approaches require vetoes for visible particles that could carry away the apparent deficit.

The NA64 electron-beam search uses a calorimetric active target for the missing-energy measurement[34]. Its absorber provides both interaction material and a measurement of the visible electromagnetic shower. In the thin-target approach used by DarkSHINE and proposed for LDMX[35], the target and the downstream calorimeter are separate. This arrangement allows a recoil trajectory to be measured before the particle enters the calorimeter.

Missing-mass measurements use a more constrained final state. In the stationary-electron approximation, Eq. (2.26) uses the incident positron and the measured photon to reconstruct the invariant mass of everything else. The method can search for a localised mass feature without detecting the invisible decay products. The photon direction enters explicitly alongside its energy. Angular acceptance, beam knowledge, and control of extra photons are therefore part of the measurement definition.

Colliding-beam experiments also access invisible final states with a measured photon. The BaBar search for $e^+e^- \to \gamma A'$ uses the single-photon signature and the known collision kinematics[36]. The fixed-target and collider versions of this reaction share a production topology, but have different centre-of-mass energies and boosts. Their useful mediator-mass ranges and photon acceptance follow from those conditions.

In a beam-dump scattering search, the production interaction and the detection interaction occur in different places. The expected event count contains a production factor, the probability that dark particles reach the downstream detector, and a scattering probability in its active material. Its coupling dependence can therefore differ from that of an invisible-production search, which infers the new particles from the recoil alone. The dark-sector coupling and mediator propagator may enter the detection probability as well as the production and decay

model.

Table 2.1 Observables and detector information used by accelerator search strategies.

Strategy	Measured information	Main experimental dependence
Missing energy	Incident energy and visible deposited energy	Calorimeter response, containment, and vetoes for escaping particles
Missing momentum	Incident and recoil momenta, with visible-energy vetoes	Track reconstruction, material effects, angular acceptance, and veto response
Missing mass	Initial four-momentum and a specified visible final state	Beam definition and visible-particle energy and angle resolution
Beam-dump scattering	Energy or recoil from a downstream dark-particle interaction	Production spectrum, transport to the detector, and scattering probability

Exposure is also defined differently across these strategies. Fixed-target experiments often quote the number of incident beam particles, whereas a collider quotes integrated collision luminosity. For a thin fixed target, the conversion requires the areal density of the scattering centres. A thick target also develops a shower, so the incident particle count alone does not describe the energy and species distributions that can produce the mediator. The electron and positron modes of the same apparatus can be compared at a matched beam count and target, but their channel cross sections must still use the appropriate nuclear or electronic density.

Chapter 3 The DarkSHINE Experiment and Detector System

The detector requirements follow from two different measurements. The tracking system compares the incident and outgoing charged-particle momenta, while the calorimeters measure the visible activity downstream of the target. A low recoil momentum supplies evidence that the beam particle has transferred a large fraction of its energy. The calorimeter veto tests whether that energy has instead appeared in photons, muons, or hadrons.

Table 3.1 Connection between the missing-energy measurement and the functions of the DarkSHINE detector subsystems.

Subsystem	Measurement	Use in the event interpretation
Tagging tracker	Incident charged-particle track and momentum	Defines the incoming particle used in the momentum comparison
Target	Known material and interaction region	Relates beam exposure and production cross section to the number of produced events
Recoil tracker	Outgoing charged-particle multiplicity and momentum	Identifies the recoil topology and measures the momentum transferred from the beam
ECAL	Electromagnetic energy and its spatial distribution	Measures the recoil electron or photon and rejects events with large visible electromagnetic energy
HCAL	Hadronic and penetrating-particle activity	Vetoes visible energy that is not adequately measured by the ECAL

3.1 Experimental Setup and Beamline

The DarkSHINE beamline extracts single electrons from the SHINE linear accelerator[44]. The accelerator operates with a 1.3 GHz microwave, providing 1.3×10^9 radio-frequency buckets per second. In standard FEL operation, a 100 pC electron bunch is loaded into one bucket out of every 1300, corresponding to a 1 MHz bunch rate. For the DarkSHINE branch, an additional low-energy laser system is used to produce a single-electron bunch; one electron is injected into each otherwise empty bucket and accelerated to 8 GeV. A dedicated kicker system[45] then distributes these single-electron bunches to the DarkSHINE beamline, sepa-

rate from the FEL branches (FEL-I, II, III). SHINE has a designed repetition rate of 1 MHz. The sensitivity studies in this thesis assume an integrated exposure of 3×10^{14} electrons on target (EOT) over ten years. The same ten-year beam count is used for positron running, with positrons on target (POT) replacing EOT.

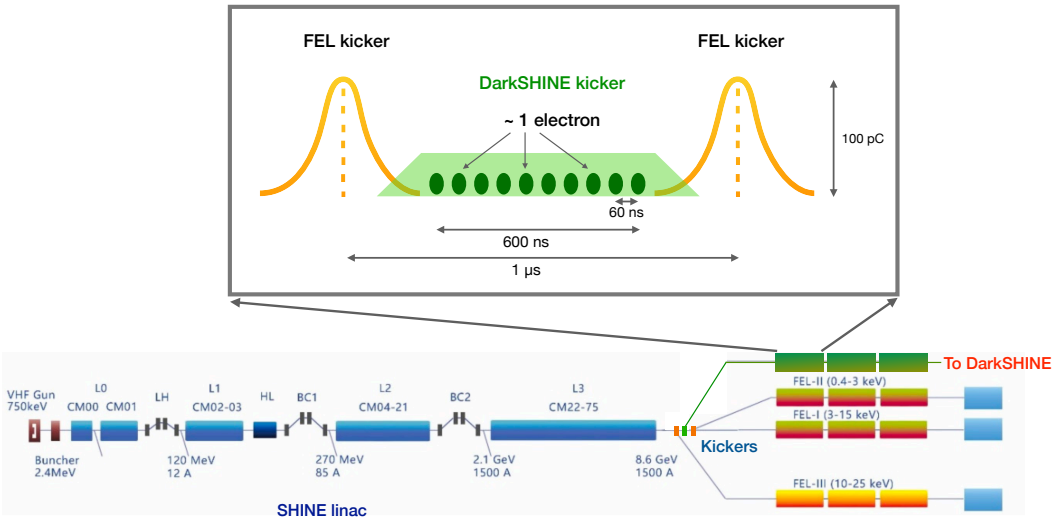


Figure 3.1 Illustration of the SHINE linac[44] and its extraction kickers[45]. The FEL kickers direct the 100 pC bunches to the light-source branches; the dedicated DarkSHINE kicker extracts the single-electron bunches to the fixed-target beamline.

The overall detector layout is shown in Fig. 3.2. Along the beam direction the main subsystems are, in order: the dipole magnet housing the tagging tracker, the tungsten target, the recoil tracker at the downstream edge of the magnet, the electromagnetic calorimeter (ECAL), and the hadronic calorimeter (HCAL). The tagging tracker reconstructs the momentum of each incident electron before it strikes the target; the recoil tracker and ECAL jointly measure the momentum and energy of the scattered electron; and the HCAL acts as a veto for muons and neutral hadrons that could otherwise mimic a missing-energy signal.

3.2 Magnetic Field System

The magnetic field bends incident electrons in the tagging region for precise momentum measurement and supplies sufficient bending power in the recoil region while retaining acceptance for low-momentum electrons ($p \sim 30$–50 MeV) that are kinematically characteristic

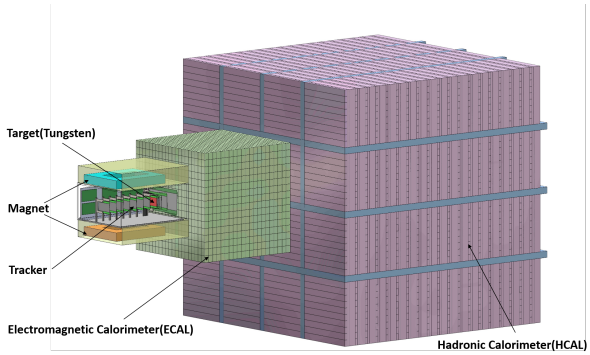


Figure 3.2 Schematic view of the DarkSHINE detector. Electrons enter from the left. The dipole magnet (yellow/cyan) encloses the tagging tracker; the recoil tracker sits at the magnet exit (1.5 T region); the tungsten target is interposed between the two trackers; ECAL follows the recoil tracker; HCAL closes the downstream end.

of dark-photon signal events.

These two requirements translate into different field profiles for the two tracker volumes. The tagging tracker demands a highly uniform 1.5 T field to ensure unambiguous track curvature measurement. The recoil tracker, by contrast, benefits from a deliberately non-uniform field that falls from 1.5 T at the upstream end to approximately 0.75 T at the downstream end; the decreasing field prevents low-momentum recoil electrons from curling before they reach the tracking layers. Beyond the recoil tracker the field drops gradually to zero inside the ECAL volume.

Square Helmholtz coils positioned above and below the tracker support box are adopted to realise this configuration (Fig. 3.3). The geometry produces a uniform field at the centre of the coil pair (tagging region) while the natural fall-off at the coil edges creates the desired gradient across the recoil tracker. Simulated field maps confirming this behaviour are shown in Fig. 3.4.

For a uniform field, the momentum component perpendicular to the field and the curvature radius obey[46]

$$p_{\perp B}\,[\mathrm{GeV}/c] \simeq 0.3\,|q/e|\,B\,[\mathrm{T}]\,R\,[\mathrm{m}]. \tag{3.1}$$

Here $p_{\perp B}$ is the momentum component perpendicular to the field. The relation explains the different curvature of the incident beam and soft recoils at a common field strength. It applies locally as a uniform-field description; trajectories in the mapped field require transport through the spatially varying field.

The measured curvature also has contributions from finite hit precision and multiple scattering. In the uniform-field approximation these can be expressed as separate contributions to the curvature variance[46],

$$\sigma_k^2 = \sigma_{k,\text{measurement}}^2 + \sigma_{k,\text{scattering}}^2, \qquad k = 1/R. \tag{3.2}$$

The sensor precision controls the measurement term. The material traversed by the particle contributes to the scattering term. Improving the precision of a single hit consequently addresses only one part of the momentum uncertainty.

The longitudinal field profile describes how the bending environment changes from one subsystem to the next. The two-dimensional map resolves its variation across the tracking aperture and along the beam. A trajectory samples the field along its own path, so the momentum reconstruction must use the field prescription appropriate to that geometry.

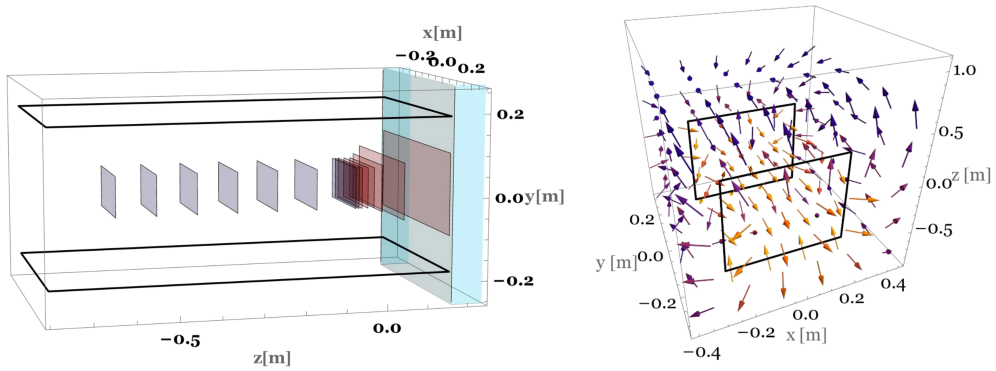


Figure 3.3 Left: tracker support assembly inside the magnet bore. Right: magnetic-field vectors generated by the square Helmholtz coils. The central region provides the nearly uniform tagging field, while the recoil tracker occupies the falling field near the coil edge.

3.3 Tracking System

3.3.1 Requirements and Sensor Technology

Momentum reconstruction for dark-photon searches requires measuring the vector difference between the incoming and outgoing electron momenta. This places tight constraints on both the spatial resolution and the momentum acceptance of the tracking detectors. A position resolution better than 10 μm is needed to achieve the required momentum precision,

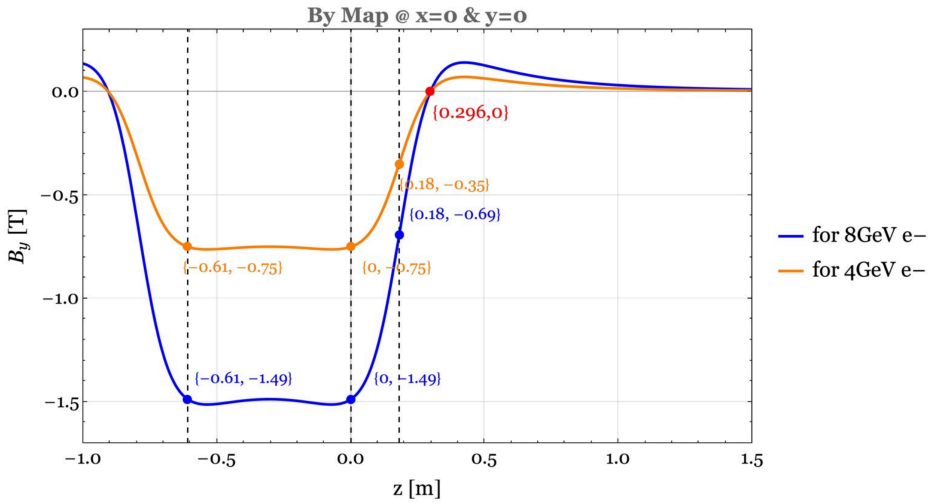


Figure 3.4 Magnetic field in the DarkSHINE tracker. Simulated field component B_y along the beam axis at $x = y = 0$, for the 8 GeV (blue) and 4 GeV (orange) field settings.

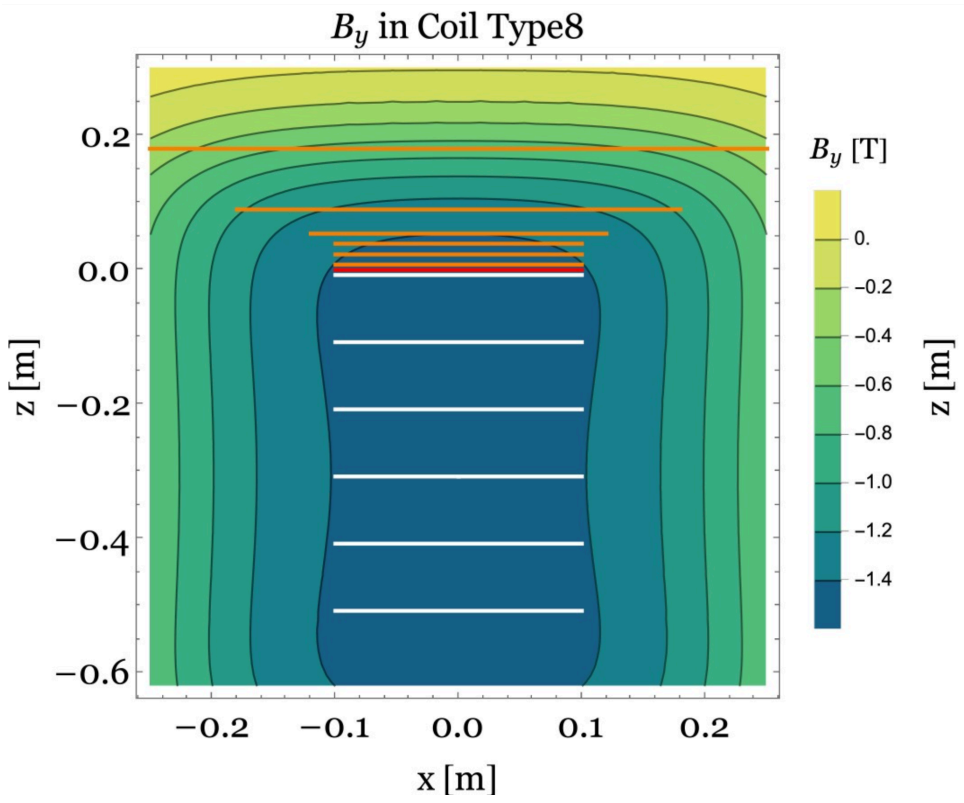


Figure 3.4 (continued) Magnetic field in the DarkSHINE tracker. B_y in the x–z plane for the 8 GeV setting, with tracker planes overlaid. The tagging region has a nearly uniform field, which decreases across the recoil tracker.

and the recoil tracker must maintain high acceptance over a broad momentum range from ~30 MeV up to ~2 GeV.

AC-LGAD (AC-coupled Low-Gain Avalanche Detector) silicon strip technology is chosen to satisfy these requirements. The internal gain of AC-LGAD sensors improves the signal-to-noise ratio without the full breakdown characteristic of avalanche photodiodes, while the AC-coupling scheme allows fine pitch with reduced inter-strip capacitance, enabling sub-10 μm spatial resolution together with nanosecond-level time resolution.

The tracking system layout is illustrated in Fig. 3.5, and the key geometric parameters of the two tracker volumes are listed in Table 3.2.

The rotated strip planes provide information in the second transverse coordinate. Each strip measurement constrains the hit position most precisely across the strip direction. Combining measurements from planes at different orientations resolves the position ambiguity along a single strip. The layout in Fig. 3.5 consequently connects the sensor's local spatial resolution to a three-dimensional trajectory. Track finding and fitting then use these measurements as described in Section 4.3.

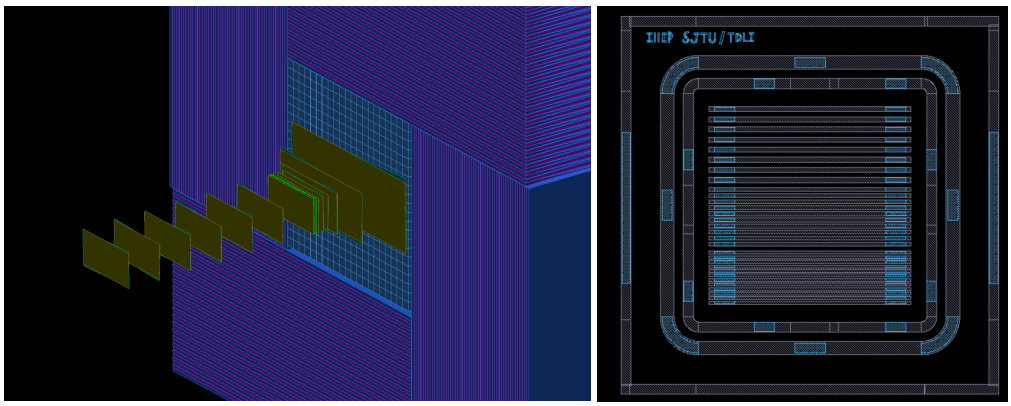


Figure 3.5 Left: Layout of the DarkSHINE tracking system showing the seven-layer tagging tracker, the tungsten target, and the six-layer recoil tracker. Adjacent layers within each station are rotated by 100 mrad to provide two-dimensional hit information. Right: AC-LGAD strip-sensor design.

3.3.2 Tagging Tracker

The tagging tracker reconstructs the full momentum vector of each 8 GeV beam electron before it reaches the target. Figure 3.6 shows the layer-number scan used to select the seven-layer configuration. Reconstruction efficiency reaches approximately 99.8% with five layers

Table 3.2 Key geometric parameters of the tracking detectors.

Node	Centre (mm)	Size (mm)			Arrangement	Comments
	z	x	y	z		
Tagging tracker	−307.8	200	400	600.2	7 layers	2nd layer: 0.1 rad rotation
Target	0	100	200	0.35	—	
Recoil tracker	+94.0	500	800	172.7	6 layers	2nd layer: 0.1 rad rotation

and changes little thereafter. The relative momentum resolution improves from 2.52% with four layers to 0.87% with seven layers. Seven layers thus provide sub-percent momentum resolution while retaining a reconstruction efficiency of 99.81%. Beyond seven layers, the resolution gain per additional layer is smaller.

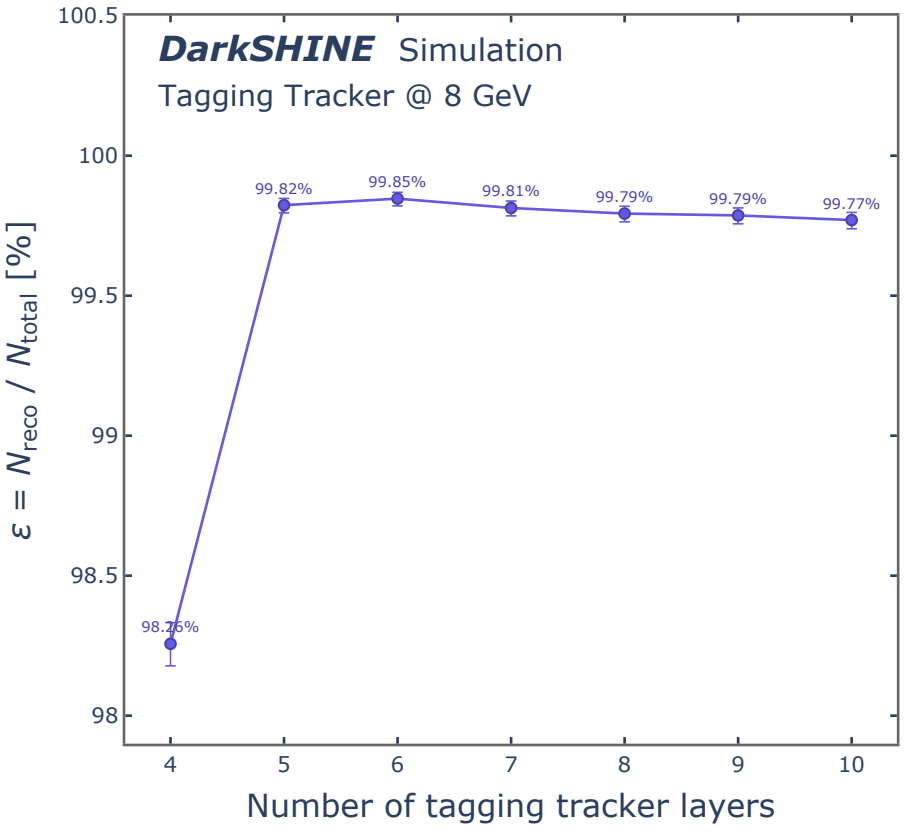


Figure 3.6 Tagging-tracker layer-number scan for 8 GeV electrons. Reconstruction efficiency $\varepsilon = N_{\mathrm{reco}}/N_{\mathrm{total}}$, where the numerator and denominator are the reconstructed and total counts. Seven layers give 99.81% efficiency.

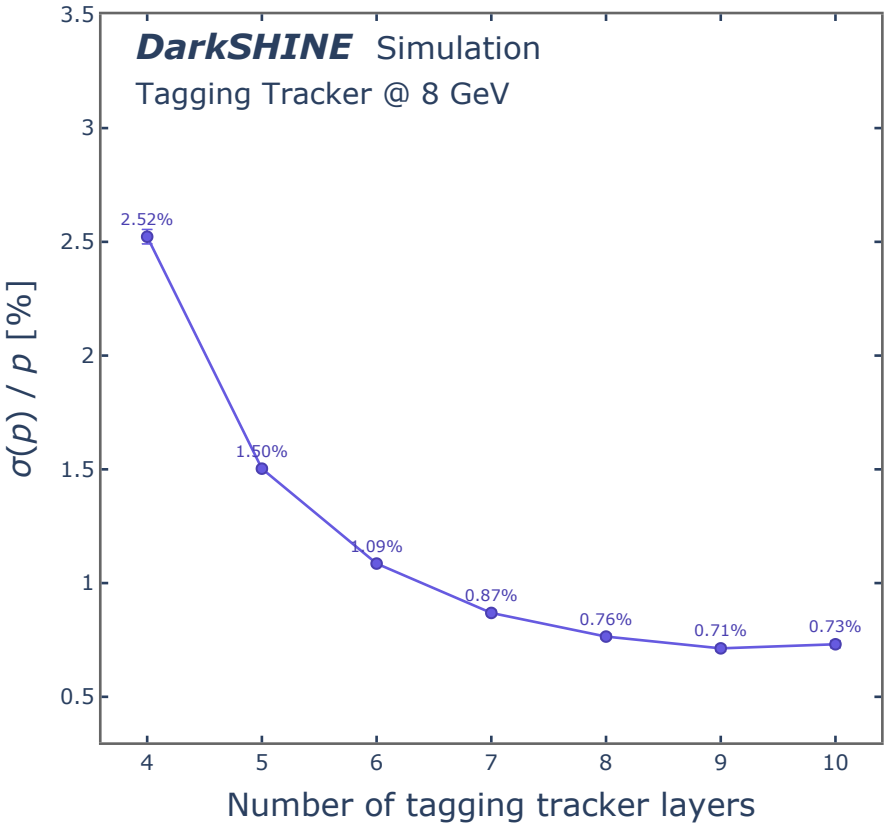


Figure 3.6 (continued) Tagging-tracker layer-number scan for 8 GeV electrons. Relative momentum resolution $\sigma(p)/p$, where p is the electron momentum. Seven layers give 0.87% resolution; further layers yield smaller improvements.

3.3.3 Recoil Tracker

After the target, the recoil tracker records the trajectories of scattered charged particles. Signal electrons typically emerge with momenta between 30 MeV and 2 GeV, whereas background events produce heavier charged particles ($\mu^{\pm}$, $\pi^{\pm}/K^{\pm}$) whose different rigidity and dE/dx allow partial discrimination. Six detector layers are deployed within the non-uniform field region, with their positions and dimensions optimised to maximise acceptance for the softest recoil electrons (Fig. 3.7).

The recoil-tracker scans address the lower end of the accepted momentum range. Their 30 MeV electron sample probes trajectories with substantial bending, for which the transverse dimensions and longitudinal positions of the layers affect whether enough measurements remain available. The layer-number scan and the length scan vary different geometric properties. Adding a layer increases the number of potential measurements, while changing the overall length also moves the measurement surfaces along the trajectory.

The ACTS framework is used for track reconstruction. In the comparison in Fig. 3.8, it outperforms the preliminary Kalman filter developed for DarkSHINE. The momentum resolution achieved by ACTS is shown as a function of recoil electron momentum.

3.3.4 Sensor Prototype Performance

The laser scans characterise the timing response of the sensor prototypes, giving time resolutions of 6.5–8.2 ps for wafer-11 and 8.8–12.3 ps for wafer-12[47]. The fabrication parameters, electrical scans, and laser measurements are discussed together in Section 3.7.1.

3.4 Tungsten Target

The target converts a fraction of the beam electrons into dark photons through a process analogous to bremsstrahlung. The expected signal yield N_{sig} is

$$N_{\mathrm{sig}} = \sigma_{A'} \times \rho \times 0.1X_0 \times L \times \frac{N_A}{M_{\mathrm{target}}} \times 10^{-36} \times \varepsilon^2, \tag{3.3}$$

where $\sigma_{A'}$ is the dark-photon production cross section for mass $m_{A'}$ after division by ε^2, expressed in pb, ρ is the target density, $0.1X_0$ is the target thickness expressed in units of the radiation length X_0, $L = 3 \times 10^{14}$ is the ten-year EOT, N_A is Avogadro's number, M_{target} is

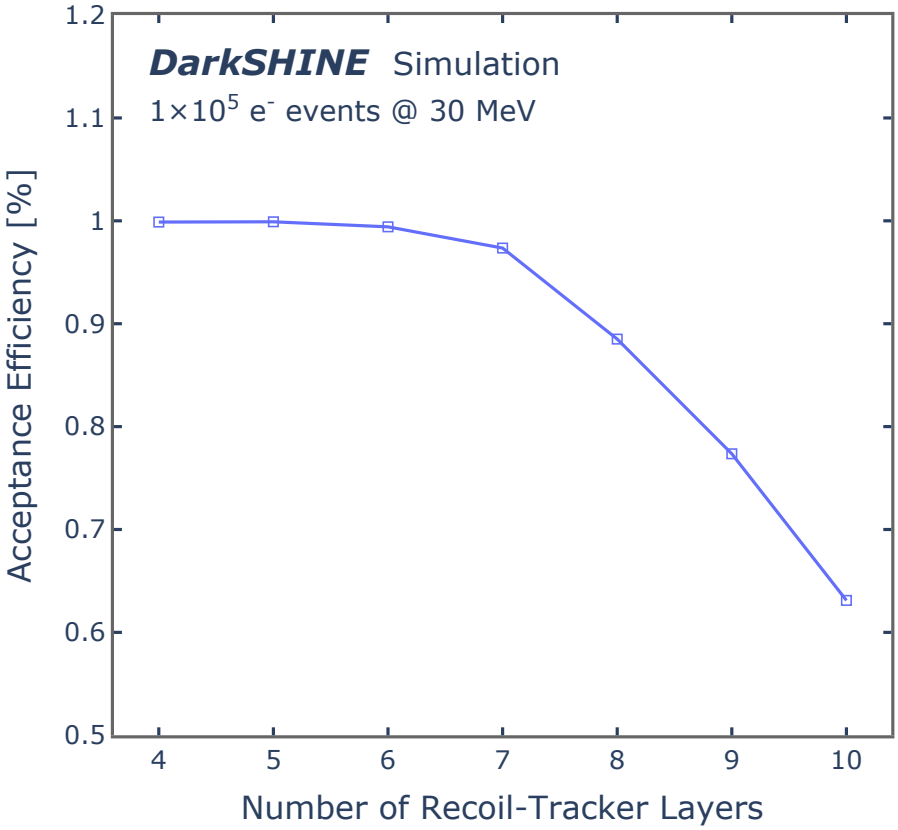


Figure 3.7 Recoil-tracker acceptance efficiency for simulated 30 MeV electrons, using 10^5 incident electrons in each scan. Variation with the number of layers, probing acceptance for strongly bent low-momentum trajectories.

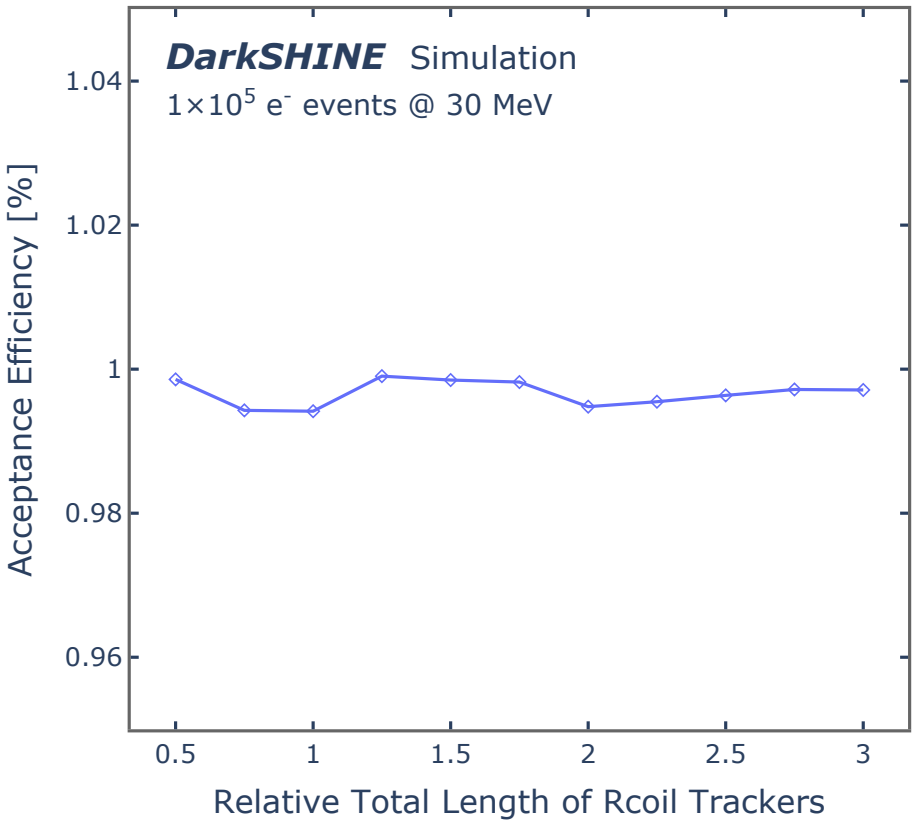


Figure 3.7 (continued) Recoil-tracker acceptance efficiency for simulated 30 MeV electrons, using 10^5 incident electrons in each scan. Variation with the total tracker length divided by its baseline value, probing acceptance for strongly bent low-momentum trajectories.

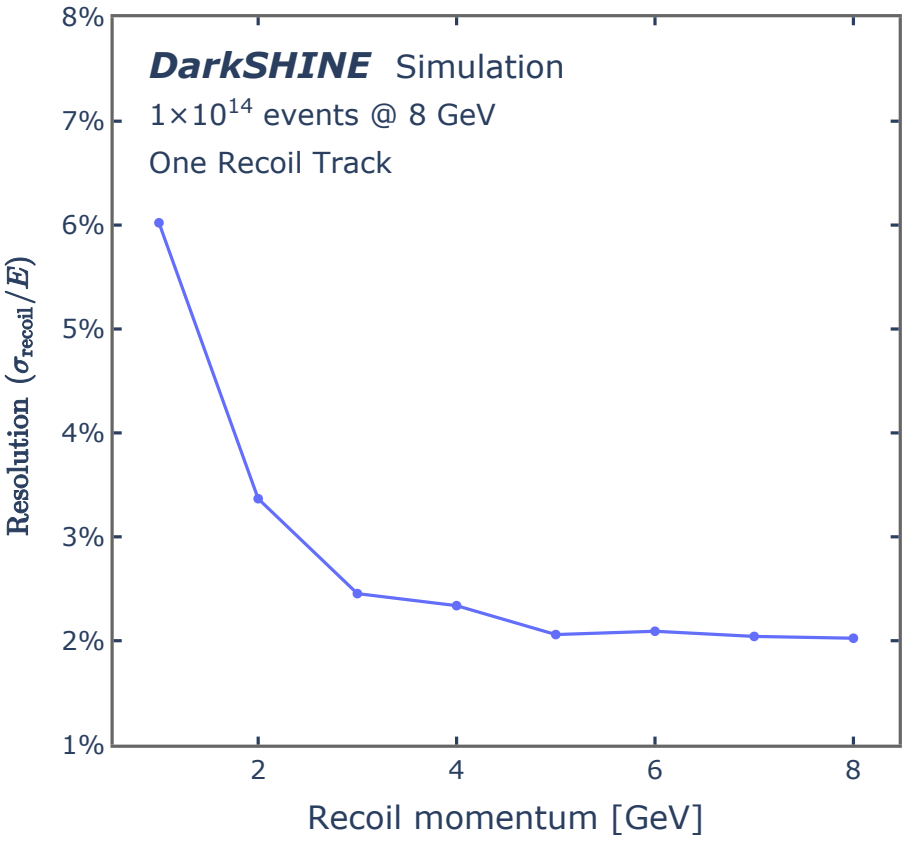


Figure 3.8 Recoil-tracker momentum reconstruction. Relative momentum resolution versus electron momentum in the 8 GeV beam sample with one recoil track.

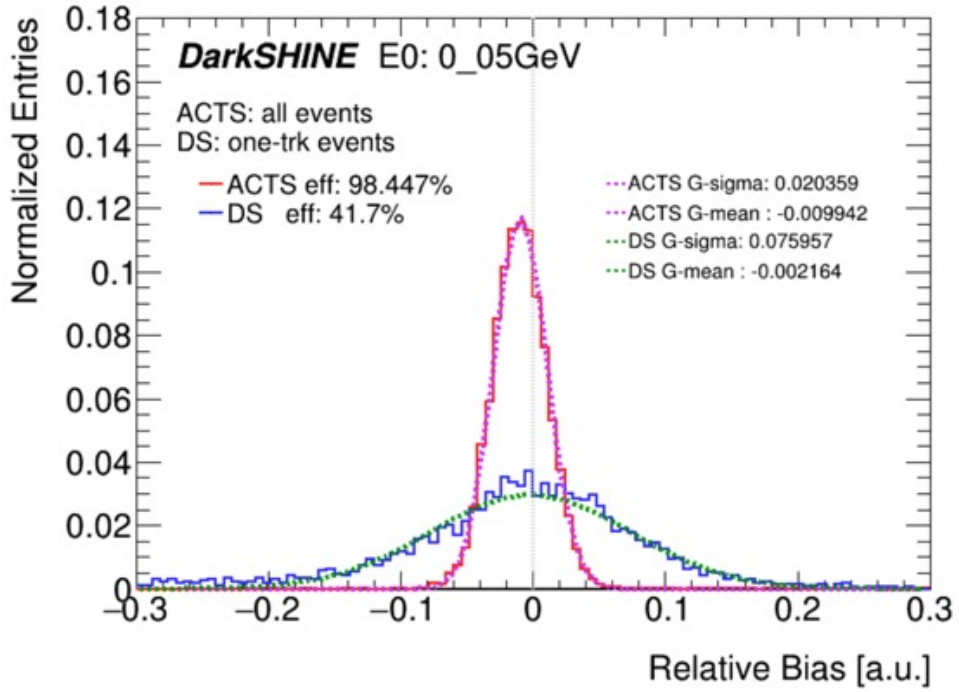


Figure 3.8 (continued) Recoil-tracker momentum reconstruction. Comparison of ACTS with the preliminary DarkSHINE Kalman filter. ACTS gives the smaller momentum spread in the displayed comparison.

the atomic mass, and ε is the kinetic mixing parameter.

The physical cross section is consequently $\varepsilon^2\sigma_{A'}$. The factor 10^{-36} converts its numerical value in pb to cm^2, matching the target density and thickness expressed in centimetre-based units. The factor L in this expression is the incident beam count. After multiplication by the nuclear areal density, it gives the target luminosity in inverse square centimetres. This convention keeps the coupling dependence explicit when yields at different values of ε are compared.

The material factors in Eq. (3.3) can be grouped into the areal density of target nuclei. Writing the physical target thickness as t,

$$n_N = \frac{\rho t N_A}{M_{\text{target}}}, \qquad t = 0.1X_0, \tag{3.4}$$

gives n_N in nuclei per unit area when consistent length and mass units are used. The product of areal density and the physical production cross section is the thin-target interaction probability. Multiplication by the incident beam count then gives the number produced before detector acceptance and selection. The corresponding selected yield includes the response

factors developed in Section 5.3.

For comparisons at a fixed fraction of a radiation length, both the physical thickness and the number of nuclei per unit area change with material. A cross-section ratio per nucleus consequently differs from a yield ratio at fixed beam exposure. Figure 3.9 displays these two questions separately. The yield comparison includes the target material through Eq. (3.4).

For a neutral target, the electron areal density is $n_e = Zn_N$. This distinction becomes necessary in positron running: nuclear dark bremsstrahlung is normalised per target nucleus, whereas annihilation is normalised per atomic electron. The same physical foil and beam count therefore define different target luminosities for these processes. The geometrical target thickness remains common to their transport simulation.

The dependence of signal yield on the choice of target material was evaluated by comparing tungsten (W), lead (Pb), aluminium (Al), and beryllium (Be), whose properties are summarised in Table 3.3. Cross-section and signal-yield ratios relative to tungsten are shown in Fig. 3.9. The Pb/W signal-yield ratio is close to unity across the full mass range, while Al and Be show increased yields only at high $m_{A'}$.

Table 3.3 Physical parameters of candidate target materials[46].

Material	Atomic number Z	Atomic mass M_{target} [g mol^{-1}]	Radiation length X_0 [cm]	Density ρ [g cm^{-3}]
W	74	184	0.35	19.30
Pb	82	207	0.56	11.35
Al	13	27	8.90	2.70
Be	4	9	35.28	1.85

In most EOT events the beam electron traverses the target without interaction. However, a small fraction undergo hard bremsstrahlung, photonuclear reactions ($\gamma N \rightarrow$ hadrons), electron-nuclear reactions, or $\gamma \rightarrow \mu^+\mu^-$ conversion. Although each branching ratio is small, the enormous accumulated EOT statistics make these rare processes the dominant backgrounds. Cross sections per atom for rare processes are compared in Fig. 3.10; after weighting by density and radiation length, the Al and Be target options yield electron-nuclear backgrounds larger than tungsten by factors of 2.5 and 5.3, respectively, and photonuclear backgrounds by factors of 4 and 11.7.

An additional class of irreducible backgrounds arises from processes that produce neutrinos in the final state: Møller + charged-current quasi-elastic scattering (CCQE), neutrino

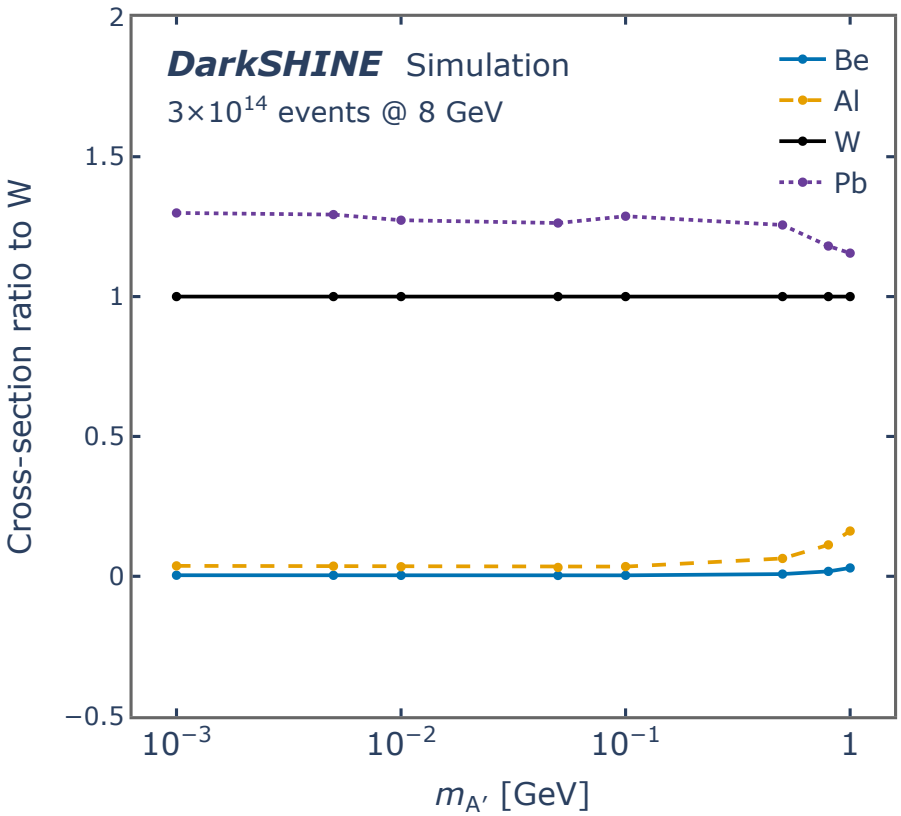


Figure 3.9 Dark-photon production for an 8 GeV electron beam in Be, Al, W, and Pb, normalised to tungsten. Cross-section ratio per atom versus dark-photon mass $m_{A'}$.

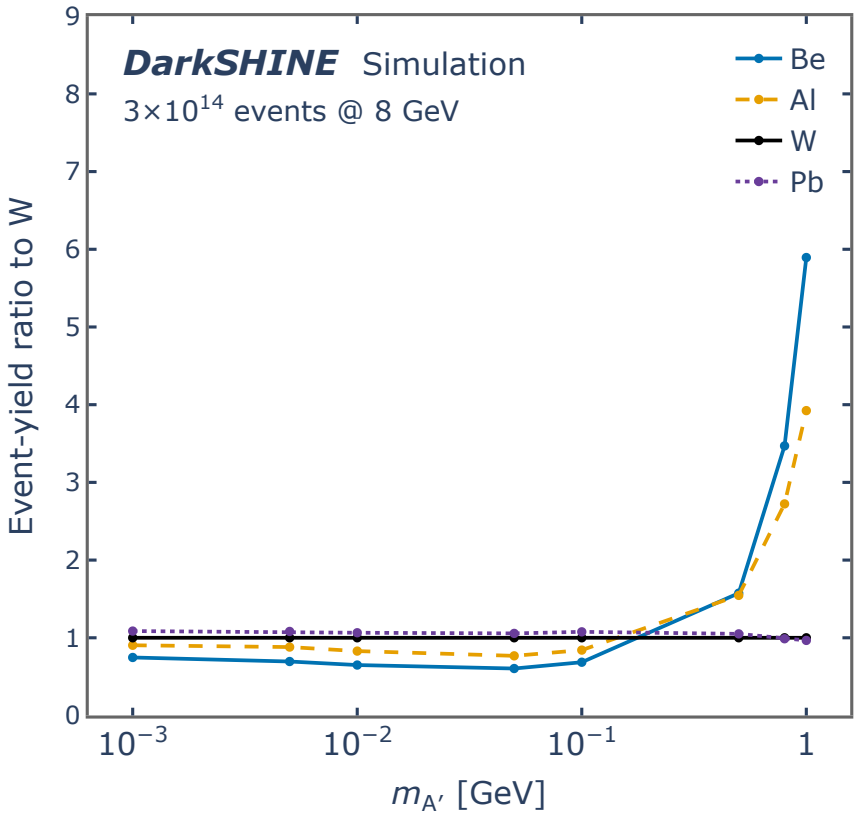


Figure 3.9 (continued) Dark-photon production for an 8 GeV electron beam in Be, Al, W, and Pb, normalised to tungsten. Signal-yield ratio for a $0.1X_0$ target and 3×10^{14} EOT at $\varepsilon = 1$, including the material-dependent nuclear areal density.

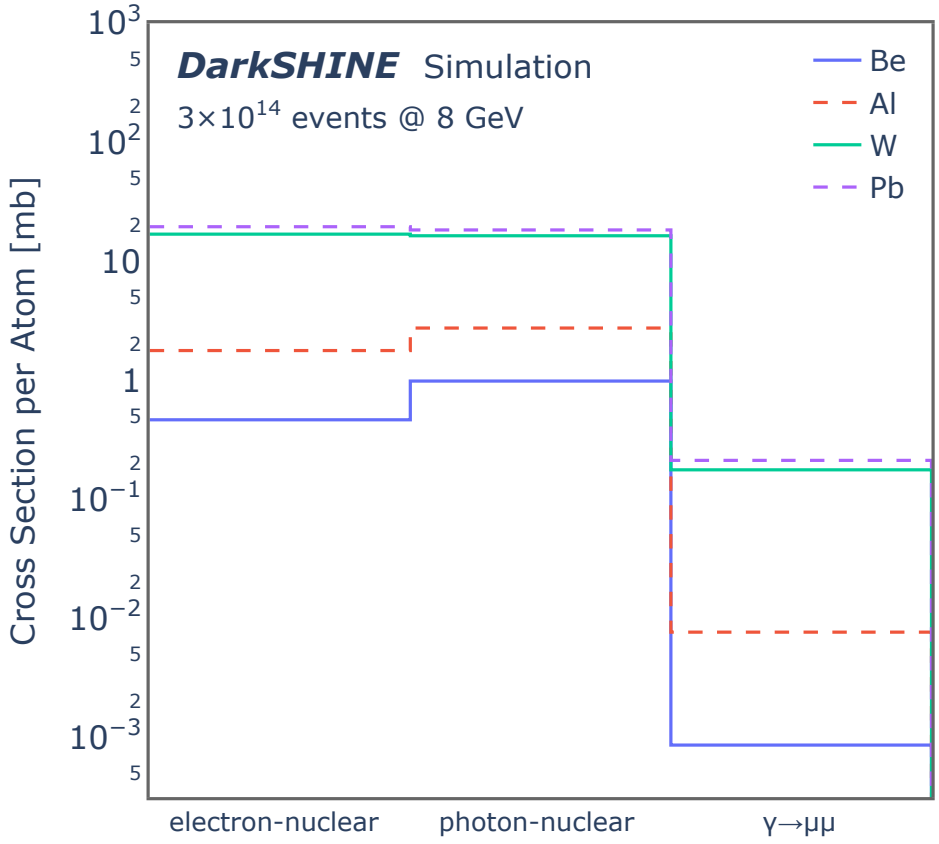


Figure 3.10 Rare-background production for an 8 GeV electron beam in Be, Al, W, and Pb. Cross sections per atom for electron-nuclear, photonuclear, and $\gamma \rightarrow \mu^+\mu^-$ processes.

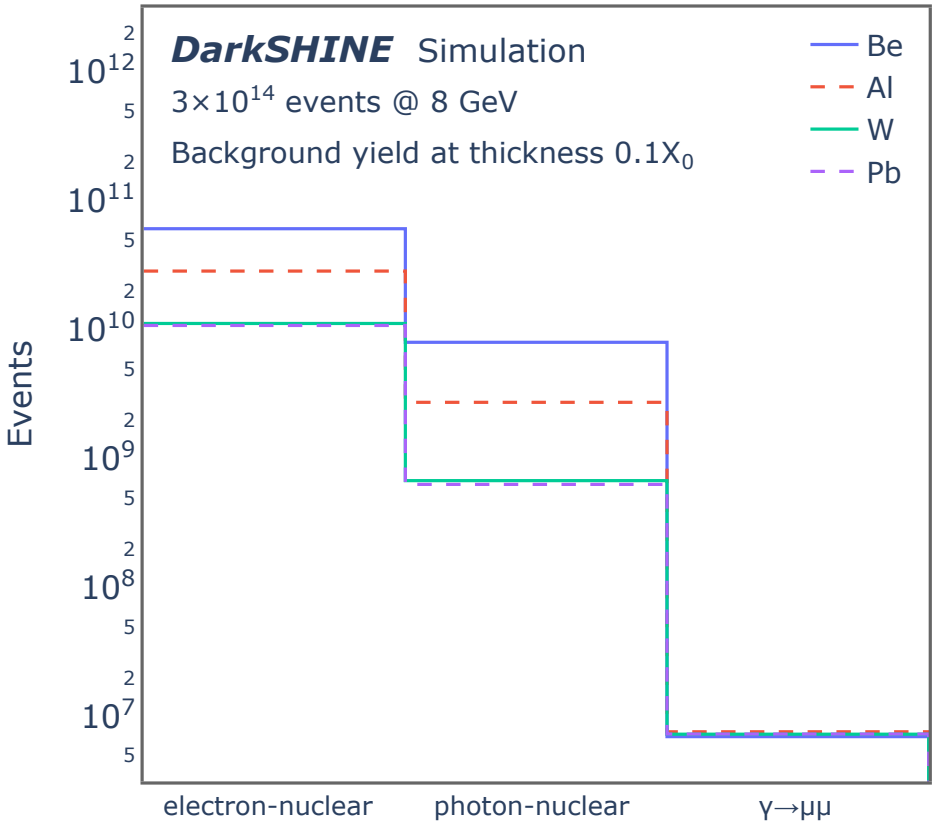


Figure 3.10 (continued) Rare-background production for an 8 GeV electron beam in Be, Al, W, and Pb. Yields for $0.1X_0$ targets and 3×10^{14} EOT, including the material density and radiation length.

pair production, bremsstrahlung + charged-current exchange, and CCQE with an exclusive π^0 final state[48]. Among these, Møller + CCQE and CCQE + π^0 depend on the target material and constitute genuine missing-energy backgrounds. Their event yields are compared in Table 3.4; the W and Pb cases are substantially smaller than Al or Be.

Table 3.4 Irreducible neutrino-induced missing-energy background yields for different target materials (3×10^{14} EOT, $E_{\text{beam}} = 10$ GeV).

Events per 3×10^{14} EOT	W	Pb	Al	Be
Møller + CCQE	3.6×10^{-4}	3.0×10^{-4}	8.6×10^{-3}	7.3×10^{-2}
CCQE with π^0	0.3–0.6	0.28–0.55	1.3–2.7	4.1–8.2

Tungsten is selected over lead for two additional practical reasons: higher density (19.3 vs. 11.4 g cm^{-3}) allows a more compact target for the same radiation-length thickness, and its melting point of 3422°C is favourable for operation under beam heating. The DarkSHINE baseline target is a 350 μm tungsten foil, corresponding to 0.1 X_0.

3.5 Electromagnetic Calorimeter

3.5.1 Reference Design

The ECAL must measure the energy of the scattered electron with sufficient resolution to distinguish the low-energy recoil characteristic of dark-photon production from the near-beam-energy deposits of ordinary bremsstrahlung events. In signal events, the dark photon carries most of the beam energy away invisibly; more than 75% of signal events with $m_{A'} =$ 1 MeV deposit less than 4 GeV in the ECAL, while inclusive backgrounds populate the region near 8 GeV (Fig. 3.11). At the same time, the ECAL volume must be large enough to contain electromagnetic showers from the heaviest dark photons, which recoil the electron to larger angles (Fig. 3.11).

LYSO crystal[49] is adopted as the active medium. Its high light yield underpins the electromagnetic energy resolution, its fast scintillation decay time (40 ns) matches the 1 MHz beam rate, and its radiation hardness is adequate for the expected dose. The ECAL is a fully segmented homogeneous calorimeter consisting of $20 \times 20 \times 11$ crystals, each of dimensions $2.5 \times 2.5 \times 4.0$ cm^3 and wrapped in aluminium reflective film. Each crystal is read out by a silicon photomultiplier (SiPM)[50-51]. The transverse extent of the array is approximately

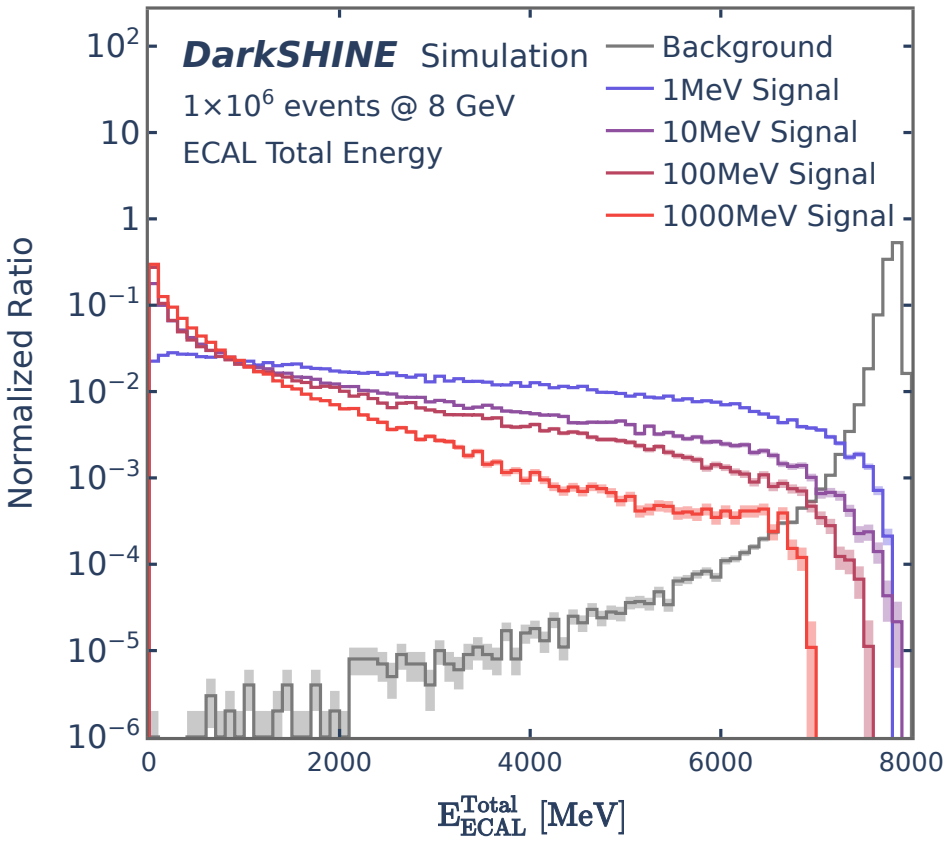


Figure 3.11 Normalised distributions for an 8 GeV electron beam, comparing inclusive background with dark-photon masses of 1, 10, 100, and 1000 MeV. Total ECAL energy. Signal events populate lower visible energies.

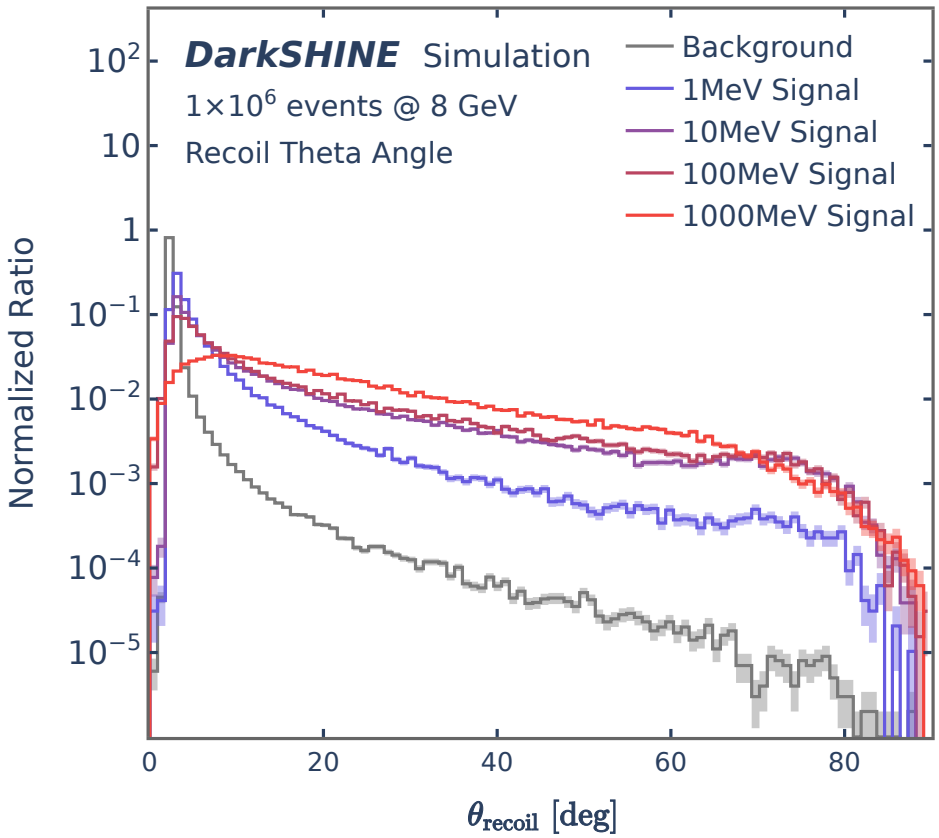


Figure 3.11 (continued) Normalised distributions for an 8 GeV electron beam, comparing inclusive background with dark-photon masses of 1, 10, 100, and 1000 MeV. Recoil-electron polar angle. Heavier mediators broaden the recoil-angle distribution.

$50\times50\,\mathrm{cm}^2$, and the depth along the beam axis is ~44 cm, corresponding to roughly 39 radiation lengths, providing near-complete electromagnetic containment[52]. The ECAL structure is illustrated in Fig. 3.12.

The transverse dimensions and the longitudinal depth address different containment requirements. The depth allows the electromagnetic shower to develop along the beam direction, while the transverse aperture must intercept particles emerging from the target at nonzero angle. The recoil-angle distribution in Fig. 3.11 is therefore relevant even for a deep calorimeter. Once a particle enters the array, the crystal segmentation records how the deposited energy is shared between cells and layers. Total energy uses the sum of these deposits; maximum-cell energy and shower-shape quantities retain their spatial distribution.

In a homogeneous calorimeter the same crystal provides the material in which the shower develops and the scintillation signal used to measure it. The optical and electronic response is nevertheless an additional measurement stage. Energy deposited in a crystal produces scintillation light, only part of which reaches the photosensor and contributes to the readout. These response stages are studied in Section 3.7.

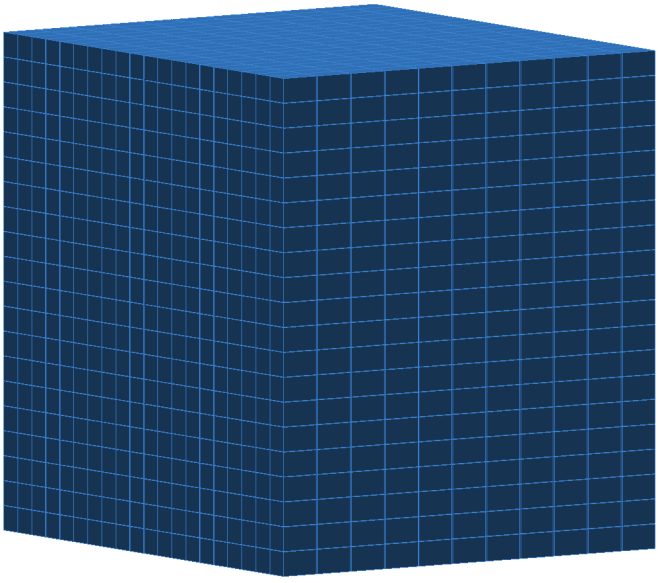


Figure 3.12 Reference DarkSHINE ECAL: a non-staggered $20\times20\times11$ array of $2.5\times2.5\times4.0\,\mathrm{cm}^3$ LYSO crystals. Each crystal is read out by a SiPM at its downstream face; the array spans approximately $50\times50\,\mathrm{cm}^2$ transversely and 44 cm in depth.

3.5.2 Expected Performance

3.5.2.1 Energy Resolution

The energy resolution is evaluated by simulating single electrons from 1 to 8 GeV incident on the ECAL alone (no target or other detectors). Results after digitisation are shown in Fig. 3.13. The stochastic term of the digitised energy resolution is better than 2%, supported by the high intrinsic light yield of LYSO. The energy response is linear across the full range, with a small downward bias from passive-material absorption and invisible energy components.

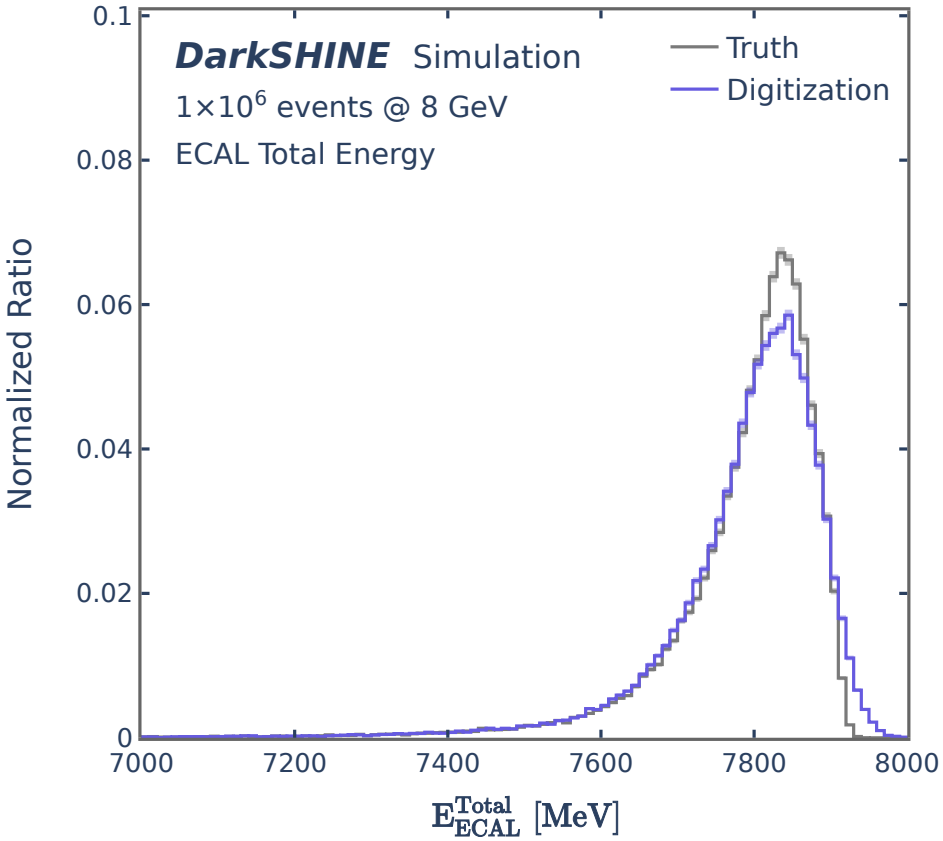


Figure 3.13 ECAL response to isolated electrons, without the target or upstream detectors. Truth denotes deposited energy; digitisation adds the readout response. Total energy for 8 GeV electrons at truth and digitised levels.

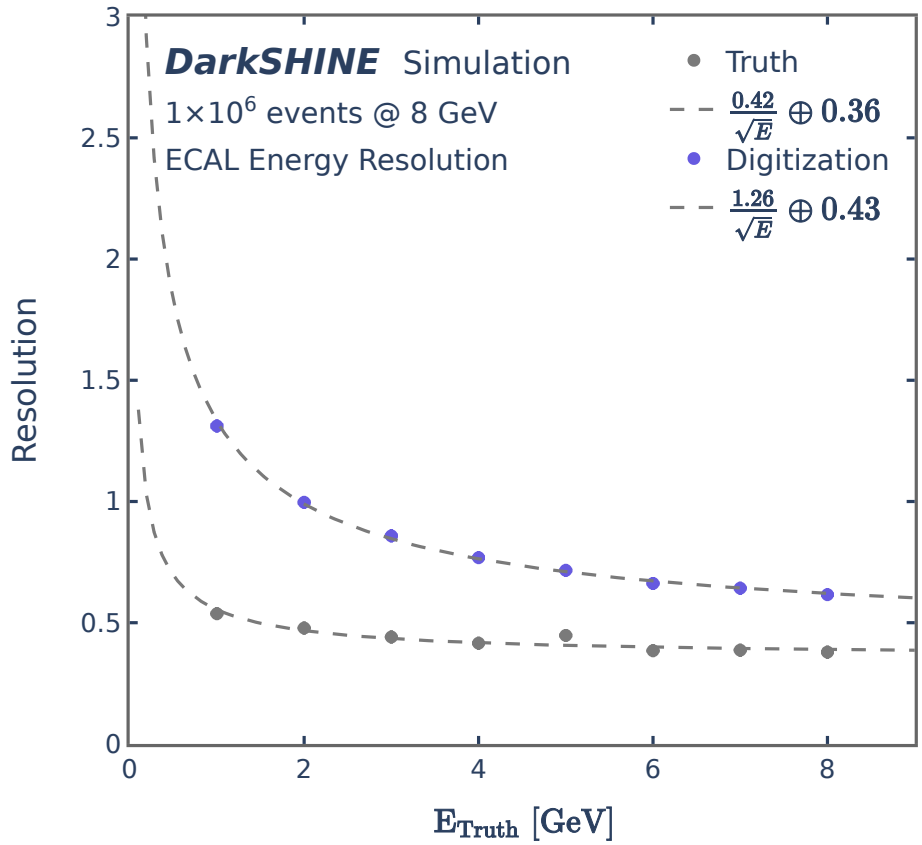


Figure 3.13 (continued) ECAL response to isolated electrons, without the target or upstream detectors. Truth denotes deposited energy; digitisation adds the readout response. Relative energy resolution in percent over the 1–8 GeV scan.

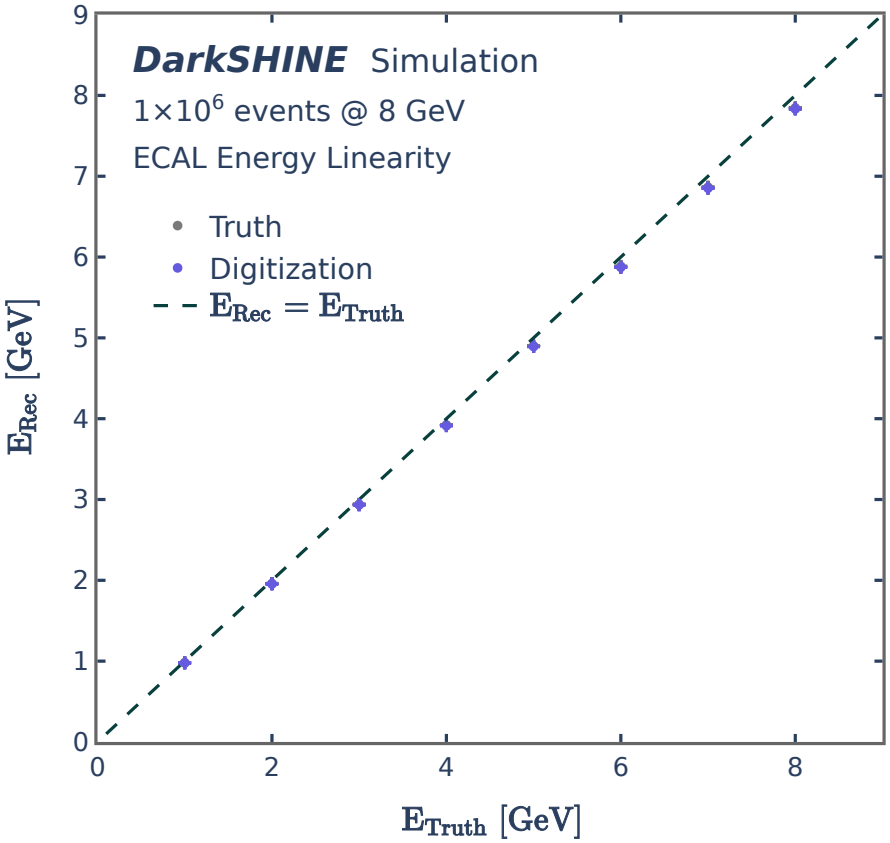


Figure 3.13 (continued) ECAL response to isolated electrons, without the target or upstream detectors. Truth denotes deposited energy; digitisation adds the readout response. Mean response versus incident electron energy.

3.5.2.2 Signal Efficiency and Shower Variables

The signal region is defined by a total ECAL energy below 2.5 GeV[53], which is determined jointly with other sub-detector thresholds to suppress all backgrounds to below the unit level in 3×10^{14} EOT. Signal efficiencies for different dark-photon masses are shown in Fig. 3.14: as $m_{A'}$ increases, the recoil electron carries less energy and efficiency rises.

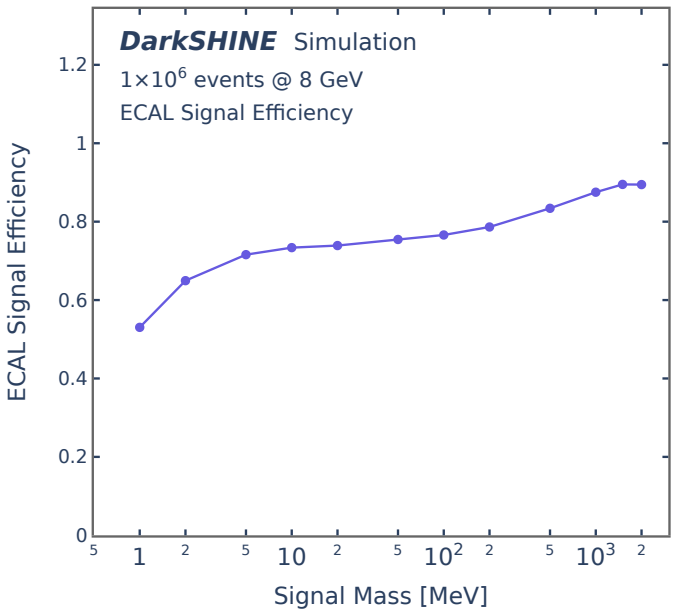


Figure 3.14 Efficiency of the total-ECAL-energy requirement $E_{\mathrm{ECAL}} < 2.5$ GeV versus dark-photon mass for an 8 GeV electron beam.

Beyond the total energy, the fine segmentation of the ECAL allows reconstruction of shower-shape variables that discriminate electromagnetic showers (signal) from hadronic showers initiated by photonuclear or electron-nuclear backgrounds. Several such variables are compared in Fig. 3.15. Compared with electromagnetic events, hadronic events characteristically produce more hits, deeper shower centroids, and broader transverse profiles.

Total energy measures the amount of visible activity; the other quantities test how that activity is distributed. Hit multiplicity counts the contributing detector cells under the hit definition. The mean energy per hit combines the energy sum with that multiplicity. Layer information retains the development along the calorimeter depth, while transverse observables describe the spread across the array. Events with similar total energies can consequently have

different values of these quantities. Their interpretation requires the cell size and hit threshold to remain fixed when samples are compared. The reconstruction definitions are given in Section 4.7.

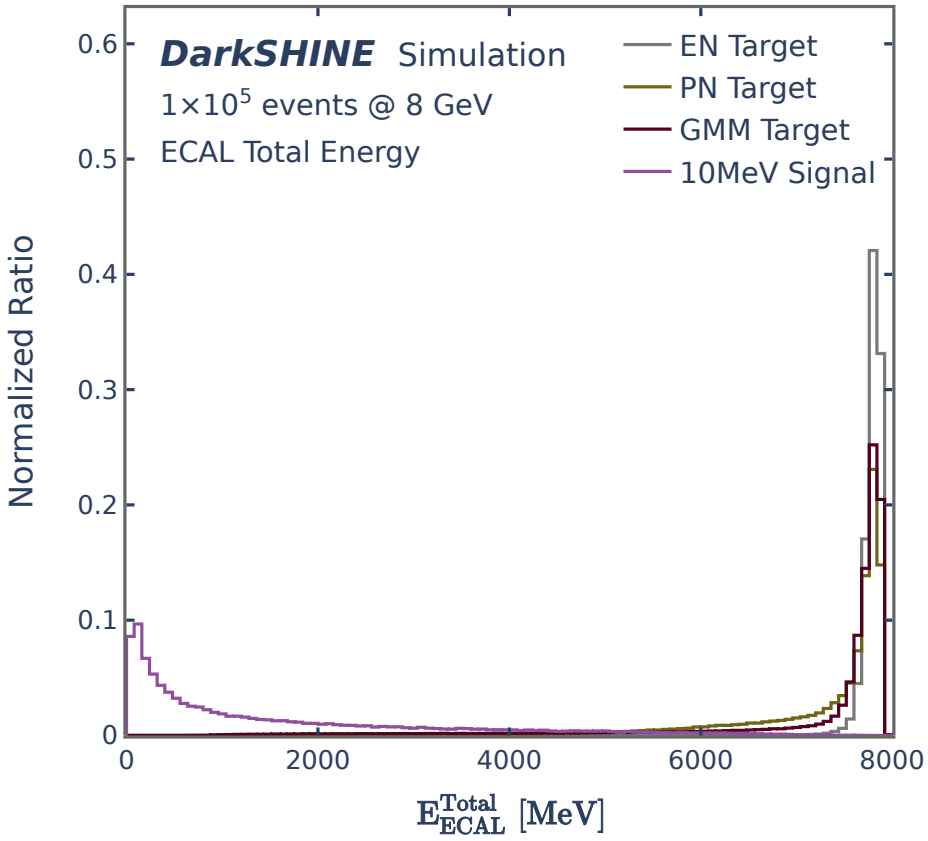


Figure 3.15 Normalised ECAL observables for an 8 GeV electron beam, comparing a 10 MeV dark-photon signal with target electron-nuclear (EN), photonuclear (PN), and photon-conversion-to-muon-pair (GMM) backgrounds. Total deposited energy.

3.6 Hadronic Calorimeter

3.6.1 Role and Physics Requirements

The HCAL vetoes neutral particles, primarily neutrons, that pass through the ECAL without producing detectable deposits and thus mimic a missing-energy signature. Neutral hadrons arise predominantly from photonuclear (PN) and electron-nuclear (EN) interactions

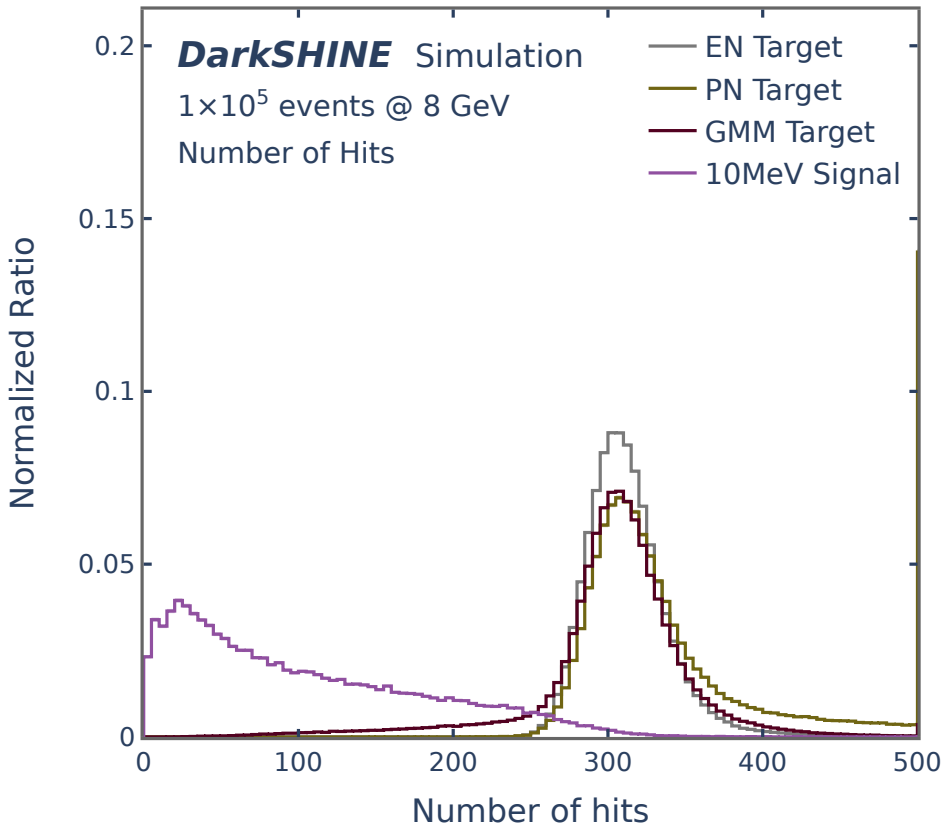


Figure 3.15 (continued) Normalised ECAL observables for an 8 GeV electron beam, comparing a 10 MeV dark-photon signal with target electron-nuclear (EN), photonuclear (PN), and photon-conversion-to-muon-pair (GMM) backgrounds. Number of fired cells.

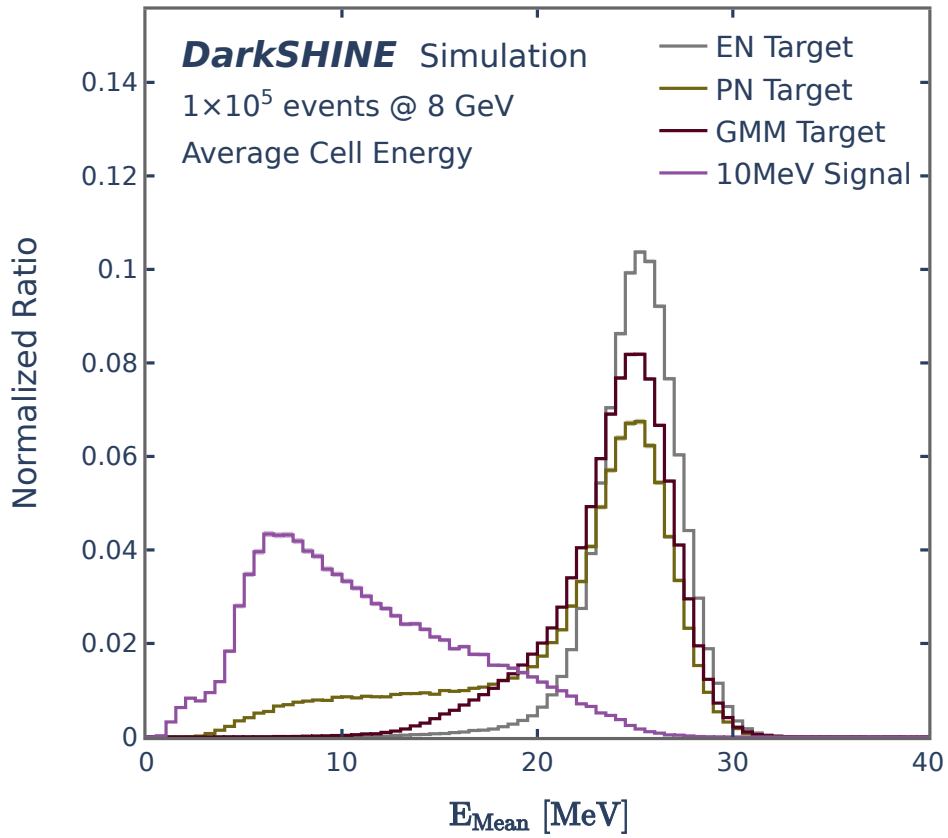


Figure 3.15 (continued) Normalised ECAL observables for an 8 GeV electron beam, comparing a 10 MeV dark-photon signal with target electron-nuclear (EN), photonuclear (PN), and photon-conversion-to-muon-pair (GMM) backgrounds. Mean energy per fired cell.

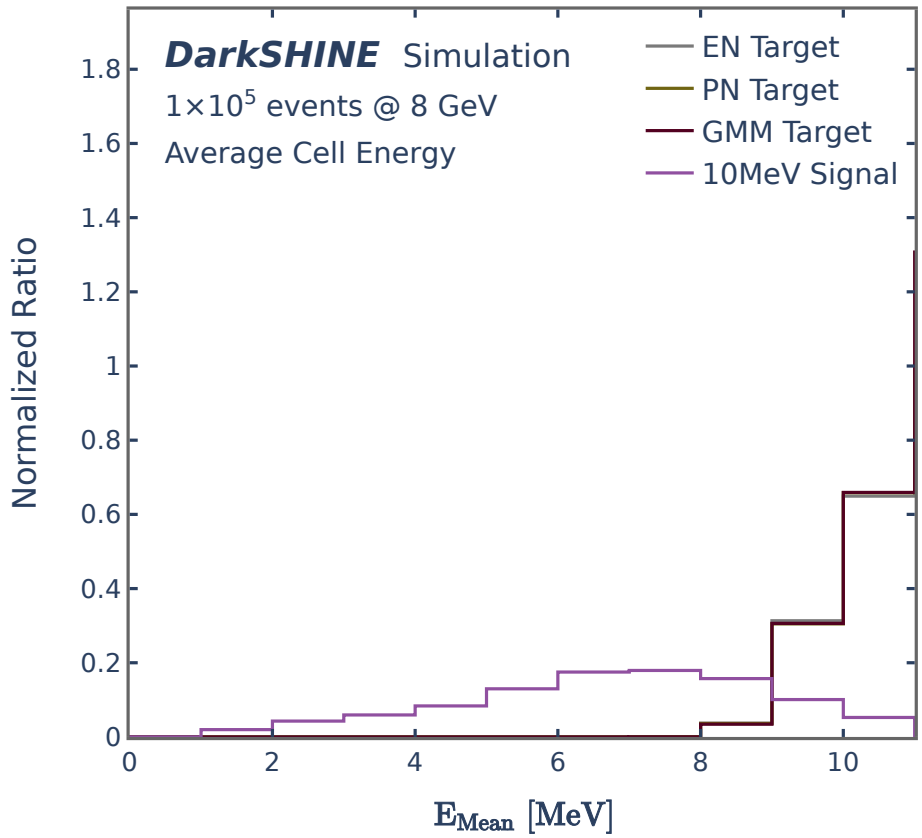


Figure 3.15 (continued) Normalised ECAL observables for an 8 GeV electron beam, comparing a 10 MeV dark-photon signal with target electron-nuclear (EN), photonuclear (PN), and photon-conversion-to-muon-pair (GMM) backgrounds. Number of layers containing a hit.

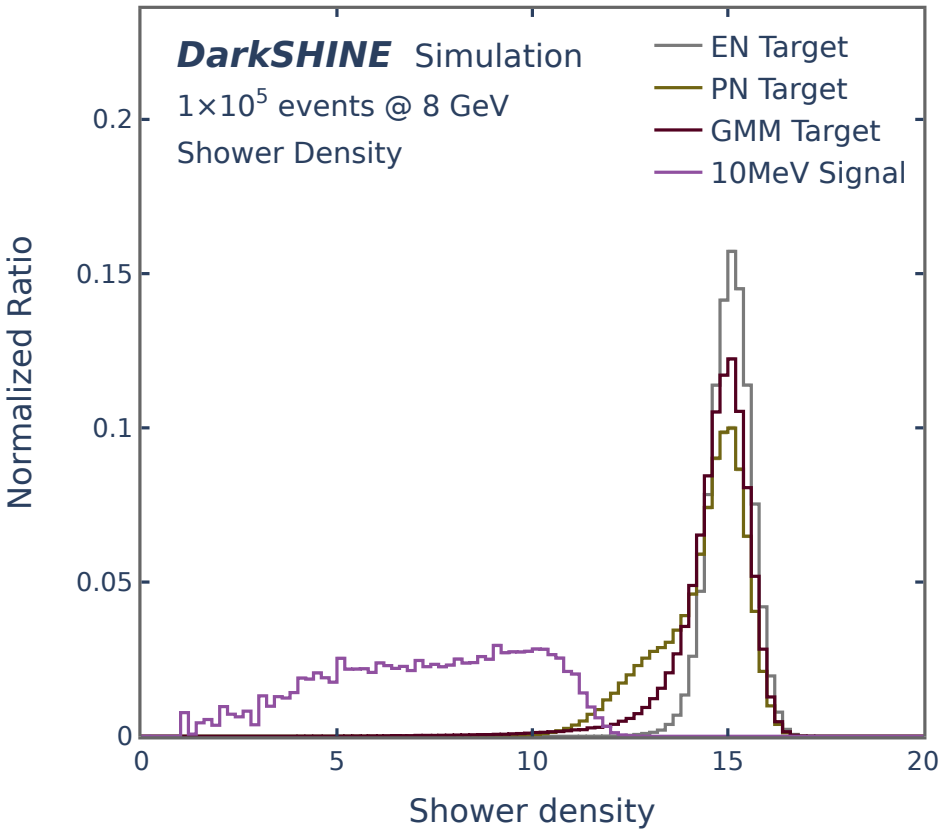


Figure 3.15 (continued) Normalised ECAL observables for an 8 GeV electron beam, comparing a 10 MeV dark-photon signal with target electron-nuclear (EN), photonuclear (PN), and photon-conversion-to-muon-pair (GMM) backgrounds. Shower density is the mean number of hits in the $3 \times 3 \times 3$ neighbourhood of each fired cell.

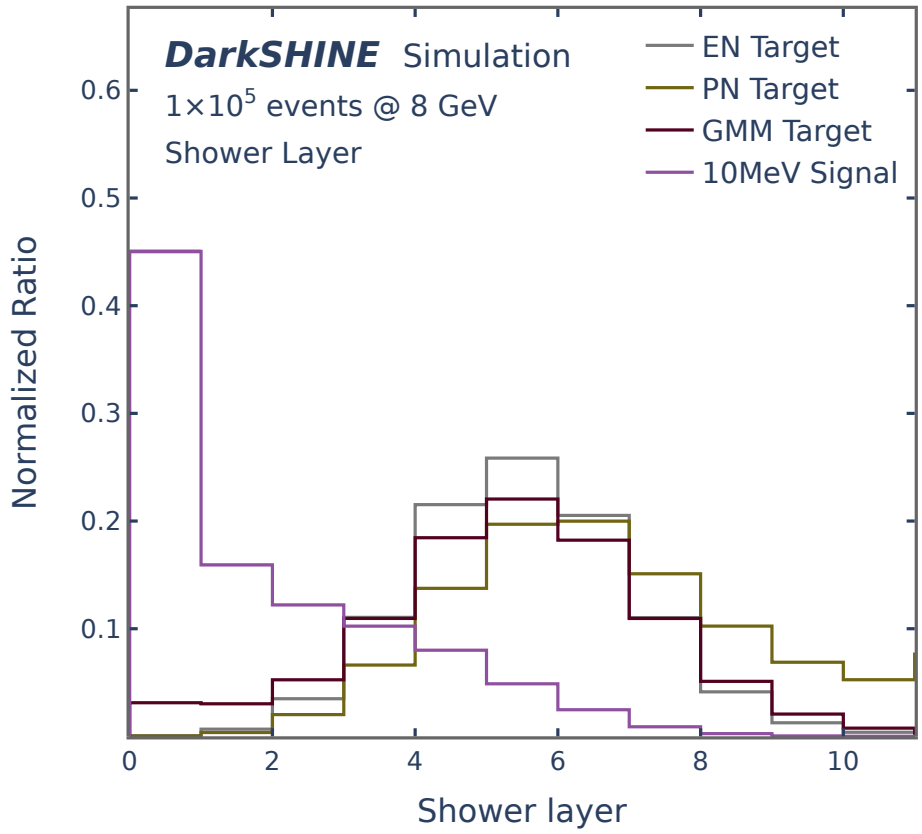


Figure 3.15 (continued) Normalised ECAL observables for an 8 GeV electron beam, comparing a 10 MeV dark-photon signal with target electron-nuclear (EN), photonuclear (PN), and photon-conversion-to-muon-pair (GMM) backgrounds. Shower layers count layers whose transverse shower widths in both coordinates are at least 5 cm.

in the target and ECAL; the dominant final-state particle types from these processes are listed in Table 3.5. Neutrons account for 64–73% of all particles produced, making neutron-veto efficiency the key design driver.

Table 3.5 Final-state particle composition from EN and PN background processes.

Process	Neutron	Proton	Pion	Kaon
Electron-nuclear	73.42%	21.52%	4.64%	0.42%
Photon-nuclear	64.95%	18.56%	14.43%	2.06%

Events passing the ECAL energy threshold $E_{\mathrm{ECAL}} < 2.5$ GeV are expected to contain around 10^6 energetic neutrons per 10^{14} EOT[53] (Fig. 3.16). For these high-energy neutrons ($E_n \gtrsim 2$ GeV) the required veto inefficiency is below 10^{-5}. For low-energy neutrons, which cannot alone account for the observed ECAL energy deficit, a veto inefficiency below 10^{-3} suffices.

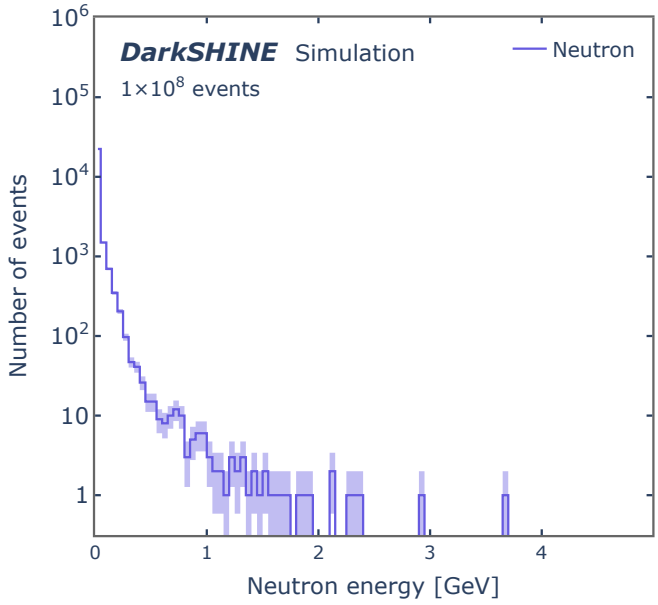


Figure 3.16 Energy distribution of neutrons surviving the $E_{\mathrm{ECAL}} < 2.5$ GeV selection. Very few neutrons above 2.5 GeV remain; under 10^{14} EOT approximately 10^6 energetic neutrons are expected.

3.6.2 Reference Conceptual Design

The HCAL is a sampling calorimeter with alternating layers of iron absorber and plastic scintillator (Fig. 3.17). The main HCAL body has a 4×4 m^2 transverse cross section, divided into sixteen 100×100 cm^2 modules to keep individual scintillator bars at a manageable length. Along the beam axis the total depth exceeds 400 cm, comprising 80 iron absorber plates (each 3 cm thick) interleaved with 162 scintillator layers (each 1 cm thick). Adjacent sensitive layers are oriented with their strip axes perpendicular (x and y alternating), providing two-dimensional spatial information without additional readout complexity (Fig. 3.18).

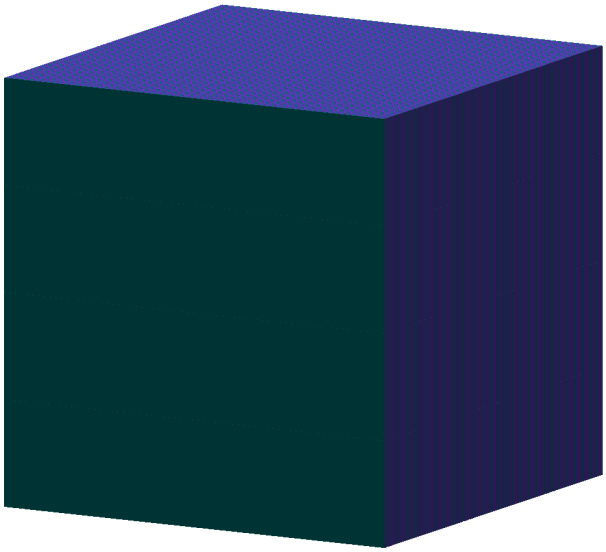


Figure 3.17 Original DarkSHINE HCAL geometry: a 4×4 m^2 sampling calorimeter divided into sixteen modules. Alternating iron absorber plates and plastic-scintillator layers give a depth exceeding 4 m for detecting penetrating hadronic activity.

Each scintillator strip ($100\times5\times1$ cm^3) is wrapped in reflective film and instrumented with wavelength-shifting (WLS) fibres that channel scintillation light to a SiPM mounted on the side of the module, keeping all readout electronics outside the active volume.

The absorber and the scintillator have separate functions in this sampling structure. The iron supplies material in which hadronic interactions develop, while charged particles crossing the scintillator produce the measured light. The energy measured in the sensitive layers is consequently a fraction of the incident hadron energy. The neutron calibration in Fig. 3.20 determines the conversion from scintillator response to reconstructed hadron energy.

Longitudinal segmentation samples activity at different depths, and the orthogonal strip

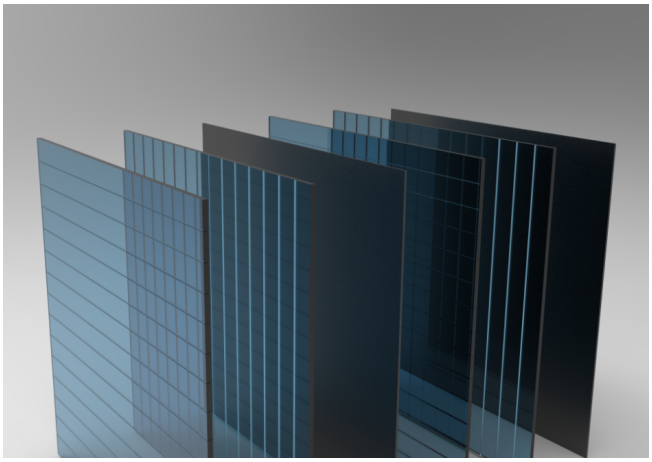


Figure 3.18 Detail of the HCAL scintillator–absorber arrangement. Translucent layers are scintillators with perpendicular strip orientations in adjacent layers; opaque layers are iron absorbers.

directions provide transverse information in successive sensitive layers. For the veto, the existence of a sufficiently large local deposit and the total activity are both useful. A particle can produce a concentrated deposit in a small number of strips or distribute its response over many layers. Retaining cell information permits these patterns to be tested in addition to their sum.

3.6.3 Expected Performance

Veto performance is characterised by injecting neutrons of known energy and applying two simultaneous cuts: total HCAL energy $E_{\mathrm{HCAL}}^{\mathrm{total}} < 30$ MeV and maximum single-cell energy $E_{\mathrm{HCAL}}^{\mathrm{MaxCell}} < 0.1$ MeV. Figure 3.19 shows the veto inefficiency as a function of HCAL depth for neutrons from 100 MeV to 3 GeV. The baseline design meets the physics requirements across the full energy range.

The simultaneous total-energy and maximum-cell conditions specify what it means for the HCAL to be quiet in this study. An event is missed only when its total activity is below the sum threshold and no individual cell exceeds the local threshold. The two observables are related, but they test different distributions of energy among the cells. Many small deposits can exceed the total-energy threshold without producing one exceptionally large cell. A local deposit can fail the single-cell requirement even when the sum is still below its allowed value.

Energy calibration is performed by comparing the true neutron energy to the energy deposited in the active scintillator layers (sampling fraction), which is linear across the relevant

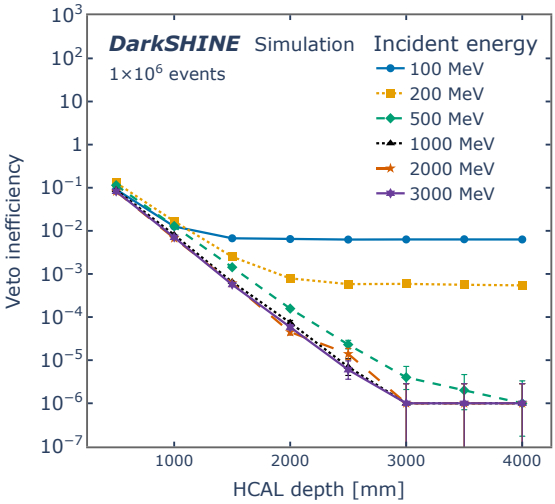


Figure 3.19 Simulated neutron-veto inefficiency versus HCAL depth for incident energies from 100 MeV to 3 GeV. A neutron is missed when both the total HCAL energy is below 30 MeV and the maximum cell energy is below 0.1 MeV. Separate curves show the energy dependence of this missed fraction as the detector depth increases.

range (Fig. 3.20). After applying the calibration scale factor derived from hadrons, the sum of calibrated ECAL and HCAL energies reconstructs the 8 GeV beam energy for inclusive events (Fig. 3.21).

3.7 Subsystem Prototyping and Instrumentation Validation

Component measurements and prototype tests help choose the detector parameters used in Monte Carlo optimisation. They show how hit positions, deposited energy, and particle identity appear in the electrical signals measured by the detector. They also identify effects that need to enter calibration and systematic-uncertainty models.

3.7.1 AC-LGAD sensor characterisation

Two batches of AC-LGAD strip-sensor prototypes were fabricated by the Institute of Microelectronics, Chinese Academy of Sciences, following IHEP designs. The strip pitch/width combinations were 100/50, 60/40, and 45/30 μm. The two batches, wafer-11 (0.01 P) and wafer-12 (10 P), differed in their n^+ implant dose, where P denotes the phosphorus-dose unit used in the prototype description.

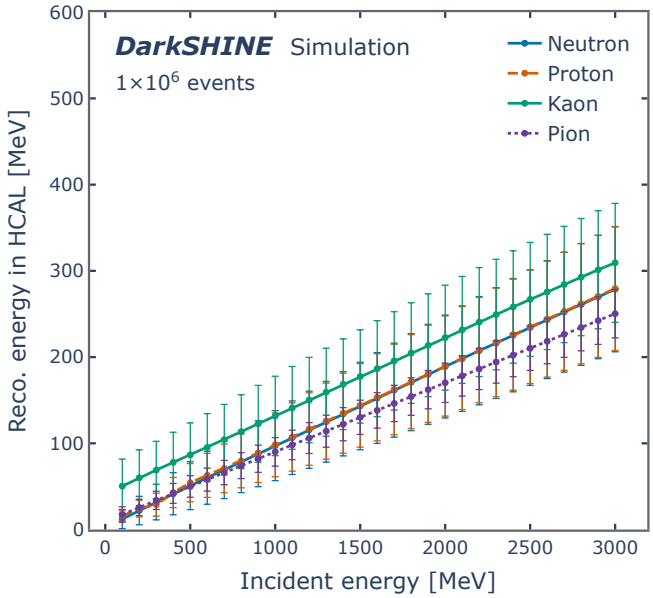


Figure 3.20 Energy deposited in the HCAL scintillator layers versus incident energy for neutrons, protons, kaons, and pions. The approximately linear response determines the hadronic calibration scale used to convert the sampling signal to reconstructed energy.

The AC-LGAD operating principle introduced in Section 3.3 combines internal charge gain with a continuous resistive layer and capacitively coupled metal electrodes. Charge generated in the depleted bulk drifts into the multiplication region, where a moderate avalanche gain improves timing. The signal then spreads laterally through the resistive sheet and couples to several neighbouring strips. The ratio of strip amplitudes carries sub-pitch position information, while the summed signal retains the deposited charge.

Higher gain improves the signal, but excessive gain increases shot noise and the probability of premature breakdown. Excessive charge sharing degrades two-track separation. Sensor thickness, implant dose, sheet resistance, electrode pitch, and coupling capacitance must therefore be chosen together. The DarkSHINE prototypes use strip geometries intended to provide precise measurement in the bending coordinate while keeping the number of electronic channels manageable. Stereo information is obtained by rotating adjacent sensor planes, as described in Section 3.3.

Prototype sensors were characterised at room temperature before the spatial measurements. Electrical tests included scans of current and capacitance as functions of applied

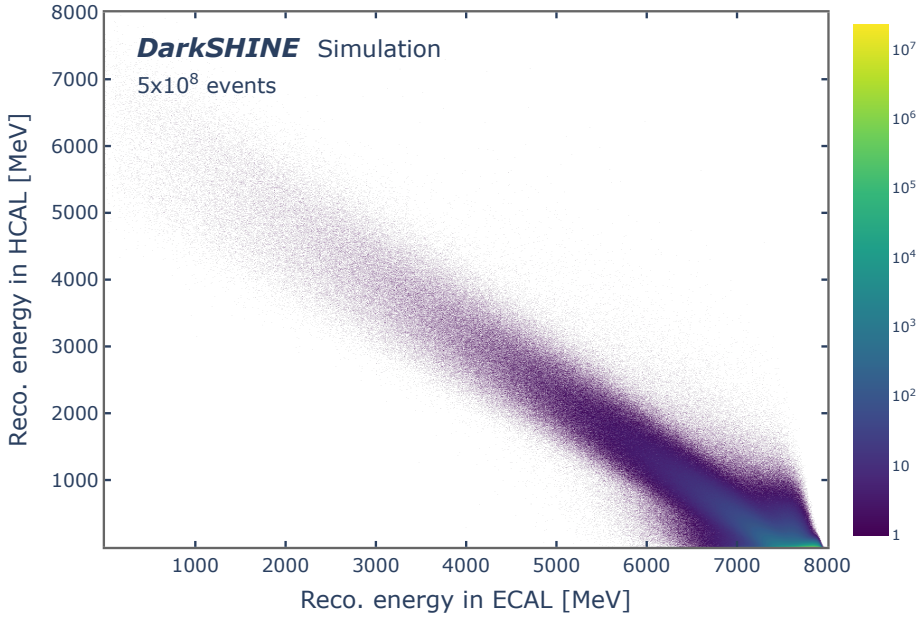


Figure 3.21 Reconstructed ECAL energy versus calibrated HCAL energy for an 8 GeV electron-beam inclusive sample. The colour scale gives the event population. Events with reduced ECAL energy can carry substantial HCAL energy; their sum reconstructs the incident beam energy.

voltage. The leakage-current curve tests junction quality and defines a safe operating range below breakdown. The capacitance curve constrains the depletion voltage and effective bulk doping. Representative measurements are shown in Figure 3.22; the variation between devices is used to choose the bias point for subsequent tests.

Position response was studied with a laser transverse-current technique (TCT), using a focused infrared laser scanned across the sensor surface. The laser provides a reproducible localised charge deposit and allows response uniformity to be mapped independently of beam-test alignment. The setup and a prototype module are shown in Figure 3.23. At each scan point the waveform amplitude and time are recorded for the illuminated strip and its neighbours. An amplitude-ratio estimator can then be calibrated against the known laser position. The residual distribution after this calibration provides a direct measurement of spatial resolution, while the variation of the residual mean reveals local non-uniformities.

The laser scans measure the time resolution as a function of test position. The reported resolutions are 6.5–8.2 ps for wafer-11 and 8.8–12.3 ps for wafer-12[47]. The timing response

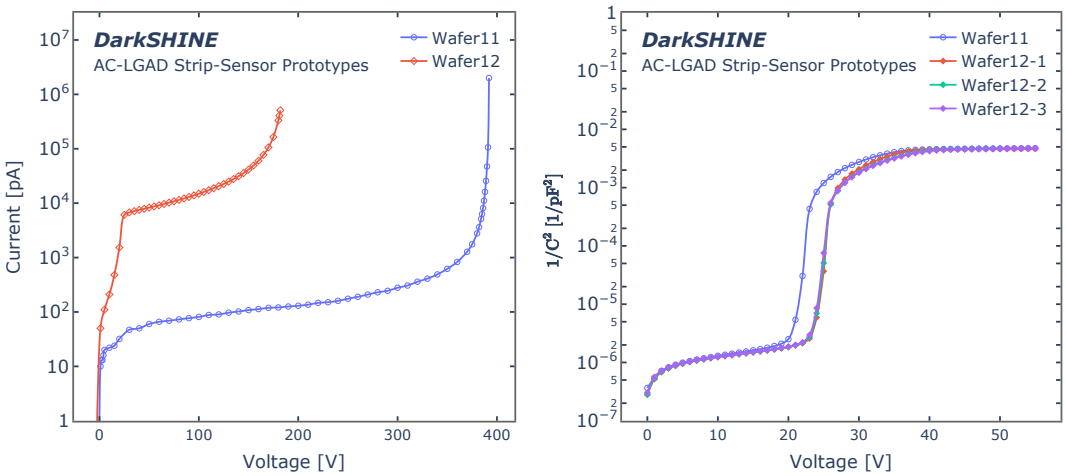


Figure 3.22 Room-temperature electrical characterisation of AC-LGAD prototypes. Left: leakage current versus bias for wafer-11 and wafer-12. Right: capacitance–voltage scans for wafer-11 and three wafer-12 devices. The scans constrain breakdown and depletion behaviour when selecting the operating bias for the subsequent laser measurements.

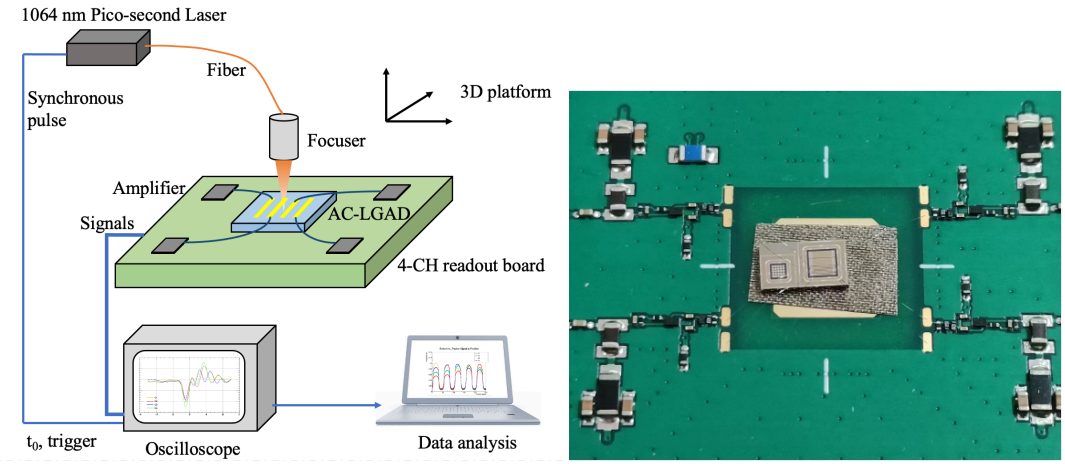


Figure 3.23 Laser characterisation of the AC-LGAD tracker sensor. Left: a focused 1064 nm laser is scanned with a three-dimensional platform; amplified strip signals and the synchronous trigger are recorded by an oscilloscope. Right: sensor module mounted on its four-channel readout board for the scan.

is approximately uniform over the scanned region, with a slight degradation near the sensor edges.

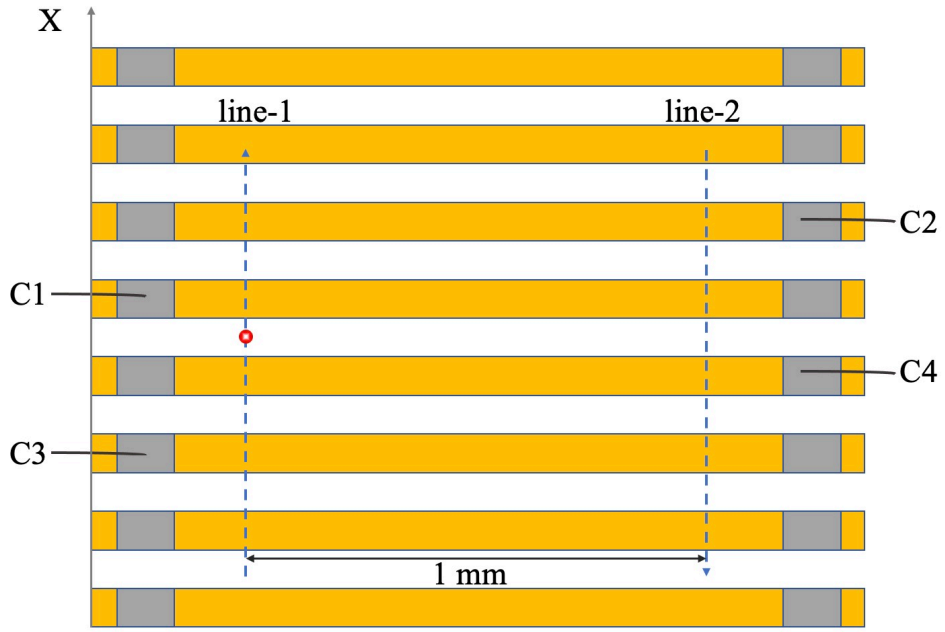


Figure 3.24 Laser measurements of representative AC-LGAD prototypes[47]. Strip geometry and the two laser scan paths.

3.7.2 LYSO crystal and SiPM response

The ECAL is a homogeneous calorimeter: the LYSO crystal is both absorber and active medium. Its short radiation length, small Molière radius, high light yield, and approximately 40 ns scintillation decay time make it suitable for a compact and fast detector[49,54]. A crystal is wrapped to improve optical isolation and light collection and is coupled to one or more silicon photomultipliers (SiPMs). The module design must provide repeatable optical contact, mechanical positioning, thermal control, and access for bias and readout services.

An SiPM consists of many Geiger-mode avalanche pixels. If N_γ incident photons illuminate N_{pix} pixels uniformly within a time short compared with pixel recovery, an idealised response is

$$N_{\text{fired}} = N_{\text{pix}} \left[1 - \exp\left(-\frac{\varepsilon_{\text{PDE}} N_\gamma}{N_{\text{pix}}} \right) \right], \tag{3.5}$$

where ε_{PDE} is the photon-detection efficiency, so that $\varepsilon_{\text{PDE}} N_\gamma$ is the mean number of primary detected photoelectrons before pixel saturation. The expression is linear only when the oc-

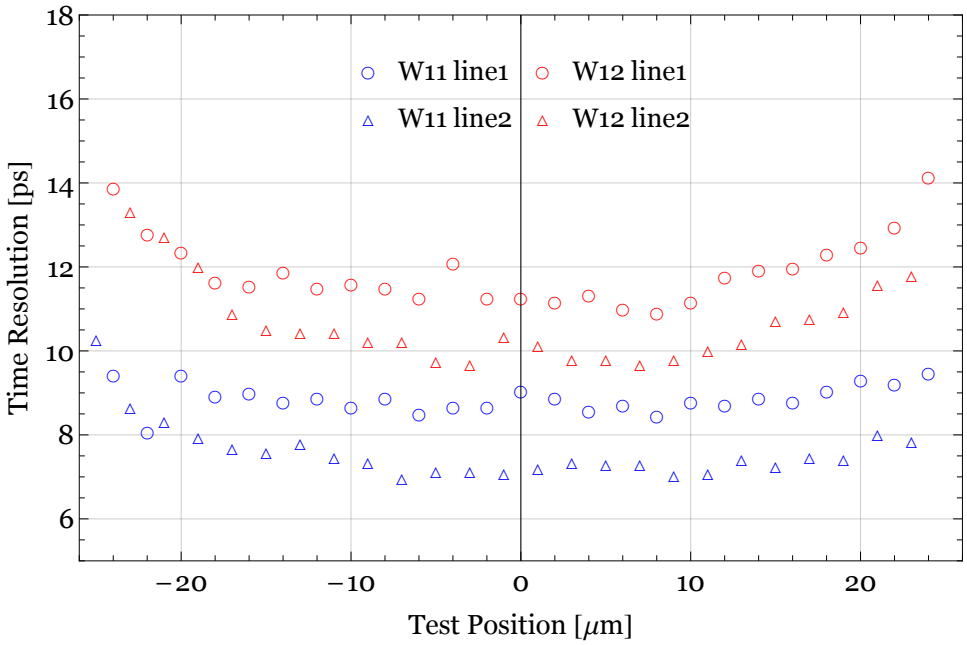


Figure 3.24 (continued) Laser measurements of representative AC-LGAD prototypes[47]. Time resolution in ps versus laser test position in μm. Blue and red curves identify wafer-11 and wafer-12; circles and triangles denote the two scan lines. Wafer-11 has better time resolution across the scanned region.

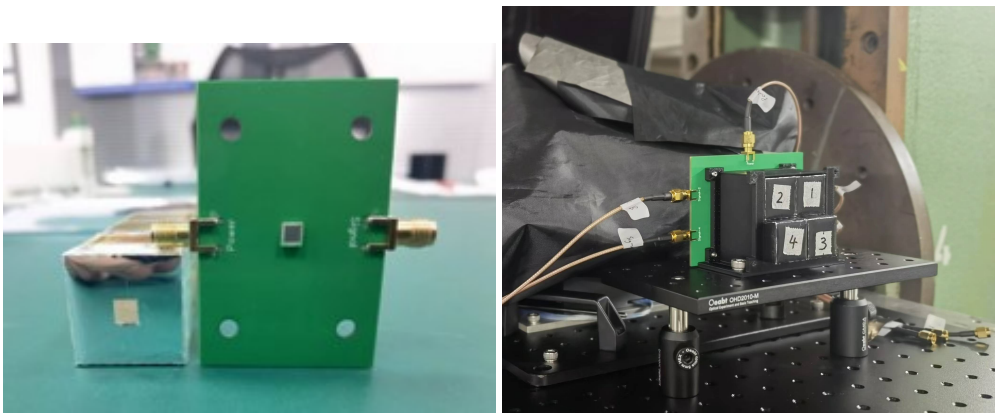


Figure 3.25 ECAL component prototypes. Left: a LYSO crystal coupled to a silicon photomultiplier (SiPM). Right: the four-channel LYSO module used to test assembly and readout before the electron-beam response measurements.

cupancy is small. Pixel recovery, optical crosstalk, afterpulsing, non-uniform illumination, and the time profile of the scintillation light modify the ideal relation. Equation (3.5) nevertheless explains why a high pixel density is required for the large LYSO light yield and why the digitisation model in Chapter 4 includes saturation explicitly.

Defining the mean occupancy as $\eta = \varepsilon_{\mathrm{PDE}} N_\gamma / N_{\mathrm{pix}}$ makes the two limits explicit. For $\eta \ll 1$, expanding the exponential gives $N_{\mathrm{fired}} \simeq \varepsilon_{\mathrm{PDE}} N_\gamma$. At large occupancy, the ideal response approaches N_{pix} because additional photons can reach pixels that have already fired. The formula describes a short pulse with negligible recovery during illumination. The software waveform model retains arrival times and reduces the gain when the same pixel is hit before its configured recovery time has elapsed. It therefore treats the time dependence separately from this static occupancy relation.

The gain is measured from the separation of peaks in a low-light charge spectrum. Once the conversion between integrated charge and photoelectrons is known, response linearity can be tested by varying the incident light. A useful calibration sequence therefore separates three quantities: electronic pedestal and noise, single-pixel gain, and the non-linear response at high occupancy. Temperature and over-voltage must be recorded because both gain and dark-count rate depend on them. These measurements determine the calibration of the photosensor response.

LYSO also contains naturally occurring ^{176}Lu. Its long-lived beta decay and subsequent gamma emissions produce a continuous low-energy spectrum. In a large crystal array the activity contributes to the pedestal environment, especially when cell thresholds are very low. The spectrum can also provide a persistent monitoring source: changes in its reconstructed peak and continuum structure reveal channel gain drifts without an external source. Figure 3.26 shows the measured intrinsic spectrum and its simulation.

3.7.3 ECAL readout chain

The readout must cover signals from low-energy deposits relevant to veto thresholds up to multi-GeV deposits in the hottest crystal, while preserving timing at the SHINE repetition rate. The prototype architecture separates the analogue preamplifier, analogue-to-digital conversion, and field-programmable gate array (FPGA) processing stages[52]. This modularity permits the analogue gain and bandwidth to be optimised independently from the data-

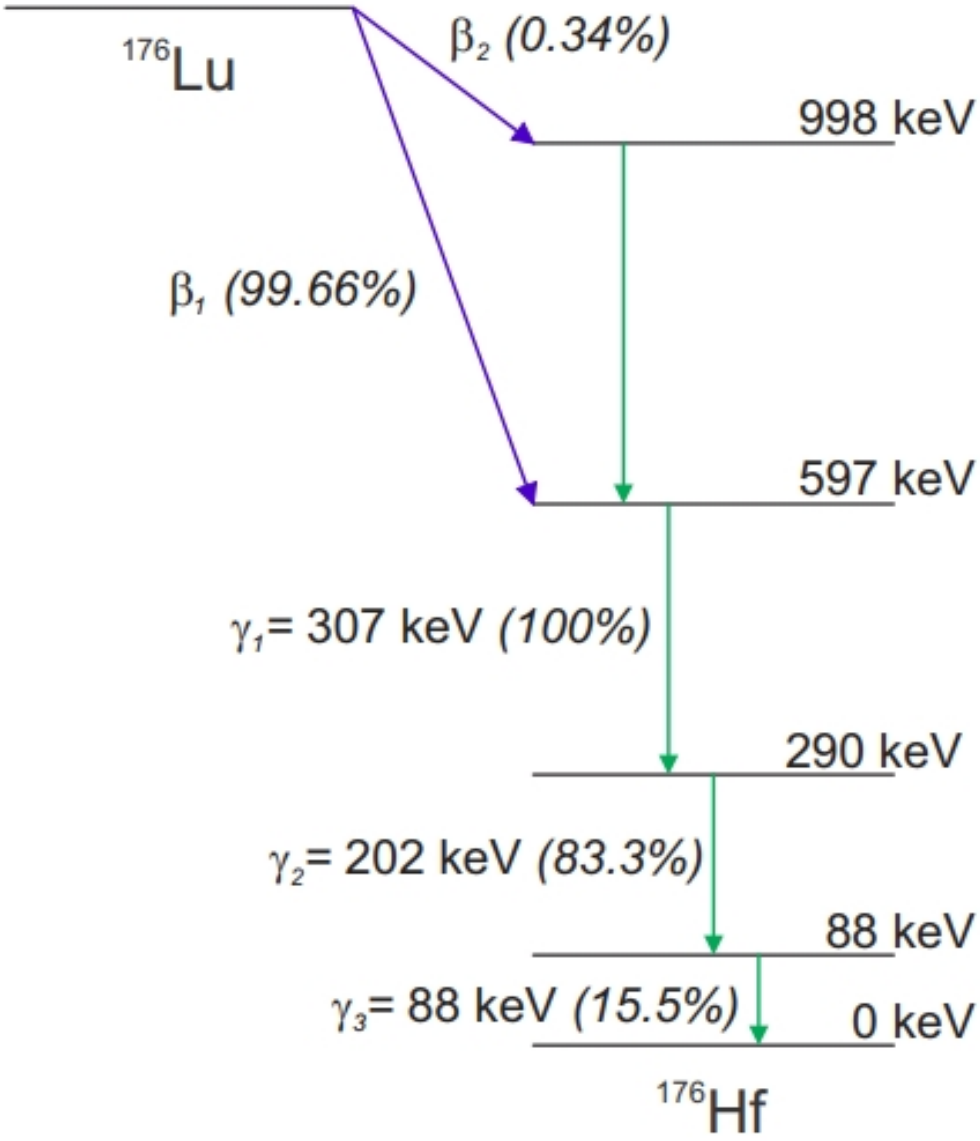


Figure 3.26 Intrinsic radioactivity of LYSO. Schematic ^{176}Lu decay chain.

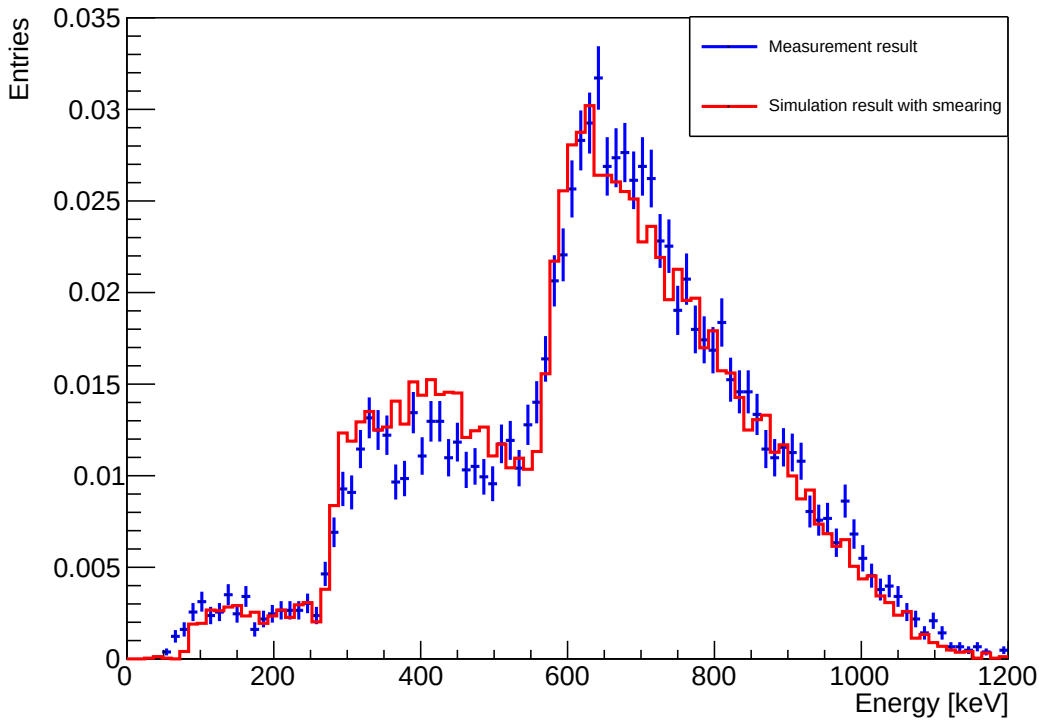


Figure 3.26 (continued) Intrinsic radioactivity of LYSO. Measured energy spectrum (blue points) and simulation with smearing (red histogram). The spectrum tests the low-energy crystal response and provides a reference for monitoring channel gain.

transport firmware.

The preamplifier converts SiPM current into a voltage pulse and shapes the bandwidth to control noise and aliasing. A high-gain path improves resolution for small signals, whereas a low-gain path or appropriate attenuation protects the dynamic range for large showers. The ADC sampling frequency must resolve the pulse shape well enough to estimate amplitude and time. The FPGA then performs baseline subtraction, digital filtering, pulse finding, timestamp assignment, and packet formatting. Buffer depth is chosen from the maximum acceptable readout latency and burst structure rather than the average data rate alone.

Electronic performance is evaluated at several levels. Pedestal runs measure the mean baseline, channel-to-channel spread, and temporal correlations. Injected pulses measure gain linearity and the impulse response. A sinusoidal input followed by a Fourier analysis probes signal-to-noise ratio, harmonic distortion, and effective number of bits. Finally, crystal signals test the complete optical and electronic chain. Figure 3.28 shows representative waveform and frequency-domain diagnostics.

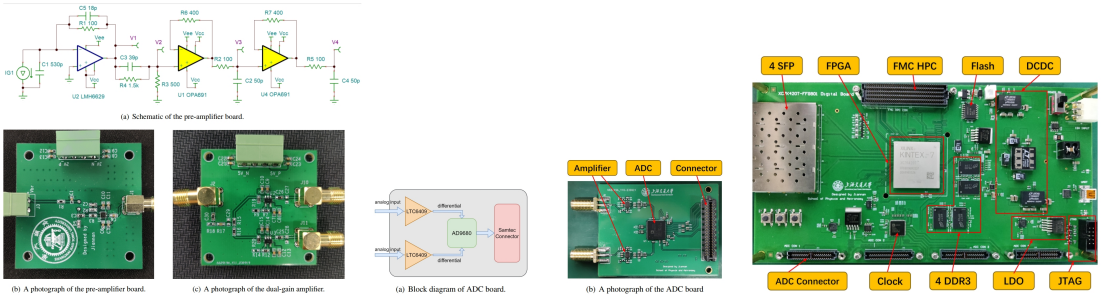


Figure 3.27 Prototype ECAL electronics chain: analogue preamplifier (left), ADC stage (centre), and FPGA processing board (right)[52].

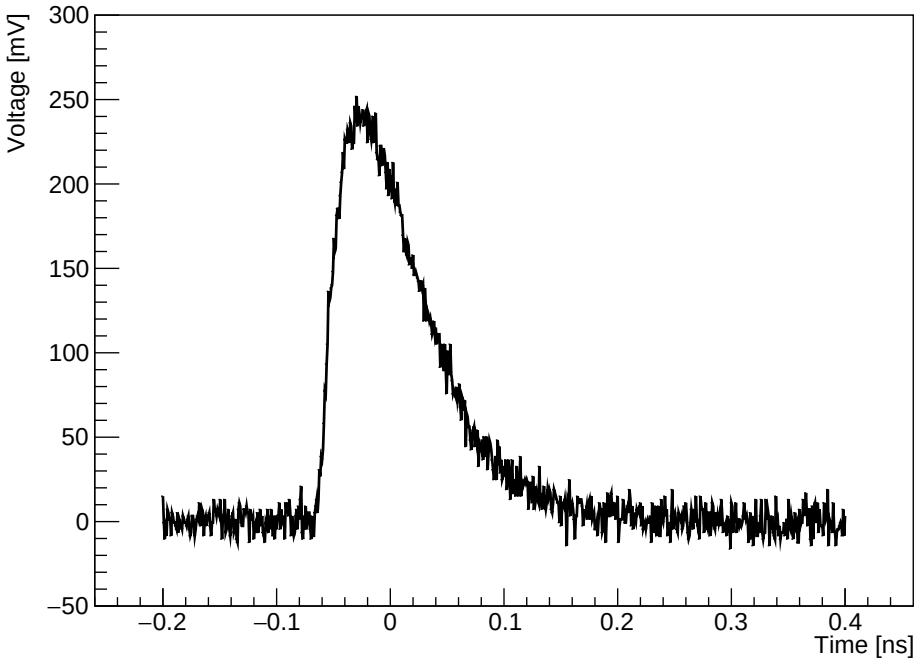


Figure 3.28 ECAL readout tests[52]. Digitised crystal waveform for 5 GeV incident electrons, used to study pulse reconstruction.

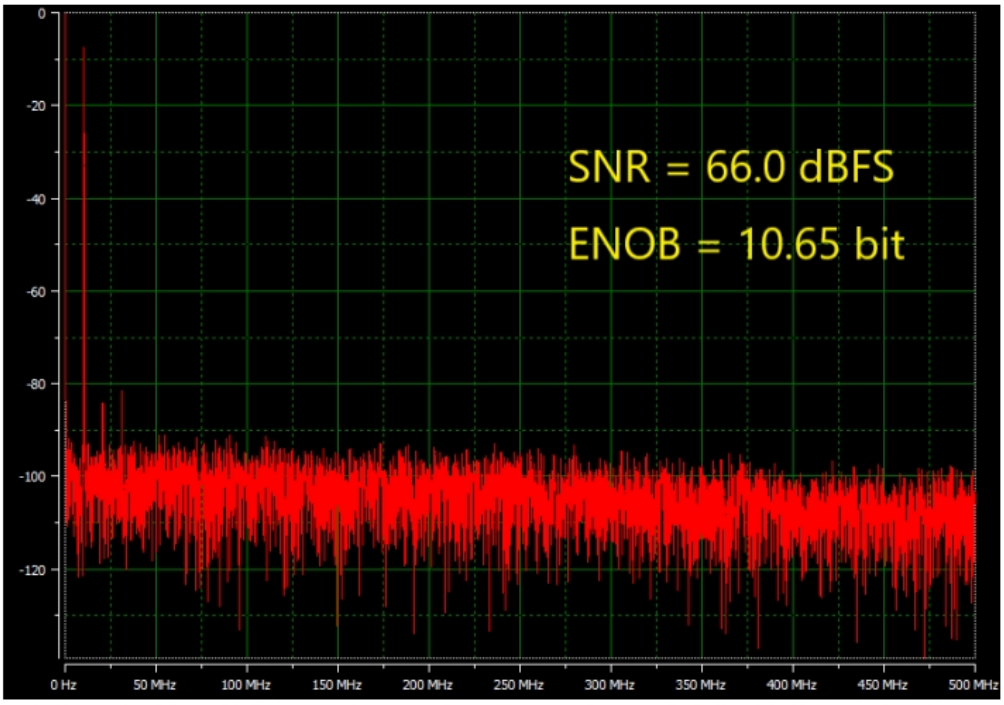


Figure 3.28 (continued) ECAL readout tests[52]. ADC frequency spectrum under sinusoidal excitation, used to measure signal-to-noise ratio (SNR) and effective number of bits (ENOB), probing noise and distortion.

For an experiment with nearly continuous beam delivery, data reduction must occur close to the front end. A practical hierarchy is to retain waveform snippets for calibration streams, while the main physics stream stores reconstructed amplitude, time, quality flags, and a limited set of diagnostic samples. Firmware thresholds must remain looser than the offline veto thresholds so that later calibration does not introduce an irreversible inefficiency. Synchronisation markers and event identifiers are propagated through the tracker and calorimeters to prevent an apparent missing-energy event from being formed by mismatched detector frames.

3.7.4 ECAL beam-test methodology

In a beam test, a crystal matrix is exposed to particles of known energy; external counters define the trigger and tracking or collimation constrains the impact point. Runs at several energies separate an energy-independent constant term from stochastic and noise contributions. If the measured resolution is parameterised as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E/\mathrm{GeV}}} \oplus \frac{b}{E/\mathrm{GeV}} \oplus c, \tag{3.6}$$

the coefficients a, b, and c describe stochastic fluctuations, electronic noise, and energy-independent non-uniformities, respectively. These coefficients depend on clustering, containment, and event selection.

The symbol $\oplus$ in Eq. (3.6) denotes a sum in quadrature. Equivalently, the square of the relative resolution is the sum of a term proportional to $1/E$, a term proportional to $1/E^2$, and a constant term, with the energy expressed in the stated units. Their different energy dependences allow a multi-energy measurement to test whether a single response description accounts for the observed widths.

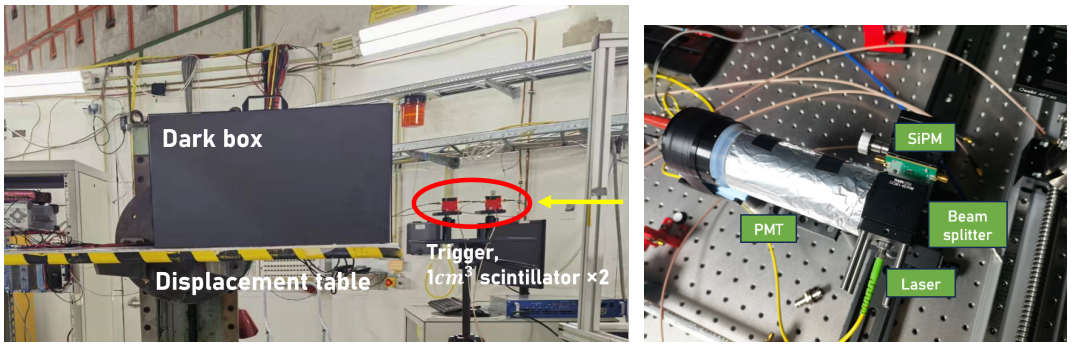


Figure 3.29 ECAL response-test arrangements. Left: electron-beam setup for the crystal module at DESY TB-22. Right: SiPM dynamic-range measurement using a laser, beam splitter, and photomultiplier tube (PMT) reference. The beam test measures the assembled crystal response; the laser setup isolates the photosensor nonlinearity.

For each energy point, the analysis begins with pedestal subtraction and channel gain equalisation. Events with incomplete triggers, abnormal baselines, or impact points near the prototype boundary are rejected. The response of the central cell tests the local dynamic range, while the cluster sum tests containment and inter-channel calibration. Agreement of the single-cell spectrum is especially sensitive to SiPM saturation; agreement of the cluster sum is additionally sensitive to passive material and optical cross-talk. Comparisons at multiple beam energies therefore constrain different parts of the digitisation model.

Figure 3.30 compares the same crystal channel at incident electron energies of 1, 2, 3, and 5 GeV. The DESY beam test used a four-channel LYSO module; the displayed channel contains a $2.5 \times 2.5 \times 5$ cm^3 crystal, with the beam directed at the centre of its end face[55]. Its measured energy is below the beam energy because a single crystal contains only part of the shower. The red and green curves show the simulated energy deposit and the digitised response before SiPM saturation, respectively. They are close because the fluctuations

in energy deposition exceed the additional single-channel resolution smearing. Both curves increasingly overestimate the measured channel energy as the beam energy rises. Including SiPM saturation, shown in blue, shifts the response to lower energy and improves the agreement in the peak position and spectral shape.

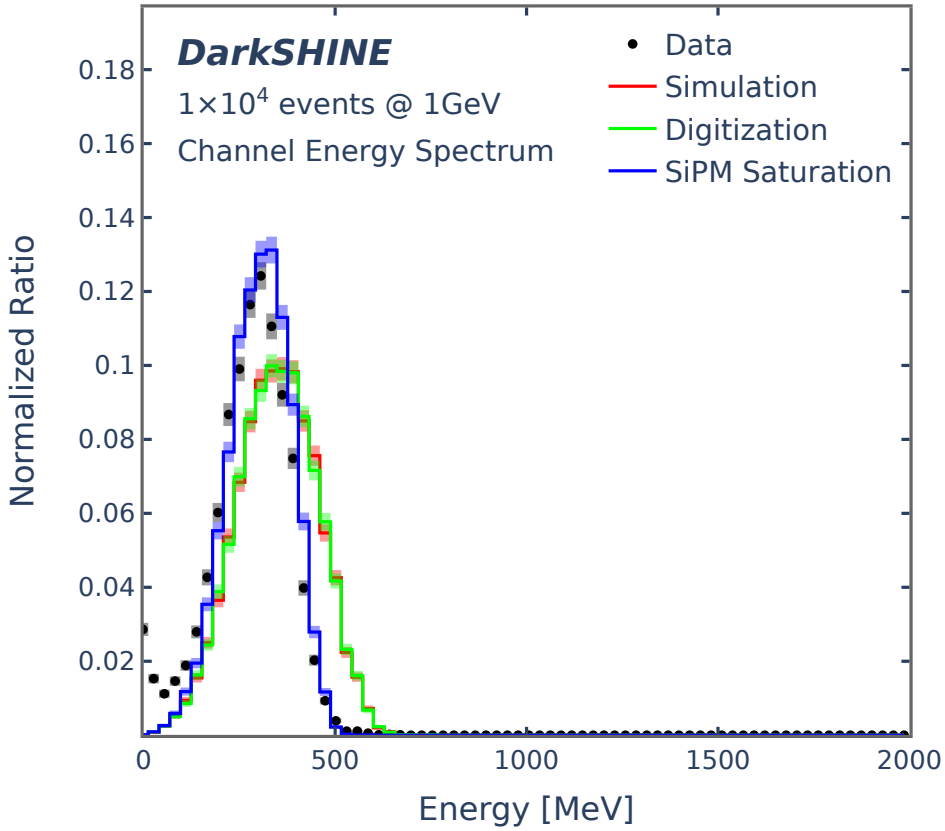


Figure 3.30 Normalised energy spectrum in the same LYSO crystal channel. Black points are beam-test data; red curves are simulated energy deposits; green curves include digitisation before SiPM saturation; blue curves include SiPM saturation. The saturation model improves agreement in peak position and spectral shape. Incident electron energy: 1 GeV. The panel is labelled with 10^4 events.

3.7.5 Radiation environment and calibration stability

The small beam spot concentrates dose in the central ECAL cells and fluence in the photosensors. Ionising dose can reduce crystal transparency and change electronics characteristics, while non-ionising energy loss increases the SiPM dark current and noise. The maps

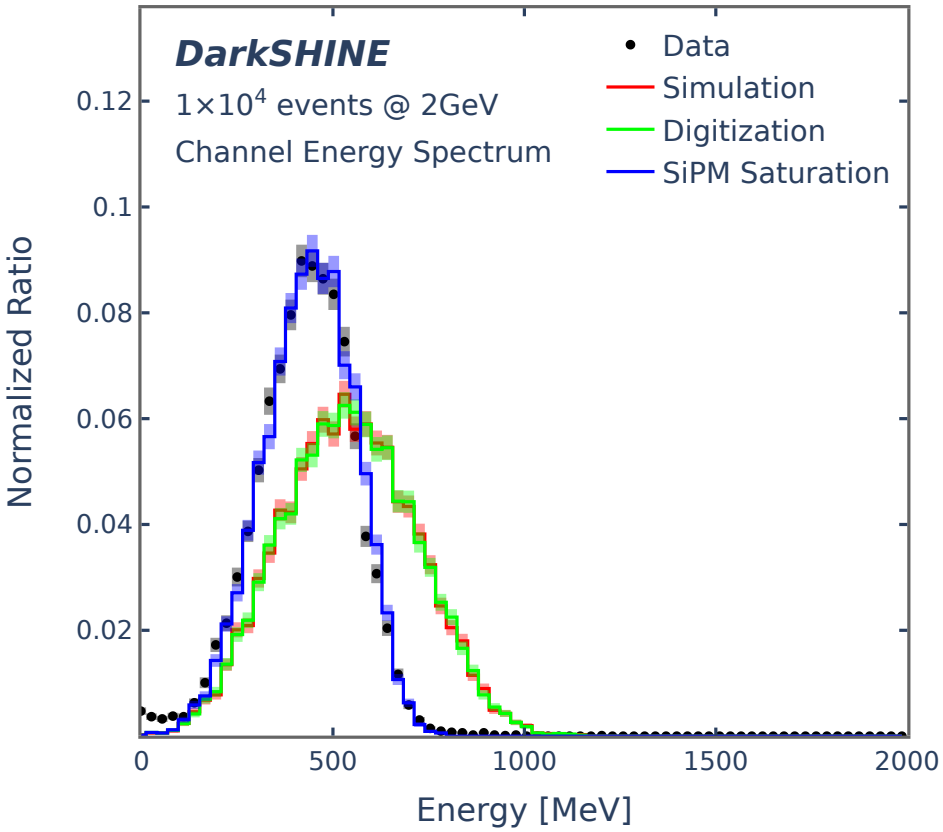


Figure 3.30 (continued) Normalised energy spectrum in the same LYSO crystal channel. Black points are beam-test data; red curves are simulated energy deposits; green curves include digitisation before SiPM saturation; blue curves include SiPM saturation. The saturation model improves agreement in peak position and spectral shape. Incident electron energy: 2 GeV. The panel is labelled with 10^4 events.

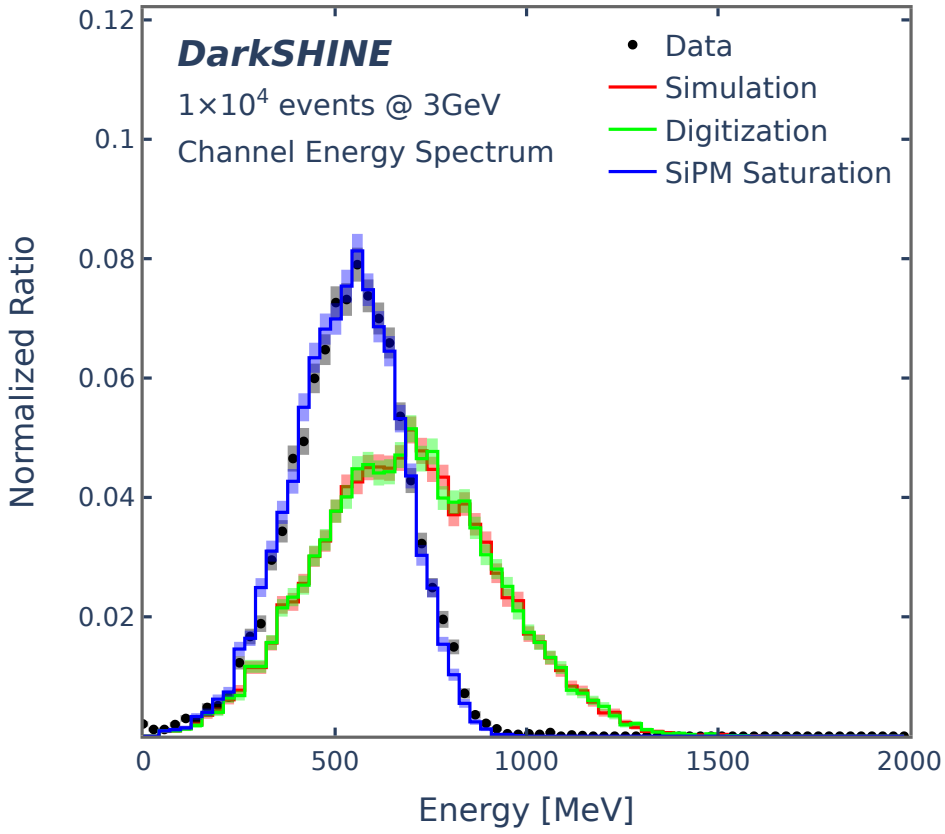


Figure 3.30 (continued) Normalised energy spectrum in the same LYSO crystal channel. Black points are beam-test data; red curves are simulated energy deposits; green curves include digitisation before SiPM saturation; blue curves include SiPM saturation. The saturation model improves agreement in peak position and spectral shape. Incident electron energy: 3 GeV. The panel is labelled with 10^4 events.

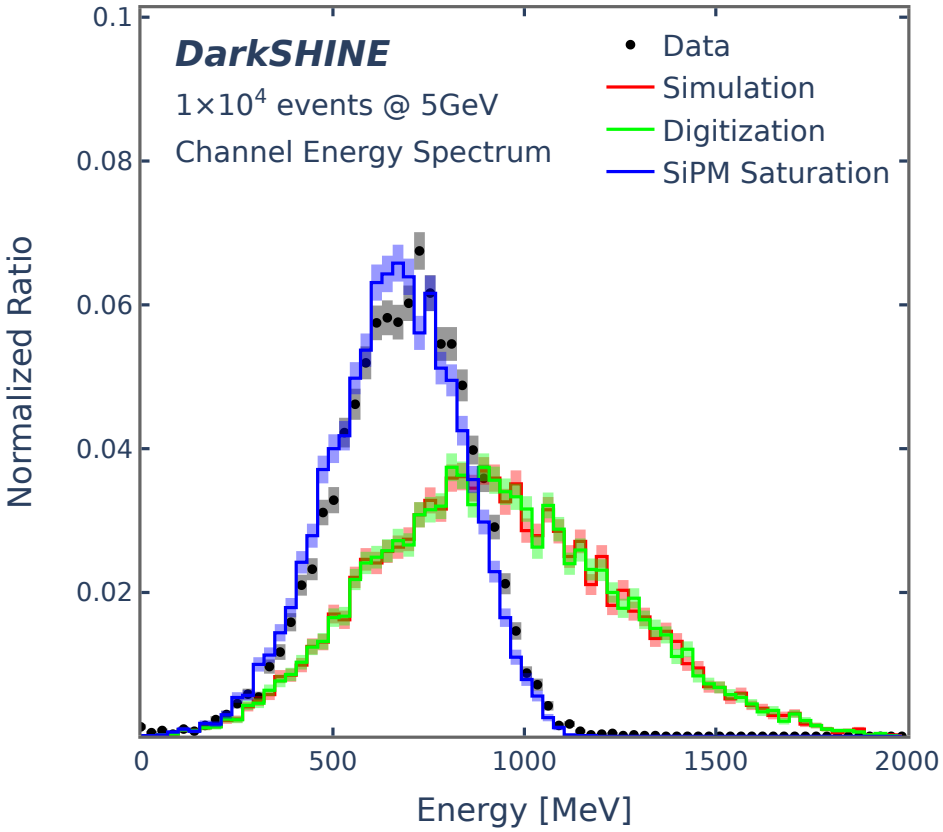


Figure 3.30 (continued) Normalised energy spectrum in the same LYSO crystal channel. Black points are beam-test data; red curves are simulated energy deposits; green curves include digitisation before SiPM saturation; blue curves include SiPM saturation. The saturation model improves agreement in peak position and spectral shape. Incident electron energy: 5 GeV. The panel is labelled with 10^4 events.

in Fig. 3.31 describe the simulated distributions of ionising dose and a non-ionising damage proxy for the design exposure. They identify the regions in which response stability requires particular attention.

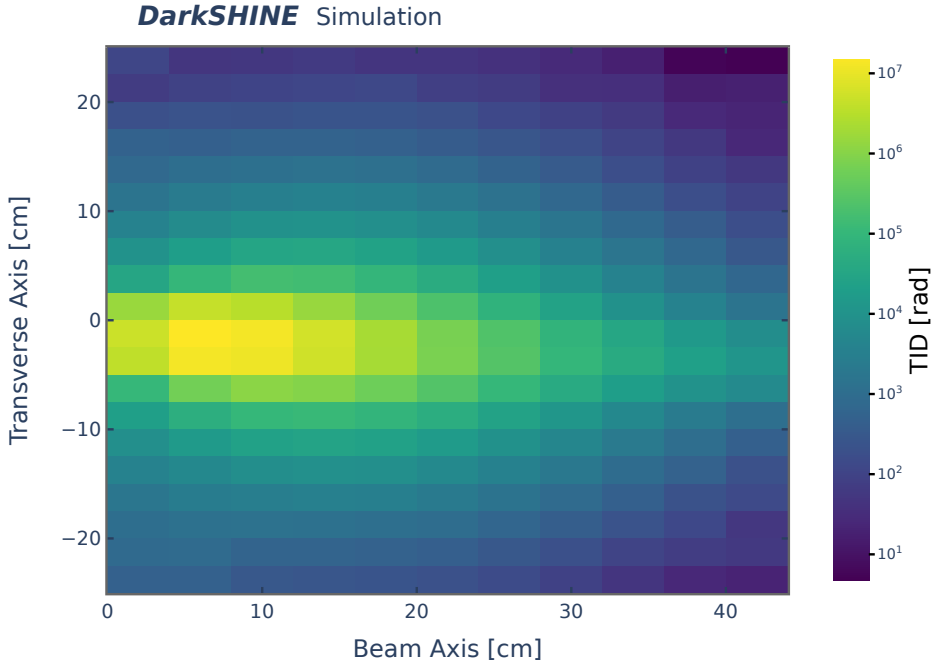


Figure 3.31 Simulated ECAL radiation exposure for 3×10^{14} EOT, versus beam-axis and transverse position. Total ionising dose (TID) in the crystals, in rad. The central concentration motivates spatially dependent monitoring and calibration.

Radiation effects would enter an operating experiment through time-dependent detector response and calibration. Crystal transparency changes the light-collection efficiency, while increasing dark current broadens the noise distribution and may force a higher zero-suppression threshold. Both effects can bias the tails of total and maximum-cell energy used in the signal selection. A calibration plan therefore requires frequent pedestal runs, optical or charge injection, monitoring with intrinsic LYSO activity, and periodic physics samples with known energy. Calibration constants must be stored with validity intervals so that simulated conditions can be matched to each data period.

3.7.6 HCAL active unit and optical readout

The HCAL uses alternating iron absorbers and plastic-scintillator strips. A charged secondary crossing a strip produces scintillation photons that are captured by a wavelength-

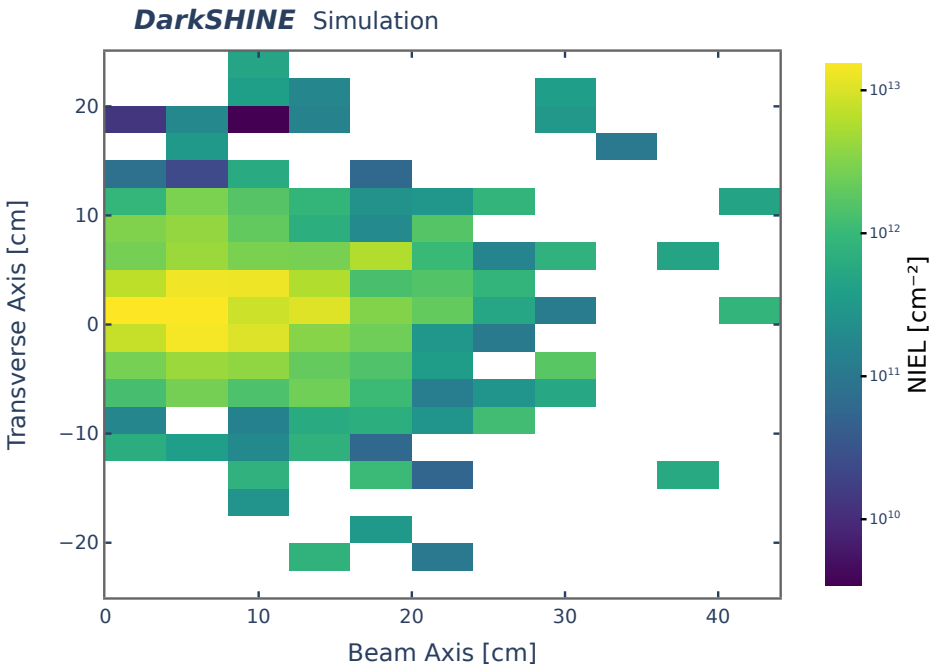


Figure 3.31 (continued) Simulated ECAL radiation exposure for 3×10^{14} EOT, versus beam-axis and transverse position. Non-ionising energy-loss (NIEL) damage proxy at the photosensors, expressed as a 1 MeV-neutron-equivalent fluence in cm^{-2}. The central concentration motivates spatially dependent monitoring and calibration.

shifting fibre and transported to a SiPM. Alternating strip orientation between layers provides coarse two-dimensional localisation and reduces projective gaps. The longitudinal segmentation is chosen to detect both prompt charged particles and spatially extended neutron-induced activity.

Several unit geometries were compared to optimise light yield and uniformity. The fibre groove changes the optical path and mechanical assembly, while scintillator and reflector materials affect attenuation and collection efficiency. Cosmic-ray tests provide a convenient minimum-ionising-particle signal for these comparisons. A telescope of trigger counters selects tracks, and the response is measured as a function of position along the strip. The mean photoelectron yield constrains veto efficiency; its spatial variation determines whether a single threshold can be used over the full strip length.

Scanning the trigger position varies the optical path from the energy deposit to the photosensor and measures light-collection uniformity along the strip. The groove and material comparisons in Fig. 3.33 measure photoelectrons per MeV as the number of fibres increases.

The measured yield rises approximately linearly between one and three fibres; the comparison also tests strip thicknesses and the HND, EJ, and HTX scintillators.

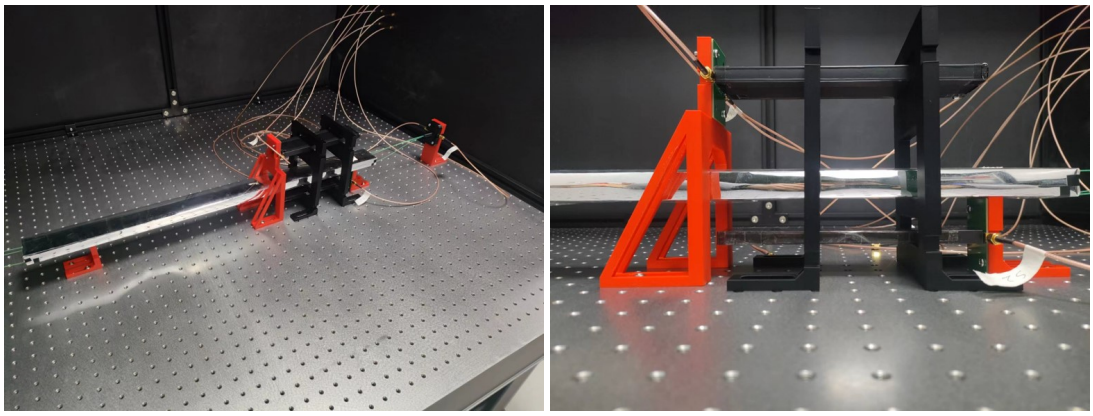


Figure 3.32 Cosmic-ray test of HCAL scintillator strips. Left: overall setup; right: close-up of the test strip and trigger assembly. Two small counters select through-going particles in coincidence, allowing the light yield to be measured at selected positions along the strip.

The complete readout chain must preserve small delayed signals following a large prompt pulse. This requirement influences recovery time, waveform window, and the handling of afterpulsing. Dark noise creates isolated hits, but a neutron candidate is normally defined using coincidences in space and time across multiple layers.

3.8 Mechanical Integration

The DarkSHINE detector is installed in an existing experimental hall, so the mechanical design must respect the floor loading capacity and hall dimensions while satisfying alignment tolerances for each sub-detector and providing maintenance access to individual components.

The tracker assembly, consisting of the tagging tracker, target, and recoil tracker, is mounted on a horizontal aluminium plate housed inside the magnet bore support box. Liquid cooling is delivered via copper tubes pressed into grooves in the plate. The square Helmholtz coils clamp above and below the support box on a rigid, non-magnetic frame with adjustable brackets that maintain coil parallelism and separation. The magnet and ECAL frames share a common rail system, allowing both to slide longitudinally for maintenance access (Fig. 3.35). The ECAL itself is mounted on a support frame on the same rails; each crystal column is coupled through a SiPM and PCB to a cooling plate behind the array. The HCAL, as the heaviest component, sits on an independent support frame separate from the rail system; its 16 mod-

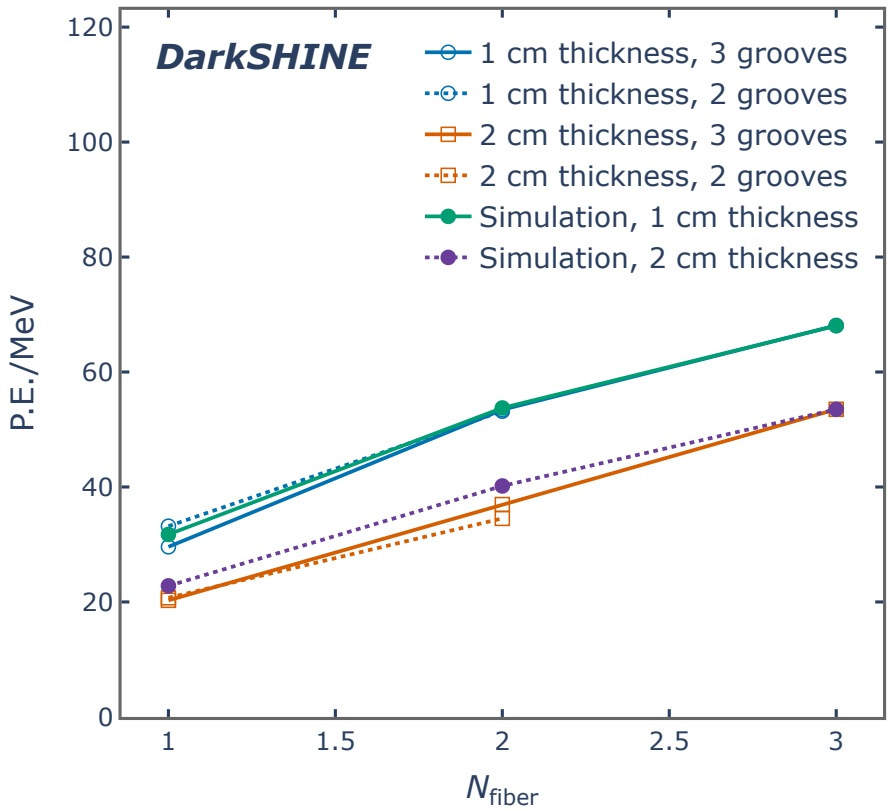


Figure 3.33 HCAL photoelectron yield per MeV versus the number of wavelength-shifting fibres N_{fiber}. Comparison of 1 and 2 cm scintillator thicknesses with two or three grooves, and simulation.

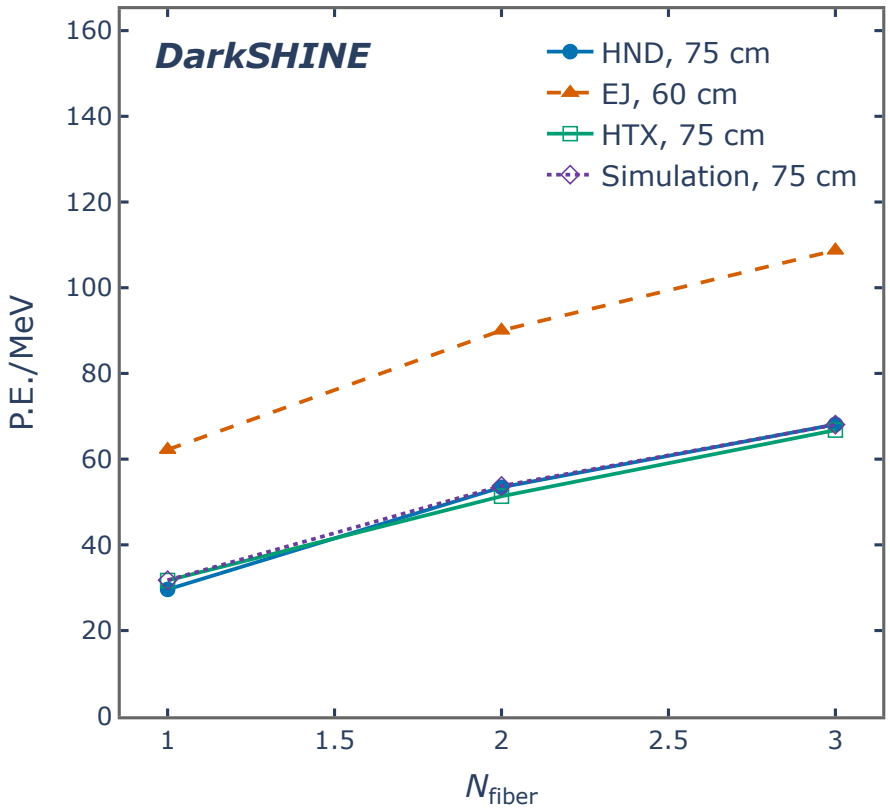


Figure 3.33 (continued) HCAL photoelectron yield per MeV versus the number of wavelength-shifting fibres N_{fiber}. Comparison of HND and HTX scintillators of 75 cm length and EJ scintillator of 60 cm length, with the 75 cm simulation for reference.

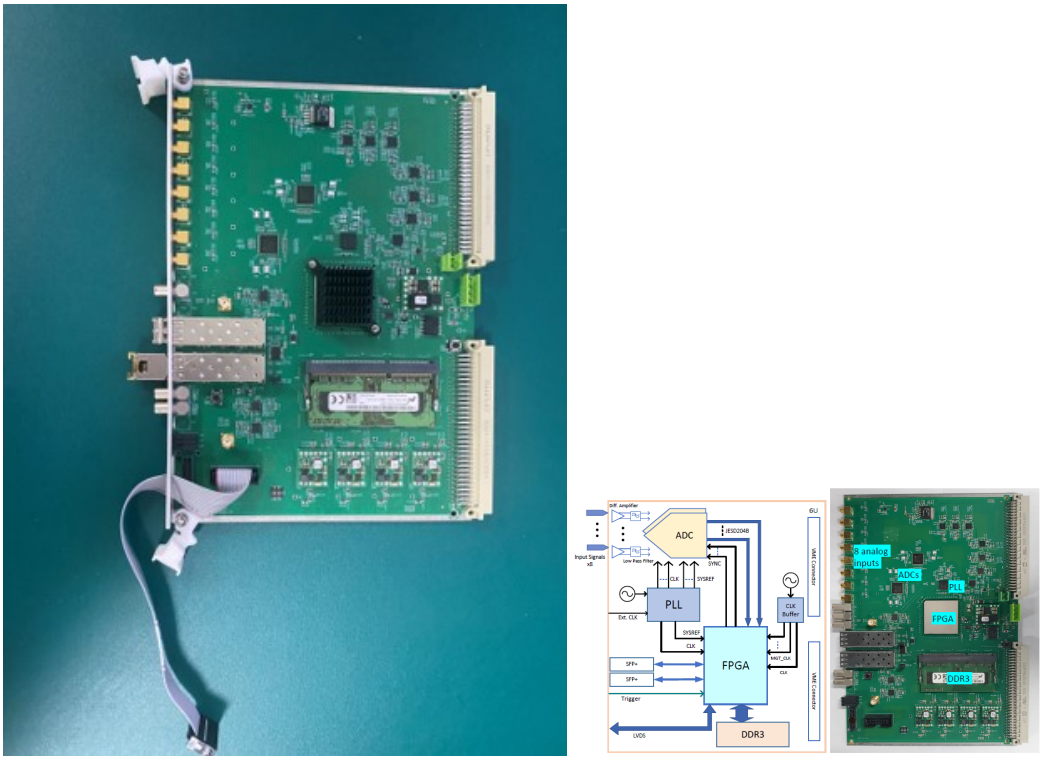


Figure 3.34 HCAL readout electronics. Left: assembled readout board. Right: block diagram and labelled board photograph showing the analogue inputs, ADCs, clock circuit, FPGA, and memory. These stages digitise the photosensor signals and buffer them for event readout.

ules are connected to the frame via dedicated brackets designed for modular replacement (Fig. 3.38).

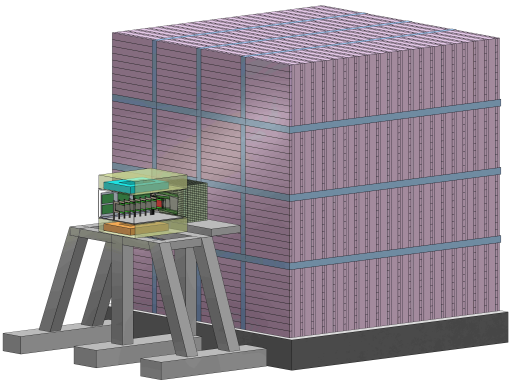


Figure 3.35 Mechanical support of the reference DarkSHINE detector. The tracker, magnet, and ECAL share rails for longitudinal access, while the sixteen-module HCAL rests on an independent load-bearing frame.

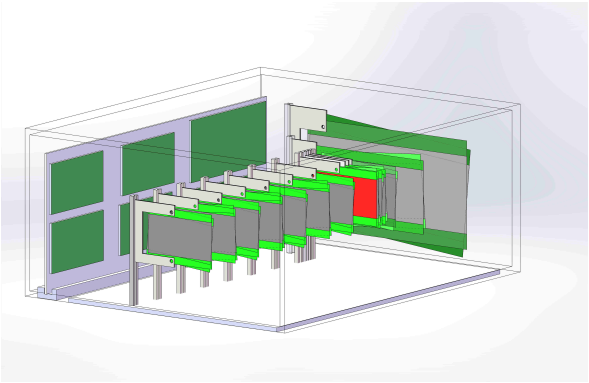


Figure 3.36 Tracker and target support assembly inside the magnet bore. An aluminium plate carries the detector elements, with copper cooling lines supplying thermal control and a surrounding box supporting the assembly.

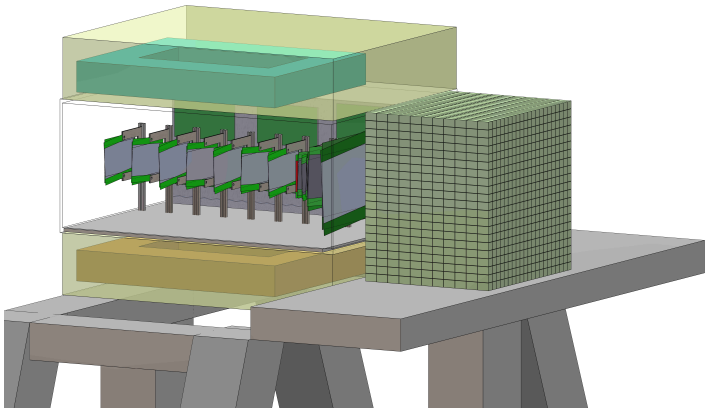


Figure 3.37 ECAL support on the rails shared with the tracker and magnet. The crystal assembly is held above the rail frame, which allows longitudinal movement for maintenance access.

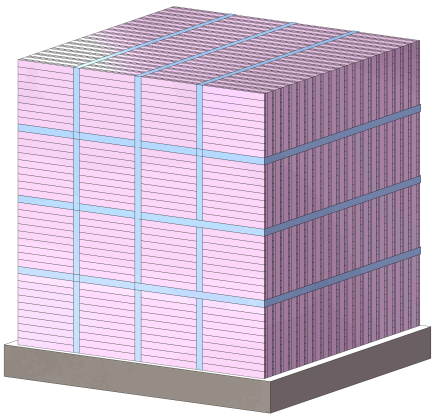


Figure 3.38 HCAL support structure. Sixteen modules are attached by dedicated brackets to an independent frame that carries the calorimeter load and permits modular replacement.

Chapter 4 Event Simulation and Reconstruction in DarkSHINE

A DarkSHINE signal is characterised by a large loss of beam energy and momentum with little visible activity downstream. Predicting the signal efficiency and rare Standard Model backgrounds requires following the interaction products through the detector and modelling their measured response. Reconstruction then converts detector hits into the track momenta and calorimeter observables used in the event selection. The software must support these studies while managing the computing cost of rare-process generation and optical-photon transport.

4.1 The DarkSHINE Simulation Software Framework

The DarkSHINE Software (DSS) is the integrated simulation and analysis framework for the DarkSHINE project. It comprises three principal components:

- `DSimu`: a detector simulation application based on `Geant4`[56] and ROOT[57], implementing the full DarkSHINE detector geometry.
- `DAna`: a modular framework for event reconstruction and data analysis, operating on the ROOT output files produced by `DSimu`.
- `DDis`: a dedicated event display tool for visualisation of simulated and reconstructed events.

The software implementation described in this chapter follows the `master` development branch of the DarkSHINE code.

The geometry and event-data interfaces support detector optimisation by exposing configurable subsystem dimensions, materials, and placement, together with the particle and hit information needed for performance analysis. Geometry scans can then be evaluated through tracking acceptance and momentum resolution, calorimeter energy response, and background-veto performance. The interfaces connect these detector studies to the reconstructed observables used in the physics selection.

4.1.1 Workflow

DSS applications use ROOT files to exchange event data. A custom C++ class, `DEvent`, stores the particle and hit collections used at each stage, from simulation through digitisation and reconstruction to analysis. Keeping these stages in one framework allows the same event to be followed through the processing chain.

`DSimu` uses `Geant4`, a C++ toolkit for particle transport through matter, to simulate energy deposition, secondary-particle production, and nuclear interactions. The detector geometry is constructed programmatically with explicitly defined materials. Sensitive-detector deposits are recorded in `DEvent` as `SimulatedHit` collections.

The current primary-source manager provides a Geant4 general particle source and a single-particle gun. It also supports sampling a momentum distribution from a ROOT histogram for the particle-gun route. These sources define the incident particle, position, direction, and energy for transport studies. Signal studies can obtain their hard-process kinematics from an external generator, as described in Section 4.5. During transport the primary and its secondaries deposit energy in the active materials. Deposits are stored in `SimulatedHit` collections, while particle records are stored in `McParticle` collections. These records provide Monte Carlo truth for studying the detector response.

In a scintillator detector, light production, collection, and photosensor response connect deposited energy to the measured signal. The software represents the calorimeter response through `CalorimeterHit` collections. These can be formed from a detailed optical response or from parameterised hit-energy smearing. Reconstruction selects the collection appropriate to its response model and retains the simulated-hit information for comparisons.

The simulation output is a ROOT file containing `DEvent` objects with the selected truth and hit collections, together with optional optical information. Optical transport increases the computing cost substantially (Section 4.1.7) and is disabled in the supplied reference configuration. Detailed optical studies and parameterised detector-response production serve different purposes. The available response paths are described in Section 4.2.

`DAna` provides reconstruction and analysis processors. Tracking processors form measurements from simulated strip hits, while calorimeter processors read the selected response-hit collections. The algorithms estimate tracks, calorimeter clusters, and event observables from these inputs. Controlled reconstruction studies can also use truth information, including

the truth-smeared seeding configuration described in Section 4.3.3. The processing stages use a common algorithm-processor interface.

The processors can be individually enabled or disabled, with dependencies defined by the selected chain. A `DAna` configuration specifies their execution order and parameters. Each processor registers the reconstructed quantities needed by the analysis. The ROOT I/O class `EventStoreAndWriter` writes these quantities to the output file.

4.1.2 Data structure: `AnaEvent` and `DEvent`

Both simulation and reconstruction use `DEvent` to store event data. It contains five principal collection classes, separating transport steps, particles, and detector hits (the base class, where applicable, is indicated after the colon):

- **DStep**
- **McParticle** : *DParticle*
- **ReconstructedParticle** : *DParticle*
- **SimulatedHit** : *DHit*
- **CalorimeterHit** : *DHit*

Different detector regions can contain hits of the same C++ class, and alternative response models can produce several calorimeter-hit collections for one event. The processor configuration selects the named collection to read.

In the current interface, `AnaEvent` is a type alias for `DEvent`. The name used by an analysis processor therefore refers to the same event type used by simulation. The public `DEvent` name selects a versioned event class, currently `DEvent_b1_5`. Collection maps are keyed by name and point to vectors of objects of the corresponding type. Figure 4.2 summarises their organisation.

The collection vectors and their elements use raw pointers. The event's initialisation and cleanup routines clear the stored elements and, when requested, the collection containers.

The executable starts from the `FTFP_BERT` physics list, registers the additional photon-physics module and step limiter, and enables dark-photon physics when signal production is selected. Optical transport is optional. The current executable registers `G4Optical Physics` when that option is enabled. The code also retains the dedicated optical and LUT modules used in detailed response studies. The response configuration determines which

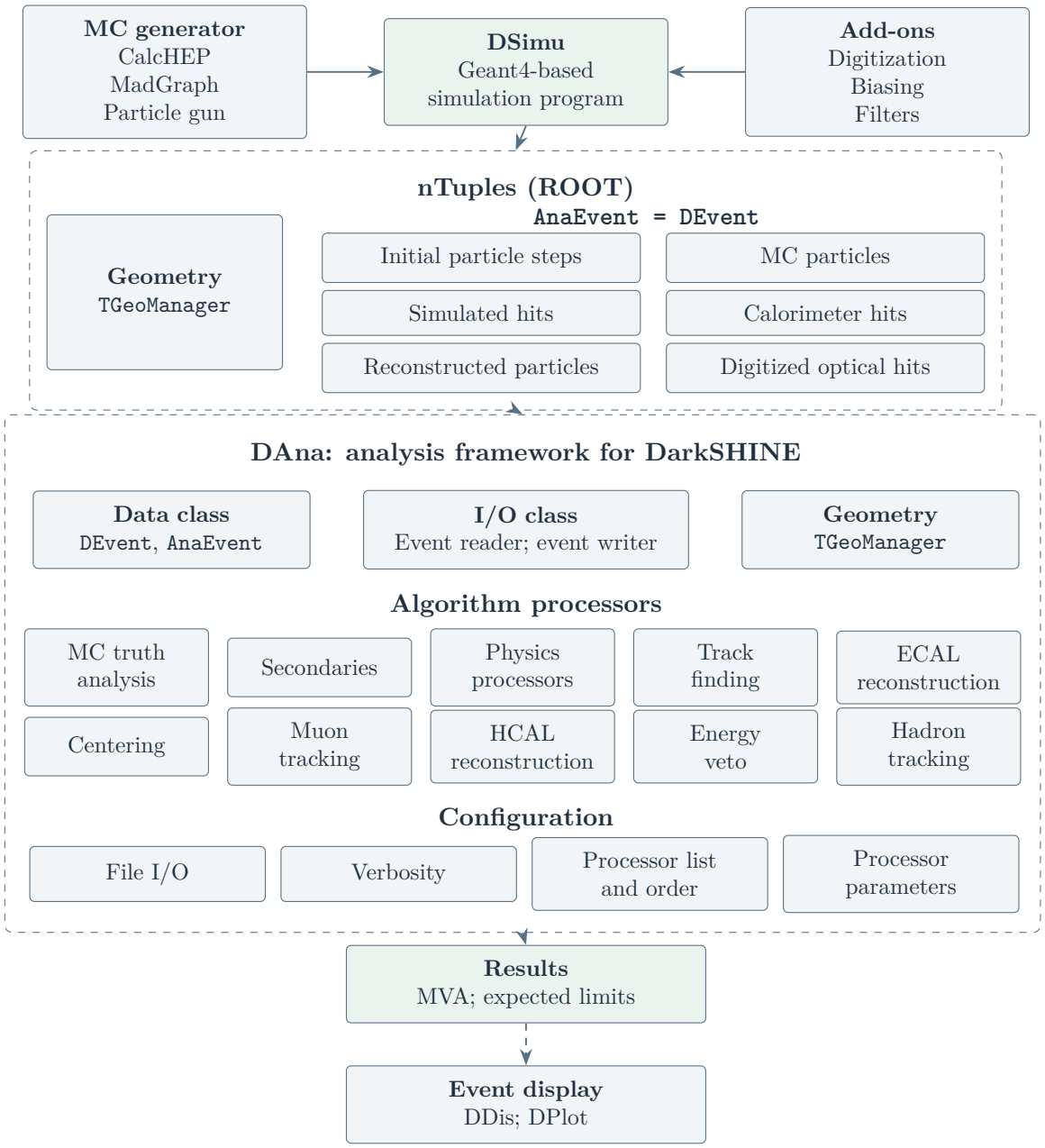


Figure 4.1 DarkSHINE Software (DSS) event-processing workflow. The Geant4-based simulation application `DSimu` writes detector hits and Monte Carlo truth to ROOT files. The reconstruction and analysis application `DAna` reads these records and writes reconstructed observables for physics analysis.

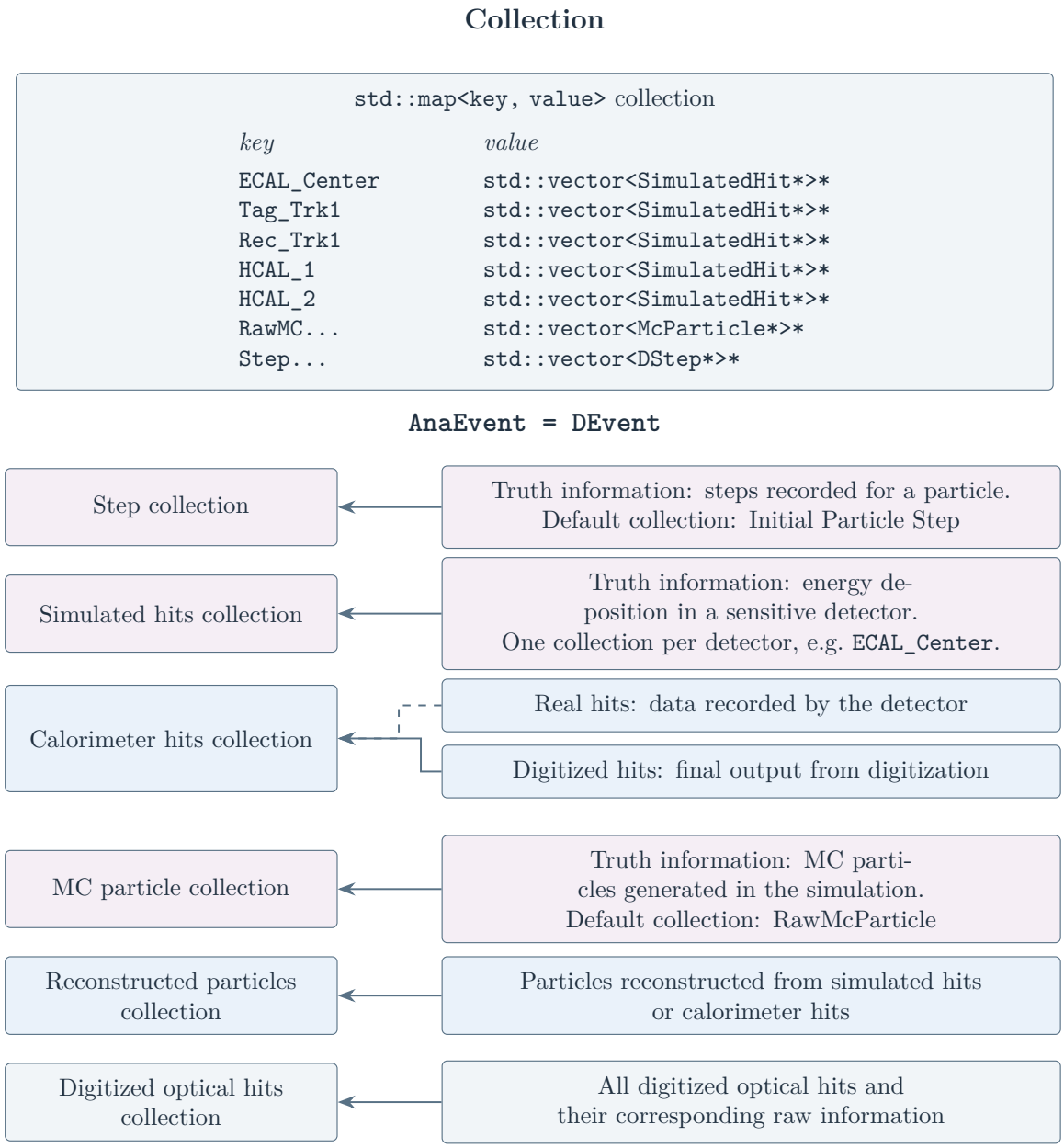


Figure 4.2 Event-data organisation in DarkSHINE Software. A `DEvent` holds named collections of transport steps (`DStep`), simulated particles (`McParticle`), reconstructed particles (`ReconstructedParticle`), detector deposits (`SimulatedHit`), and calorimeter-response hits (`CalorimeterHit`). Simulation and reconstruction use this common event representation.

route is used; the optical benchmarks below describe those studies.

When executed without arguments, `DSimu` runs in interactive (graphical) mode; with command-line arguments, it runs in batch mode. The configuration file specified after the `-y` flag is a YAML file controlling the overall program configuration, while the file after `-m` specifies `Geant4` macro commands (e.g. particle gun settings and number of events).

4.1.3 `DSimu` configuration overview

The YAML configuration file controls several categories of settings: global variables, ROOT output management, `DEvent` collection options, event biasing, event filters, optical photon settings, and detector geometry parameters. The principal options in each category are described below.

Global variables. Table 4.1 summarises the global configuration options, which control geometry export, overlap checking, and a shortcut for dark photon signal production that automatically enables the relevant biasing and filtering.

Table 4.1 Global configuration variables in `DSimu`.

Name	Description	Default
save_geometry	Save geometry as **TGeoManager** in the ROOT file	false
check_overlaps	Enable overlap checking in `Geant4` (computationally expensive)	false
signal_production	Enable dark photon signal production; automatically activates biasing for DMProcessDMBrem, records all MC particles, and applies a process filter on the target region	false

ROOT output management. Table 4.2 lists the output file, tree name, and job run number. The event count is supplied through `general_particle_source.beam_on` or a command-line override.

Table 4.2 ROOT output manager options.

Name	Description	Example
outfile_Name	Output ROOT file name	`dp_out.root`
tree_Name	Tree name in the output file	`Dark_Photon`
Run_Number	Job run number used in event identification	0

DEvent collections. The options in table 4.3 control which Monte Carlo truth information is persisted. Saving all MC particles increases the output size because a single 8 GeV electromagnetic shower can produce up to $O(10^4)$ particles.

Table 4.3 `DEvent` collection options.

Name	Description	Default
save_all_mcp	Save all MC particles	false
save_MC	Save selected MC particles	true
save_initial_particle_step	Save initial particle steps	true
RawMCCollection_Name	Name of MC particle collection	`RawMCParticle`
InitialParticleStep-Collection_Name	Name of initial particle step collection	`Initial_Particle_Step`

Event biasing. Cross-section biasing increases the sampling rate of rare processes that would otherwise require very large inclusive samples. The options in Table 4.4 select a process, the target or ECAL region, and a cross-section scale factor. Only a minimum energy threshold is applied; there is no upper energy cutoff, as the processes of interest remain relevant across the kinematic range above threshold.

Table 4.4 Event biasing options.

Name	Description	Example
if_bias	Master switch for biasing	false
if_bias_target	Enable biasing on the target	false
if_bias_ECAL	Enable biasing in the ECAL	false
BiasProcess	Physics process to be biased	`GammaToMuPair`
BiasFactor	Cross-section scale factor	10^{20}
BiasEmin	Minimum energy threshold [Energy, Unit]	[4, `GeV`]

Event filters. Filters are used in conjunction with biasing to reduce disk usage when generating rare events. Filters can be applied to the primary track (Stage 0) or to secondary tracks and their descendants (Stage 1). The filtering logic is implemented in the `UserSteppingAction`, with `UserStackingAction` managing the track scheduling across three stages: Stage 0 processes only the primary track (pushing secondaries to a waiting stack); Stage 1 processes secondaries within the region of interest; and Stage 2 processes all remaining secondaries if the event passes all filters. If any filter condition is violated, the event is aborted. Table 4.5 summarises the filter configuration.

Table 4.5 Event filter options. Each filter entry specifies the particle PDG ID (or process name), energy range, distance range, and inclusion/exclusion flags for Stage 0 and Stage 1.

Name	Description	Example
if_filter	Master switch for event filters	false
particle_filters_ parameters	Particle filter list: [PDG ID, $E_{\min}$, unit, $E_{\max}$, unit, $d_{\min}$, unit, $d_{\max}$, unit, fInclude, fStage0, fStage1]	See text
process_filters_ parameters	Process filter list (analogous format)	See text

Detector geometry. The software geometry shown in figure 4.3 implements the subsystem layout described in Section 3.1. Table 4.6 summarises its default parameters[55]. Relative to the reference geometry in Chapter 3, the principal changes concern the calorimeters. The ECAL uses $21\times21\times11$ LYSO(Ce) crystals ($2.5\times2.5\times4.0\,\mathrm{cm}^3$ each) read out by SiPMs, arranged in a staggered configuration where successive layers are offset by half a crystal width to eliminate inter-crystal gaps[54]. The HCAL features an optimised non-uniform absorber design: the first 70 layers use 1 cm-thick iron absorbers to efficiently capture low-energy neutrons, while the remaining 18 layers use 5 cm-thick absorbers for high-energy neutron containment, achieving approximately $10\,\lambda$ of total interaction length in a 1.5 m × 1.5 m transverse footprint[58]. A side-HCAL, consisting of four scintillator panels surrounding the ECAL, is incorporated to veto particles exiting the ECAL at large angles. The user can selectively enable or disable each sub-detector. Two magnetic fields along the y-axis are applied in the tagging and recoil tracker regions, respectively, and are independently configurable. Additional tracker layers can be appended by specifying their size and position. The ECAL and HCAL parameters are also fully configurable.

4.1.4 `DSimu` software architecture

`DSimu` follows the `Geant4` state machine shown in Figure 4.4. The **Control** and **FilterManager** classes are singleton instances, initialised before the `Geant4` kernel.

4.1.5 Processing an event through the framework

Initialisation establishes the materials, geometry, field, physics processes, and configuration before the event loop begins. Event processing starts from a fresh event record and a newly generated or imported primary state. During transport, user actions inspect tracks

Table 4.6 Summary of the default DarkSHINE detector parameters, incorporating the updated ECAL[54] and optimised HCAL[58] designs.

Component	Material	Key dimensions
Tagging tracker	AC-LGAD Si strips (7 layers)	$20.1 \times 10\,\mathrm{cm}^2$ per layer
Target	Tungsten ($0.1\,X_0$)	$20 \times 10\,\mathrm{cm}^2 \times 350\,\mu\mathrm{m}$
Recoil tracker	AC-LGAD Si strips (6 layers)	transverse size increases from 20.1×10 to $50.1 \times 20\,\mathrm{cm}^2$
ECAL	LYSO(Ce) crystals	21×21×11 cells (2.5×2.5×4.0 cm^3), staggered
HCAL (main)	Fe / plastic scintillator	1.5×1.5 m^2; 70 layers × 1 cm + 18 × 5 cm Fe
HCAL (side)	Fe / plastic scintillator	4 configurable panels surrounding the ECAL
Tagging B	—	1.5 T ($-y$ direction)
Recoil B	—	mapped field with a characteristic strength of ~0.75 T

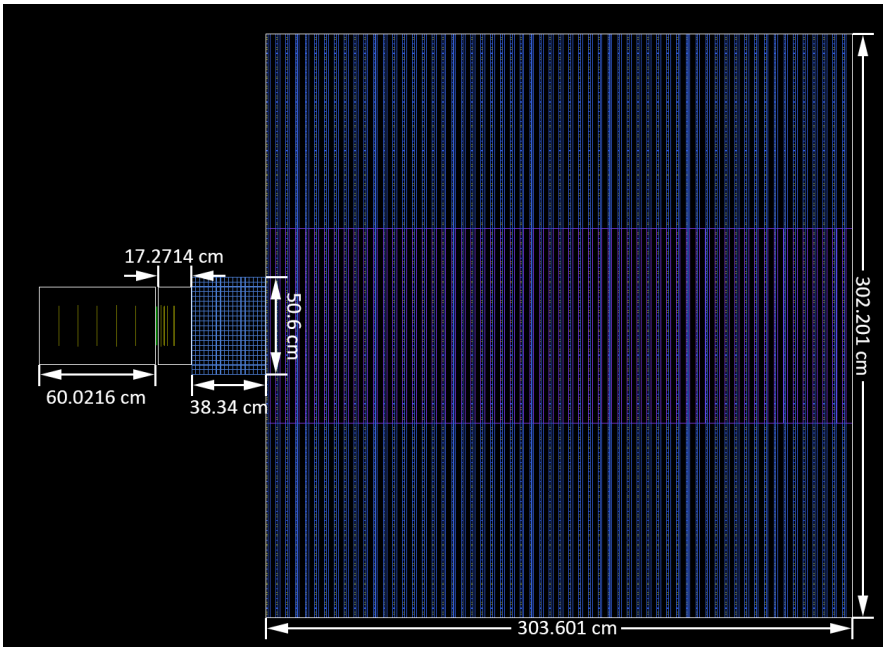


Figure 4.3 Overview of the default detector geometry in `DSimu`. Yellow lines indicate the tagging and recoil trackers, the green line represents the tungsten target, the light blue region is the ECAL, and the dark blue region is the HCAL with crossed X–Y geometry.

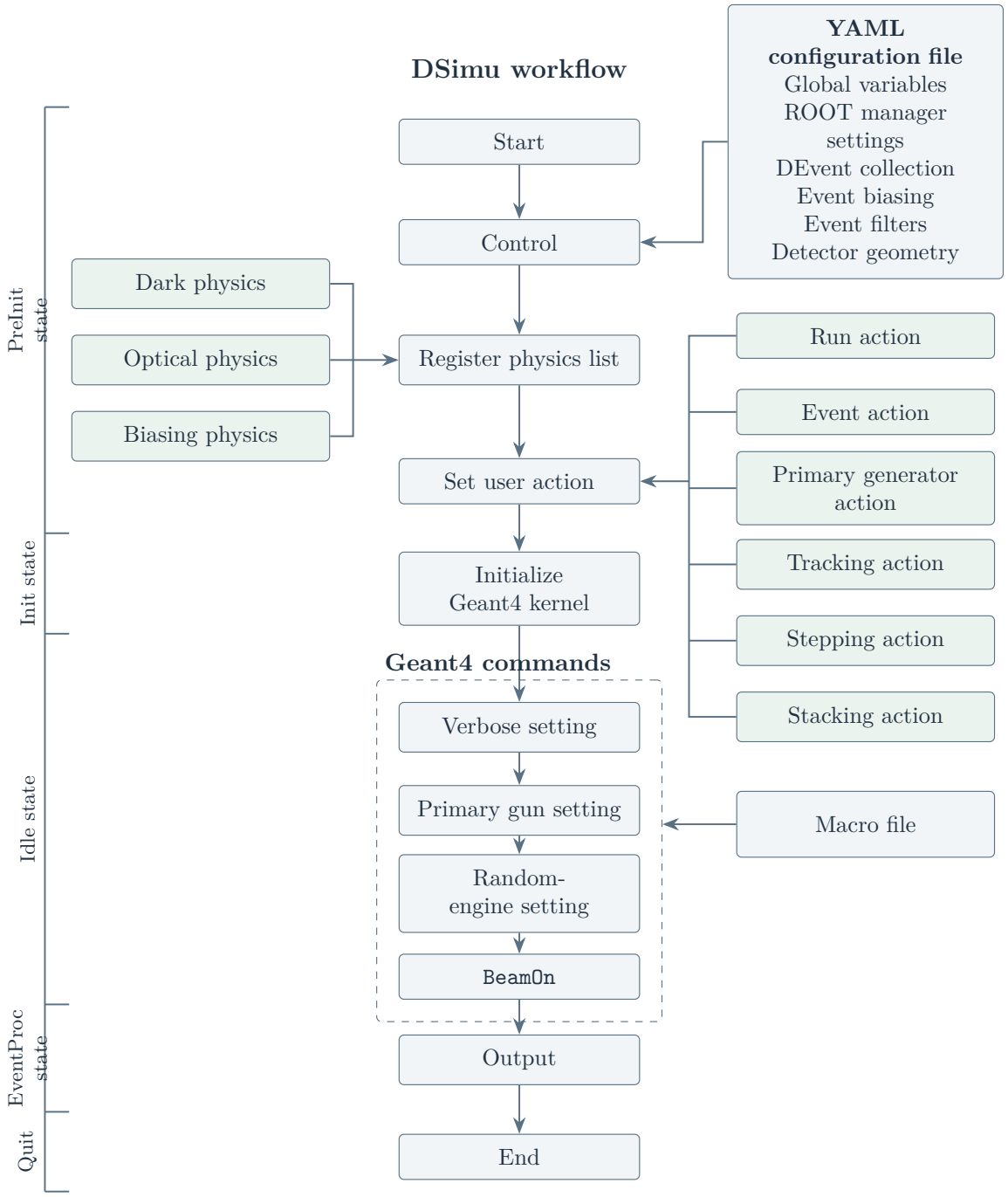


Figure 4.4 Initialisation and event loop in the DarkSHINE simulation application `DSimu`. The singleton `Control` and `FilterManager` objects are created before the Geant4 kernel. Geometry and physics are initialised before transport, during which user actions apply filters and assemble the event output.

and steps, the sensitive detectors accumulate hits, and filters can terminate events that fail the production definition. The event output is assembled from the enabled collections.

Within `DAna`, the `AnaProcessor` interface defines methods for the beginning of processing, event initialisation, event processing, event checking, and the end of processing: `Begin`, `InitEvt`, `ProcessEvt`, `CheckEvt`, and `End`. A processor registers its parameters and output variables, retrieves its input collections, and updates its event quantities.

The analysis writer connects registered variables to ROOT tree branches. Scalars record event quantities such as total energy or track multiplicity; vectors retain the corresponding per-object quantities. Registration specifies the storage address used when the tree is filled. The event loop updates those values before writing the entry.

4.1.6 Staged transport and filter decisions

Staged transport postpones the processing of tracks until they are needed for a production decision. In `StackingAction`, a new track is assigned either to the urgent stack, which can be transported immediately, or to the waiting stack. At Stage 0, tracks with parent identifier zero are urgent and secondary tracks wait. At Stage 1, the region-of-interest test determines which waiting tracks become urgent. At the following stage all tracks are eligible for transport. When the urgent stack is exhausted, `NewStage` increments the stage and requests reclassification. `PrepareNewEvent` resets the stage to zero for the next event.

The region of interest is specified by longitudinal intervals associated with the filter definitions. The track-classification test checks track positions against these intervals to schedule transport. Particle and process filters inspect the resulting steps.

The particle and process filters inspect different properties. A particle filter searches the secondaries produced in a step for a specified PDG identifier and secondary total-energy range. Before this search it requires the pre-step kinetic energy and the magnitude of the step's kinetic-energy change to reach the configured minimum, and checks the post-step longitudinal position. A process filter checks the process defining the step and the kinetic-energy change across it. Both use stage switches to select whether the primary or secondary transport stage is inspected. The process-name comparison also permits the configured process name to be found inside a wrapped name, as occurs when a biasing wrapper is active.

An exclusion filter can reject an event as soon as a forbidden match is observed. The

filter manager stores whether each required match has been found and checks these flags at the configured inclusion-check stage. This is one shared stage; enabling an inclusion requirement for Stage 1 moves its check to Stage 2. With several inclusion filters, every required flag must be set. Per-event initialisation clears the flags, the hard-bremsstrahlung status, and the accumulated ECAL-energy estimate.

The loose energy vetoes provide another opportunity for early termination. The stepping action accumulates the ECAL energy deposited during transport and compares it with the configured production veto. An event whose accumulated deposit already exceeds the upper bound cannot return below that bound by adding further deposits. Its remaining shower need not be transported for that production task.

The computational saving follows from the order of these decisions. A primary that fails a Stage 0 requirement can be rejected before its waiting secondaries are transported. A secondary-process requirement can be evaluated before releasing the remainder of a shower. The retained event still contains the transport needed by its production definition.

Table 4.7 Roles of biasing, staged transport, and filtering in rare-process production. Their effects on the sample must be distinguished when normalising the selected events.

Operation	What it controls	Consequence for interpretation
Cross-section biasing	Sampling frequency of a selected interaction	Requires consistent physical-rate normalisation
Staged transport	Order in which tracks are processed	Allows an earlier production decision
Inclusion filter	Presence of a specified particle or process	Defines a restricted generated sample
Exclusion filter	Absence of a specified particle or process	Removes events outside the production definition
Offline selection	Reconstructed event quantities	Determines the analysis acceptance within the generated domain

4.1.7 Simulation performance

Table 4.8 summarises the computational performance of `DSimu` for representative event types, measured on a single core of an Intel Xeon E5-2680v4 processor at 2.4 GHz. Inclusive 8 GeV electron events produce extensive electromagnetic showers, resulting in $O(10^4)$ secondary particles and a correspondingly large per-event processing time. Enabling optical

photon tracking increases the simulation time by approximately two orders of magnitude; the lookup-table (LUT) approach described in section 4.2 reduces this overhead by a factor of ∼50, making optical simulation feasible for large-scale production. The output ROOT file size is approximately 20–50 kB per inclusive event without optical photons, increasing to ∼200 kB per event with full optical tracking enabled. The memory footprint of `DSimu` is approximately 500 MB for the default detector geometry.

Table 4.8 Simulation performance benchmarks for `DSimu` with an 8 GeV electron beam on the default detector geometry (single-core, Intel Xeon E5-2680v4).

Configuration	Events/min	File size/event
Inclusive (no optical)	∼30	∼30 kB
Inclusive (full optical)	∼0.3	∼200 kB
Inclusive (LUT optical)	∼15	∼50 kB
Signal (dark photon, no optical)	∼50	∼20 kB

4.2 Hit Digitisation

Optical simulation and digitisation translate energy deposits in the ECAL and HCAL into readout signals. The detailed response includes scintillation light yield and decay time, photon transport, SiPM saturation and pixel recovery, electronic noise, finite dynamic range, and ADC discretisation.

For an energy deposit E_{dep} and scintillation yield Y, the mean number of emitted scintillation photons is proportional to YE_{dep}. Let η_{col} denote the light-collection probability and η_{PDE} the photon-detection efficiency of the sensor. In a linear response model the mean detected count is

$$\langle N_{\mathrm{det}} \rangle = YE_{\mathrm{dep}}\,\eta_{\mathrm{col}}\eta_{\mathrm{PDE}}. \tag{4.1}$$

The factors describe different parts of the detector. Collection depends on the crystal, wrapping, interfaces, and emission position. PDE describes the conversion at the photosensor. Pixel occupancy and recovery then determine how the detected photons contribute to the electrical signal. The product in Eq. (4.1) describes the low-occupancy response; the time-dependent digitisation model accounts for the subsequent sensor response.

4.2.1 Optical-photon simulation and lookup table

Optical photons in `Geant4` undergo scintillation and Cherenkov production, absorption, reflection, refraction, and scattering at material boundaries. A single energetic interaction can generate approximately 10^5 optical photons, making full optical tracking prohibitively expensive for large physics samples. A lookup-table (LUT) method is therefore provided for scintillation photons, which account for more than 98% of the optical yield in LYSO. Detection efficiency and transit-time distributions are precomputed as functions of the photon-production position and direction and are sampled during event simulation.

The LUT was generated from 10^4 single-crystal events containing 6.4×10^8 photons. Its detection efficiency agrees with full optical simulation to within 5%, as shown in Figure 4.5. The measured processing times are approximately 200 s per event for full optical tracking and 4 s per event for the LUT method, compared with approximately 2 s without optical photons. The LUT therefore provides a speed-up of about 50 relative to full tracking.

4.2.2 Electronics response and calibration

The electronics model converts detected photons into SiPM and ADC waveforms. It supports both on-the-fly and deferred digitisation and stores three response levels: the detected-photon counts, the waveform after the SiPM response, and the fully digitised ADC samples. The model includes the single-photoelectron pulse shape, pixel gain loss and recovery, pedestal fluctuations, electronic noise, timing jitter, ADC resolution, and saturation. The reference parameters for the illustrated electronics-response study are summarised in Table 4.9. They specify that response study; production configurations can select their own settings.

Table 4.9 Reference parameters of the calorimeter electronics model.

Parameter	Value
Number of SiPM pixels	16×3600
Pixel recovery time constant	50 ns
Sampling interval	1.25 ns
Waveform time window	100 ns
ADC resolution	12 bits
ADC baseline	1750 counts
Clock jitter	0.05 ns
Aperture jitter	0.1 ps

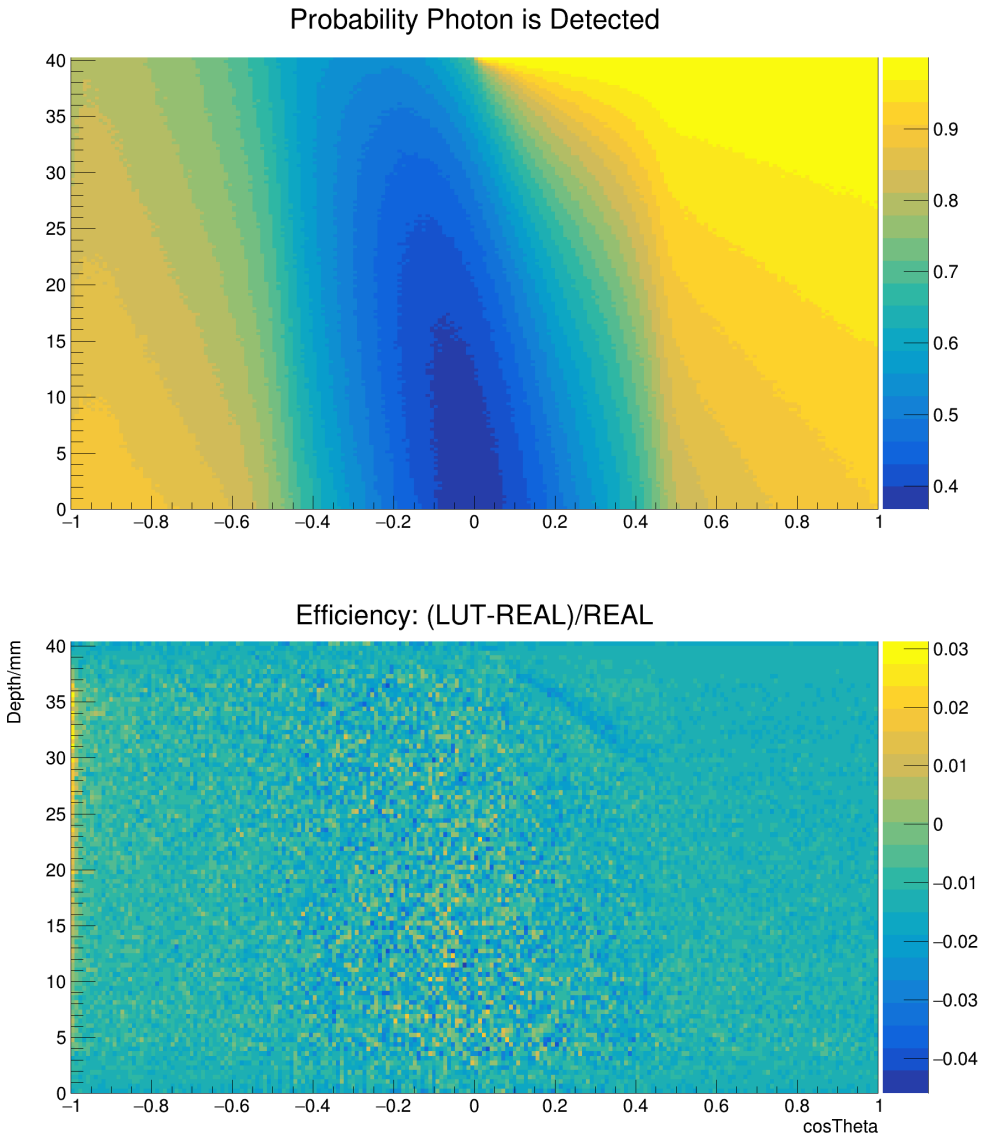


Figure 4.5 Optical-photon detection efficiency in a single ECAL crystal. The first panel shows its dependence on scintillation depth and the cosine of the emission angle. The second shows the relative difference $(\mathrm{LUT}-\mathrm{REAL})/\mathrm{REAL}$**, where LUT denotes the lookup-table response and REAL denotes full optical-photon tracking.**

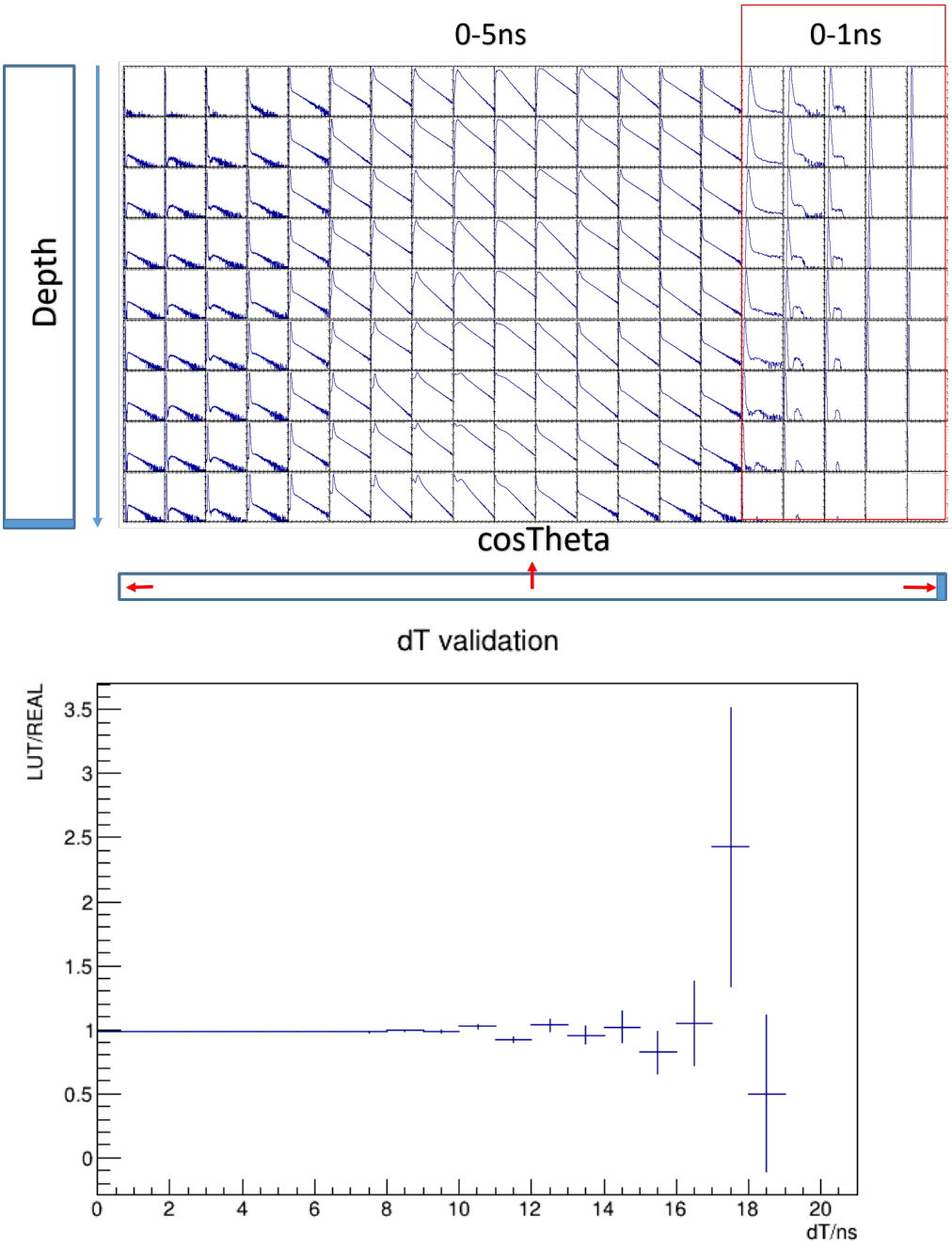


Figure 4.6 Optical-photon transit times in a single ECAL crystal. Here $\mathrm{d}T$ is the time between photon production and arrival at the photosensor, in ns. The first panel shows the transit-time distributions in bins of scintillation depth and emission angle. The second shows the ratio of the marginal $\mathrm{d}T$ distributions: LUT is the lookup-table simulation and REAL is full optical-photon tracking. A ratio of one indicates agreement.

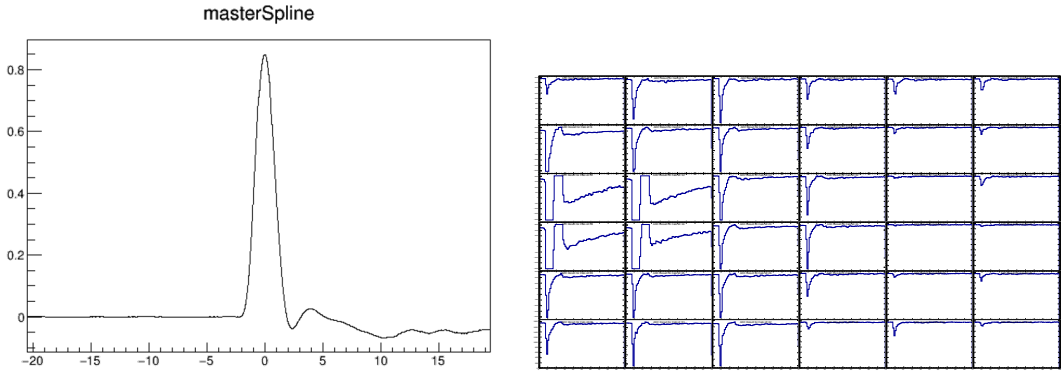


Figure 4.7 Inputs and output of the ECAL electronics-response study. Left: reference single-photoelectron pulse shape used for the silicon-photomultiplier response. Right: digitised waveforms in the central 6×6 ECAL cells, illustrating overflow and undershoot after the sensor and readout response.

Calibration relates the stored optical response to an energy estimator. In the reference ECAL study, the pre-digitised photon-count response was compared with truth energy using 0.1% of the nominal LYSO photon yield. The reported response fit gives a coefficient $c_{\mathrm{fit}} = 0.731 \pm 0.001$. Figure 4.8 shows the corresponding response and its linear fit. The conversion implemented in the current processor is specified in Section 4.2.4.

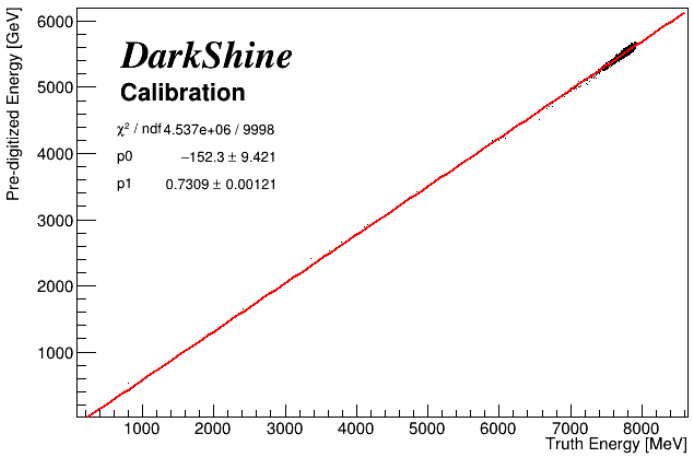


Figure 4.8 Pre-digitised ECAL response versus truth energy in the optical-response study using 0.1% of the nominal LYSO scintillation yield. The red line is a linear fit, with slope $c_{\mathrm{fit}} = 0.731 \pm 0.001$.

4.2.3 Waveform formation and the energy estimator

The single-photoelectron pulse fixes the time dependence of a pixel contribution. If photon detections occur at times t_j, a representation of the analogue signal is

$$V(t) = V_{\mathrm{ped}} + \sum_j g_j h(t - t_j) + n(t), \tag{4.2}$$

where h is the reference pulse, g_j is its response scale, V_{ped} is the pedestal, and n is electronic noise. The scale g_j can depend on the history of the pixel receiving the photon. Thus a sum of pulse shapes can describe a non-linear light-to-signal response: the number and amplitude of the contributions need not remain proportional to the incident light.

Pixel recovery introduces a time scale in addition to the scintillation decay time and the optical transit time. Two photons reaching the same pixel within its recovery period need not produce the charge of two fully recovered pixels. A single integrated photon count loses this ordering information. The waveform model retains photon arrival times so that the sensor response can depend on their separation. The finite acquisition window also limits the contributions included in the signal.

Sampling evaluates the analogue waveform at a sequence of times separated by the configured sampling interval. Timing jitter changes the sampling positions; ADC discretisation changes the values assigned to the sampled voltages. Their effects differ near a rising edge and near a flat pedestal. A baseline offset can shift all samples, while a timing shift changes the sampled pulse shape. The waveform and ADC parameters consequently specify more information than a single resolution applied to the final event energy.

For an integrated estimator, the pedestal-subtracted signal can be written schematically as

$$Q_{\mathrm{win}} = \Delta t \sum_{k \in \mathrm{window}} (V_k - V_{\mathrm{ped}}), \qquad E_{\mathrm{rec}} = C_Q Q_{\mathrm{win}}, \tag{4.3}$$

with calibration constant C_Q. Its value and units depend on the signal convention, sampling interval, and integration window. The photon-count and waveform-integral estimators describe different response levels and require their respective calibration factors.

4.2.4 Parameterised response and stored optical information

The current `Digitizer` processor retains an optical-response path and a `FastSmear` path. They supply different representations of the calorimeter measurement. The optical

path reads a stored `DigiForm` collection, which contains photon-count and waveform information. The parameterised path starts from simulated hits and forms new calorimeter-hit collections with a chosen resolution model. These paths can be configured within the same framework without requiring every physics sample to contain optical photons.

In the stored-optical path, the code defines a conversion multiplier

$$C_{\mathrm{opt}} = \frac{C_{\mathrm{cal}}}{Y_{\mathrm{nom}} F_Y}, \tag{4.4}$$

where C_{cal} is the configurable `Calibration_Factor`, Y_{nom} is `Nominal_Yield`, and F_Y is read from the stored response record. The processor forms estimators from photon counts and waveform integrals using this multiplier. The present optical-hit output uses the scaled photon count; the waveform-based quantities are also calculated in the processor.

The parameterised path loops over the simulated calorimeter hits. For a hit energy E, the implemented width is

$$\sigma_E(E) = \sqrt{A^2 E + B^2 E^2 + C^2}. \tag{4.5}$$

The coefficient units follow the energy convention of the parameterisation. A Gaussian value centred on E is drawn with this width and is restricted to non-negative energy. The output hit retains the input cell identifiers and position. The operation changes the energy response assigned to a transported deposit while preserving its location in the event.

Several parameter sets can be applied to the same input collection. They are written under names ending in `_FS0`, `_FS1`, and so on. The supplied parameter sets include a zero-width response and alternatives with non-zero resolution. Separate switches select processing for the ECAL and for the main and side HCAL. The `RecECAL` input-collection list can then select several of these outputs, retaining the correspondence between an event and each response representation.

Smearing individual cells changes the total energy through their sum and the largest-cell energy through their ordering. Clipping negative draws to zero changes the response near a single-cell threshold.

4.3 Track Reconstruction

Track reconstruction proceeds in two stages. Track finding associates tracker hits with charged-particle candidates, and track fitting estimates their position, direction, momentum,

and covariance. The legacy low-multiplicity chain uses a greedy finder based on the Kasa circle fit[59]. In each iteration, hit combinations containing at most one hit per layer are tested, the candidate with the best fit quality is retained, and its hits are removed before the next iteration.

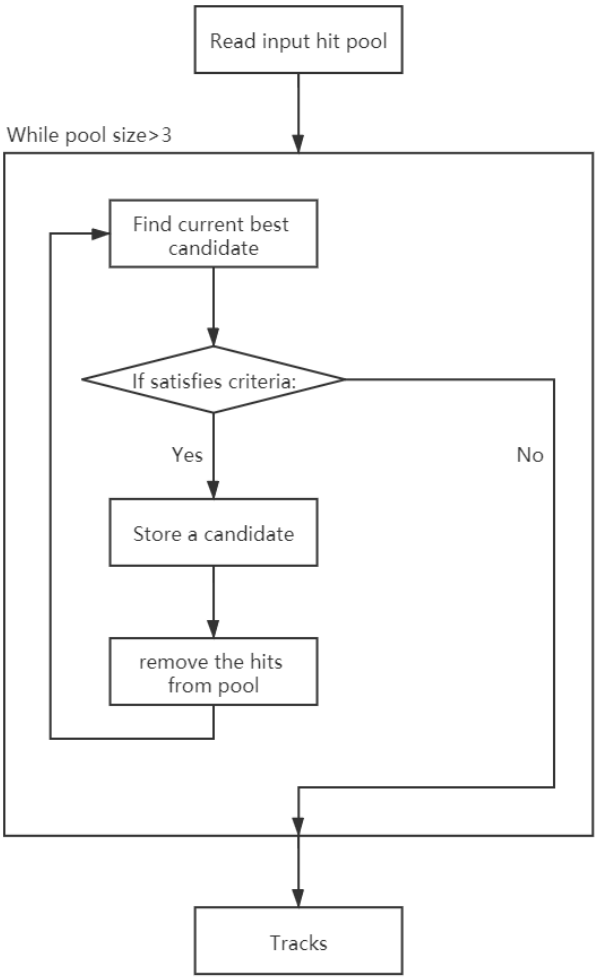


Figure 4.9 Greedy track finding in the DarkSHINE tracker. Hit combinations with at most one hit per layer are ranked with a Kasa circle fit. The best accepted candidate is retained, its hits are removed, and the search repeats until no further candidate passes the finder requirements.

4.3.1 Candidate construction and the circle pre-fit

Track finding determines which hits belong together before a momentum fit can be performed. With n_ℓ available hits in layer ℓ, the number of combinations choosing one hit in each of L layers is $\prod_{\ell=1}^{L} n_\ell$. This dependence explains why hit multiplicity matters even for

a detector with few layers. A sparse single-track event presents a much smaller assignment problem than an event with the same number of layers and several hits in each one.

For a field approximately along y, the bending projection is the x–z plane. Write a circle in that plane as

$$x_i^2 + z_i^2 + ax_i + bz_i + c = 0. \tag{4.6}$$

The Kasa pre-fit determines a, b, c by minimising the sum of squared algebraic residuals. With a matrix M whose rows are $(x_i, z_i, 1)$ and a vector d whose entries are $-(x_i^2 + z_i^2)$, the least-squares equations are

$$M^{\mathrm{T}} M \begin{pmatrix} a \\ b \\ c \end{pmatrix} = M^{\mathrm{T}} d. \tag{4.7}$$

The fitted centre and radius follow from completing the square:

$$x_c = -\frac{a}{2}, \qquad z_c = -\frac{b}{2}, \qquad R = \frac{1}{2}\sqrt{a^2 + b^2 - 4c}. \tag{4.8}$$

This formulation reduces a circle estimate to sums over the hit coordinates. It is suitable for comparing many candidate combinations before carrying out a fit that includes the material and field description.

The legacy greedy finder ranks the resulting candidates using a radial goodness. For distances r_i from the fitted centre, the implemented quantity is

$$G = 1 - \sqrt{\frac{\sum_i (R - r_i)^2}{N R^2}}. \tag{4.9}$$

Hits lying on a common fitted circle give $G = 1$; radial deviations lower the value. The finder also tests consistency in the non-bending coordinate.

Once a candidate passes the finder requirements, its hits are removed from the pool and the search repeats. This is the greedy step: an accepted assignment constrains the choices available for later tracks. The method is therefore sensitive to ambiguous or shared-hit configurations. Its role in the described low-multiplicity chain is to provide track candidates and initial curvature estimates.

Two fitting approaches are implemented. The GenFit-based Kalman filter propagates a track state and covariance between detector layers while accounting for magnetic-field curvature, energy loss, and multiple scattering[60-61]. An analytical Riemann fit provides a faster

alternative for simple trajectories. The Kasa fit is used primarily for candidate ranking because it is fast but does not provide the complete material treatment required for the final momentum estimate.

4.3.2 State propagation and measurement updates

A fitted track contains a state estimate and its uncertainty at a specified reference surface. Denote the state by $\mathbf{x}_k$ and its covariance by P_k at detector layer k. Propagation from the preceding layer gives a predicted state $\mathbf{x}_{k|k-1}$ and covariance

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^{\mathrm{T}} + Q_k, \tag{4.10}$$

where F_k is the linearised transport Jacobian and Q_k represents process noise. The field determines the curved trajectory used by the propagation. Material changes its mean momentum and introduces fluctuations.

For measured coordinates $\mathbf{m}_k$ with covariance V_k, let $h_k(\mathbf{x})$ be the predicted measurement and H_k its local Jacobian. The innovation and its covariance are

$$\mathbf{r}_k = \mathbf{m}_k - h_k(\mathbf{x}_{k|k-1}), \tag{4.11}$$

$$S_k = H_k P_{k|k-1} H_k^{\mathrm{T}} + V_k. \tag{4.12}$$

The Kalman gain and updated state can then be written

$$K_k = P_{k|k-1} H_k^{\mathrm{T}} S_k^{-1}, \tag{4.13}$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + K_k \mathbf{r}_k. \tag{4.14}$$

The GenFit interface constructs planar measurements from the hit positions and their two-dimensional covariance matrices. The initial curvature provides a momentum seed, while the material and field interfaces supply the information for propagation.

For inclusive 8 GeV electron events with typically one to three tracks in each tracker region, the combined finding and fitting chain requires approximately 10 μs per event on one core of an Intel Xeon E5-2680v4 processor. Preliminary single-track and low-multiplicity samples give a finding efficiency above 95% for tracks with at least four hits and a fake rate below 2%.

4.3.3 ACTS integration in the reconstruction framework

ACTS is integrated through two interfaces. The independent `ActsSequencer` processor exchanges data between a DarkSHINE event and an ACTS algorithm sequence. The

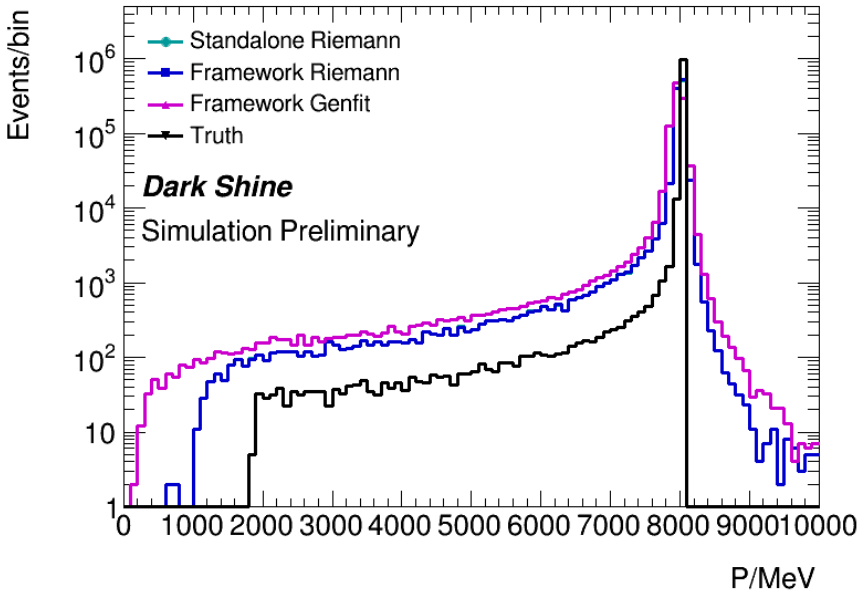


Figure 4.10 Momentum spectra from standalone Riemann fitting (cyan), Riemann fitting within the framework (blue), and GenFit Kalman fitting within the framework (magenta), together with the simulated truth spectrum (black). The vertical axis shows events per bin on a logarithmic scale.

existing `Tracking` processor also offers ACTS-related choices for seeding, finding, and fitting while retaining its own hit pools and track objects. These interfaces have separate configuration and output paths. The processor list selects whether the independent sequencer runs, while method parameters select the optional components within `Tracking`.

4.3.3.1 Geometry and measurement conversion

The sequencer creates separate contexts for the tagging and recoil trackers. It reads their layer positions, dimensions, strip counts, thicknesses, and stereo angles from the detector information loaded by the analysis. These quantities configure an ACTS telescope geometry. This geometry supplies the navigation and measurement surfaces.

The interface changes coordinates according to

$$(x, y, z)_{\mathrm{ACTS}} = (z, -x, -y)_{\mathrm{DSS}}. \tag{4.15}$$

Positions and momenta use this mapping, with explicit unit conversions where required. The magnetic-field adapter transforms an ACTS position back to the DarkSHINE frame, queries the shared field description, and transforms the resulting field vector into the ACTS frame. A constant-field option is also available. Using the same transformation for geometry, mo-

menta, and field preserves the sign and direction of the fitted curvature.

For each tracker region, the simulated-hit reader takes the two strip collections. It pairs entries at the same longitudinal layer and reconstructs the second transverse coordinate using the stereo angle. Candidate positions outside the implemented tracker-length check are removed. The converted ACTS hit retains the particle identifier and momentum contribution from the simulated record. The subsequent ACTS digitisation algorithm constructs measurements, source links, and maps to the contributing particles and simulated hits. Gaussian smearing of the local measurement coordinates is configured within the context.

4.3.3.2 Seeding and combinatorial track finding

The reference sequencer configuration uses truth-smeared seeds. Its particle reader imports truth tracks crossing the tracker region and particles created inside it. A truth selector chooses the seed particles, after which the smearing stage constructs initial track parameters. The configured position, direction, and relative-momentum spreads determine the parameter perturbations and starting covariance.

The combinatorial Kalman filter receives the initial parameters together with measurements, source links, tracking geometry, and field provider. Its measurement selection is controlled by a compatibility threshold and a maximum number of candidates on a surface. Allowing several candidates retains alternative hit associations as the trajectory is extended. The implementation requests shared-hit information and provides an option for backward propagation. The resulting tracks are converted into trajectories, followed by adaptive multi-vertex finding. The trajectory converter reads the track-finding output in this sequence; the separately defined Kalman-refitting routine is not part of the configured execution path.

A space-point seeding alternative includes a space-point maker, a seed finder, and an initial-parameter estimation configuration. The parameter-estimation stage still needs to be registered in the executable sequence, and this route awaits validation.

4.3.3.3 Optional components in the existing processor

The `Tracking` interface selects seeding, finding, and fitting separately. Its ACTS seed adapter estimates circle parameters from hit triplets and returns the native DarkSHINE seed objects. The optional finder then refits a circle with the Kasa method and extends the candidate through intermediate layers by selecting compatible hits near the predicted position.

This finder follows the circle-extrapolation implementation in the adapter. It is separate from the combinatorial Kalman filter used by `ActsSequencer`.

The optional `DActsFitter` uses an ACTS magnetic-field stepper and propagator to transport parameters between hit surfaces, with local Kalman measurement updates in the DarkSHINE adapter. It writes momentum components and fit-quality quantities back into the existing track object. Extrapolation through the common fitter interface then supplies the downstream reference information.

4.3.3.4 Output and the calorimeter interface

The independent sequencer writes tagging and recoil track counts, momenta, directions, fit quality, and trajectory summaries. The latter include the number of measurements, outliers, and holes. Tracks are ordered by absolute momentum before their parameter vectors are filled. Recoil-vertex positions are converted back into the DarkSHINE frame. The momentum extraction and ECAL propagation use an electron-charge convention.

Recoil trajectories are propagated to the ECAL entrance to provide reference positions and momenta. These are exposed through the `Acts_ECal_seed_` output variables. The ECAL processor can read them for track-assisted clustering, applying its track-quality selection before passing reference points to the clustering algorithm. This dependence requires the sequencer to run before the calorimeter processor when this matching path is selected. The existing `Tracking` processor writes its own `ECal_seed_` variables; the two sets of names identify separate input paths.

4.4 ECAL Cluster Reconstruction

The staggered ECAL supplies three-dimensional cell energies for shower reconstruction. The current `RecECAL` processor calculates both whole-calorimeter summaries and topological clusters. The summaries retain the observables used by the baseline missing-energy analysis, while the cluster outputs describe spatially separated deposits and optional track associations. For cell energies E_i at positions (x_i, y_i, z_i), the primary observables are

$$E_{\text{total}} = \sum_i E_i, \qquad E_{\max} = \max_i E_i, \tag{4.16}$$

$$f_n = \frac{\sum_{i=1}^{n} E_{\max,i}}{E_{\text{total}}}, \qquad \langle \xi \rangle = \frac{\sum_i E_i \xi_i}{E_{\text{total}}}, \tag{4.17}$$

where $\xi \in \{x, y, z\}$. Second moments and a lateral moment are also calculated to characterise the shower width and depth[62].

Studies with inclusive, photonuclear, and $\gamma \to \mu^+\mu^-$ samples show that the total energy, the fraction contained in the hottest cells, the lateral moment, and the transverse second moments provide useful discrimination. These quantities are stored as reconstructed observables and can be combined in later cut-based or multivariate analyses. The implemented neighbour-based clustering and track-assisted path are described in Section 4.7.2.

4.5 Signal Generation and Generator Validation

The detector simulation accepts externally generated events as well as processes implemented directly in `Geant4`. Matrix-element generators calculate the hard interaction and allow scans over model parameters. A `Geant4`-native process samples the production point during transport, retaining its correlations with energy loss and multiple scattering in the target. The two routes therefore provide different checks of signal modelling. They are compared in overlapping phase space before being used for an acceptance calculation.

4.5.1 External event-generator chain

For dark-photon bremsstrahlung, CalcHEP is used to evaluate the leading-order matrix element and generate events for specified values of $m_{A'}$ and ϵ[63]. The nuclear form factor suppresses momentum transfers for which the coherent approximation is no longer valid[43]. Event records are converted to a standard interchange format and passed to the transport simulation. Each event retains the generator weight, incoming beam definition, hard-process identifier, and final-state four-vectors. These metadata are copied into the event record so that selection efficiencies can be recomputed without rerunning the generator.

The cross section changes rapidly with mediator mass and scales as ϵ^2 in the on-shell invisible benchmark. Samples can therefore be generated at a convenient reference coupling and reweighted to a different value:

$$w(\epsilon) = w(\epsilon_{\mathrm{ref}}) \frac{\epsilon^2}{\epsilon_{\mathrm{ref}}^2}, \tag{4.18}$$

provided that the width remains narrow, the invisible branching fraction is unchanged, and the kinematic distribution is not modified. Reweighting in mass is not performed because

both phase space and the nuclear form factor change. Instead, separate mass points are simulated and efficiencies are interpolated only after checking that the selection response varies smoothly.

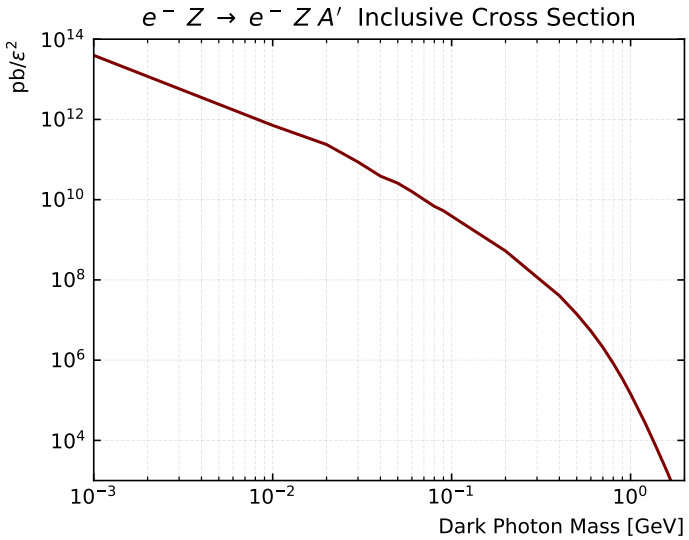


Figure 4.11 Inclusive dark-bremsstrahlung cross section $e^- Z \to e^- Z A'$ on tungsten, computed with CalcHEP for an 8 GeV electron beam. The vertical axis is σ/ϵ^2 in pb; ϵ is the kinetic-mixing parameter. The horizontal axis is the dark-photon mass $m_{A'}$ in GeV.

Generator-level distributions provide the first validation layer. Energy and angle conservation are checked event by event, and one-dimensional projections are compared between mass points. The dark photon is typically forward and carries a large fraction of the beam energy; the recoil electron or positron is correspondingly softer and has a larger deflection. Figure 4.12 shows the principal kinematic variables used to validate the interface. The transverse separation at a downstream plane is particularly useful because it connects the generator calculation to tracker and calorimeter acceptance.

The event-interface validation includes less visible failure modes. Units are checked explicitly when converting between generator and `Geant4` conventions. The beam direction and coordinate handedness are verified with single-particle samples. Particle status codes are translated so that stable generator particles become transport tracks without duplicating intermediate particles. Finally, the sum of transported four-momenta is compared with the generator record before detector effects are enabled.

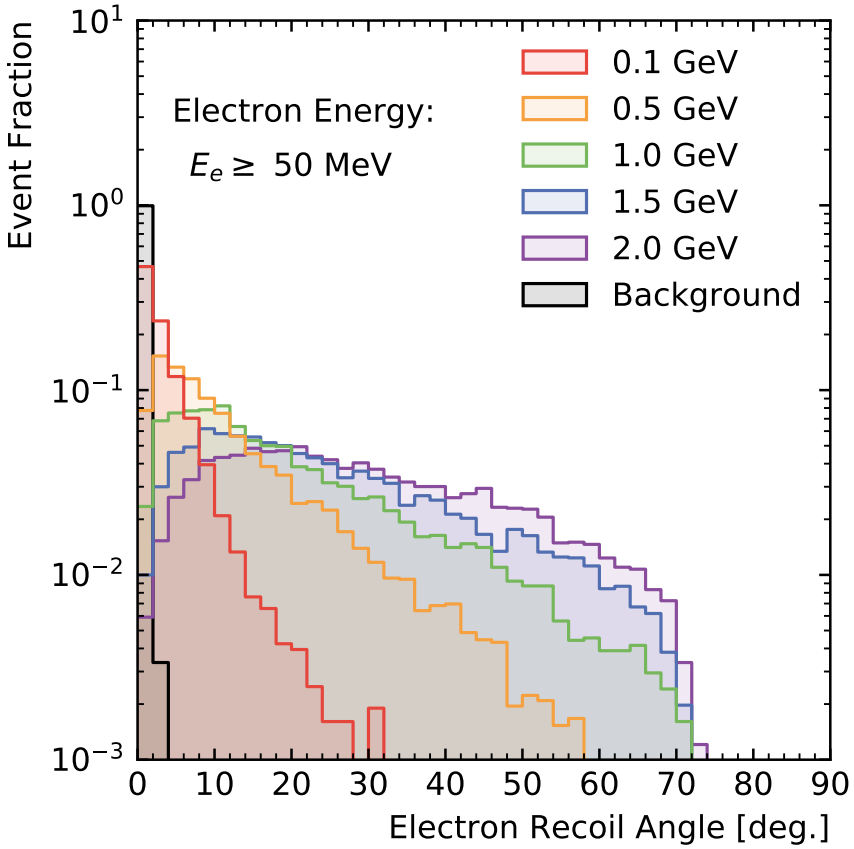


Figure 4.12 CalcHEP dark-bremsstrahlung samples for an 8 GeV electron beam. Signal curves correspond to the dark-photon masses given in the legend. Event fractions compare distribution shapes. Recoil-electron angle, with an ordinary-interaction background sample for comparison.

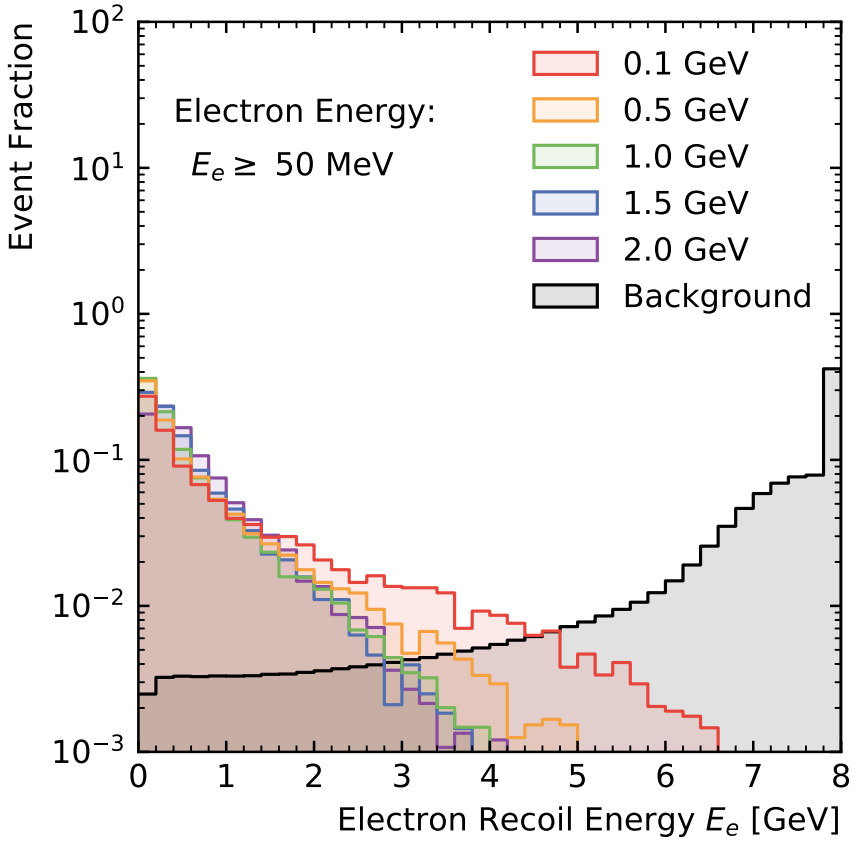


Figure 4.12 (continued) CalcHEP dark-bremsstrahlung samples for an 8 GeV electron beam. Signal curves correspond to the dark-photon masses given in the legend. Event fractions compare distribution shapes. Recoil-electron energy, with an ordinary-interaction background sample for comparison.

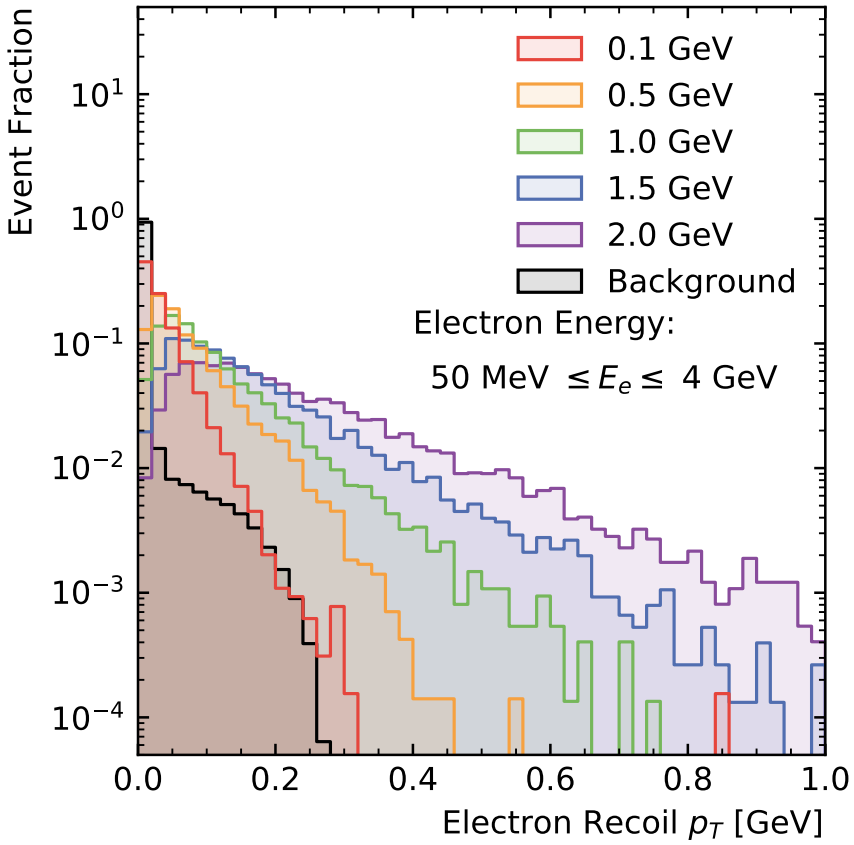


Figure 4.12 (continued) CalcHEP dark-bremsstrahlung samples for an 8 GeV electron beam. Signal curves correspond to the dark-photon masses given in the legend. Event fractions compare distribution shapes. Recoil-electron transverse momentum, with an ordinary-interaction background sample for comparison.

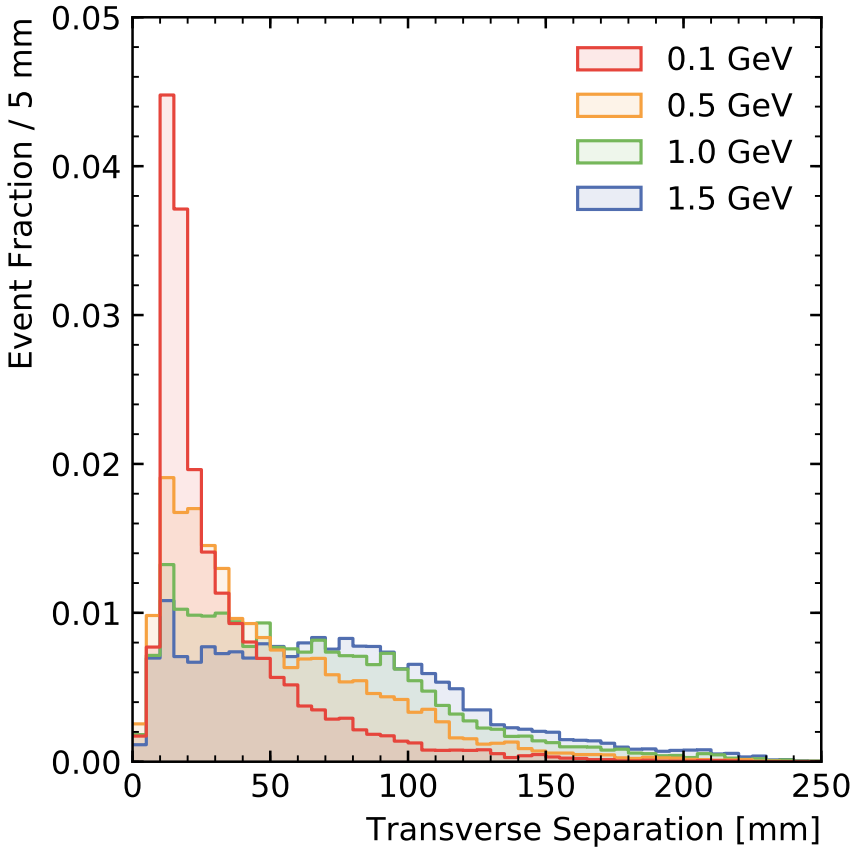


Figure 4.12 (continued) CalcHEP dark-bremsstrahlung samples for an 8 GeV electron beam. Signal curves correspond to the dark-photon masses given in the legend. Event fractions compare distribution shapes. Projected transverse separation.

4.5.2 Transport-native signal process

The `Geant4`-native implementation samples dark bremsstrahlung as a discrete process during transport. Its differential model is based on the improved Weizsäcker–Williams description. For a beam lepton with energy E_0, candidate values of

$$x = \frac{E_{A'}}{E_0}, \qquad \theta_{A'} \tag{4.19}$$

are drawn from an envelope distribution and accepted with probability proportional to the differential cross section. The recoil nucleus absorbs the remaining momentum transfer, while the outgoing lepton and mediator are inserted into the `Geant4` track stack. Sampling the process inside the target automatically includes the distribution of production depth and the incident-particle energy at that point.

An accept–reject generator is valid only if its envelope exceeds the target density throughout the sampled domain. The implementation therefore records the maximum ratio of target to proposal density during validation runs. Boundary samples are included near $x \to 1$, small emission angle, and the kinematic endpoint. A ratio above unity indicates that the envelope condition has failed and the resulting sample is invalid. The integrated rate is separately compared with the reference cross section.

Figure 4.13 compares the transported energy fraction for external and native implementations. The current native model is harder at high $E_{A'}/E_{\mathrm{beam}}$ than the external-generator baseline. This difference can affect signal acceptance whenever the analysis selects the high-missing-energy tail. It therefore identifies a source of generator-model dependence in the acceptance calculation.

4.6 Rare Standard Model Processes

The background prediction for a missing-energy search depends on rare tails of the inclusive electromagnetic sample. DarkSHINE uses cross-section biasing and staged filters to generate enough events in those tails. Their physical normalisation requires suitable event weights, comparisons with unbiased samples, and external validation wherever data exist.

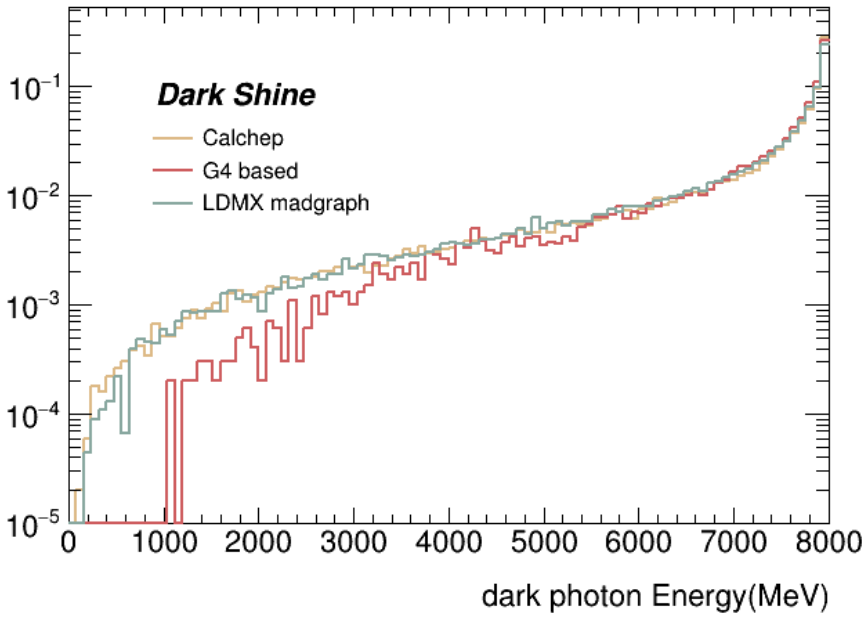


Figure 4.13 Dark-photon energy distributions from CalcHEP, MadGraph5, and the Geant4-native generator for an 8 GeV electron beam. Generated spectra versus dark-photon energy $E_{A'}$, in MeV. The Geant4-native model gives a harder high-energy tail, changing the recoil kinematics relevant to missing-energy selections.

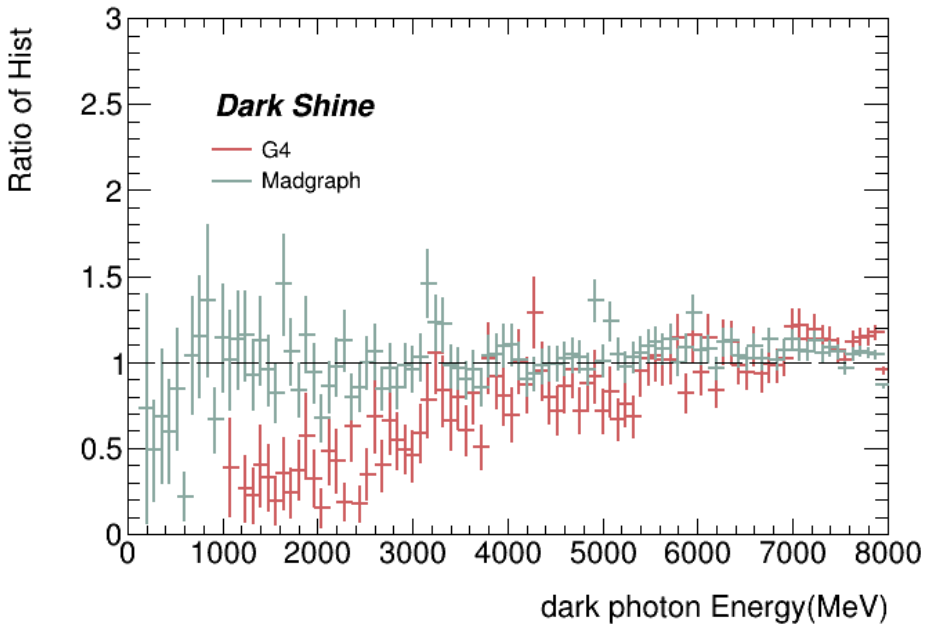


Figure 4.13 (continued) Dark-photon energy distributions from CalcHEP, MadGraph5, and the Geant4-native generator for an 8 GeV electron beam. Ratios to the CalcHEP spectrum versus dark-photon energy. The Geant4-native model gives a harder high-energy tail.

4.6.1 Photonuclear and electronuclear interactions

In a photonuclear interaction, a real photon is absorbed by a nucleus and initiates a hadronic cascade. Electronuclear interactions are related processes initiated by a virtual photon from the beam lepton. They can mimic a dark-photon event when a hard bremsstrahlung photon transfers most of the beam energy to neutral hadrons and the remaining visible particles evade the calorimeter veto. The background rate thus depends on both the hard-interaction model and the detector response to neutrons and other hadrons.

The transport model combines parameterised cross sections with intranuclear cascade, precompound, evaporation, and high-energy string descriptions over different energy ranges. Transitions are checked for discontinuities in inclusive rates and final-state multiplicities. Validation observables include recoil momentum, neutron multiplicity, angular distributions, and visible-energy fraction.

Validation against electron-scattering data is performed by reproducing the beam and target definition and selecting the same recoil-particle phase space. The CLAS setup used for one such comparison is illustrated in Figure 4.14. The comparison is made after detector acceptance and binning are applied to the simulation.

Figure 4.15 compares the transport-process composition of the recoil-proton sample and its momentum spectrum with data. The bias factor changes the sampling frequency while retaining the underlying physical model. Residual disagreement with the data therefore contributes to uncertainty in the predicted background yield.

4.6.2 Muon-pair conversion

Conversion of a hard photon into a muon pair in the nuclear field is much rarer than ordinary pair conversion but is dangerous when one or both muons traverse poorly instrumented regions. Its threshold and nuclear form-factor dependence differ from electron-pair conversion, so it is implemented as a dedicated process[65-67]. The relevant validation quantities include the pair invariant mass, energy sharing, opening angle, and momentum transferred to the nucleus.

The two-muon momentum correlation determines whether both particles can be vetoed by the HCAL. A very asymmetric event may contain one easily identified muon and a much softer partner, while a small opening angle can place both tracks in the same detector re-

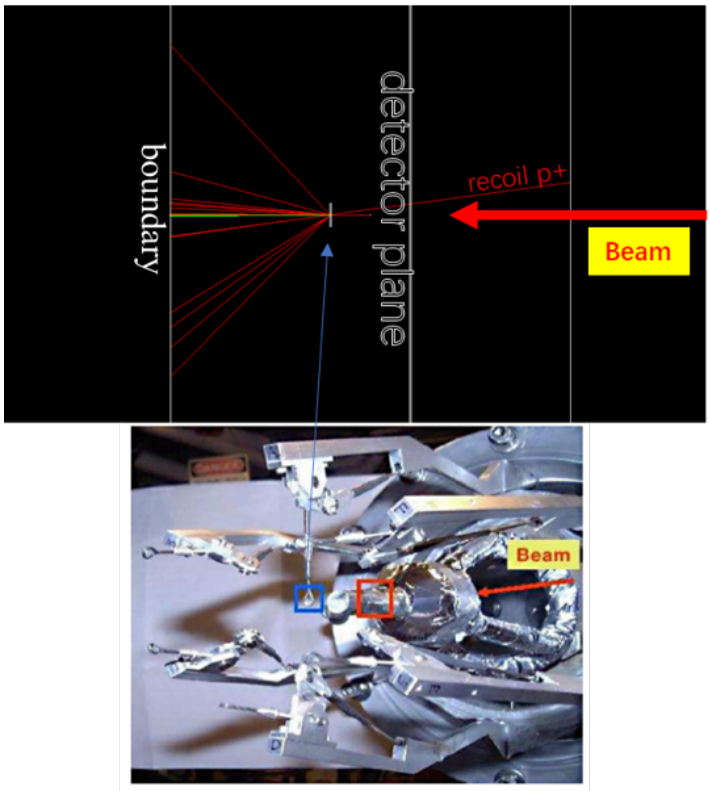


Figure 4.14 CLAS experimental configuration used to validate the Geant4 photonuclear model with recoil-proton measurements[64]. The simulated beam, target, and accepted recoil phase space follow the measurement for the comparison of proton momentum spectra.

gion. Nuclear recoil controls the apparent missing transverse momentum. These features are illustrated in Figure 4.16.

4.6.3 Biasing weights and effective statistics

Suppose a physical process with interaction probability p is enhanced by a factor B. In the dilute limit the biased probability is approximately Bp, and an event containing the biased interaction receives weight $1/B$. In a transport simulation with competing processes and finite path lengths, the exact likelihood ratio also contains survival-probability corrections. The Geant4 generic biasing interface provides transport weights for this purpose. The framework can record a track weight in the event output.

For a selected sample with weights w_j, the predicted yield and its Monte Carlo variance are estimated as

$$\hat{N} = \sum_{j \in \mathrm{sel}} w_j, \qquad \widehat{\mathrm{Var}}_{\mathrm{MC}}(\hat{N}) = \sum_{j \in \mathrm{sel}} w_j^2. \tag{4.20}$$

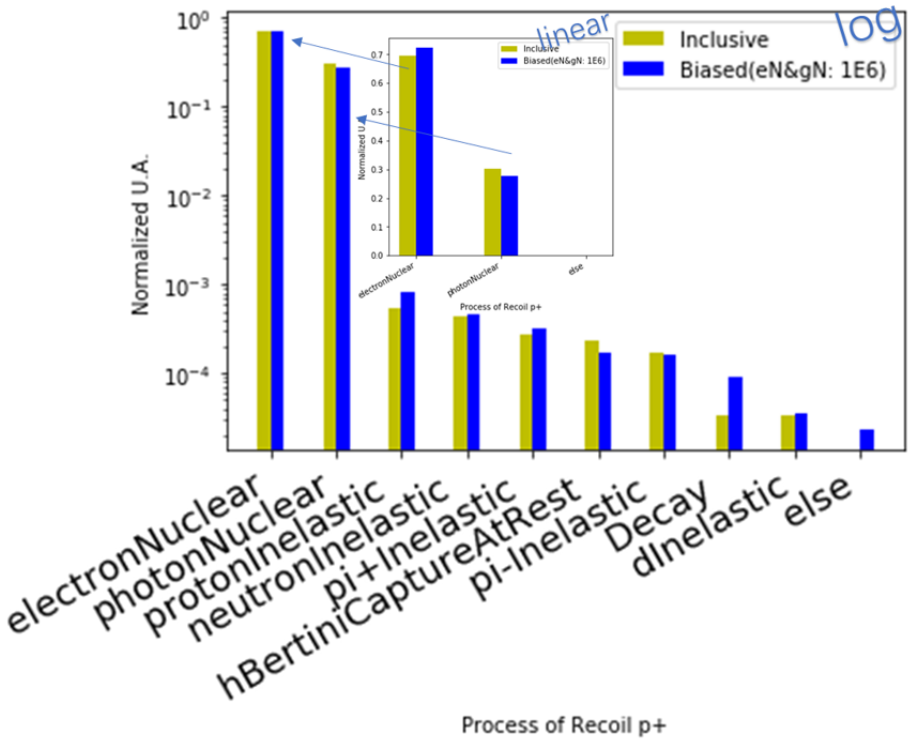


Figure 4.15 Geant4 recoil-proton validation using the CLAS configuration. Normalised process composition in inclusive simulation and a sample with electronuclear and photonuclear cross sections enhanced by 10^6. The inset uses a linear vertical scale.

The corresponding effective number of events is

$$N_{\text{eff}} = \frac{(\sum_j w_j)^2}{\sum_j w_j^2}. \tag{4.21}$$

These expressions follow from weighted event counting. In a histogram bin, an event contributes w_j to the content and w_j^2 to the usual sum-of-squared-weights variance estimate. Combining independent simulated events adds these contributions. If every event has the same weight w, the content is Nw, its variance estimate is Nw^2, and $N_{\text{eff}} = N$. With unequal weights, the same total content can have a larger variance because fewer events carry most of the weight.

For non-negative weights, the Cauchy–Schwarz inequality gives

$$\left(\sum_{j=1}^{N} w_j\right)^2 \le N \sum_{j=1}^{N} w_j^2, \qquad N_{\text{eff}} \le N. \tag{4.22}$$

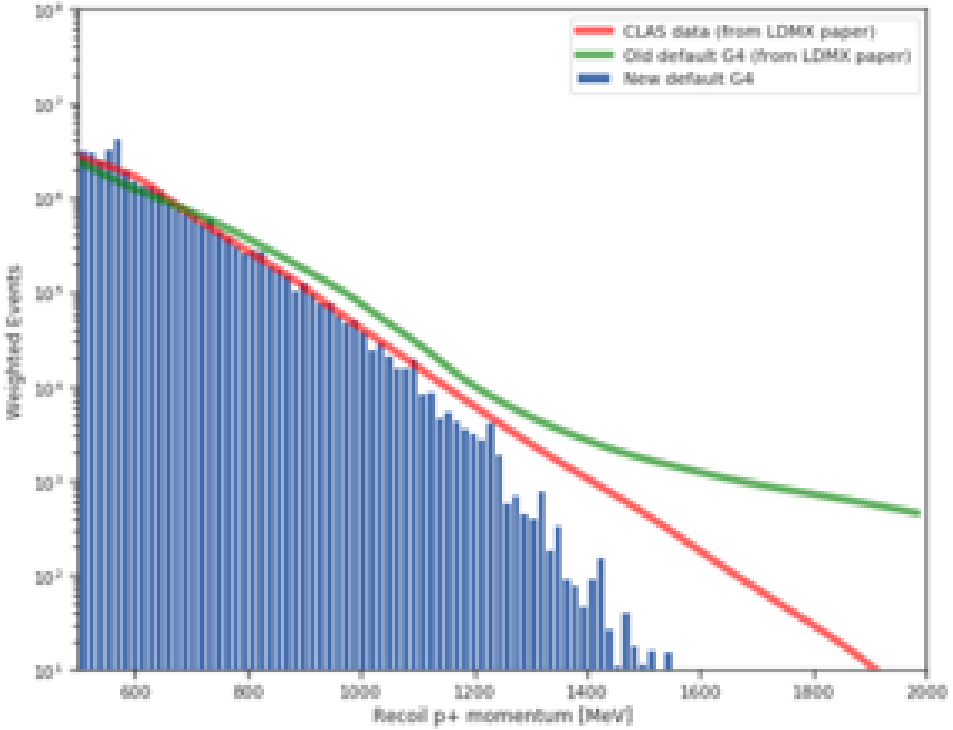


Figure 4.15 (continued) Geant4 recoil-proton validation using the CLAS configuration. Weighted recoil-proton momentum spectra from the old and new default Geant4 models, compared with CLAS reference data[64]. The dedicated simulation contains 2×10^8 electrons and uses a bias factor of 10^6.

Equality holds when all weights are equal. Under this counting estimate, the relative Monte Carlo standard deviation is

$$\frac{\sqrt{\sum_j w_j^2}}{\sum_j w_j} = \frac{1}{\sqrt{N_{\text{eff}}}}. \tag{4.23}$$

The reason that $1/B$ is only a limiting weight can be illustrated with a single process of constant interaction rate κ per unit length. In a uniform material, the density for its first interaction at path length ℓ is $\kappa \exp(-\kappa\ell)$. Increasing the rate by B gives $B\kappa \exp(-B\kappa\ell)$. The likelihood ratio is

$$w(\ell) = \frac{1}{B} \exp[(B-1)\kappa\ell]. \tag{4.24}$$

The exponential accounts for the changed probability of reaching the interaction point without an earlier interaction. The approximation $w = 1/B$ follows when this correction is negligible.

A selection changes which weights enter the sum. Consequently, N_{eff} must be evaluated after the selection of interest. A sample with satisfactory effective statistics before the

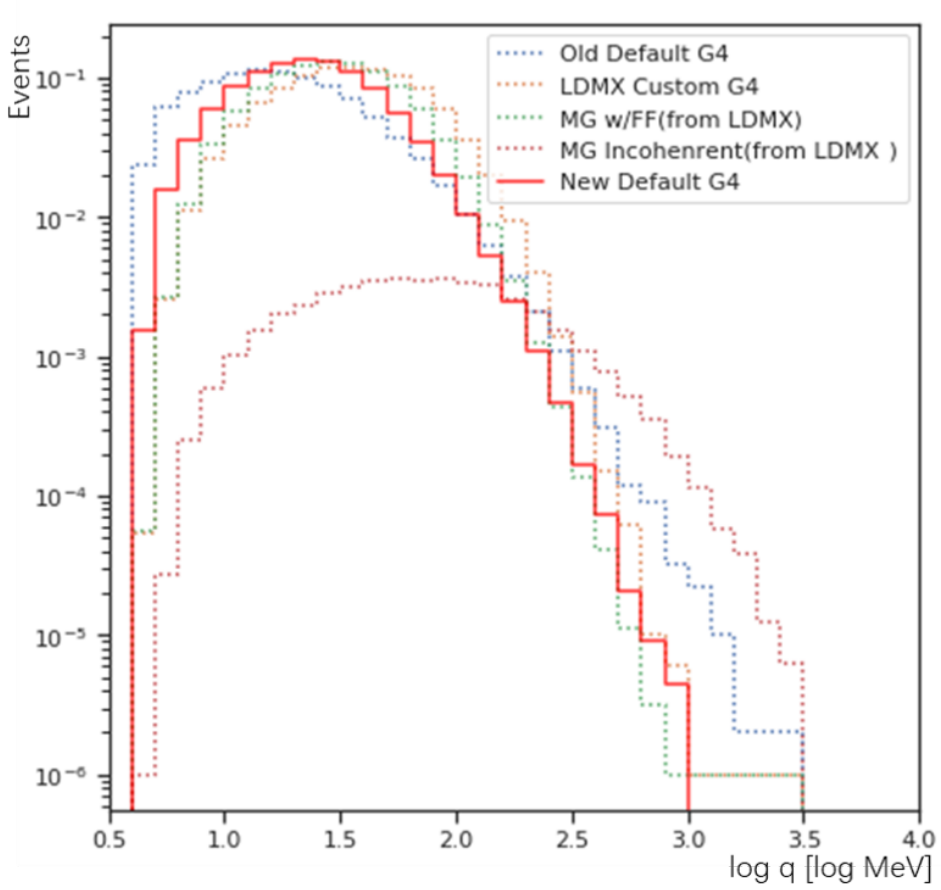


Figure 4.16 Validation of photon conversion to a muon pair in a nuclear field, $\gamma Z \rightarrow \mu^{+}\mu^{-}Z$. Nuclear momentum transfer q, plotted as $\log_{10}(q/\mathrm{MeV})$, for default and custom Geant4 models and MadGraph calculations.

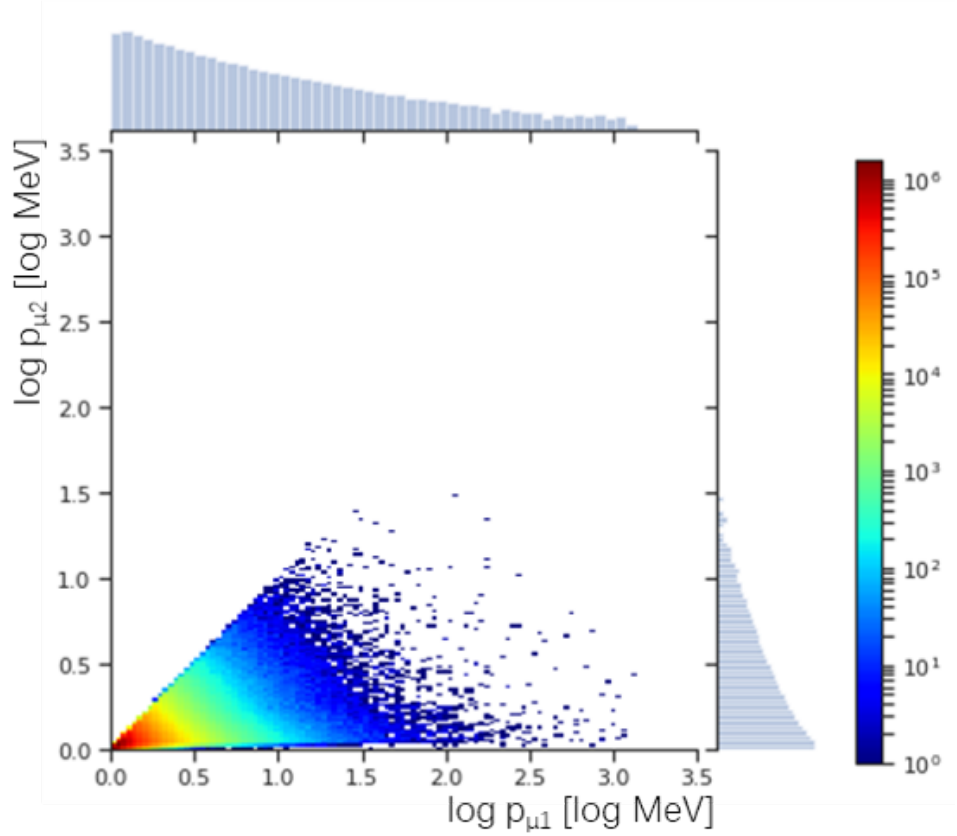


Figure 4.16 (continued) Validation of photon conversion to a muon pair in a nuclear field, $\gamma Z \to \mu^+\mu^- Z$. Correlation of the two muon momentum magnitudes, with both coordinates shown as $\log_{10}(p/\mathrm{MeV})$. One-dimensional projections appear above and beside the main panel.

calorimeter veto can have poor statistics after it. Scaling every event by an exposure factor C gives a yield multiplied by C and a Monte Carlo variance multiplied by C^2; it leaves N_{eff} unchanged.

Biasing is validated in a control region where an unbiased inclusive sample retains enough events. The comparison proceeds in three steps: the integrated weighted rate, the shape of variables used by the selection, and the ancestry and production-region composition are checked separately. Bias factors are also varied; a stable prediction over a reasonable range provides a useful implementation check.

4.7 Detailed Calorimeter Feature Reconstruction

The ECAL observables introduced in Section 4.4 describe both deposited energy and shower topology. Even in a cut-based analysis, the longitudinal and transverse profiles help explain why a background passes or fails a selection.

4.7.1 Energy-ordering and spatial moments

Let the calibrated energy and centre of cell i be E_i and $\mathbf{r}_i = (x_i, y_i, z_i)$. The energy fraction in the n hottest cells is

$$f_n = \frac{\sum_{i=1}^{n} E_{(i)}}{\sum_j E_j}, \tag{4.25}$$

where $E_{(1)} \geq E_{(2)} \geq \cdots$ is the ordered energy sequence. Small n probes the shower core; large n measures how quickly the cluster approaches full containment. Values at different n are correlated because their cell sums overlap.

Energy-weighted centroids and covariance elements are

$$\bar{r}_a = \frac{1}{E_{\text{tot}}} \sum_i E_i r_{i,a}, \tag{4.26}$$

$$C_{ab} = \frac{1}{E_{\text{tot}}} \sum_i E_i (r_{i,a} - \bar{r}_a)(r_{i,b} - \bar{r}_b). \tag{4.27}$$

Diagonalising the covariance matrix gives principal shower axes and widths without assuming that the shower is aligned with the detector axes. For a staggered crystal geometry, the moments are defined using physical cell centres rather than integer indices.

Figure 4.17 compares the total and maximum-cell energy for inclusive and dedicated rare-process samples. These variables have direct physical interpretations and are retained in the positron signal-region definitions. The two energies are strongly correlated.

The hottest-cell fractions and lateral moments probe whether an event contains a compact electromagnetic shower or diffuse hadronic activity. Representative distributions are shown in Figure 4.18.

For non-negative cell energies, f_n cannot decrease when n increases and cannot exceed unity. Once all retained cells are included, it equals unity. The difference $f_{n+1} - f_n$ is the next ordered cell energy divided by the total energy.

For one coordinate x, the uncentred and central second moments satisfy

$$\frac{\sum_i E_i x_i^2}{E_{\text{tot}}} = \frac{\sum_i E_i (x_i - \bar{x})^2}{E_{\text{tot}}} + \bar{x}^2. \tag{4.28}$$

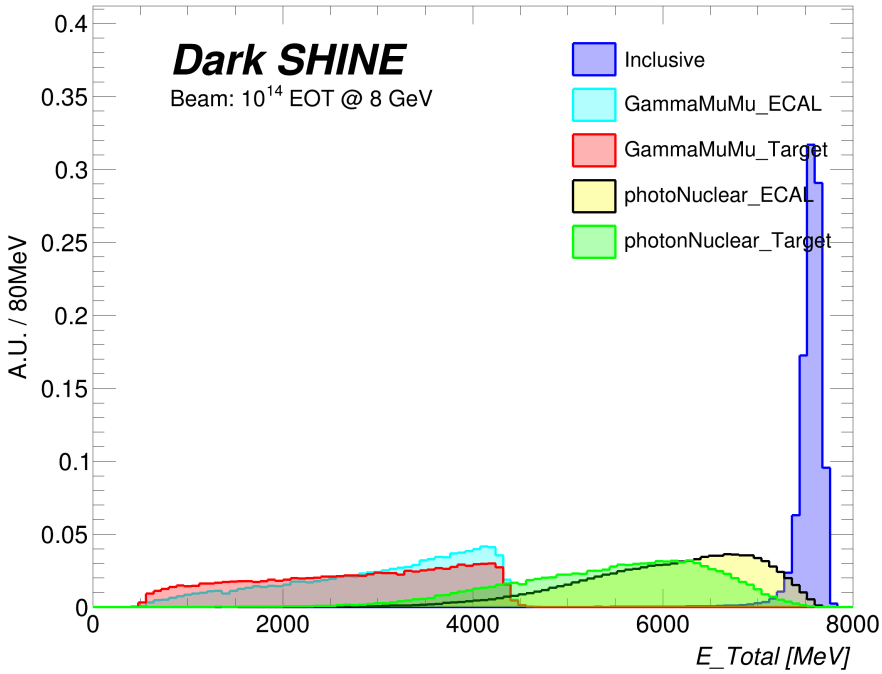


Figure 4.17 Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. Total deposited energy, in MeV.

An uncentred second moment can increase when a compact shower moves away from the coordinate origin, even if its width is unchanged. The central moment describes its spread about the energy centroid.

4.7.2 Topological clusters and track-assisted reconstruction

The calorimeter-only implementation is `TopoCluster_Analysis`. Its track-assisted counterpart is `PFTopoCluster_Analysis`. The `RecECAL` processor selects between them with its `TrackMatch` setting. Both operate on named calorimeter-hit collections and retain the whole-calorimeter energy and shape summaries. The cluster outputs add information about the organisation of those deposits.

The clustering code first converts calorimeter hits to helper objects and orders them by energy. It builds a position map and a neighbour map from the detector information held by `AnaData`. For the staggered geometry, the neighbour relation uses the stored geometry map.

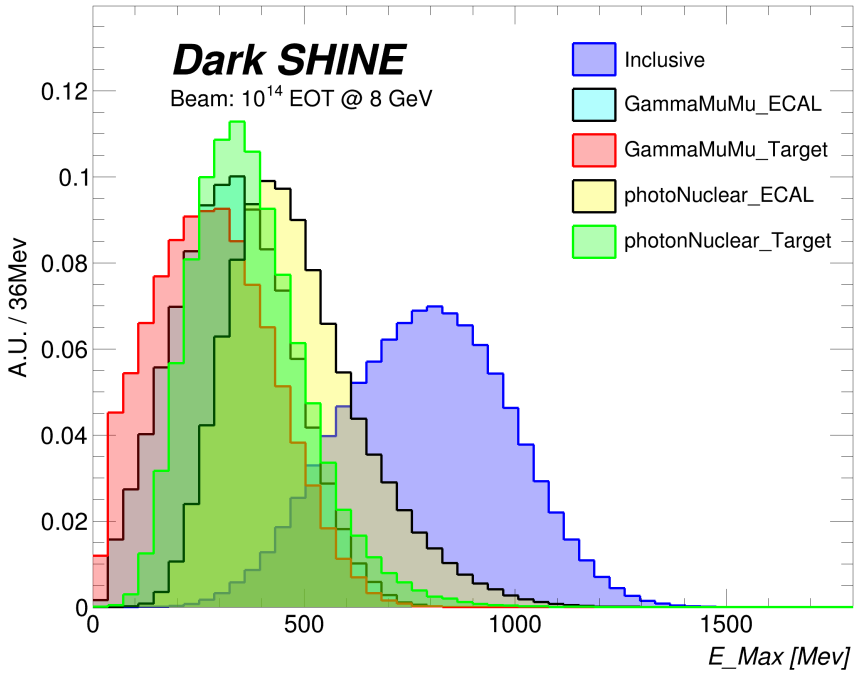


Figure 4.17 (continued) Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. Largest single-cell deposit, in MeV.

Three threshold parameters govern initial cluster formation. For cell energy E_i and configured noise scale E_{noise}, define $u_i = E_i/E_{\text{noise}}$. An unassigned cell can start a cluster when $u_i > S$. Neighbours with $u_i > P$ can be included, while those also satisfying $u_i > N$ can continue the expansion. Here S, N, and P are the configured `EThres_S`, `EThres_N`, and `EThres_P` values. The distinction permits a lower-energy boundary cell to contribute without necessarily extending the cluster through further neighbours.

The highest-energy free hit is considered first. It receives a parent-cluster identifier and enters a queue of cells from which the neighbours are explored. An eligible unassigned neighbour receives the same parent identifier. It is added to the queue when it passes the continuation threshold. After expansion, the assigned cells are removed from the free-hit list. The process repeats until the highest-energy remaining hit no longer passes the seed threshold. This construction separates disconnected deposits while keeping the energy ordering explicit.

The splitting stage searches within each parent cluster for candidate shower cores. Its

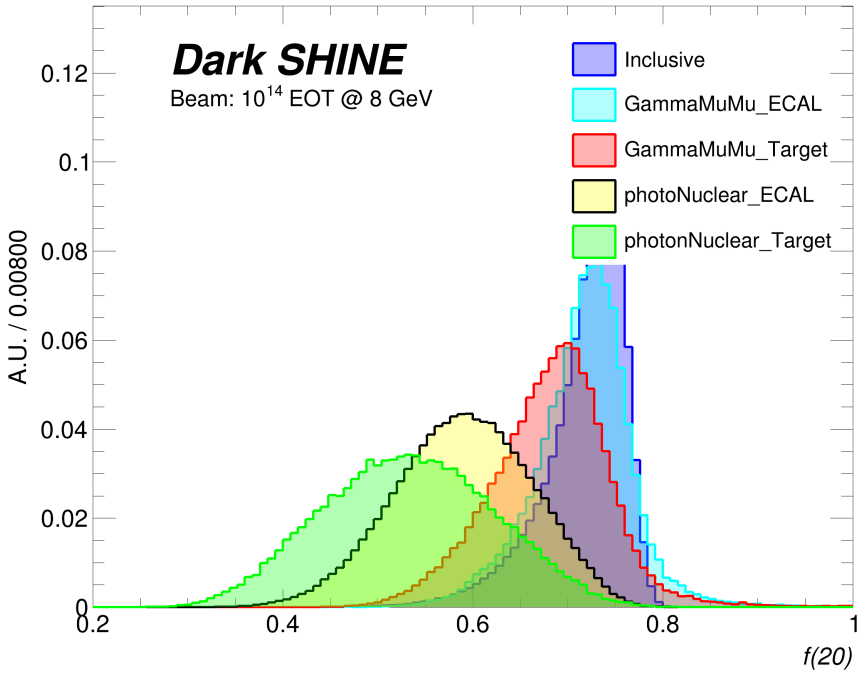


Figure 4.18 Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. Fraction of total energy in the 20 highest-energy cells.

tests use an absolute energy parameter, `Critical_E`, and a neighbour-count requirement, `Critical_N`. The accepted cores define subcluster identifiers. Expansion from these cores records cells associated with one or more subclusters. Shared cells can then contribute fractions of their energy to the competing subclusters. The weighting helper uses the form

$$w_1 = \frac{E_1}{E_1 + E_2 \exp(d_1 - d_2)}, \tag{4.29}$$

where the subcluster-energy estimates are E_1, E_2 and the distances to their reference positions are scaled by the configured electromagnetic length. The helper combines energy and distance information when assigning a shared deposit. Parent and subcluster identifiers retain the relationship between the connected deposit and its split components.

For each parent cluster and subcluster, the implementation calculates energy, position, spatial widths, and cell counts. It also estimates a direction and projects the cluster to the ECAL entrance. Where the longitudinal information does not support the direction estimate, the code retains a fallback based on the cluster position. The output includes the parent

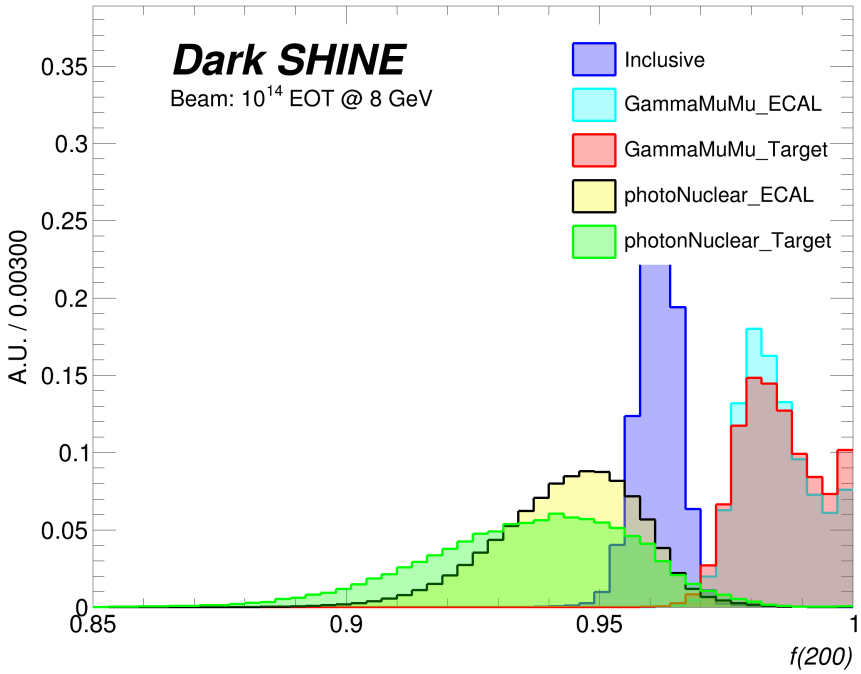


Figure 4.18 (continued) Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. Fraction of total energy in the 200 highest-energy cells.

identifier and the number of subclusters so that analyses can distinguish a connected shower region from its resolved components.

The track-assisted path supplies reference positions and momenta at the ECAL entrance. With reconstructed matching selected, the processor reads the ACTS extrapolation variables and applies its track-quality requirement before passing the reference points to the clustering code. A truth-matching option is also available for simulation studies.

The matching implementation compares clusters and subclusters with the track reference states, retaining the selected track identifiers. Its compatibility quantity combines a spatial distance from the trajectory with an energy comparison.

When the input collection is empty, `RecECAL` has an explicit branch that records zero activity and cluster counts, or skips the collection if requested. Some shape quantities are marked undefined. This treatment preserves the distinction between no visible activity and a reconstructed compact shower.

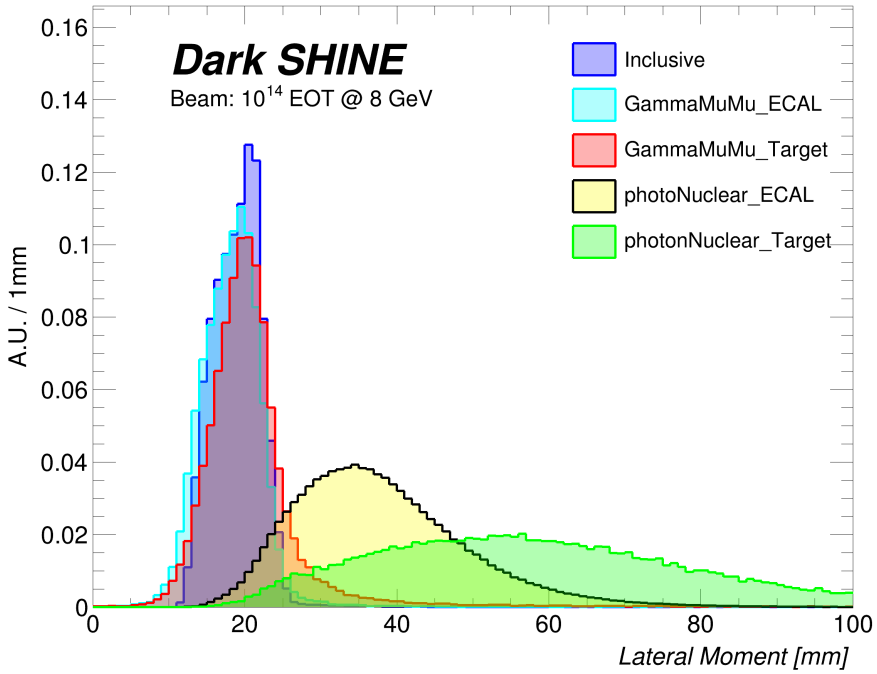


Figure 4.18 (continued) Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. Energy-weighted mean transverse distance from the shower centroid, in mm.

4.8 Software Quality and Reproducible Production

Production records should include: the source-code revision; compiler and external-library versions; geometry and material configuration; physics list and biasing settings; calibration constants; generator configuration and cross section; random seeds; and input-file checksums. A human-readable run summary is useful, but machine-readable metadata are required for automated consistency checks.

The simulation configuration provides a random seed and an option to restore a stored random state. The run number and requested event count enter event identification.

Validation can follow the boundaries between framework components. A small transport sample tests geometry construction, activated processes, and the saved collections. Reconstruction checks use the same named inputs as the intended production and inspect object counts and output quantities. A physics comparison then tests the distributions relevant to the selection.

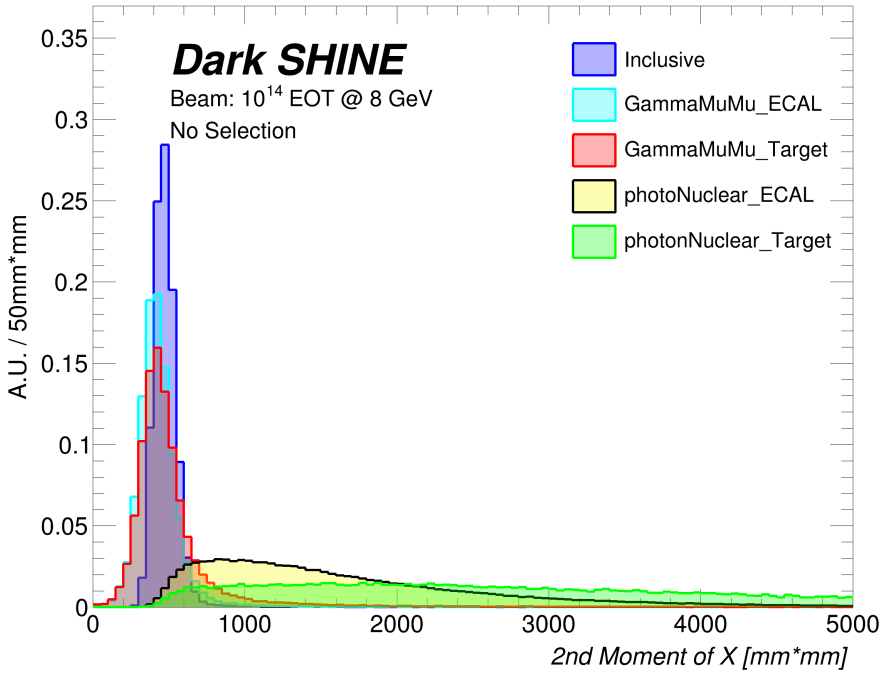


Figure 4.19 Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. The uncentred second moment is $\langle \xi^2 \rangle = \sum_i E_i \xi_i^2 / \sum_i E_i$ about the detector-coordinate origin. $\langle x^2 \rangle$, in mm^2.

A pilot sample permits the production settings to be checked before a large campaign. Geometry overlaps, missing collections, non-finite quantities, and unexpected event or file counts can be inspected at this stage. For biased samples, the requested process and production region must also match the filter definition.

4.9 Chapter Summary

The DarkSHINE software carries events from configurable Geant4 transport through digitisation, track and calorimeter reconstruction, and analysis output using a common event model. Rare-process biasing and filtering concentrate computing time on the backgrounds relevant to missing-energy selections. For the optical response, the LUT reduces the measured processing time from about 200 s to 4 s per event.

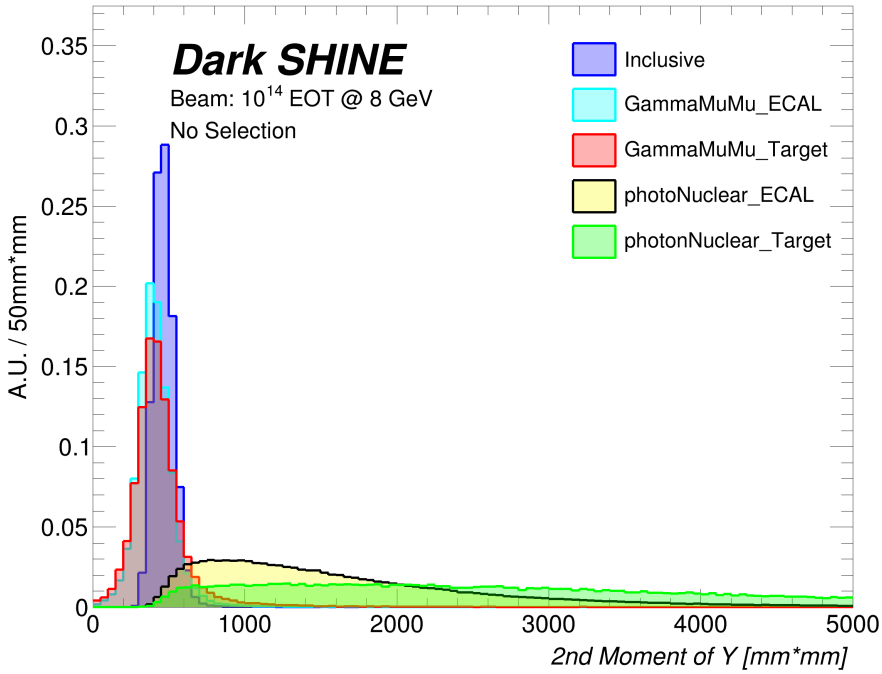


Figure 4.19 (continued) Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. The uncentred second moment is $\langle \xi^2 \rangle = \sum_i E_i \xi_i^2 / \sum_i E_i$ about the detector-coordinate origin. $\langle y^2 \rangle$, in mm^2.

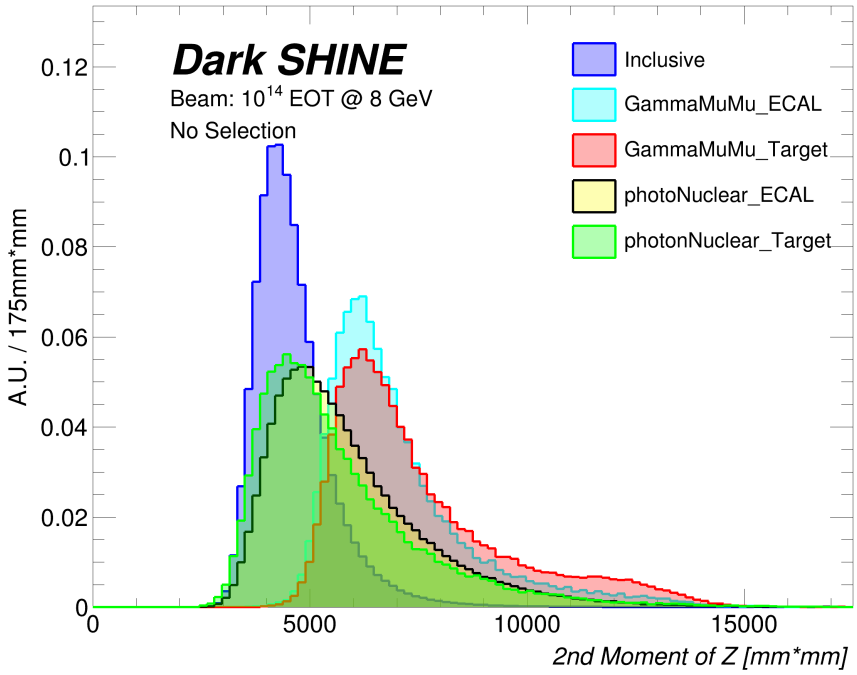


Figure 4.19 (continued) Unit-area ECAL distributions for an 8 GeV electron beam. Samples comprise inclusive interactions and dedicated photonuclear and $\gamma \to \mu^+\mu^-$ processes in the target/tracker or ECAL. The uncentred second moment is $\langle \xi^2 \rangle = \sum_i E_i \xi_i^2 / \sum_i E_i$ about the detector-coordinate origin. $\langle z^2 \rangle$ along the beam, in mm^2.

Chapter 5 Positron-Beam-Mode Analysis

The DarkSHINE detector and simulation framework also permit a study of positron running. Positrons can annihilate on atomic electrons as well as radiate dark photons through bremsstrahlung. The visible particles left by these processes differ, so the same apparatus can search for several event topologies. We use these differences to define the event categories and evaluate their combined sensitivity. The simulated samples and projected limits follow the positron-mode study[68].

5.1 Physics scope and simulated samples

The benchmark uses an 8 GeV positron beam and $N_{e^+} = 3 \times 10^{14}$ positrons on target (POT), corresponding to the ten-year exposure benchmark. The simulation campaign uses the `baseline1` detector configuration in the `positron` branch of the DarkSHINE software. This is the reference design described in Chapter 3 and illustrated in Figure 3.2, retained from the earlier published electron-beam projection[69]. Comparisons with electron running use the same beam count and target to match the integrated luminosity.

Three signal mechanisms are retained: nuclear dark bremsstrahlung, $e^+N \to e^+NA'$; non-resonant annihilation, $e^+e^- \to \gamma A'$; and resonant (s-channel) annihilation, $e^+e^- \to A'^* \to \chi\bar{\chi}$. Here N denotes a target nucleus. In the annihilation notation $e^+e^-N \to (\gamma)A'N$, the nucleus is a spectator and the electron belongs to the material. Invisible decay $A' \to \chi\bar{\chi}$ is assumed to dominate.

For an electron initially at rest, the available annihilation energy is

$$s = 2m_e(E_{e^+} + m_e), \qquad \sqrt{s} \simeq 90.4 \text{ MeV}. \tag{5.1}$$

Consequently, the on-shell $\gamma A'$ final state has a mediator-mass endpoint near 90.4 MeV, while resonant production is enhanced around that scale. Motion of bound electrons broadens the idealised stationary-electron picture. For the s-channel contribution, the cross section includes the finite-width dark-photon propagator and its off-shell contribution to the $\chi\bar{\chi}$ final state. The resonance is associated with the propagator pole, while the final-state threshold is $s \geq 4m_\chi^2$. The on-shell description applies to an undecayed A' produced as an external

particle.

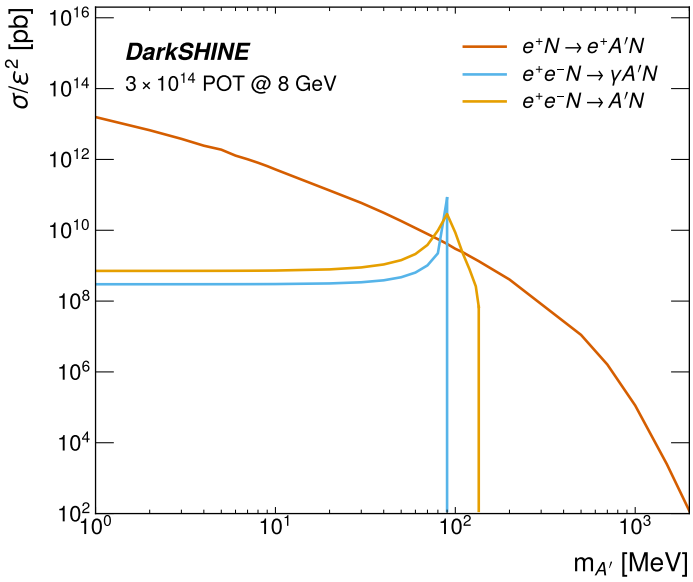


Figure 5.1 Production cross sections divided by the squared kinetic mixing ε^2 for an 8 GeV positron beam on tungsten. Curves show nuclear dark bremsstrahlung, non-resonant annihilation with an accompanying photon, and s-channel annihilation to $\chi\bar{\chi}$ through a finite-width dark-photon propagator, versus mediator pole mass $m_{A'}$. Annihilation on an electron at rest has centre-of-mass energy $\sqrt{s} \simeq 90.4$ MeV, setting the non-resonant endpoint and characteristic resonant scale. Curves from Ref.[68].

5.1.1 Signal generation and visible final states

Signal production combines CalcHEP calculations[63] with the Geant4 dark-photon processes. Transport and detector response are evaluated with the Geant4-based DarkSHINE simulation. The grid contains 30 dark-bremsstrahlung masses, 18 non-resonant masses, and 23 resonant masses. Each channel–mass combination has 10^5 generated events. The simulated mass points are listed in Table 5.1.

The visible remnant determines which observables are useful. Dark bremsstrahlung normally leaves a recoil positron whose momentum is reduced by the invisible energy transfer. Non-resonant annihilation instead leaves a photon, with no outgoing positron at the annihilation vertex. Resonant annihilation has neither of these visible remnants. These are production-level expectations. A recoil positron can escape acceptance or fail reconstruction, and a photon can leave too little detectable energy. A generated mechanism therefore need not enter the measured category that most closely resembles its ideal final state.

Table 5.1 Mediator masses used for the detector-level efficiency calculation, in MeV. Each entry corresponds to 10^5 generated events in the indicated production mode[68].

production mechanism of signal	dark photon mass (MeV)
bremsstrahlung emission	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 135, 140, 200, 500, 700, 1000, 1500, 2000
non-resonant annihilation	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90
resonant annihilation	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 135

5.1.2 Standard Model production and sample normalisation

Standard Model transport uses the Geant4 FTFP-BERT physics list. The rare processes of interest are photonuclear interactions (PN), positron-nuclear interactions (retaining the sample label EN), and photon conversion to muon pairs (GMM). PN and GMM photons arise from hard bremsstrahlung or annihilation, with $E_\gamma > 2$ GeV in the dedicated sample definitions. Target/tracker and ECAL production locations are treated separately. An additional target-annihilation sample covers annihilation outside these rare-process categories, including the ordinary multiphoton final states $e^+e^- \to \gamma\gamma(\gamma)$.

Biasing enhances the frequency of rare processes during generation. The physical normalisation is calculated from the process cross sections and the dedicated biasing studies. The equivalent unbiased beam count for a sample is

$$N_{\mathrm{POT}}^{\mathrm{eff}} = \frac{N_{\mathrm{attempt}}\epsilon_{\mathrm{gen}}}{\mathcal{B}}, \tag{5.2}$$

where $\mathcal{B}$ is the process branching ratio and ϵ_{gen} is the generator-filter acceptance. Thus generation attempts and accepted events are distinct bookkeeping quantities. The numerical normalisations in Table 5.2 follow this convention.

A loose production filter removes ordinary interactions with little apparent missing energy or momentum:

$$|P_{\mathrm{tag}} - P_{\mathrm{rec}}| > 1\ \mathrm{GeV}, \tag{5.3}$$

$$E_{\mathrm{ECAL}}^{\mathrm{total}} < 7\ \mathrm{GeV}. \tag{5.4}$$

These conditions reduce transport and storage costs; the final analysis thresholds are specified separately below. An unbiased inclusive sample of 2.5×10^9 POT provides an independent check of the biased production in populated regions.

Table 5.2 Dedicated background production for an 8 GeV positron beam and its equivalent unbiased positrons on target (POT). PN denotes photonuclear interactions, EN positron-nuclear interactions, and GMM photon conversion to muon pairs; target and ECAL production are listed separately. Generator-filter efficiencies are percentages; $N_{\mathrm{POT}}^{\mathrm{eff}} = N_{\mathrm{attempt}}\epsilon_{\mathrm{gen}}/\mathcal{B}$ uses their fractional values and the process branching ratio $\mathcal{B}$. Inputs follow Ref.[68].

Process	Branching ratio	Generation attempts	ϵ_{gen} (%)	Effective POT
PN ECAL	1.85×10^{-4}	1×10^{8}	27.38%	1.48×10^{11}
Annihil Target	2.93×10^{-5}	1×10^{8}	99.00%	3.38×10^{12}
EN ECAL	4.27×10^{-6}	1×10^{8}	1.38%	3.23×10^{11}
GMM ECAL	1.35×10^{-6}	1×10^{8}	27.27%	2.02×10^{13}
PN Target	1.32×10^{-6}	1×10^{8}	6.59%	4.99×10^{12}
EN Target	5.28×10^{-7}	1×10^{8}	1.47%	2.78×10^{12}
GMM Target	1.39×10^{-8}	1×10^{7}	6.57%	4.73×10^{13}

5.2 From final-state topology to signal regions

Three mutually exclusive categories are used: SR1 contains a recoil positron and missing energy ($e^+ + E_{\mathrm{miss}}$); SR2 contains a photon-like ECAL deposit and missing energy ($\gamma + E_{\mathrm{miss}}$); SR3 contains missing energy with no reconstructed recoil track and very little ECAL activity (E_{miss}). Every category requires exactly one tagging track. The recoil-track multiplicity separates SR1 from the other two, while complementary maximum-cell-energy inequalities separate SR2 from SR3. Their orthogonality is therefore a property of the measured selection, independent of signal origin.

We denote the upstream and downstream track counts by $N_{\mathrm{trk}}^{\mathrm{tag}}$ and $N_{\mathrm{trk}}^{\mathrm{rec}}$, and their momentum magnitudes by P_{tag} and P_{rec}. For either calorimeter, the superscript "total" denotes the summed deposited energy and "MaxCell" the largest single-cell energy. Table 5.3 gives the working points used for all mass hypotheses.

Table 5.3 Detector-level requirements defining SR1 ($e^+ + E_{\mathrm{miss}}$), SR2 ($\gamma + E_{\mathrm{miss}}$), and SR3 ($E_{\mathrm{miss}}$). The complementary recoil-track and ECAL single-cell requirements make the categories disjoint[68].

Selection Criteria	$e^+ + E_{\mathrm{miss}}$	$\gamma + E_{\mathrm{miss}}$	E_{miss}
N_{trk}^{tag} (Tagging tracks)	=1	=1	=1
N_{trk}^{rec} (Recoil tracks)	**=1**	**=0**	**=0**
$P_{tag} - P_{rec}$ (Missing momentum)	> 4 GeV	–	–
E_{ECAL}^{total} (Total ECAL energy)	< 2.5 GeV	< 2 GeV	< 2 GeV
$E_{ECAL}^{MaxCell}$ (Max ECAL cell)	–	≥ 1 **MeV**	< 1 **MeV**
E_{HCAL}^{total} (Total HCAL energy)	< 0.1 GeV	< 0.1 GeV	< 0.1 GeV
$E_{HCAL}^{MaxCell}$ (Max HCAL cell)	< 1 MeV	< 1 MeV	< 1 MeV

5.2.1 Tracking and calorimeter discrimination

The scalar momentum requirement for SR1 is $\Delta P = P_{\text{tag}} - P_{\text{rec}} > 4$ GeV. The 1 GeV value used in sample production is a loose preselection, as seen in Figure 5.2. Requiring one tagging track and the appropriate recoil multiplicity also removes backgrounds with additional reconstructed charged particles.

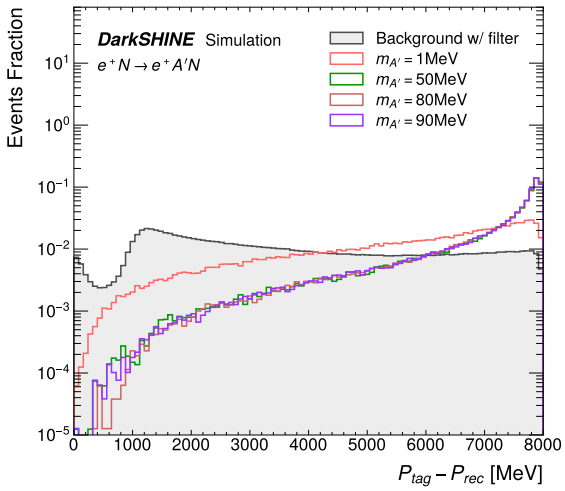


Figure 5.2 Scalar momentum difference $\Delta P = P_{\text{tag}} - P_{\text{rec}}$ between upstream tagging and downstream recoil tracks for an 8 GeV positron beam. Dark-bremsstrahlung signals with mediator masses 1, 50, 80, and 90 MeV are compared with inclusive background after the loose production filter. The recoil-positron category SR1 requires $\Delta P > 4$ GeV, above the 1 GeV production preselection.

The ECAL upper bounds limit the visible electromagnetic energy to 2.5 GeV in SR1 and 2 GeV in the two zero-recoil-track regions. In non-resonant annihilation the photon becomes soft near the mediator-mass endpoint. This explains why a final state containing a photon can still satisfy a low total-energy cut. The 90 MeV sample in Figure 5.3 illustrates this behaviour: its ECAL energy is concentrated below 200 MeV, consistently with the corresponding distribution in Figure 5.4.

For zero-recoil-track events, the largest ECAL cell distinguishes an observable photon deposit from an almost empty calorimeter. SR2 requires at least 1 MeV in one cell; SR3 requires every cell to remain below 1 MeV. Figure 5.5 shows how the two annihilation samples populate this observable. The HCAL provides a further veto: its total energy must be below 100 MeV and its largest cell below 1 MeV. Penetrating muons and hadronic activity are strongly reduced by these requirements.

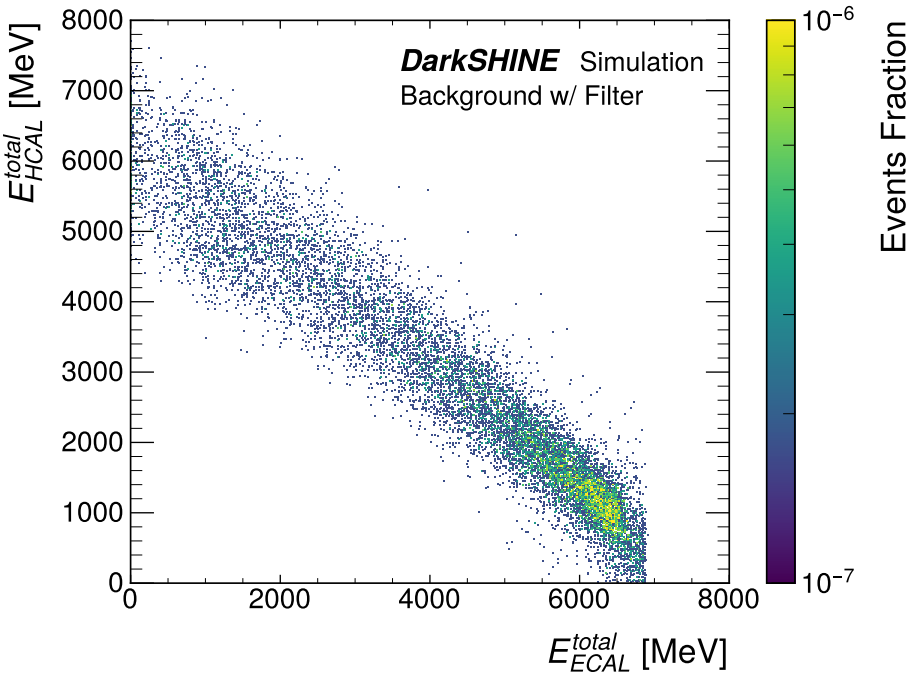


Figure 5.3 Correlations of summed ECAL and HCAL energies after the tracking requirements in the 8 GeV positron simulation. Inclusive Standard Model background over the full energy range.

The ECAL high-gain digitisation model gives an equivalent electronic noise of about 2.7 keV per channel[70-71]. The prospective study does not include complete digitisation of HCAL electronic noise and pileup. Its 1 MeV HCAL single-cell requirement is consequently an assumed achievable performance that requires prototype validation.

5.2.2 Logical separation of the measured categories

The category assignment can be expressed using indicator functions. Let $\mathbb{I}[C]$ equal one when a condition C holds and zero otherwise, and write the common tagging and HCAL requirements as

$$H = \mathbb{I}[N_{\mathrm{trk}}^{\mathrm{tag}} = 1]\,\mathbb{I}[E_{\mathrm{HCAL}}^{\mathrm{total}} < 100\ \mathrm{MeV}]\,\mathbb{I}[E_{\mathrm{HCAL}}^{\mathrm{MaxCell}} < 1\ \mathrm{MeV}]. \tag{5.5}$$

The complete region indicators are then

$$I_1 = H\,\mathbb{I}[N_{\mathrm{trk}}^{\mathrm{rec}} = 1]\,\mathbb{I}[\Delta P > 4\ \mathrm{GeV}]\,\mathbb{I}[E_{\mathrm{ECAL}}^{\mathrm{total}} < 2.5\ \mathrm{GeV}], \tag{5.6}$$

$$I_2 = H\,\mathbb{I}[N_{\mathrm{trk}}^{\mathrm{rec}} = 0]\,\mathbb{I}[E_{\mathrm{ECAL}}^{\mathrm{total}} < 2\ \mathrm{GeV}]\,\mathbb{I}[E_{\mathrm{ECAL}}^{\mathrm{MaxCell}} \geq 1\ \mathrm{MeV}], \tag{5.7}$$

$$I_3 = H\,\mathbb{I}[N_{\mathrm{trk}}^{\mathrm{rec}} = 0]\,\mathbb{I}[E_{\mathrm{ECAL}}^{\mathrm{total}} < 2\ \mathrm{GeV}]\,\mathbb{I}[E_{\mathrm{ECAL}}^{\mathrm{MaxCell}} < 1\ \mathrm{MeV}]. \tag{5.8}$$

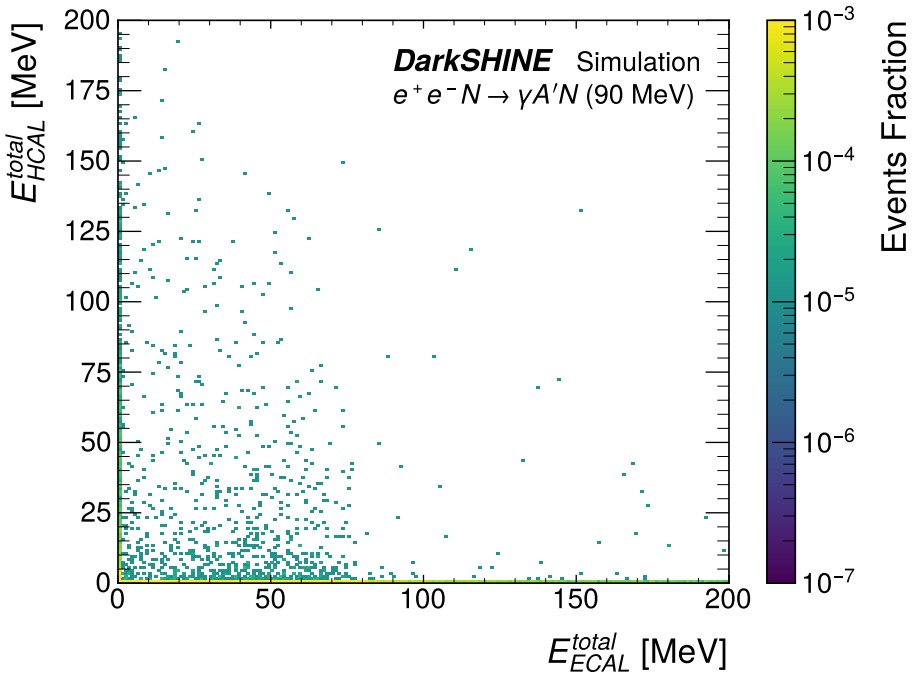


Figure 5.3 (continued) Correlations of summed ECAL and HCAL energies after the tracking requirements in the 8 GeV positron simulation. Non-resonant annihilation signal $e^+e^- \to \gamma A'$ at $m_{A'} = 90$ MeV, with 0–200 MeV on both axes. Near the annihilation endpoint the accompanying photon is soft.

The recoil multiplicity makes $I_1 I_2 = I_1 I_3 = 0$. For the two zero-track categories, the complementary maximum-cell conditions give $I_2 I_3 = 0$. In particular, an event exactly at the 1 MeV boundary belongs to SR2 if it passes the other requirements. Thus $I_1 + I_2 + I_3$ is either zero or one.

5.2.3 Efficiency and migration between categories

The nine channel–region efficiencies in Figures 5.6 to 5.8 are all needed for the signal prediction. A non-resonant photon outside the ECAL acceptance, or below the single-cell threshold, can place an annihilation event in SR3. Such migration remains part of the signal. Likewise, a dark-bremsstrahlung event can populate a zero-recoil-track region when its charged remnant is not reconstructed. The SR3 efficiency for this production mode reaches approximately 18% at the upper end of the simulated mass grid, so the off-diagonal efficiencies can contribute substantially to the selected yield.

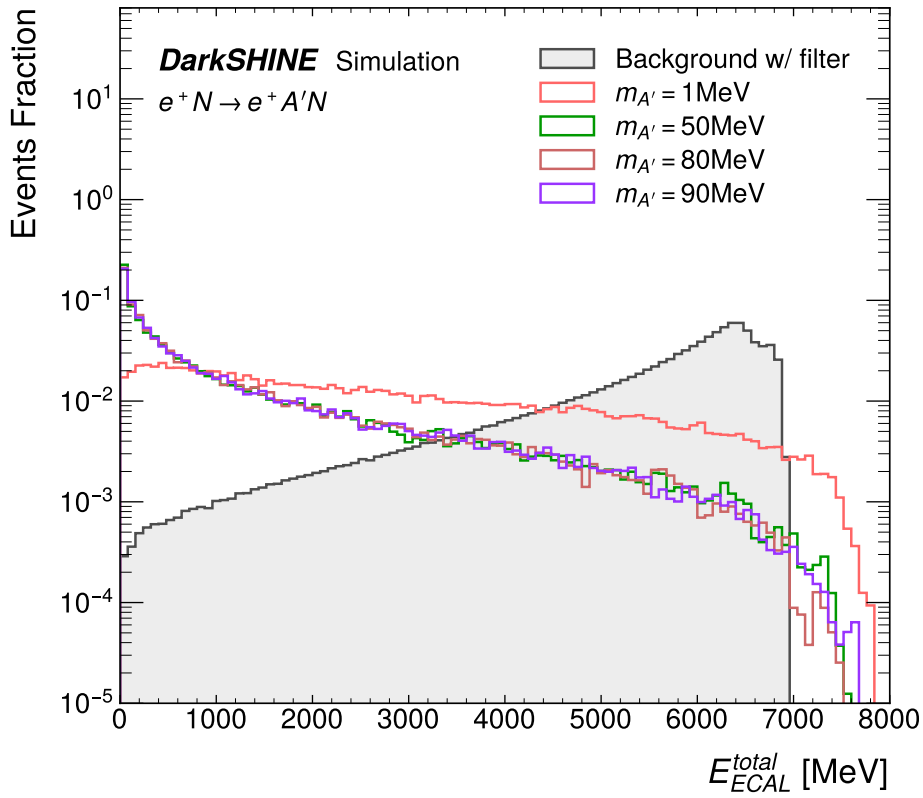


Figure 5.4 Summed ECAL energy for an 8 GeV positron beam. Signals with masses indicated in the legend are compared with the filtered inclusive Standard Model background. The background production filter requires total ECAL energy below 7 GeV. Dark-bremsstrahlung signals.

5.2.4 Background rejection through the selection sequence

Tables 5.4 to 5.6 give raw event counts at successive selection stages. No inclusive or dedicated background event passes a complete signal-region selection.

The cutflow columns describe nested event sets. If C_j denotes the events remaining after the first j analysis requirements, then $C_{j+1} \subseteq C_j$. A conditional survival fraction for one step is

$$a_j = \frac{|C_j|}{|C_{j-1}|}, \qquad \frac{|C_k|}{|C_0|} = \prod_{j=1}^{k} a_j, \tag{5.9}$$

when the denominators are nonzero.

In SR1, the recoil-track and missing-momentum requirements already distinguish the dedicated backgrounds. PN ECAL retains 4,641,895 events after the one-recoil condition and 3,114,535 after the momentum requirement. For EN ECAL the corresponding counts

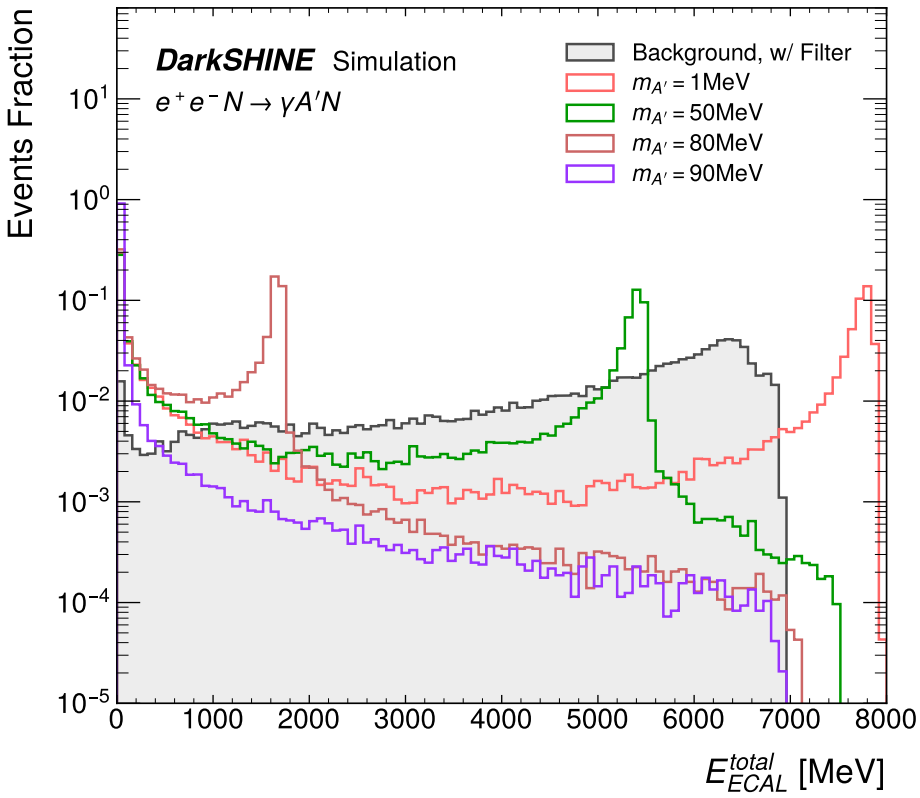


Figure 5.4 (continued) Summed ECAL energy for an 8 GeV positron beam. Signals with masses indicated in the legend are compared with the filtered inclusive Standard Model background. The background production filter requires total ECAL energy below 7 GeV. Non-resonant annihilation signals. The 90 MeV sample lies near the 90.4 MeV endpoint and leaves a soft photon.

are 99,812 and 58. The momentum condition therefore acts very differently on the selected samples. The HCAL requirements then reduce the PN ECAL row to 88,871 events before the final ECAL cut. GMM ECAL reaches six events after the total-HCAL requirement, remains at six after the single-cell requirement, and becomes empty after the ECAL requirement.

The zero-recoil condition in SR2 selects a much smaller subset of PN ECAL than the one-recoil condition in SR1. It retains 138,513 PN ECAL events after the common tagging selection. The maximum-ECAL-cell lower bound leaves this number unchanged, while the HCAL requirements reduce it to 2,552. The final total-ECAL condition again has no survivors. In the target-annihilation sample, the maximum-cell requirement changes 20,649 events to 20,569, after which the HCAL conditions reduce the count to 241. Thus the photon-category requirement and the hadronic veto test different aspects of the surviving events.

SR3 applies the total-ECAL requirement before the HCAL conditions and the maximum-

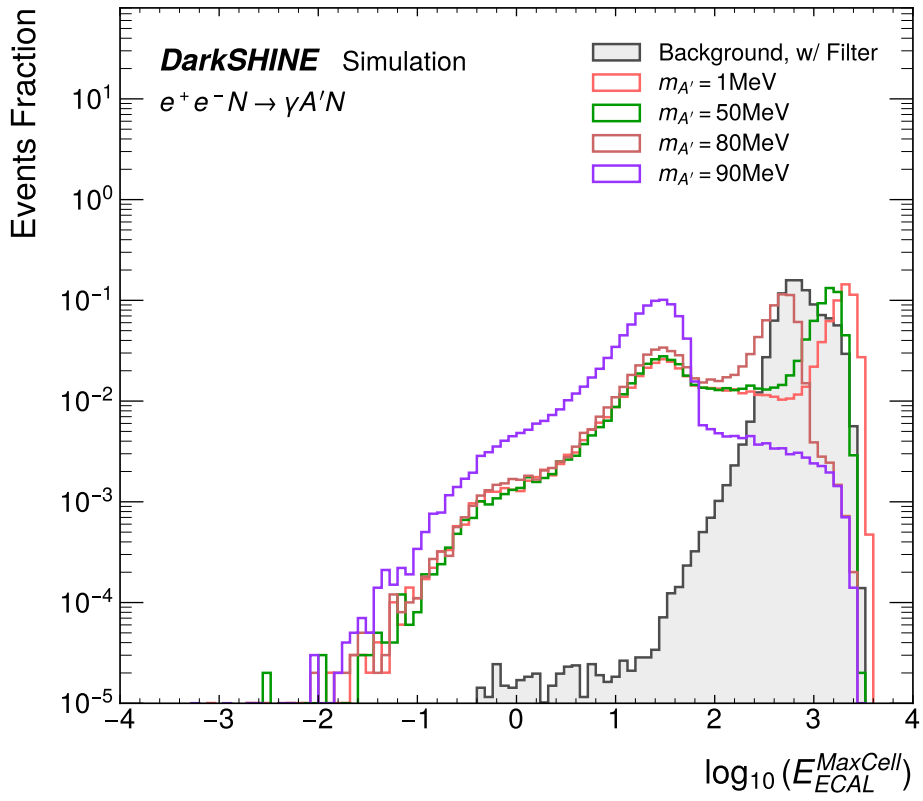


Figure 5.5 Largest single-cell ECAL energy on a logarithmic scale for an 8 GeV positron beam. Signals at the masses given in the legend are compared with the filtered inclusive background. Non-resonant annihilation. The photon energy depends on mediator mass.

cell upper bound. For PN ECAL, the total-energy cut leaves 120,147 of the 138,513 zero-recoil events; the subsequent HCAL cuts leave 1,096, and the maximum-cell cut removes those remaining events. In the inclusive row, the corresponding sequence is 9,240, 8,390, 61, 11, and zero.

The ordering differs between the SR2 and SR3 tables because their displayed sequences end with different ECAL requirements. A comparison of intermediate columns must therefore keep the preceding selection fixed.

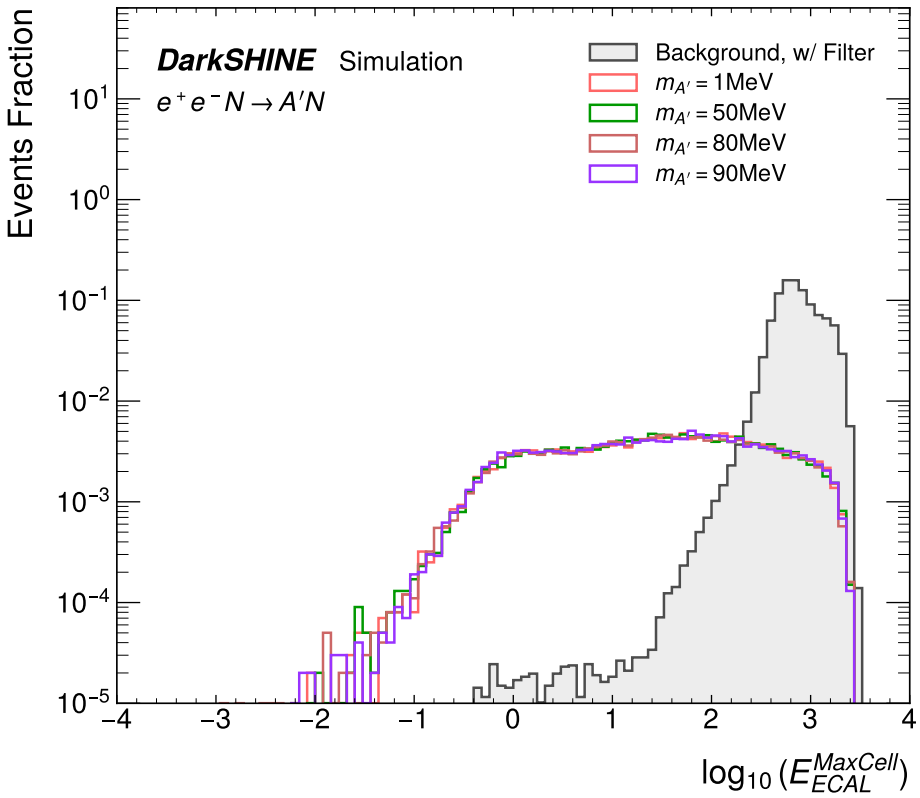


Figure 5.5 (continued) Largest single-cell ECAL energy on a logarithmic scale for an 8 GeV positron beam. Signals at the masses given in the legend are compared with the filtered inclusive background. Resonant annihilation. Invisible events have little visible activity; small deposits can arise from interactions in tracking material.

5.3 Luminosity and selected signal yields

5.3.1 Target luminosity and the production normalisation

A common normalisation links the generated cross sections to the event categories. For a production mechanism c, let n_c be the number of scattering centres per unit transverse area. Dark bremsstrahlung uses the areal density of tungsten nuclei; the annihilation mechanisms use that of atomic electrons. The corresponding integrated target luminosity is

$$\mathcal{L}_c = N_{e^+} n_c . \tag{5.10}$$

The POT benchmark specifies the incident beam count, while $\mathcal{L}_c$ has inverse-area units. Multiplying $\mathcal{L}_c$ by a cross section produces a dimensionless expected count. For the same tungsten target, the electron areal density is Z times the nuclear areal density.

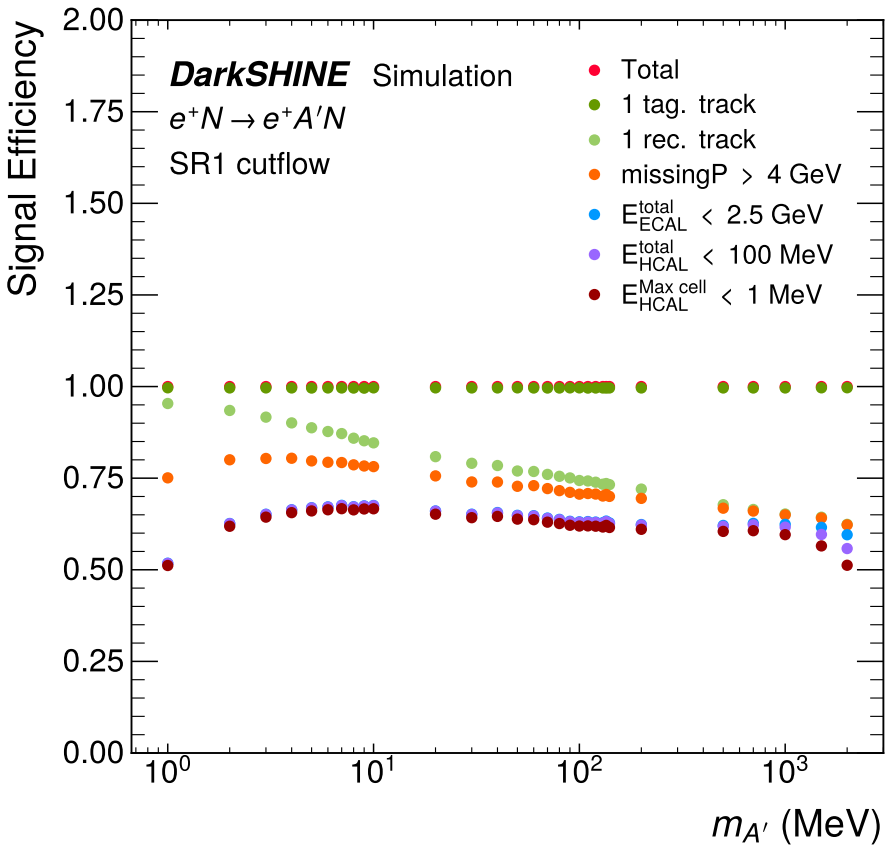


Figure 5.6 Acceptance times selection efficiency for the recoil-positron category SR1 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR1 selection. Dark bremsstrahlung.

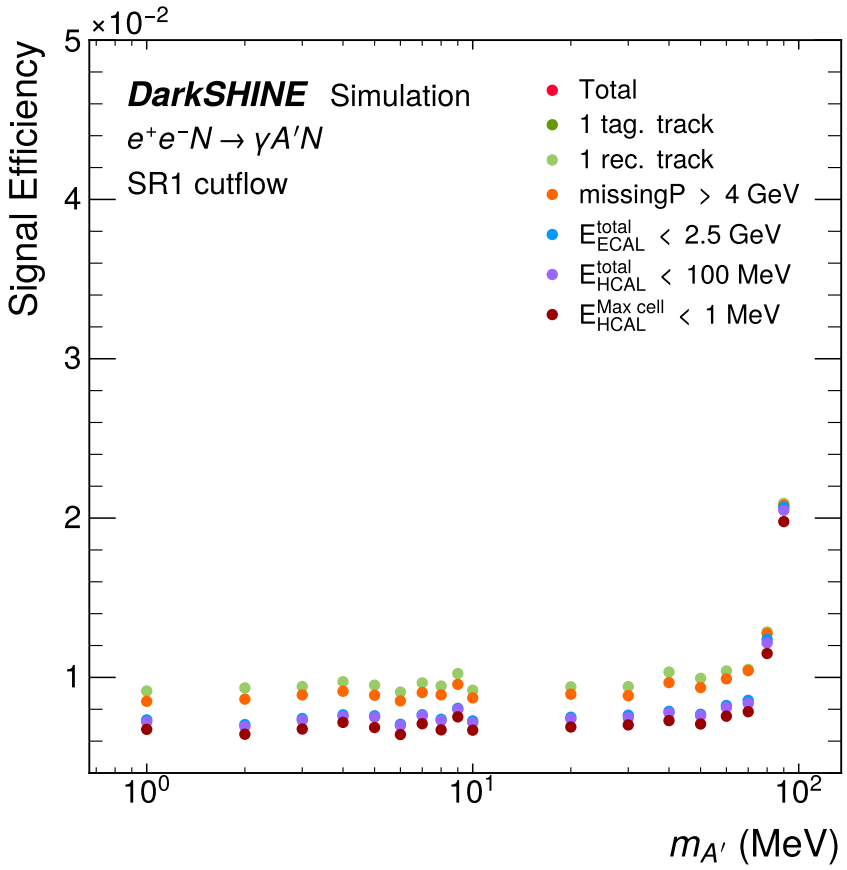


Figure 5.6 (continued) Acceptance times selection efficiency for the recoil-positron category SR1 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR1 selection. Non-resonant annihilation.

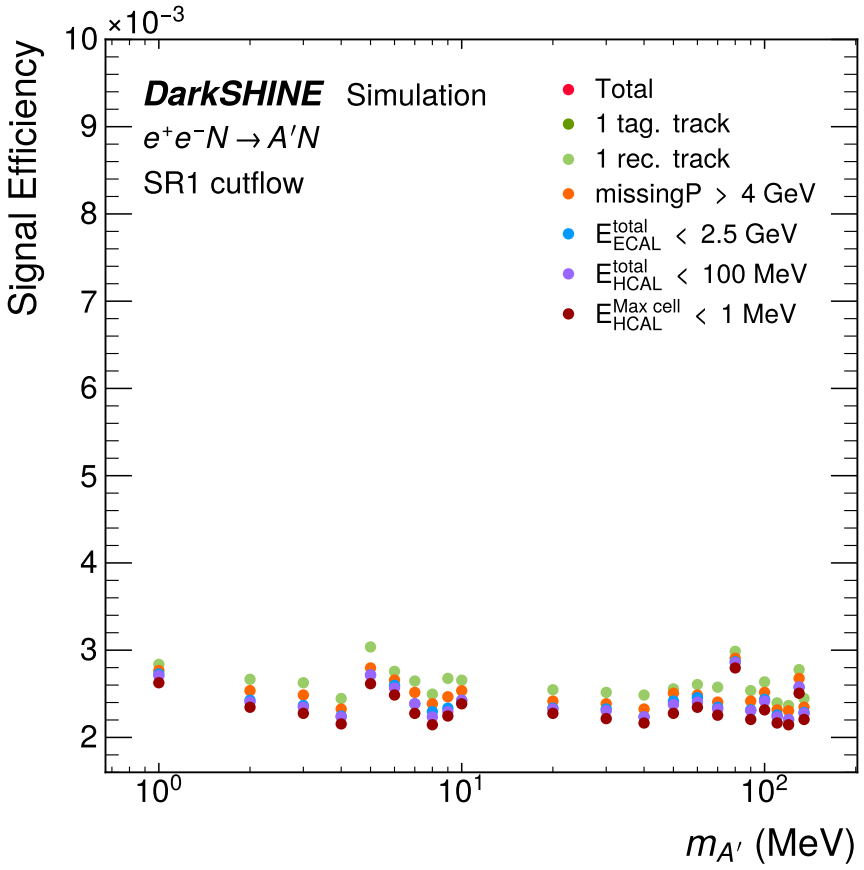


Figure 5.6 (continued) Acceptance times selection efficiency for the recoil-positron category SR1 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR1 selection. Resonant annihilation. The small feature around 5 MeV occurs at an efficiency of roughly 3×10^{-3} and is attributed to finite signal-sample statistics.

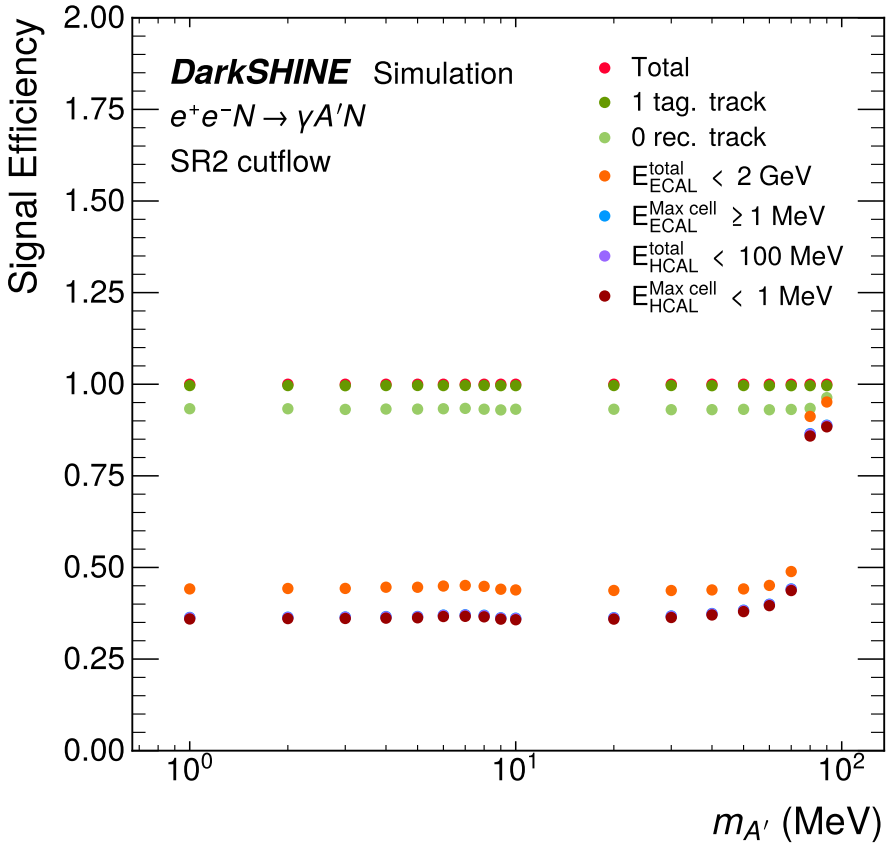


Figure 5.7 Acceptance times selection efficiency for the photon category SR2 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR2 selection. Non-resonant annihilation.

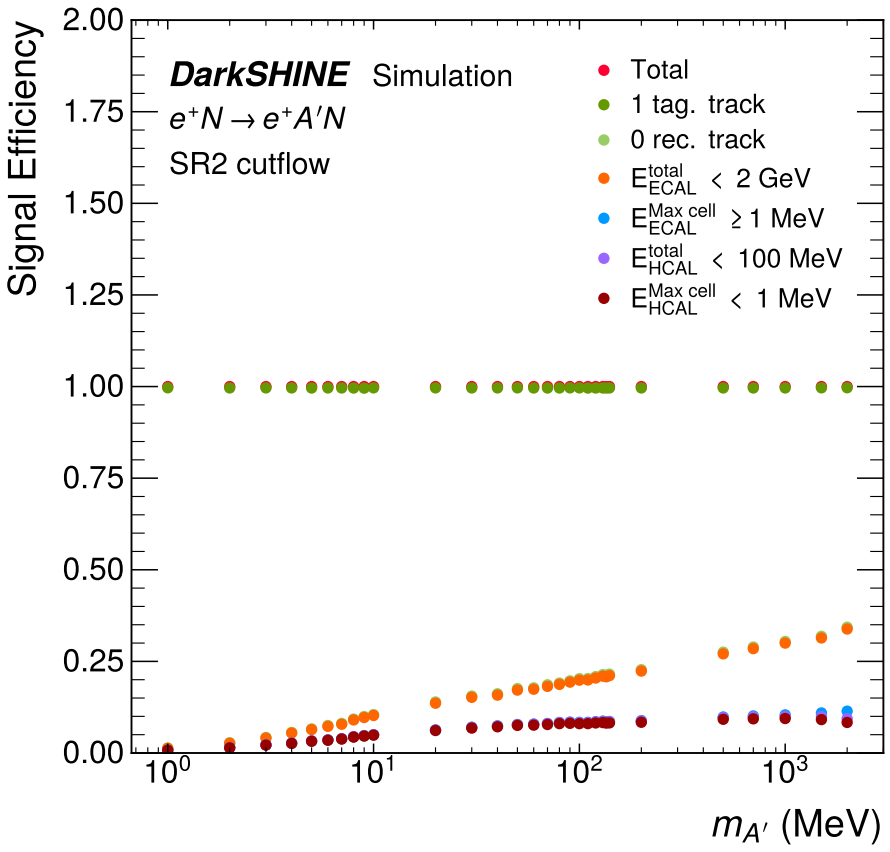


Figure 5.7 (continued) Acceptance times selection efficiency for the photon category SR2 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR2 selection. Dark bremsstrahlung. This contribution quantifies migration into the zero-recoil-track photon category.

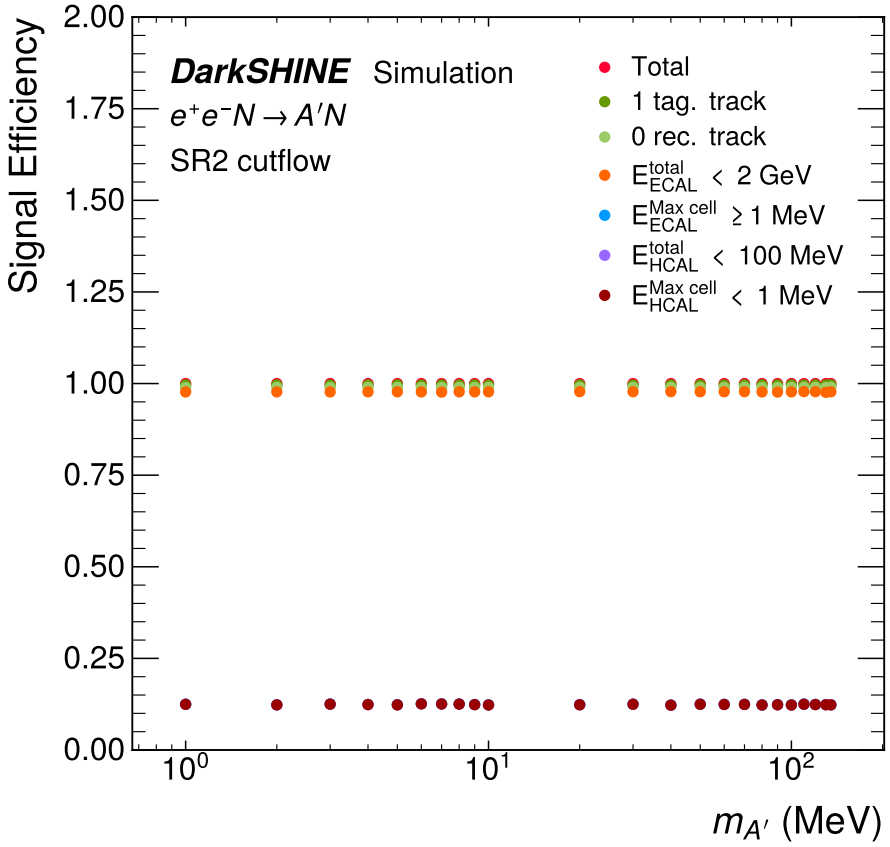


Figure 5.7 (continued) Acceptance times selection efficiency for the photon category SR2 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR2 selection. Resonant annihilation. This contribution quantifies migration into the zero-recoil-track photon category.

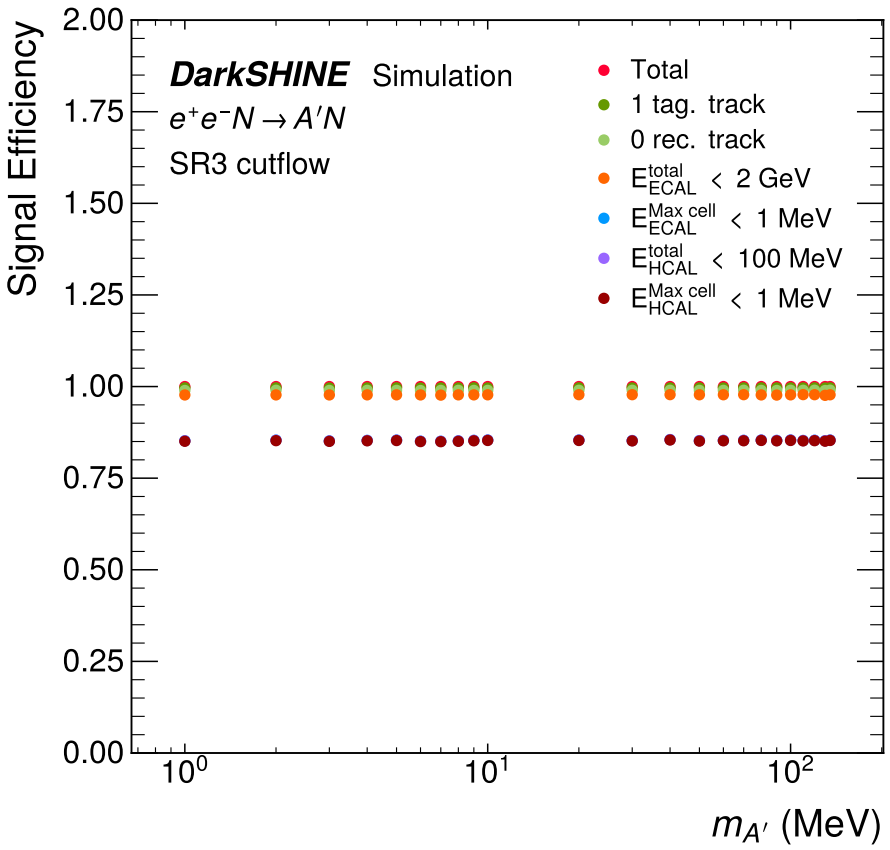


Figure 5.8 Acceptance times selection efficiency for the low-activity category SR3 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR3 selection. All three production mechanisms contribute to this category in the combined fit. Resonant annihilation.

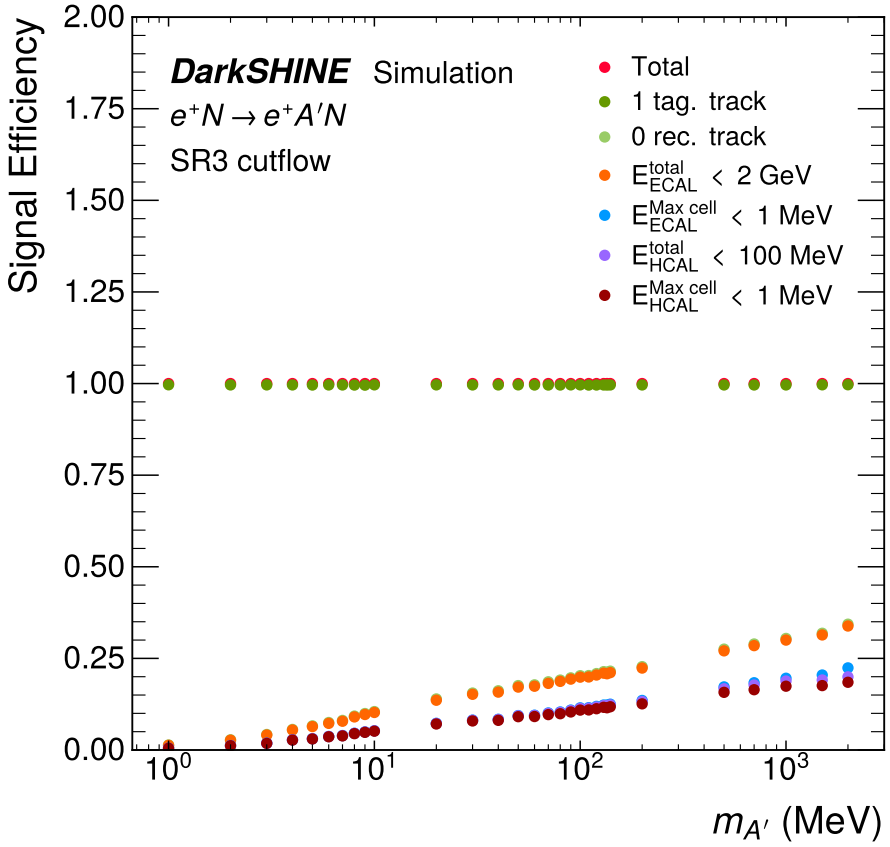


Figure 5.8 (continued) Acceptance times selection efficiency for the low-activity category SR3 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR3 selection. All three production mechanisms contribute to this category in the combined fit. Dark bremsstrahlung.

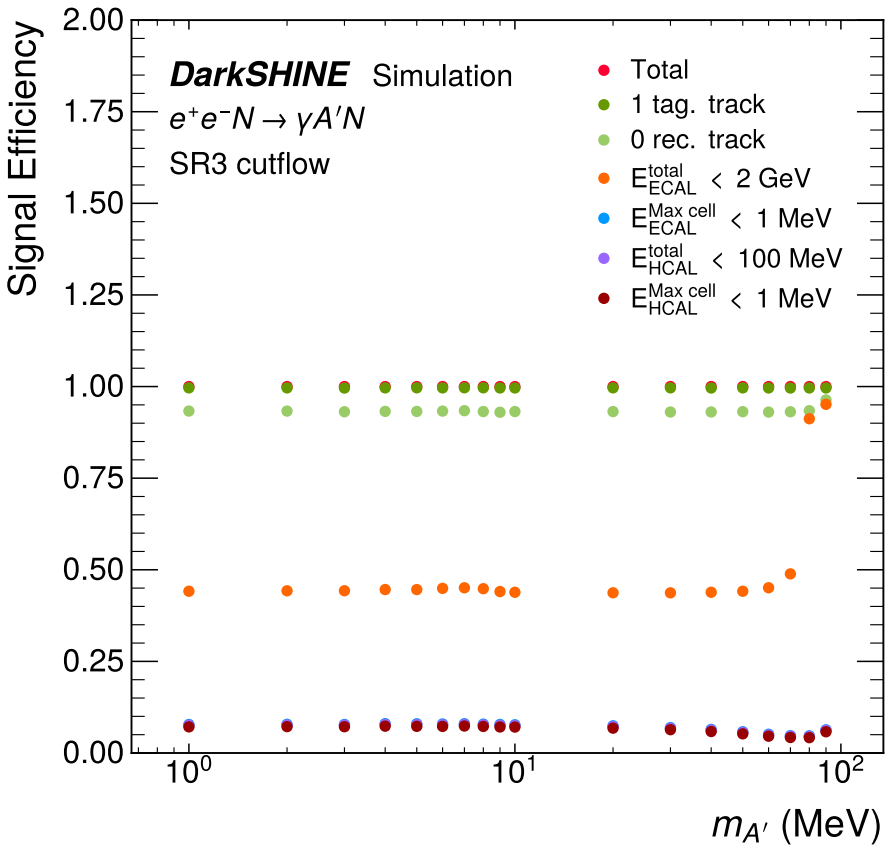


Figure 5.8 (continued) Acceptance times selection efficiency for the low-activity category SR3 in the 8 GeV positron simulation, versus mediator mass. Each fraction uses the 10^5 events generated at that channel–mass point as its denominator; the numerator passes the full SR3 selection. All three production mechanisms contribute to this category in the combined fit. Non-resonant annihilation.

Table 5.4 Unweighted background cutflow for the recoil-positron category SR1 in the 8 GeV positron simulation. PN denotes photonuclear interactions, EN positron-nuclear interactions, and GMM photon conversion to muon pairs. The upper block lists effective positrons on target (POT), generation attempts, and counts after the production filter. The lower block gives cumulative counts after each analysis requirement. Counts follow Ref.[68].

Sample	Effective POT	Attempts	Filter
PN ECAL	1.48×10^{11}	1×10^{8}	5829159
PN Target	4.99×10^{12}	1×10^{8}	5623324
EN Target	2.78×10^{12}	1×10^{8}	1385618
Annihil Target	3.38×10^{12}	1×10^{8}	289221
GMM ECAL	2.02×10^{13}	1×10^{8}	9333419
EN ECAL	3.23×10^{11}	1×10^{8}	124934
GMM Target	4.73×10^{13}	1×10^{7}	663639
Inclusive	2.5×10^{9}	2.5×10^{9}	1472414

Sample	$N_{trk}^{tag} = 1$	$N_{trk}^{rec} = 1$	$\Delta P >$ 4 GeV	E_{HCAL}^{total} < 0.1 GeV	$E_{HCAL}^{MaxCell}$ < 1 MeV	E_{ECAL}^{total} < 2.5 GeV
PN ECAL	5794358	4641895	3114535	272207	88871	0
PN Target	5386830	455540	429175	11831	2459	0
EN Target	1326519	117549	114845	977	206	0
Annihil Target	288943	5595	5466	201	60	0
GMM ECAL	9331661	8842313	6123627	6	6	0
EN ECAL	124241	99812	58	5	2	0
GMM Target	663576	20	0	0	0	0
Inclusive	1470635	1359999	570317	7175	2077	0

For a foil of physical thickness t, mass density ρ, and molar mass M_{W}, these two areal densities are

$$n_N = \frac{\rho t}{M_{\mathrm{W}}} N_A, \qquad n_e = Z n_N, \tag{5.11}$$

where N_A is Avogadro's constant. The product ρt is the target mass per unit area. Dividing by the molar mass gives the number of moles per unit area, and multiplication by N_A gives the number of atoms per unit area. Thus, when ρ, t, and M_{W} are expressed in $\mathrm{g\,cm^{-3}}$, cm, and $\mathrm{g\,mol^{-1}}$, respectively, n_N has units of $\mathrm{cm^{-2}}$. A cross section expressed in $\mathrm{cm^2}$ then produces a dimensionless interaction probability per incident positron in the thin-target expression. The target thickness quoted as $0.1X_0$ must be converted to a physical length before it is used in this formula.

The expected number of invisible hard-process events is

$$u_c(m, \epsilon) = \mathcal{L}_c \, \sigma_c(m, \epsilon) \mathrm{Br}_{\mathrm{inv}}. \tag{5.12}$$

Table 5.5 Unweighted background cutflow for the photon category SR2 in the 8 GeV positron simulation. PN denotes photonuclear interactions, EN positron-nuclear interactions, and GMM photon conversion to muon pairs. The upper block lists effective positrons on target (POT), generation attempts, and counts after the production filter. The lower block gives cumulative counts after each analysis requirement. Counts follow Ref.[68].

Sample	Effective POT	Attempts	Filter
PN ECAL	1.48×10^{11}	1×10^{8}	5829159
Annihil Target	3.38×10^{12}	1×10^{8}	289221
PN Target	4.99×10^{12}	1×10^{8}	5623324
EN Target	2.78×10^{12}	1×10^{8}	1385618
GMM ECAL	2.02×10^{13}	1×10^{8}	9333419
EN ECAL	3.23×10^{11}	1×10^{8}	124934
GMM Target	4.73×10^{13}	1×10^{7}	663639
Inclusive	2.5×10^{9}	2.5×10^{9}	1472414

Sample	$N_{trk}^{tag} = 1$	$N_{trk}^{rec} = 0$	$E_{ECAL}^{MaxCell}$ ≥ 1 MeV	E_{HCAL}^{total} < 0.1 GeV	$E_{HCAL}^{MaxCell}$ < 1 MeV	E_{ECAL}^{total} < 2 GeV
PN ECAL	5794358	138513	138513	8051	2552	0
Annihil Target	288943	20649	20569	768	241	0
PN Target	5386830	5391	5391	162	40	0
EN Target	1326519	1633	1633	8	0	0
GMM ECAL	9331661	67950	67950	0	0	0
EN ECAL	124241	0	0	0	0	0
GMM Target	663576	0	0	0	0	0
Inclusive	1470635	9240	9208	178	53	0

Here $\sigma_c \mathrm{Br}_{\mathrm{inv}}$ denotes the invisible-production cross section. For an external on-shell A', it is the production cross section times the invisible branching fraction. For the s-channel contribution, it is evaluated directly for $e^+e^- \to A'^* \to \chi\bar{\chi}$, including the finite-width propagator. The cross section and luminosity determine this number before the analysis selection. The 10^5 Monte Carlo events generated at each channel–mass point sample the fraction entering each reconstructed category. They determine the detector response to the production rate supplied by Equation (5.12).

5.3.2 The response matrix and its normalisation

To retain migrations explicitly, define

$$R_{rc}(m) = \frac{N_{rc}^{\mathrm{selected}}(m)}{N_{c}^{\mathrm{generated}}(m)}, \tag{5.13}$$

Table 5.6 Unweighted background cutflow for the low-activity category SR3 in the 8 GeV positron simulation. PN denotes photonuclear interactions, EN positron-nuclear interactions, and GMM photon conversion to muon pairs. The upper block lists effective positrons on target (POT), generation attempts, and counts after the production filter. The lower block gives cumulative counts after each analysis requirement. Counts follow Ref.[68].

Sample	Effective POT	Attempts	Filter
PN ECAL	1.48×10^{11}	1×10^{8}	5829159
PN Target	4.99×10^{12}	1×10^{8}	5623324
EN Target	2.78×10^{12}	1×10^{8}	1385618
GMM ECAL	2.02×10^{13}	1×10^{8}	9333419
Annihil Target	3.38×10^{12}	1×10^{8}	289221
GMM Target	4.73×10^{13}	1×10^{7}	663639
EN ECAL	3.23×10^{11}	1×10^{8}	124934
Inclusive	2.5×10^{9}	2.5×10^{9}	1472414

Sample	$N_{trk}^{tag} = 1$	$N_{trk}^{rec} = 0$	E_{ECAL}^{total} < 2.0 GeV	E_{HCAL}^{total} < 0.1 GeV	$E_{HCAL}^{MaxCell}$ < 1 MeV	$E_{ECAL}^{MaxCell}$ < 1 MeV
PN ECAL	5794358	138513	120147	3922	1096	0
PN Target	5386830	5391	4176	66	14	0
EN Target	1326519	1633	1554	2	0	0
GMM ECAL	9331661	67950	67950	0	0	0
Annihil Target	288943	20649	1111	0	0	0
GMM Target	663576	0	0	0	0	0
EN ECAL	124241	0	0	0	0	0
Inclusive	1470635	9240	8390	61	11	0

where r labels SR1–SR3 and c labels the three generated mechanisms. Each column is obtained from an independently simulated channel sample. Because the regions are disjoint, $\sum_r R_{rc} \leq 1$; events failing every region supply the remaining fraction.

With the channel order bremsstrahlung, non-resonant annihilation, and resonant annihilation, the response can be written explicitly as

$$\mathbf{R}(m) = \begin{pmatrix} R_{1,\mathrm{brem}} & R_{1,\mathrm{nonres}} & R_{1,\mathrm{res}} \\ R_{2,\mathrm{brem}} & R_{2,\mathrm{nonres}} & R_{2,\mathrm{res}} \\ R_{3,\mathrm{brem}} & R_{3,\mathrm{nonres}} & R_{3,\mathrm{res}} \end{pmatrix}. \tag{5.14}$$

All elements are dimensionless. A column follows one generated mechanism through the three measured categories and describes how its events are retained or lost. A row identifies the production rates that must be added to predict the yield in one measured category.

For each channel, define the total selected fraction and the rejected fraction as

$$\eta_c = \sum_{r=1}^{3} R_{rc}, \qquad R_{0c} = 1 - \eta_c. \tag{5.15}$$

The index 0 collects events that fail all three regions. Including it makes each column sum to unity.

5.3.3 Combining channels and changing the reference coupling

For invisible branching fraction $\mathrm{Br}_{\mathrm{inv}}$, the selected yield is

$$s_r(m,\epsilon) = \sum_c \mathcal{L}_c\, \sigma_c(m,\epsilon)\, \mathrm{Br}_{\mathrm{inv}}\, R_{rc}(m). \tag{5.16}$$

The benchmark takes $\mathrm{Br}_{\mathrm{inv}} \simeq 1$. At fixed masses and dark coupling the production rate scales as ϵ^2, allowing one common signal strength to scale the full matrix of contributions. The sum over channels is carried out separately for each signal region. It therefore includes events reconstructed outside the region most closely associated with their production topology, using the same signal strength for all contributions.

In vector notation the prediction is

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \mathbf{R} \begin{pmatrix} u_{\mathrm{brem}} \\ u_{\mathrm{nonres}} \\ u_{\mathrm{res}} \end{pmatrix}. \tag{5.17}$$

Summing the region predictions gives

$$s_{\mathrm{all}} = \sum_{r=1}^{3} s_r = \sum_c u_c \eta_c. \tag{5.18}$$

Each selected event enters this sum once because the regions are disjoint. The lost fraction R_{0c} contributes to the produced count but not to s_{all}. This sum is useful for checking the normalisation, although the statistical analysis retains the region information and its separate background predictions.

At fixed masses and dark-sector parameters, the benchmark can be expressed relative to a reference coupling:

$$\sigma_c(m,\epsilon) = \frac{\epsilon^2}{\epsilon_{\mathrm{ref}}^2}\sigma_c(m,\epsilon_{\mathrm{ref}}), \qquad s_r(m,\epsilon) = \frac{\epsilon^2}{\epsilon_{\mathrm{ref}}^2} s_r(m,\epsilon_{\mathrm{ref}}). \tag{5.19}$$

The second equality uses the same detector response and invisible branching fraction at both couplings. It is the scaling used to relate the common signal strength in the likelihood to ϵ^2. The channel proportions at a given mass remain fixed under this rescaling.

5.3.4 Example calculations with the benchmark yield table

At 90 MeV and $\epsilon^2 = 10^{-12}$, the non-resonant sample gives 1.96 produced events and 1.74 events in SR2 (Table 5.7). Their ratio,

$$R_{2,\mathrm{nonres}}(90\ \mathrm{MeV}) \simeq \frac{1.74}{1.96}, \tag{5.20}$$

is the generated-to-SR2 efficiency inferred from the rounded table entries. The same generated sample contributes 0.0389 events to SR1 and 0.114 events to SR3.

Reading the table by region gives a different calculation. The SR2 prediction at 90 MeV adds the bremsstrahlung, non-resonant, and resonant entries,

$$s_2|_{m=90\ \mathrm{MeV},\ \epsilon^2=10^{-12}} \simeq 0.000111 + 1.74 + 0.0870. \tag{5.21}$$

This sum agrees with the quoted total of 1.82 to the precision of the displayed values. It includes resonant events in the photon category and bremsstrahlung events without the recoil-track topology of SR1.

The response efficiency and the selected channel composition have different normalisations. For a region with nonzero signal yield, its predicted fraction from channel c is

$$F_{c|r} = \frac{u_c R_{rc}}{s_r}, \qquad \sum_c F_{c|r} = 1. \tag{5.22}$$

The efficiency R_{rc} starts from events of a specified generated channel. The composition $F_{c|r}$ starts from the predicted signal in a specified measured region and includes the relative production rates. For example, $1.74/1.96$ characterises the non-resonant SR2 efficiency at 90 MeV, while $1.74/1.82$ characterises its contribution to the SR2 signal mixture.

5.3.5 Background exposure and the role of effective POT

The background analysis uses normalisation factors calculated from the physical process cross sections and the biasing studies. For each dedicated sample, Equation (5.2) expresses this normalisation as an equivalent unbiased beam exposure, incorporating the process branching ratio and generator-filter acceptance. These analysis factors are calculated independently of the transport-level weight stored in the simulation event record; that field is not used to normalise the present analysis.

The exposure multiplier associated with a background sample p is

$$k_p = \frac{N_{e^+}}{N^{\text{eff}}_{\text{POT},p}}. \tag{5.23}$$

It expresses the ratio of the desired physical exposure to the exposure represented by that sample. A process yield normalised to $N^{\text{eff}}_{\text{POT},p}$ is multiplied by k_p to obtain the benchmark yield.

5.4 Background prediction beyond the simulated statistics

The benchmark beam count exceeds the inclusive sample size by five orders of magnitude. Even the biased samples have no survivors after the full cuts, so a direct count cannot supply the nominal background prediction. The analysis instead models survival fractions in populated sidebands and continues the fitted functions toward the signal-region thresholds, following the approach used for electron running[69]. This continuation is an assumption on which the projected limits depend.

The survival curves are cumulative: relaxing an upper energy threshold includes the events accepted at every tighter value. Each fitted curve is evaluated at the nominal signal-region threshold.

5.4.1 Dedicated rare-process estimates

The initial comparison in Figure 5.9 uses branching-ratio weights to put the dedicated samples on the same scale as inclusive production. It includes the target-annihilation category as well as the six PN, EN, and GMM categories. The agreement in populated sidebands tests the normalisation and biased production.

For SR1 and SR2, the extrapolation variable x is the total-ECAL-energy threshold expressed in MeV. After the other requirements are applied, each contributing sample is fitted with

$$f(x) = \exp(a + b \ln x), \tag{5.24}$$

using separately fitted parameters. Branching-ratio normalisation and the benchmark POT convert the survival fractions to expected event counts. SR1 receives EN and PN contributions, yielding a nominal total of 0.03 events below 2.5 GeV (Figure 5.10). GMM is omitted

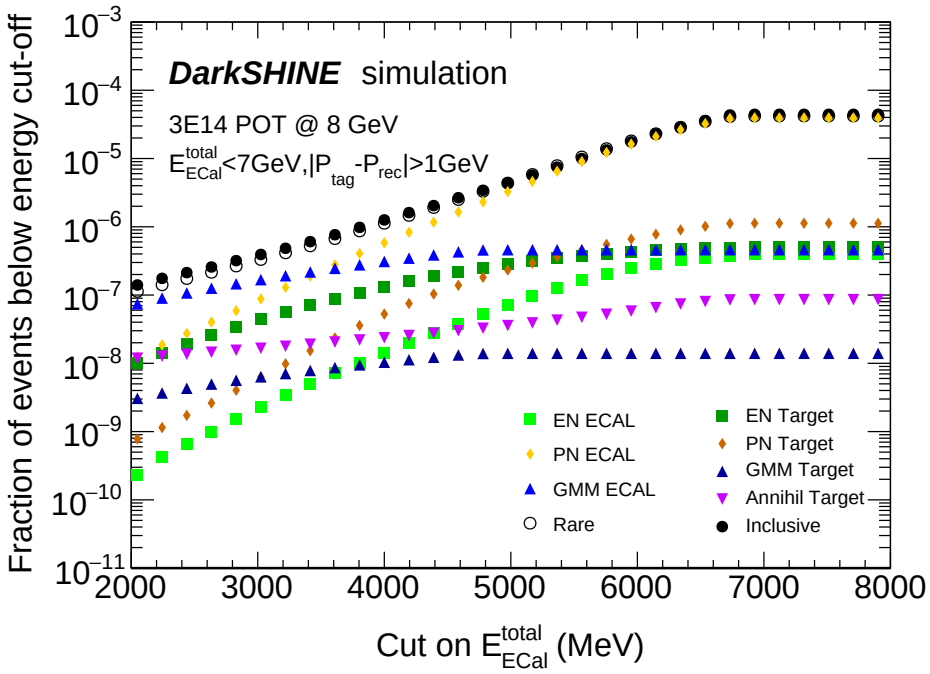


Figure 5.9 Background survival fraction versus the total electromagnetic-calorimeter energy threshold in the 8 GeV positron simulation. The loose production preselection requires $|P_{\mathrm{tag}} - P_{\mathrm{rec}}| > 1$ GeV and total ECAL energy below 7 GeV. The inclusive sample is compared with the normalised dedicated rare-process samples, including target annihilation. Their agreement is tested where both samples contain events.

from this nominal estimate because no event remains after the relevant vetoes in the available simulation; the unmeasured tail is a limitation of that approximation.

For the numerical threshold variable used here, the fitted expression is also $f(x) = e^a x^b$. Its logarithm is linear in $\ln x$, so b describes the slope in the logarithmic threshold coordinate and a fixes the normalisation in the chosen units. The value of x in the fit is the threshold expressed in MeV. Changing that numerical convention changes the fitted normalisation. The 2.5 GeV SR1 working point is therefore evaluated as 2500 in the fit variable.

In SR2 the zero-recoil-track condition suppresses target EN, and the HCAL vetoes remove GMM in the available samples. PN therefore controls the modeled residual. The two contributions in Figure 5.11 give 1.3×10^{-4} expected events at the 2 GeV total-energy threshold.

SR3 requires a joint treatment of the ECAL total and maximum-cell energies. Writing

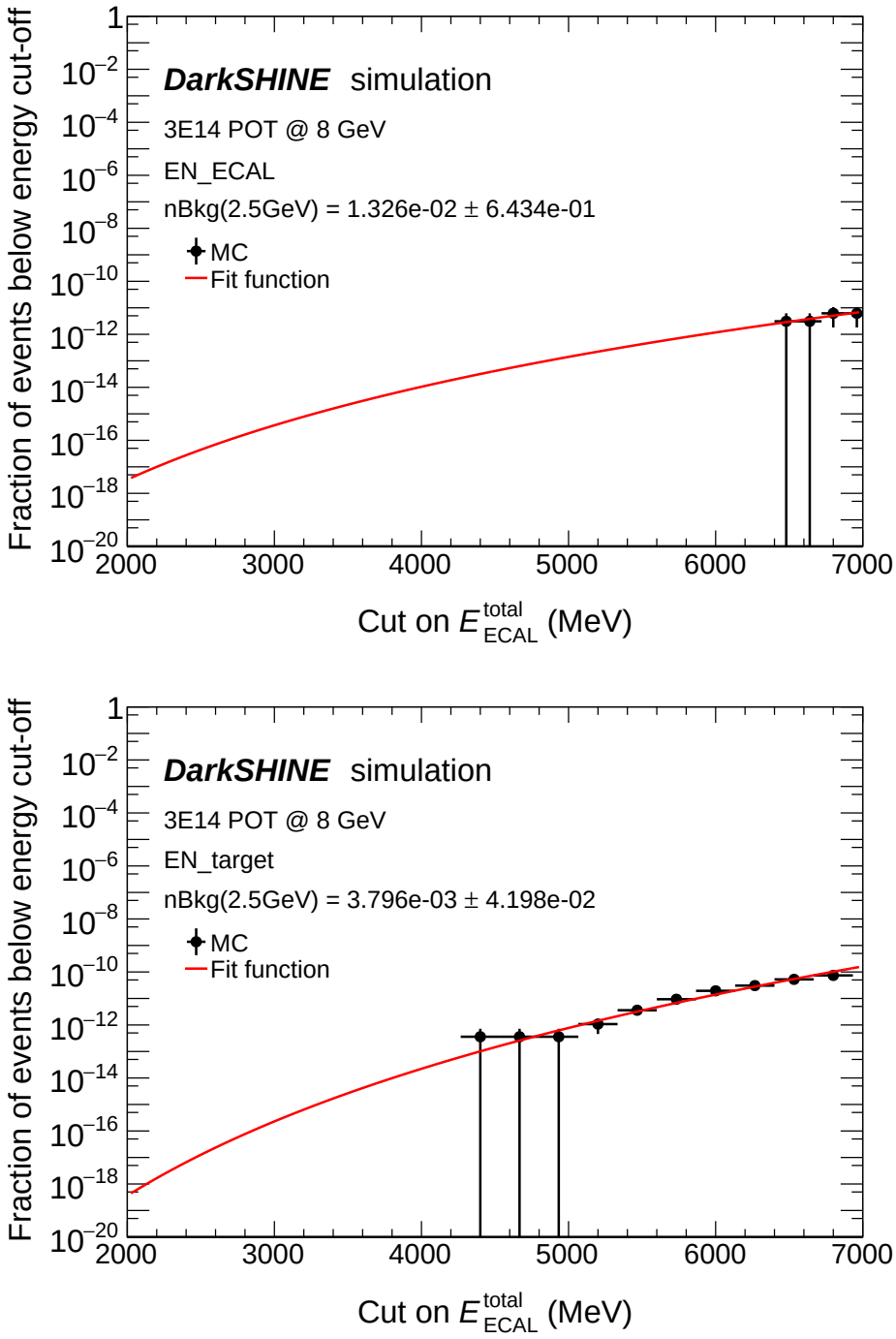
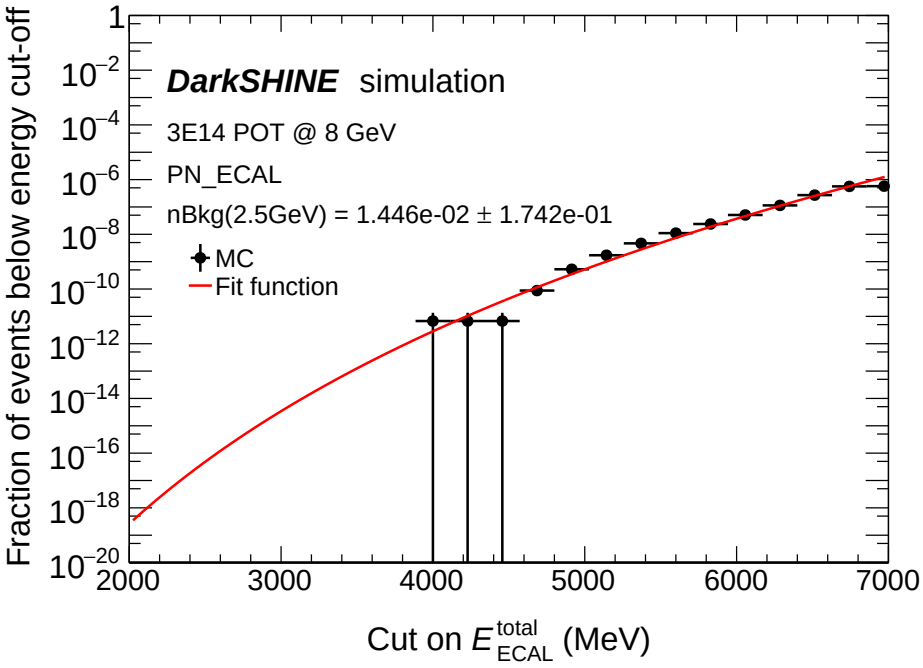


Figure 5.10 Positron-nuclear (EN) background survival fractions for the recoil-positron category SR1 in the 8 GeV positron simulation, for ECAL (upper) and target (lower) production. Black points show simulated cumulative fractions after all other SR1 requirements, versus the total-ECAL-energy threshold x, in MeV. Red curves are independent fits of $f(x) = \exp(a + b \ln x)$, continued to the nominal threshold $x = 2500$ MeV. Yield annotations use 3×10^{14} POT. Photonuclear contributions are shown in the continuation.



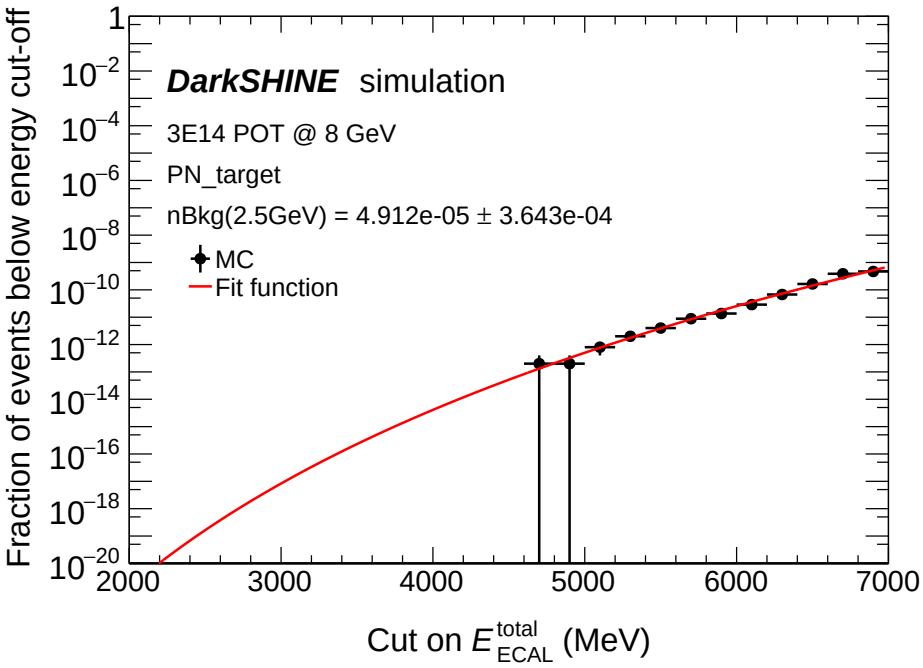


Figure 5.10 (continued) Photonuclear (PN) background survival fractions for the recoil-positron category SR1 in the 8 GeV positron simulation, for ECAL (upper) and target (lower) production. Black points show simulated cumulative fractions after all other SR1 requirements, versus the total-ECAL-energy threshold x, in MeV. Red curves are independent fits of $f(x) = \exp(a + b \ln x)$, continued to the nominal threshold $x = 2500$ MeV. Yield annotations use 3×10^{14} POT.

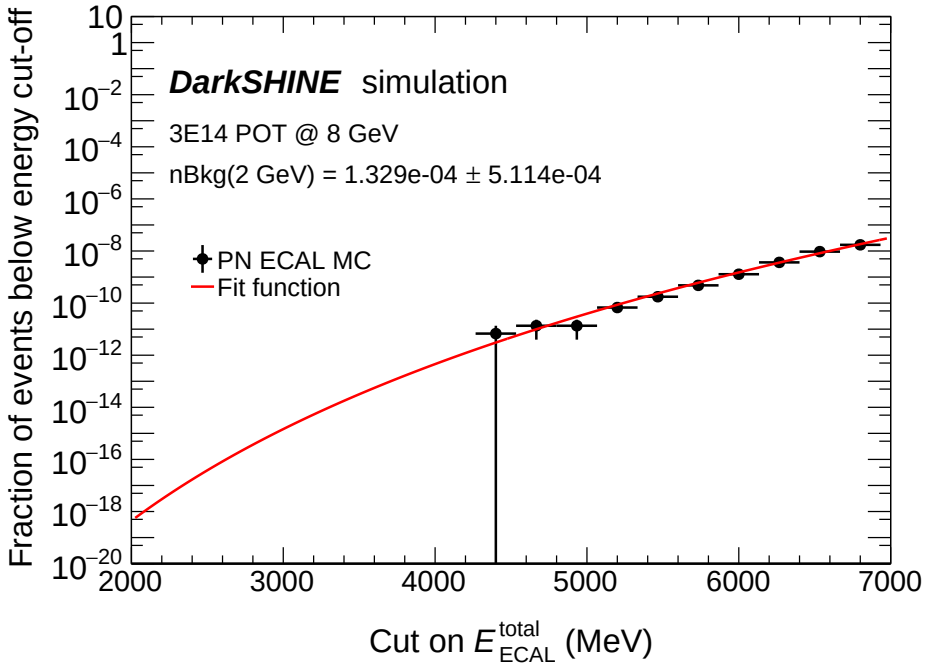


Figure 5.11 Photonuclear (PN) background survival fractions for the photon category SR2 in the 8 GeV positron simulation. Black points show cumulative simulated fractions after all other SR2 requirements, versus the total-ECAL-energy threshold x, in MeV. Red curves are fits of $f(x) = \exp(a + b \ln x)$, extrapolated to $x = 2000$ MeV. Yield annotations use 3×10^{14} POT; their uncertainties come from fitting and sample statistics and are not included as nuisance parameters in the sensitivity calculation. Interactions in the ECAL.

their thresholds as y and x, respectively, both in MeV, the fitted surface is

$$f(x, y) = a\,\{\exp[1 - by\exp(-c(x - d))] - 1\}\,. \tag{5.25}$$

The parameters are obtained independently for PN ECAL and PN target. Evaluating the continuation at $x = 1$ MeV and $y = 2000$ MeV gives a nominal combined yield of 1.6×10^{-4} events. The two-dimensional tail extrapolation in Figure 5.12 keeps both thresholds explicit, so the total-energy requirement is included when evaluating the maximum-cell-energy tail.

5.4.2 Cross-check with the inclusive sample

The inclusive sample checks the trend of the survival fraction over the range where simulated events remain. For SR1 and SR2, all selections except total ECAL energy are applied before fitting $g(x) = \exp(a + bx)$, with independent parameters for each region. For SR3 the fit uses Equation (5.25), again with parameters obtained from that sample. The corresponding

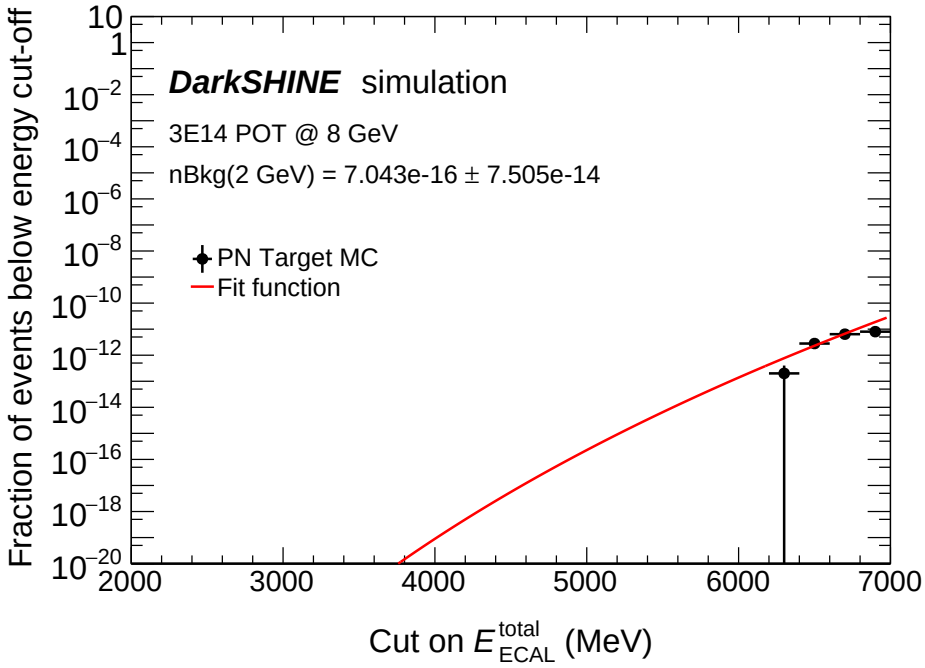


Figure 5.11 (continued) Photonuclear (PN) background survival fractions for the photon category SR2 in the 8 GeV positron simulation. Black points show cumulative simulated fractions after all other SR2 requirements, versus the total-ECAL-energy threshold x, in MeV. Red curves are fits of $f(x) = \exp(a + b \ln x)$, extrapolated to $x = 2000$ MeV. Yield annotations use 3×10^{14} POT; their uncertainties come from fitting and sample statistics and are not included as nuisance parameters in the sensitivity calculation. Interactions in the target.

comparisons appear in Figures 5.13 to 5.15.

The inclusive one-dimensional fits use $\exp(a + bx)$, an exponential in the numerical energy threshold. The dedicated estimates use the power law $\exp(a + b \ln x)$. The inclusive fits provide a cross-check in the populated sidebands; the final background inputs are the dedicated estimates.

5.5 Expected sensitivity and its physical interpretation

5.5.1 Statistical construction used for the projection

For each mediator mass, the signal prediction combines the production rates with all nine efficiencies through Equation (5.16). A single signal strength μ scales all production mechanisms. The reported projection uses a binned profile-likelihood analysis combining

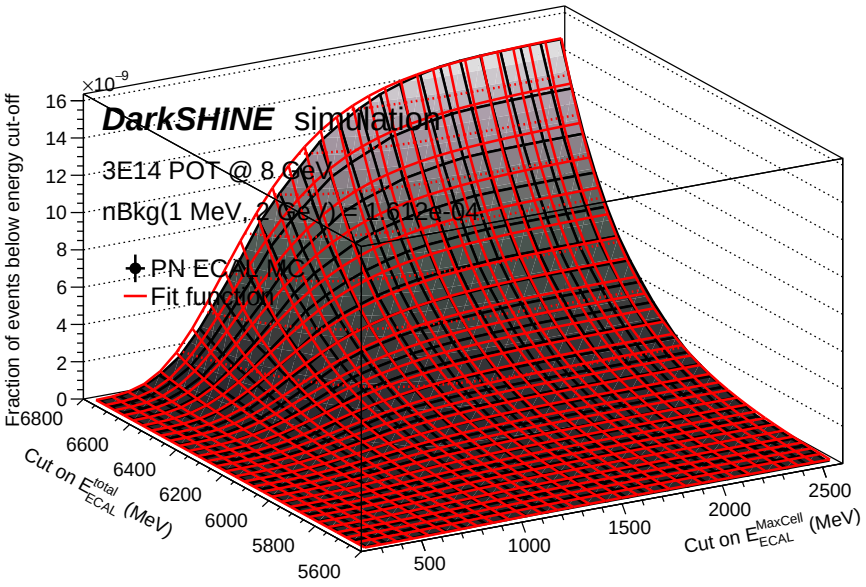


Figure 5.12 Photonuclear (PN) background survival fractions and fitted surfaces for the low-activity category SR3 in the 8 GeV positron simulation. After the other SR3 requirements, joint survival is fitted with $f(x,y) = a\{\exp[1 - by\exp(-c(x-d))] - 1\}$, where x and y are the single-cell and total-ECAL-energy thresholds, in MeV. The nominal point $(x,y) = (1, 2000)$ MeV lies beyond the populated sidebands. Interactions in the ECAL.

SR1, SR2, and SR3, with test statistic constructed from

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})}, \qquad \mu \simeq \frac{\epsilon^2}{\epsilon^2_{\text{ref}}}. \tag{5.26}$$

Here $\boldsymbol{\theta}$ collects nuisance parameters and the hatted quantities are conditional or unconditional maximum-likelihood estimates. In the procedure reported in Ref.[68], background shape uncertainties are uncorrelated between bins, and background normalisation uncertainties are based on Poisson counting statistics with the stated $\sqrt{N_{\text{bkg}}}$ prescription. The expected 90% confidence-level limits use CL_s and the asymptotic approximation under the background-only hypothesis; the applicability of this approximation has been checked for the present analysis.

The fitting and sample statistical uncertainties shown with the sideband yields are not included as nuisance parameters. The uncertainty in extrapolating the fitted functions beyond the populated range is difficult to quantify in full. The extrapolated background model provides a reasonable working estimate for this prospective study at the current detector-design stage.

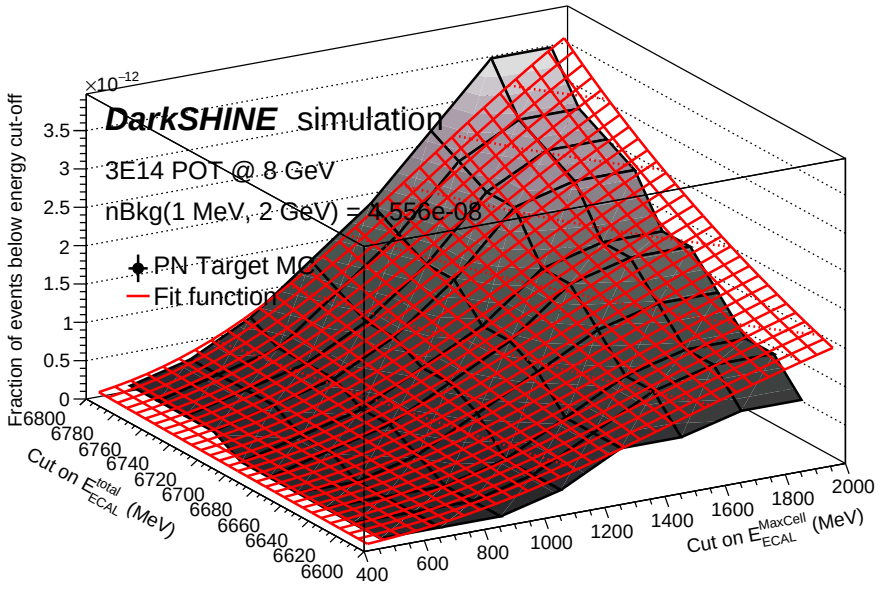


Figure 5.12 (continued) Photonuclear (PN) background survival fractions and fitted surfaces for the low-activity category SR3 in the 8 GeV positron simulation. After the other SR3 requirements, joint survival is fitted with $f(x,y) = a\{\exp[1 - by\exp(-c(x-d))] - 1\}$, where x and y are the single-cell and total-ECAL-energy thresholds, in MeV. The nominal point $(x,y) = (1, 2000)$ MeV lies beyond the populated sidebands. Interactions in the target.

The upper limit on the signal strength, $\mu_{90}(m)$, gives the coupling limit

$$\epsilon^2_{90}(m) = \mu_{90}(m)\epsilon^2_{\text{ref}}. \tag{5.27}$$

5.5.2 Relative importance of the production mechanisms

To see how much each production mechanism contributes, we retain one component at a time while keeping all three signal regions and the same background model. This comparison includes events that migrate away from the region associated with their production topology. Figure 5.16 makes this comparison. At 90 MeV the individual limits on ϵ^2 are 1.34×10^{-9} for dark bremsstrahlung, 7.29×10^{-13} for non-resonant annihilation, and 1.97×10^{-12} for resonant annihilation. Their combination reaches 5.33×10^{-13}.

The benchmark yields in Table 5.7 use fixed $\epsilon^2 = 10^{-12}$. At 10 MeV the selected signal is concentrated in SR1; at 90 MeV, SR2 and SR3 dominate. Totals are summed before rounding and can differ slightly from sums of the displayed entries.

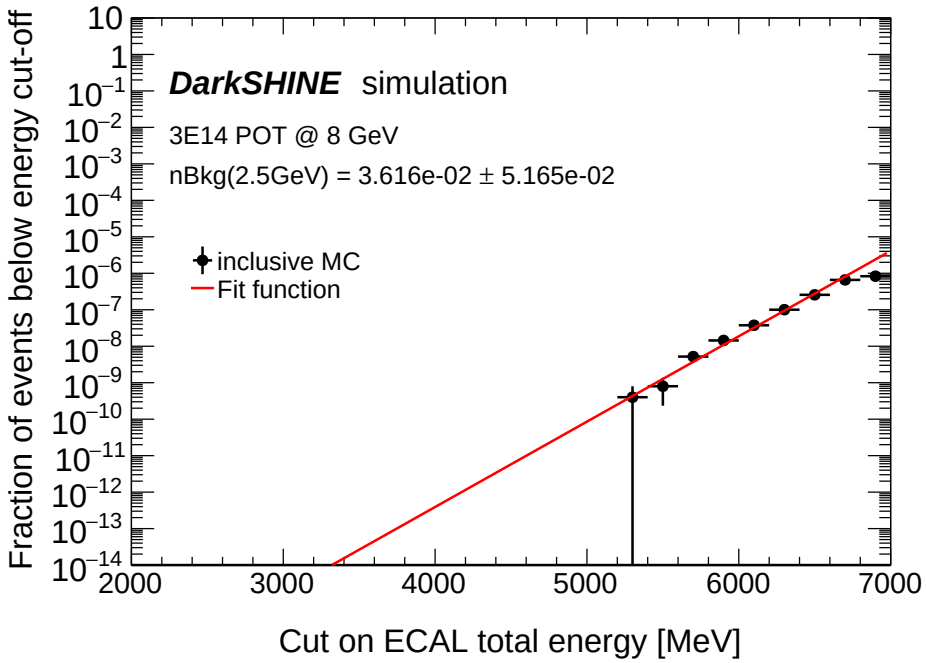


Figure 5.13 Inclusive-background survival fraction for the recoil-positron category SR1 in the 8 GeV positron simulation. All SR1 requirements except total ECAL energy are applied before scanning its upper threshold x, in MeV. Black points show the simulated cumulative fractions; the red curve is a fit of $g(x) = \exp(a + bx)$. The sample contains 2.5×10^9 POT and no event at the final 2.5 GeV threshold.

Table 5.7 Produced and selected signal counts at $\epsilon^2 = 10^{-12}$ for the benchmark 3×10^{14} POT. SR columns include migration. Numerical inputs are reproduced from Ref.[68] to illustrate the luminosity and efficiency factors in Equation (5.16).

$m_{A'}$ (MeV)	Production mode	N_{prod}	SR1	SR2	SR3
10	$e^+N \to e^+NA'$	0.173	0.115	0.00843	0.00889
	$e^+e^-N \to \gamma A'N$	0.00743	4.97×10^{-5}	0.00265	0.000526
	$e^+e^-N \to A'N$	0.0179	4.27×10^{-5}	0.00220	0.0153
	Total	0.198	0.116	0.0133	0.0247
50	$e^+N \to e^+NA'$	0.00605	0.00386	0.000456	0.000550
	$e^+e^-N \to \gamma A'N$	0.0116	8.19×10^{-5}	0.00439	0.000605
	$e^+e^-N \to A'N$	0.0351	8.00×10^{-5}	0.00438	0.0299
	Total	0.0528	0.00403	0.00923	0.0311
90	$e^+N \to e^+NA'$	0.00137	0.000850	0.000111	0.000142
	$e^+e^-N \to \gamma A'N$	1.96	0.0389	1.74	0.114
	$e^+e^-N \to A'N$	0.705	0.00156	0.0870	0.601
	Total	2.67	0.0413	1.82	0.715

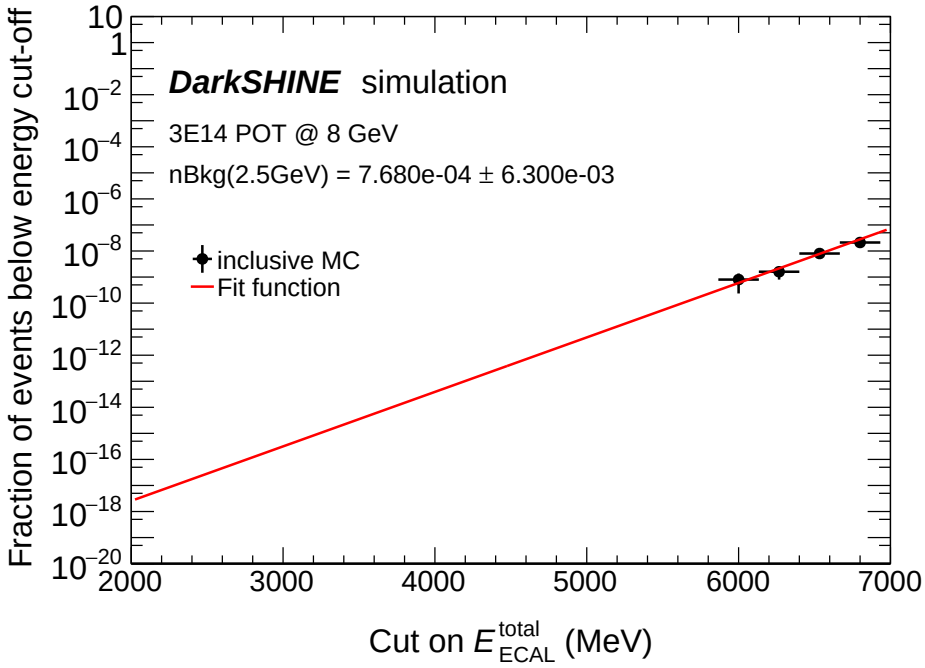


Figure 5.14 Inclusive-background survival fraction for the photon category SR2 in the 8 GeV positron simulation. Black points show cumulative fractions from 2.5×10^9 POT after all other SR2 requirements, versus the total-ECAL- energy upper threshold x, in MeV. The red curve is a fit of $g(x) = \exp(a + bx)$. The displayed yield annotation is evaluated at 2.5 GeV and scaled to 3×10^{14} POT, while the nominal SR2 threshold is 2 GeV. Its quoted uncertainty comes from fitting and sample statistics and is not included as a nuisance parameter in the sensitivity calculation.

The three benchmark masses in Table 5.7 separate changes in production from changes in selection. At 10 MeV, the produced bremsstrahlung count of 0.173 exceeds the two annihilation counts, and its selected contribution is largest in SR1. At 50 MeV, the resonant produced count of 0.0351 exceeds the non-resonant count of 0.0116 and the bremsstrahlung count of 0.00605. SR3 then contains the largest total selected count, 0.0311. At 90 MeV, the non-resonant produced count reaches 1.96 and its SR2 contribution is 1.74, which accounts for most of the 1.82 events predicted in that region.

The resonant contribution at 90 MeV is also retained in both zero-recoil categories: 0.0870 events enter SR2 and 0.601 enter SR3. The latter exceeds the non-resonant SR3 contribution of 0.114, even though non-resonant production provides the strongest individual limit after all three regions are used.

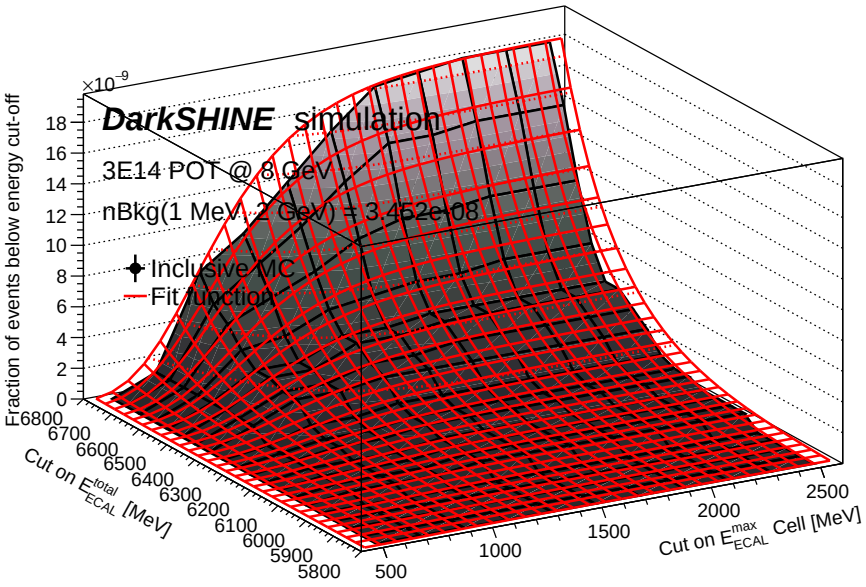


Figure 5.15 Inclusive-background survival in the two-threshold ECAL plane for the low-activity category SR3 in the 8 GeV positron simulation. The coordinates are upper thresholds on total and largest-cell ECAL energy, in MeV, after all other SR3 requirements. The red wireframe is the fitted joint-survival surface $f(x,y) = a\{\exp[1-by\exp(-c(x-d))]-1\}$, with x the single-cell threshold and y the total-energy threshold. It is not a selection boundary. No inclusive event passes the nominal thresholds $(x,y) = (1, 2000)$ MeV.

5.5.3 Comparison with other searches and thermal targets

The result in the $(m_{A'}, \epsilon^2)$ plane is shown in Figure 5.17. Around 90 MeV the combined positron projection is about two orders of magnitude stronger in ϵ^2 than the earlier DarkSHINE electron projection at the same integrated luminosity. Dark bremsstrahlung remains particularly relevant at lower masses, while the annihilation channels account for the sharper improvement near the scale in Equation (5.1).

For comparison with thermal-relic targets the same benchmark is mapped to

$$y = \epsilon^2 \alpha_D \left(\frac{m_\chi}{m_{A'}}\right)^4, \qquad m_{A'} = 3m_\chi, \quad \alpha_D = 0.5. \tag{5.28}$$

Figure 5.18 uses m_χ as the horizontal coordinate. The 90 MeV mediator scale therefore corresponds to a dark-matter mass near 30 MeV for this mass ratio. Scalar, Majorana, and pseudo-Dirac curves give model-dependent thermal-relic benchmarks. Changing the beam energy moves the annihilation scale approximately as $\sqrt{E_{e^+}}$, motivating studies with several

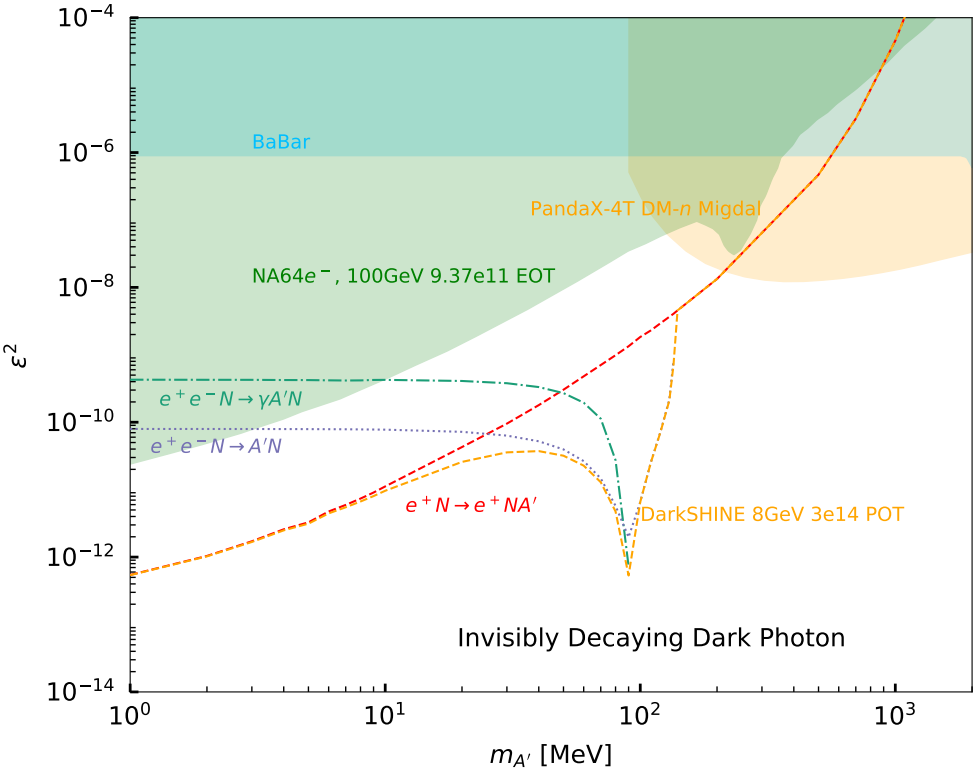


Figure 5.16 Expected 90% confidence-level limits on squared kinetic mixing ϵ^2 versus mediator mass for an 8 GeV positron beam and 3×10^{14} positrons on target. Separate dark-bremsstrahlung, non-resonant-annihilation, and resonant-annihilation projections are compared with their combination. Every fit retains all three signal regions and the same background model. Curves and comparison constraints follow Ref.[68].

positron energies.

For the chosen mass ratio and dark coupling, the coordinate transformation is particularly simple:

$$m_\chi = \frac{m_{A'}}{3}, \qquad y = \frac{\epsilon^2}{162}. \tag{5.29}$$

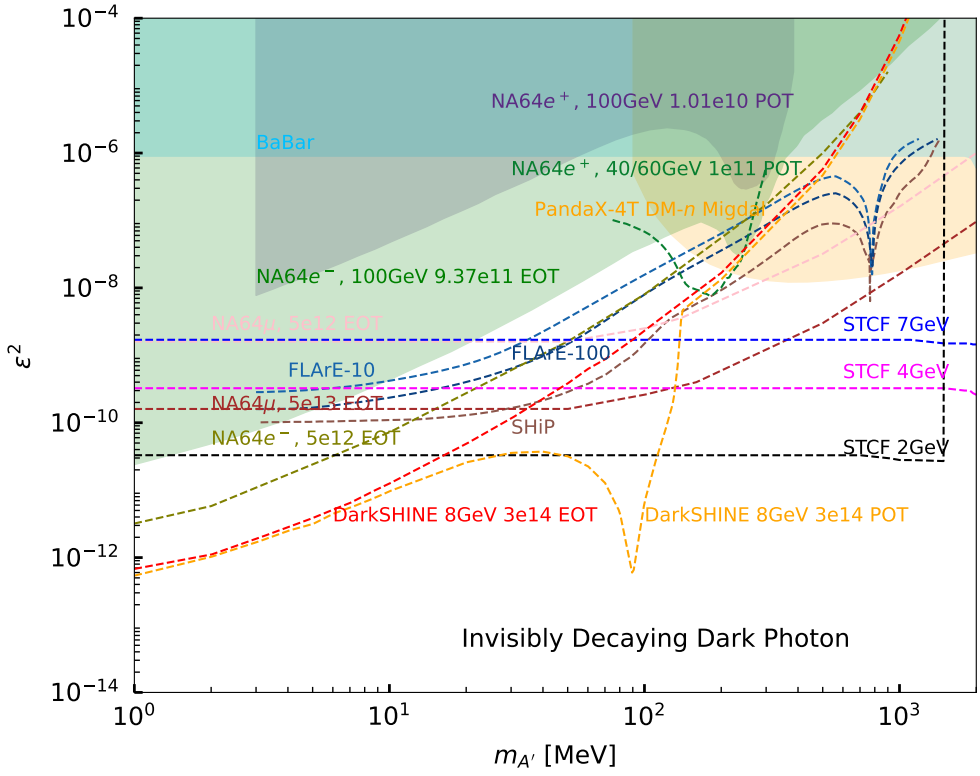


Figure 5.17 Expected 90% confidence-level sensitivity to squared kinetic mixing ϵ^2 versus mediator mass for an 8 GeV positron beam and 3×10^{14} positrons on target (yellow), with electron-beam DarkSHINE projections labelled by luminosity. The comparison includes existing NA64, PandaX-4T, and BaBar constraints[33-34,36], the NA64 positron result and 40/60 GeV projection[72], and STCF, SHiP, and FLArE projections[73-74]. Figure from Ref.[68].

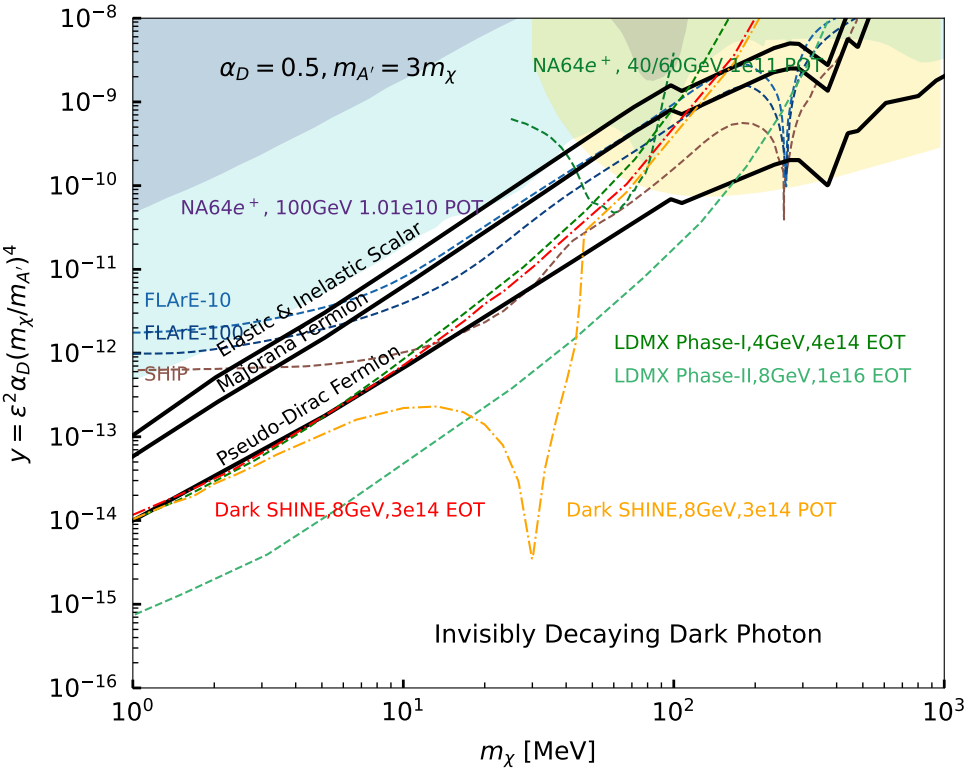


Figure 5.18 Expected 90% confidence-level DarkSHINE sensitivity and thermal-relic targets in the (m_χ, y) plane, where m_χ is the dark-matter mass and $y = \epsilon^2 \alpha_D (m_\chi/m_{A'})^4$. The benchmark fixes $m_{A'}/m_\chi = 3$ and $\alpha_D = 0.5$; the 8 GeV positron projection uses 3×10^{14} positrons on target. Existing constraints and thermal targets follow Refs.[33-34,36-37,75-80]; NA64 positron, LDMX, SHiP, and FLArE comparisons follow Refs.[35,72,74]. Figure from Ref.[68].

5.6 Summary

The positron analysis connects three production mechanisms to three exclusive reconstructed event categories through a full efficiency matrix. Both annihilation mechanisms are essential near the 8 GeV beam's characteristic annihilation mass; non-resonant annihilation gives the strongest individual projected limit at 90 MeV. With the benchmark luminosity and assumed backgrounds, their combination reaches $\epsilon^2 = 5.33 \times 10^{-13}$ there. Varying the beam energy would shift the annihilation scale and extend the study to other dark-photon masses.

Chapter 6 Summary and Outlook

This thesis presents the DarkSHINE simulation and reconstruction software and applies it to a prospective positron-beam search for invisibly decaying dark photons.

The author is one of the original designers of the DarkSHINE software framework and led the development of `DSimu`, the `DEvent` event model, interfaces for detector-geometry optimisation and performance analysis, rare-process biasing and filtering, ACTS integration, and the positron-beam analysis. The reconstruction environment also incorporates collaboration work on `DAna`, digitisation, other tracking algorithms, calorimeter reconstruction, and generator validation. These components provide the detector-level observables used in the search.

6.1 Simulation and reconstruction

The DarkSHINE software suite provides a configurable description of the experiment, from primary-particle generation and Geant4 transport to detector digitisation, reconstruction, and analysis output. A common event-data model allows truth, simulated-hit, digitised-hit, and reconstructed-particle collections to be exchanged between applications. Cross-section biasing and staged event filters make it possible to produce the large rare-process samples required by a missing-energy experiment. For detailed calorimeter studies, full optical-photon transport is available. The lookup-table approach reduces the measured processing time from about 200 s to 4 s per event, with photon-detection efficiencies agreeing to within 5%. The same geometry and response model can then be used in detector optimisation and analysis. For detector optimisation, the geometry and analysis interfaces provide the inputs for comparing layouts through tracking acceptance, momentum resolution, calorimeter energy response, and background rejection.

6.2 Positron-beam search

The positron-beam study extends the conventional dark-bremsstrahlung search with non-resonant and resonant positron–electron annihilation. Three orthogonal signal regions were

defined using recoil-track multiplicity, missing momentum, total calorimeter energy, and maximum-cell energy. Full detector simulation evaluates the efficiency of each production mechanism in every region, including migration between categories. Annihilation improves the reach around the resonance accessible with an 8 GeV beam. Dedicated biased samples supply the background sidebands for extrapolation, while an independent inclusive sample tests agreement where simulated events remain.

The produced and selected counts in the benchmark tables explain why the relative importance of the mechanisms changes with mass. At the lower displayed mass the selected signal is concentrated in the recoil category. Near the characteristic annihilation scale, the two zero-recoil-track categories supply most of the selected yield. The individual-mechanism sensitivity comparison retains all three regions, so it includes these migrations when identifying the contribution of a process.

At $m_{A'} = 90$ MeV, non-resonant annihilation gives the strongest individual projected limit; combining all mechanisms reaches $\epsilon^2 = 5.33 \times 10^{-13}$ for the ten-year exposure of 3×10^{14} POT, approximately two orders of magnitude beyond the earlier electron-beam projection at matching integrated luminosity[68].

The comparison with thermal targets adds the assumptions $m_{A'} = 3m_{\chi}$ and $\alpha_D = 0.5$. With this mass ratio, a mediator near 90 MeV corresponds to a dark-matter mass near 30 MeV. The change of variables allows the production projection to be compared with the specified relic-density curves.

6.3 Assumptions of the projection

The final signal-region cuts leave no simulated background survivors. The residual yields are estimated by extrapolating fitted survival trends from populated sidebands to the nominal thresholds. Agreement between inclusive and dedicated samples tests the normalisation and distributions in those sidebands, while the more distant tails depend on the fitted continuation. Their extrapolation uncertainty is difficult to quantify fully with the available statistics. The simulated detector response, dedicated background samples, and inclusive cross-checks provide a reasonable basis for this prospective estimate at the current design stage. The asymptotic approximation used for the limit calculation has been checked and is applicable to this analysis.

The positron projection assumes an achievable 1 MeV HCAL maximum-cell threshold. The HCAL response does not include a complete treatment of electronic noise and pileup, so the attainable veto performance remains an experimental requirement.

6.4 Further work

Larger rare-process samples would extend the populated background tails and allow a fuller assessment of the extrapolation uncertainty. Pseudo-experiments could provide an additional check of the statistical model. Beam tests under operating conditions relevant to the veto thresholds would constrain the digitisation and reconstruction response.

Simulations at additional positron-beam energies would extend the annihilation resonance scan to other dark-photon masses.

Bibliography

[1] ZWICKY F. Die Rotverschiebung von extragalaktischen Nebeln[J]. Helvetica Physica Acta, 1933, 6: 110-127.

[2] FUSCO-FEMIANO R, HUGHES J P. Nonpolytropic Model for the Coma Cluster[J/OL]. Astrophysical Journal, 1994, 429: 545. DOI: 10.1086/174342.

[3] COLLESS M, DUNN A M. Structure and dynamics of the coma cluster[J/OL]. Astrophys. J., 1996, 458: 435. arXiv: astro-ph/9508070. DOI: 10.1086/176827.

[4] VAN ALBADA T S, BAHCALL J N, BEGEMAN K, et al. Distribution of dark matter in the spiral galaxy NGC 3198.[J/OL]. Astrophysical Journal, 1985, 295: 305-313. DOI: 10.1086/163375.

[5] FREEMAN K C. On the Disks of Spiral and S0 Galaxies[J/OL]. Astrophysical Journal, 1970, 160: 811. DOI: 10.1086/150474.

[6] ROBERTS M S, WHITEHURST R N. The rotation curve and geometry of M31 at large galactocentric distances.[J/OL]. Astrophysical Journal, 1975, 201: 327-346. DOI: 10.1086/153889.

[7] RUBIN V C, FORD W K, Jr., THONNARD N. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc).[J/OL]. Astrophysical Journal, 1980, 238: 471-487. DOI: 10.1086/158003.

[8] BOSMA A. 21-cm line studies of spiral galaxies. II. The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types.[J/OL]. Astronomical Journal, 1981, 86: 1825-1846. DOI: 10.1086/113063.

[9] BULLOCK J S, BOYLAN-KOLCHIN M. Small-Scale Challenges to the ΛCDM Paradigm[J/OL]. Ann. Rev. Astron. Astrophys., 2017, 55: 343-387. arXiv: 1707.04256 [astro-ph.CO]. DOI: 10.1146/annurev-astro-091916-055313.

[10] OH S H, de BLOK W J G, BRINKS E, et al. DARK AND LUMINOUS MATTER IN THINGS DWARF GALAXIES[J/OL]. The Astronomical Journal, 2011, 141(6): 193. http://dx.doi.org/10.1088/0004-6256/141/6/193. DOI: 10.1088/0004-6256/141/6/193.

[11] OH S H, et al. High-resolution mass models of dwarf galaxies from LITTLE THINGS[J/OL]. Astron. J., 2015, 149: 180. arXiv: 1502.01281 [astro-ph.GA]. DOI: 10.1088/0004-6256/149/6/180.

[12] BOYLAN-KOLCHIN M, BULLOCK J S, KAPLINGHAT M. The Milky Way's bright satellites as an apparent failure of LCDM[J/OL]. Mon. Not. Roy. Astron. Soc., 2012, 422: 1203-1218. arXiv: 1111.2048 [astro-ph.CO]. DOI: 10.1111/j.1365-2966.2012.20695.x.

[13] BOYLAN-KOLCHIN M, BULLOCK J S, KAPLINGHAT M. Too big to fail? The puzzling darkness of massive Milky Way subhaloes[J/OL]. Mon. Not. Roy. Astron. Soc., 2011, 415: L40. arXiv: 1103.0007 [astro-ph.CO]. DOI: 10.1111/j.1745-3933.2011.01074.x.

[14] BROOKS A. Re-Examining Astrophysical Constraints on the Dark Matter Model[J/OL]. Annalen Phys., 2014, 526(7-8): 294-308. arXiv: 1407.7544 [astro-ph.CO]. DOI: 10.1002/andp.201400068.

[15] SPERGEL D N, STEINHARDT P J. Observational evidence for selfinteracting cold dark matter [J/OL]. Phys. Rev. Lett., 2000, 84: 3760-3763. arXiv: astro-ph/9909386. DOI: 10.1103/PhysRevLett.84.3760.

[16] VOGELSBERGER M, MARINACCI F, TORREY P, et al. Cosmological Simulations of Galaxy Formation[J/OL]. Nature Rev. Phys., 2020, 2(1): 42-66. arXiv: 1909.07976 [astro-ph.GA]. DOI: 10.1038/s42254-019-0127-2.

[17] ROBERTSON A, MASSEY R, EKE V. What does the Bullet Cluster tell us about self-interacting dark matter?[J/OL]. Mon. Not. Roy. Astron. Soc., 2017, 465(1): 569-587. arXiv: 1605.04307 [astro-ph.CO]. DOI: 10.1093/mnras/stw2670.

[18] CLOWE D, BRADAC M, GONZALEZ A H, et al. A direct empirical proof of the existence of dark matter[J/OL]. Astrophys. J. Lett., 2006, 648: L109-L113. arXiv: astro-ph/0608407. DOI: 10.1086/508162.

[19] ALAM S, et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample[J/OL]. Mon. Not. Roy. Astron. Soc., 2017, 470(3): 2617-2652. arXiv: 1607.03155 [astro-ph.CO]. DOI: 10.1093/mnras/stx721.

[20] AGHANIM N, et al. Planck 2018 results. VI. Cosmological parameters[J/OL]. Astron. Astrophys., 2020, 641: A6. arXiv: 1807.06209 [astro-ph.CO]. DOI: 10.1051/0004-6361/201833910.

[21] MANGANO G, SERPICO P D. A robust upper limit on $N_{\rm eff}$ from BBN, circa 2011[J/OL]. Physics Letters B, 2011, 701(3): 296-299. http://dx.doi.org/10.1016/j.physletb.2011.05.075. DOI: 10.1016/j.physletb.2011.05.075.

[22] ADE P A R, et al. Planck 2015 results. XIII. Cosmological parameters[J/OL]. Astron. Astrophys., 2016, 594: A13. arXiv: 1502.01589 [astro-ph.CO]. DOI: 10.1051/0004-6361/201525830.

[23] BERLIN A, BLINOV N. Thermal Dark Matter Below an MeV[J/OL]. Phys. Rev. Lett., 2018, 120(2): 021801. arXiv: 1706.07046 [hep-ph]. DOI: 10.1103/PhysRevLett.120.021801.

[24] AIOLA S, et al. The Atacama Cosmology Telescope: DR4 Maps and Cosmological Parameters [J/OL]. JCAP, 2020, 12: 047. arXiv: 2007.07288 [astro-ph.CO]. DOI: 10.1088/1475-7516/2020/12/047.

[25] DUTCHER D, et al. Measurements of the E-mode polarization and temperature-E-mode correlation of the CMB from SPT-3G 2018 data[J/OL]. Phys. Rev. D, 2021, 104(2): 022003. arXiv: 2101.01684 [astro-ph.CO]. DOI: 10.1103/PhysRevD.104.022003.

[26] CARR B, RAIDAL M, TENKANEN T, et al. Primordial black hole constraints for extended mass functions[J/OL]. Phys. Rev. D, 2017, 96(2): 023514. arXiv: 1705.05567 [astro-ph.CO]. DOI: 10.1103/PhysRevD.96.023514.

[27] ZUMALACARREGUI M, SELJAK U. Limits on stellar-mass compact objects as dark matter from gravitational lensing of type Ia supernovae[J/OL]. Phys. Rev. Lett., 2018, 121(14): 141101. arXiv: 1712.02240 [astro-ph.CO]. DOI: 10.1103/PhysRevLett.121.141101.

[28] BRINGMANN T, DEPTA P F, DOMCKE V, et al. Towards closing the window of primordial black holes as dark matter: The case of large clustering[J/OL]. Phys. Rev. D, 2019, 99(6): 063532. arXiv: 1808.05910 [astro-ph.CO]. DOI: 10.1103/PhysRevD.99.063532.

[29] BODDY K K, GLUSCEVIC V, POULIN V, et al. Critical assessment of CMB limits on dark matter-baryon scattering: New treatment of the relative bulk velocity[J/OL]. Phys. Rev. D, 2018, 98(12): 123506. arXiv: 1808.00001 [astro-ph.CO]. DOI: 10.1103/PhysRevD.98.123506.

[30] ARCHIDIACONO M, HANNESTAD S, MIRIZZI A, et al. Axion hot dark matter bounds after Planck[J/OL]. JCAP, 2013, 10: 020. arXiv: 1307.0615 [astro-ph.CO]. DOI: 10.1088/1475-7516/2013/10/020.

[31] IRŠIČ V, et al. New Constraints on the free-streaming of warm dark matter from intermediate and small scale Lyman-α forest data[J/OL]. Phys. Rev. D, 2017, 96(2): 023522. arXiv: 1702.01764 [astro-ph.CO]. DOI: 10.1103/PhysRevD.96.023522.

[32] ARMENGAUD E, PALANQUE-DELABROUILLE N, YÈCHE C, et al. Constraining the mass of light bosonic dark matter using SDSS Lyman-α forest[J/OL]. Mon. Not. Roy. Astron. Soc., 2017, 471(4): 4606-4614. arXiv: 1703.09126 [astro-ph.CO]. DOI: 10.1093/mnras/stx1870.

[33] HUANG D, et al. Search for Dark-Matter–Nucleon Interactions with a Dark Mediator in PandaX-4T[J/OL]. Phys. Rev. Lett., 2023, 131(19): 191002. arXiv: 2308.01540 [hep-ex]. DOI: 10.1103/PhysRevLett.131.191002.

[34] ANDREEV Y M, et al. Search for Light Dark Matter with NA64 at CERN[J/OL]. Phys. Rev. Lett., 2023, 131(16): 161801. arXiv: 2307.02404 [hep-ex]. DOI: 10.1103/PhysRevLett.131.161801.

[35] ÅKESSON T, et al. Current Status and Future Prospects for the Light Dark Matter eXperiment [C/OL]//Snowmass 2021. 2022. arXiv: 2203.08192 [hep-ex].

[36] LEES J P, et al. Search for Invisible Decays of a Dark Photon Produced in e^+e^- Collisions at BaBar [J/OL]. Phys. Rev. Lett., 2017, 119(13): 131804. arXiv: 1702.03327 [hep-ex]. DOI: 10.1103/PhysRevLett.119.131804.

[37] BATELL B, POSPELOV M, RITZ A. Exploring Portals to a Hidden Sector Through Fixed Targets [J/OL]. Phys. Rev. D, 2009, 80: 095024. arXiv: 0906.5614 [hep-ph]. DOI: 10.1103/PhysRevD.80.095024.

[38] HOLDOM B. Two U(1)'s and Epsilon Charge Shifts[J/OL]. Phys. Lett. B, 1986, 166: 196-198. DOI: 10.1016/0370-2693(86)91377-8.

[39] Del AGUILA F, COUGHLAN G D, QUIROS M. Gauge Coupling Renormalization With Several U(1) Factors[J/OL]. Nucl. Phys. B, 1988, 307: 633. DOI: 10.1016/0550-3213(88)90266-0.

[40] ANDREAS S, NIEBUHR C, RINGWALD A. New Limits on Hidden Photons from Past Electron Beam Dumps[J/OL]. Phys. Rev. D, 2012, 86: 095019. arXiv: 1209.6083 [hep-ph]. DOI: 10.11

03/PhysRevD.86.095019.

[41] IZAGUIRRE E, KRNJAIC G, SCHUSTER P, et al. Testing GeV-Scale Dark Matter with Fixed-Target Missing Momentum Experiments[J/OL]. Phys. Rev. D, 2015, 91(9): 094026. arXiv: 1411.1404 [hep-ph]. DOI: 10.1103/PhysRevD.91.094026.

[42] BATTAGLIERI M, et al. Light dark matter searches with positrons[J/OL]. Eur. Phys. J. A, 2021, 57(8): 253. arXiv: 2105.04540 [hep-ex]. DOI: 10.1140/epja/s10050-021-00524-6.

[43] BJORKEN J D, ESSIG R, SCHUSTER P, et al. New Fixed-Target Experiments to Search for Dark Gauge Forces[J/OL]. Phys. Rev. D, 2009, 80: 075018. arXiv: 0906.0580 [hep-ph]. DOI: 10.1103/PhysRevD.80.075018.

[44] ZHAO Z, WANG D, YANG Z H, et al. SCLF: An 8-GeV CW SCRF Linac-Based X-Ray FEL Facility in Shanghai[C/OL]//38th International Free-Electron Laser Conference. 2018: MOP055. DOI: 10.18429/JACoW-FEL2017-MOP055.

[45] NOSOCHKOV Y, BEUKERS T, FRY A, et al. Dark Sector Experiments at LCLS-II (DASEL) Accelerator Design[C/OL]//8th International Particle Accelerator Conference. 2017. DOI: 10.18429/JACoW-IPAC2017-TUPIK121.

[46] WORKMAN R L, et al. Review of Particle Physics[J/OL]. PTEP, 2022, 2022: 083C01. DOI: 10.1093/ptep/ptac097.

[47] LIU K, LI M, ZHANG J, et al. Performance of AC-LGAD Strip Sensor Designed for the DarkSHINE Experiment[J/OL]. 2023. arXiv: 2310.13926 [physics.ins-det].

[48] IZAGUIRRE E, KRNJAIC G, SCHUSTER P, et al. Testing GeV-scale Dark Matter with Fixed-target Missing Momentum Experiments[J/OL]. Phys. Rev. D, 2015, 91: 094026. arXiv: 1411.1404 [hep-ex]. DOI: 10.1103/PhysRevD.91.094026.

[49] KIMBLE T, CHOU M, CHAI B H T. Scintillation properties of LYSO crystals[C/OL]//2002 IEEE Nuclear Science Symposium Conference Record: vol. 3. 2002: 1434-1437. DOI: 10.1109/NSSMIC.2002.1239590.

[50] KLANNER R. Characterisation of SiPMs[J/OL]. Nucl. Instrum. Meth. A, 2019, 926: 36-56. arXiv: 1809.04346 [physics.ins-det]. DOI: 10.1016/j.nima.2018.11.083.

[51] SIMON F. Silicon Photomultipliers in Particle and Nuclear Physics[J/OL]. Nucl. Instrum. Meth. A, 2019, 926: 85-100. arXiv: 1811.03877 [physics.ins-det]. DOI: 10.1016/j.nima.2018.11.042.

[52] GUO Y, LI S, LIU K, et al. Design of High-speed Readout Electronics for the DarkSHINE Electromagnetic Calorimeter[Z/OL]. 2024. arXiv: 2407.20723 [physics.ins-det].

[53] CHEN J, et al. Prospective study of light dark matter search with a newly proposed DarkSHINE experiment[J/OL]. Sci. China Phys. Mech. Astron., 2023, 66: 211062. DOI: 10.1007/s11433-022-1983-8.

[54] ZHAO Z, LIU Q, et al. Design of a LYSO Crystal Electromagnetic Calorimeter for DarkSHINE

Experiment[J/OL]. 2024. arXiv: 2407.17800 [physics.ins-det].

[55] CHEN J, et al. DarkSHINE Baseline Design Report: Physics Prospects and Detector Technologies [J/OL]. 2024. arXiv: 2411.09345 [physics.ins-det].

[56] AGOSTINELLI S, et al. GEANT4—a simulation toolkit[J/OL]. Nucl. Instrum. Meth. A, 2003, 506: 250-303. DOI: 10.1016/S0168-9002(03)01368-8.

[57] BRUN R, RADEMAKERS F. ROOT: An object oriented data analysis framework[J/OL]. Nucl. Instrum. Meth. A, 1997, 389: 81-86. DOI: 10.1016/S0168-9002(97)00048-X.

[58] WANG Z, YUAN R, LIU H Q, et al. Design of hadronic calorimeter for DarkSHINE experiment [J/OL]. Nucl. Sci. Tech., 2024, 35: 148. arXiv: 2311.01780 [physics.ins-det]. DOI: 10.1007/s41365-024-01502-5.

[59] KÅSA I. A circle fitting procedure and its error analysis[J/OL]. IEEE Trans. Instrum. Meas., 1976, IM-25: 8-14. DOI: 10.1109/TIM.1976.6312298.

[60] RAUCH J, SCHLÜTER T. GENFIT — a Generic Track-Fitting Toolkit[J/OL]. J. Phys. Conf. Ser., 2015, 608: 012042. arXiv: 1410.3698 [physics.ins-det]. DOI: 10.1088/1742-6596/608/1/012042.

[61] HÖPPNER C, NEUBERT S, KETZER B, et al. A Novel Generic Framework for Track Fitting in Complex Detector Systems[J/OL]. Nucl. Instrum. Meth. A, 2010, 620: 518-525. arXiv: 0911.1008 [hep-ex]. DOI: 10.1016/j.nima.2010.03.136.

[62] DIJCK E A. Neutral Particle Identification in the P̄ANDA Electromagnetic Calorimeter[D]. Rijksuniversiteit Groningen, 2009.

[63] BELYAEV A, CHRISTENSEN N D, PUKHOV A. CalcHEP 3.4 for collider physics within and beyond the Standard Model[J/OL]. Comput. Phys. Commun., 2013, 184: 1729-1769. arXiv: 1207.6082 [hep-ph]. DOI: 10.1016/j.cpc.2013.01.014.

[64] HEN O. High-Momentum Nucleons in Nuclei[D]. Tel-Aviv University, 2015.

[65] KRASS A S. Photoproduction of muon pairs[J/OL]. Phys. Rev., 1965, 138: B1268-B1273. DOI: 10.1103/PhysRev.138.B1268.

[66] TSAI Y S. Pair Production and Bremsstrahlung of Charged Leptons[J/OL]. Rev. Mod. Phys., 1974, 46: 815. DOI: 10.1103/RevModPhys.46.815.

[67] SAKAKI Y, NAMITO Y, SANAMI T, et al. Implementation of muon pair production in PHITS and verification[J/OL]. Nucl. Instrum. Meth. A, 2020, 977: 164323. arXiv: 2004.00212 [physics.acc-ph]. DOI: 10.1016/j.nima.2020.164323.

[68] ZHU X L, ALFANDA H M, LI S, et al. Prospective study of light dark matter search in positron beam mode with DarkSHINE experiment initiative[J/OL]. 2026. arXiv: 2609.05900 [hep-ex].

[69] CHEN J, et al. Prospective study of light dark matter search with a newly proposed DarkSHINE experiment[J/OL]. Sci. China Phys. Mech. Astron., 2023, 66(1): 211062. DOI: 10.1007/s11433-022-1983-8.

[70] ZHAO Z, et al. Design of a LYSO Crystal Electromagnetic Calorimeter for Dark SHINE Experiment[J/OL]. Nucl. Sci. Tech., 2025, 36: 41. arXiv: 2407.17800 [hep-ex]. DOI: 10.1007/s41365-024-01618-8.

[71] GUO Y, et al. Design of High-speed Readout Electronics for the DarkSHINE Electromagnetic Calorimeter[J/OL]. JINST, 2025, 20(01): P01016. arXiv: 2407.20723 [physics.ins-det]. DOI: 10.1088/1748-0221/20/01/P01016.

[72] ANDREEV Y M, et al. Probing light dark matter with positron beams at NA64[J/OL]. Phys. Rev. D, 2024, 109(3): L031103. arXiv: 2308.15612 [hep-ex]. DOI: 10.1103/PhysRevD.109.L031103.

[73] ZHANG Y, ZHANG W T, SONG M, et al. Probing invisible decay of dark photon at BESIII and future STCF via monophoton searches[J/OL]. Phys. Rev. D, 2019, 100(11): 115016. arXiv: 1907.07046 [hep-ph]. DOI: 10.1103/PhysRevD.100.115016.

[74] BATELL B, FENG J L, TROJANOWSKI S. Detecting Dark Matter with Far-Forward Emulsion and Liquid Argon Detectors at the LHC[J/OL]. Phys. Rev. D, 2021, 103: 075023. arXiv: 2101.10338 [hep-ph]. DOI: 10.1103/PhysRevD.103.075023.

[75] DENIVERVILLE P, POSPELOV M, RITZ A. Observing a light dark matter beam with neutrino experiments[J/OL]. Phys. Rev. D, 2011, 84: 075020. arXiv: 1107.4580 [hep-ph]. DOI: 10.1103/PhysRevD.84.075020.

[76] BATELL B, ESSIG R, SURUJON Z. Strong Constraints on Sub-GeV Dark Sectors from SLAC Beam Dump E137[J/OL]. Phys. Rev. Lett., 2014, 113(17): 171802. arXiv: 1406.2698 [hep-ph]. DOI: 10.1103/PhysRevLett.113.171802.

[77] AGUILAR-AREVALO A A, et al. Dark Matter Search in a Proton Beam Dump with MiniBooNE [J/OL]. Phys. Rev. Lett., 2017, 118(22): 221803. arXiv: 1702.02688 [hep-ex]. DOI: 10.1103/PhysRevLett.118.221803.

[78] AKIMOV D, et al. First Probe of Sub-GeV Dark Matter beyond the Cosmological Expectation with the COHERENT CsI Detector at the SNS[J/OL]. Phys. Rev. Lett., 2023, 130(5): 051803. arXiv: 2110.11453 [hep-ex]. DOI: 10.1103/PhysRevLett.130.051803.

[79] ESSIG R, MANALAYSAY A, MARDON J, et al. First Direct Detection Limits on sub-GeV Dark Matter from XENON10[J/OL]. Phys. Rev. Lett., 2012, 109: 021301. arXiv: 1206.2644 [astro-ph.CO]. DOI: 10.1103/PhysRevLett.109.021301.

[80] BANERJEE D, et al. Search for vector mediator of Dark Matter production in invisible decay mode [J/OL]. Phys. Rev. D, 2018, 97(7): 072002. arXiv: 1710.00971 [hep-ex]. DOI: 10.1103/PhysRevD.97.072002.

[81] WANG Z, ZHU X, KHODA E E, et al. Probing Electroweak Phase Transition at CEPC via Exotic Higgs Decays with 4b Final States[J/OL]. Letters in High Energy Physics, 2023, 2023: 436. DOI: 10.31526/lhep.2023.436.

[82] KOZACZUK J, RAMSEY-MUSOLF M J, SHELTON J. Exotic Higgs boson decays and the electroweak phase transition[J/OL]. Physical Review D, 2020, 101(11). http://dx.doi.org/10.1103/Phy

sRevD.101.115035. DOI: 10.1103/physrevd.101.115035.

[83] KE G, MENG Q, FINLEY T, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree[J]. Advances in Neural Information Processing Systems, 2017, 30: 3146-3154.

[84] COWAN G, CRANMER K, GROSS E, et al. Asymptotic formulae for likelihood-based tests of new physics[J/OL]. The European Physical Journal C, 2011, 71(2): 1554(2011-02-09). https://doi.org/10.1140/epjc/s10052-011-1554-0. DOI: 10.1140/epjc/s10052-011-1554-0.

[85] READ A L. Presentation of search results: the CLs technique[J/OL]. Journal of Physics G: Nuclear and Particle Physics, 2002, 28(10): 2693. https://dx.doi.org/10.1088/0954-3899/28/10/313. DOI: 10.1088/0954-3899/28/10/313.

Appendix A Exotic Higgs decays at the CEPC

The thermal evolution of the Higgs field connects collider observables with possible histories of the early Universe. For the measured Higgs mass, the Standard Model electroweak transition is a crossover. An additional scalar can change the thermal potential and allow a strong first-order transition. Such a transition is relevant to electroweak baryogenesis, which also requires the additional ingredients needed to generate the baryon asymmetry. The CEPC study asks whether Higgs decays can test light-singlet scenarios compatible with this thermal history.

During my doctoral studies, I participated in this CEPC analysis of light scalars produced through Higgs decays to a visible four-bottom-quark final state, in collaboration with the coauthors of Ref.[81]. The numerical projections and simulation figures below follow that study's CEPC exposure, detector response, and running assumptions.

A.1 Physics question and scalar benchmark

A.1.1 Vacuum conditions and physical masses

Let H be the Standard Model Higgs doublet and S a real gauge singlet. The renormalisable potential used as the starting point is

$$\begin{aligned} V(H,S) = & -\mu_H^2 H^\dagger H + \lambda (H^\dagger H)^2 + \frac{a_1}{2} H^\dagger H S + \frac{a_2}{2} H^\dagger H S^2 \\ & + b_1 S + \frac{b_2}{2} S^2 + \frac{b_3}{3} S^3 + \frac{b_4}{4} S^4. \end{aligned} \tag{A.1}$$

A shift of the singlet field allows its zero-temperature vacuum expectation value to be set to zero while retaining the linear term $b_1 S$. Writing $H = (0, (v + h_0)/\sqrt{2})^\mathsf{T}$ and $S = s_0$, with $v = 246$ GeV, the doublet and singlet fluctuations mix. We use the convention

$$\begin{pmatrix} h \\ s \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} h_0 \\ s_0 \end{pmatrix}, \qquad m_h = 125\ \text{GeV}, \quad m_s < m_h/2. \tag{A.2}$$

The observed Higgs h is mostly doublet, and the light scalar s is mostly singlet. The benchmark fixes $|\sin\alpha| = 0.01$. The original analysis denotes the same small doublet component of the light state by $|\cos\theta|$. This mixing permits prompt visible decays of s in the scenario studied.

The relation between the potential and the mass eigenstates follows by expanding around $(h_0, s_0) = (0, 0)$. At this point, stationarity gives

$$\mu_H^2 = \lambda v^2, \qquad b_1 = -\frac{a_1 v^2}{4}. \tag{A.3}$$

The linear coefficient therefore remains part of the potential even though the chosen vacuum has no singlet expectation value. Setting both b_1 and the singlet expectation value to zero would impose an additional restriction on the parameters.

The quadratic terms define the curvature matrix

$$\mathcal{M}^2 = \begin{pmatrix} 2\lambda v^2 & a_1 v/2 \\ a_1 v/2 & b_2 + a_2 v^2/2 \end{pmatrix}. \tag{A.4}$$

Its off-diagonal entry mixes the doublet and singlet fluctuations. Applying the rotation in Eq. (A.2), and writing $c_\alpha = \cos\alpha$ and $s_\alpha = \sin\alpha$, gives

$$\begin{aligned} 2\lambda v^2 &= c_\alpha^2 m_h^2 + s_\alpha^2 m_s^2, \\ \frac{a_1 v}{2} &= c_\alpha s_\alpha (m_h^2 - m_s^2), \\ b_2 + \frac{a_2 v^2}{2} &= s_\alpha^2 m_h^2 + c_\alpha^2 m_s^2. \end{aligned} \tag{A.5}$$

These identities explain the parameter counting. Specifying the two masses, the mixing, and v fixes λ, a_1, and the combination $b_2 + a_2 v^2/2$; stationarity then fixes μ_H^2 and b_1. The parameters a_2, b_3, and b_4 still control scalar interactions. They can vary while the masses and mixing remain fixed, with a compensating change in b_2 when a_2 changes.

The dimensions of these parameters also distinguish their roles. λ, a_2, and b_4 are dimensionless; a_1 and b_3 have mass dimension one; and b_2 has mass dimension two. The scan variable b_3/v puts the cubic singlet interaction on a dimensionless scale. All terms in Eq. (A.1) then have mass dimension four.

A.1.2 Exotic width and visible scalar decays

The coupling relevant to the collider signature can be defined directly through $V \supset g_{hss} h s^2/2$. With canonically normalised fields, its two-body partial width is

$$\Gamma(h \to ss) = \frac{|g_{hss}|^2}{32\pi m_h} \sqrt{1 - \frac{4m_s^2}{m_h^2}}. \tag{A.6}$$

The factor for identical final-state particles is included in this expression. The rate depends on the scalar interactions as well as on the mixing angle.

This independence can be seen in the cubic part of the potential before diagonalisation:

$$V_3 = \lambda v h_0^3 + \frac{a_1}{4} h_0^2 s_0 + \frac{a_2 v}{2} h_0 s_0^2 + \frac{b_3}{3} s_0^3. \tag{A.7}$$

Using the inverse rotation, $h_0 = c_\alpha h - s_\alpha s$ and $s_0 = s_\alpha h + c_\alpha s$, the coefficient of $hs^2/2$ becomes

$$\begin{aligned} g_{hss} = {} & 6\lambda v c_\alpha s_\alpha^2 + \frac{a_1}{2} s_\alpha (s_\alpha^2 - 2c_\alpha^2) \\ & + a_2 v c_\alpha (c_\alpha^2 - 2s_\alpha^2) + 2b_3 s_\alpha c_\alpha^2. \end{aligned} \tag{A.8}$$

In the zero-mixing limit this expression tends to $a_2 v$. The portal quartic interaction can thus produce an exotic Higgs width even when direct production and decay of an individual singlet through mixing are suppressed. The nonzero mixing used in the simulation supplies the visible decay of the light state.

To see how the width becomes a branching fraction, suppose the scalar extension introduces no additional Higgs decay modes beyond ss. The couplings of h to Standard Model particles carry the doublet factor c_α, so

$$\begin{aligned} \Gamma_h &= c_\alpha^2 \Gamma_{\mathrm{SM}}(m_h) + \Gamma(h \to ss), \\ \mathrm{Br}(h \to ss) &= \frac{\Gamma(h \to ss)}{c_\alpha^2 \Gamma_{\mathrm{SM}}(m_h) + \Gamma(h \to ss)}. \end{aligned} \tag{A.9}$$

For the light scalar, decay to a Standard Model final state X instead scales as

$$\begin{aligned} \Gamma(s \to X) &= s_\alpha^2 \Gamma_{\mathrm{SM}}(m_s; X), \\ \mathrm{Br}(s \to X) &= \frac{\Gamma_{\mathrm{SM}}(m_s; X)}{\sum_Y \Gamma_{\mathrm{SM}}(m_s; Y)}. \end{aligned} \tag{A.10}$$

The second expression assumes that these are the only open decay channels. The common mixing factor cancels from the branching fractions, but remains in the total width and hence in the lifetime. Reducing the mixing can therefore leave the relative decay fractions unchanged while moving the decay vertex away from the production point. The prompt-decay assumption is an independent requirement of the reconstruction used here.

Equation (A.6) also contains the threshold factor $\beta_s = \sqrt{1 - 4m_s^2/m_h^2}$. At a fixed cubic coupling, the exotic width decreases as m_s approaches $m_h/2$. This phase-space effect concerns the model's predicted decay rate. It is distinct from the mass dependence of the selection efficiency, which enters when that rate is converted into a selected event count.

A.1.3 Thermal interpretation of the collider target

The thermal interpretation follows the parameter scans of Ref.[82]. At fixed light-scalar mass and mixing, the original analysis considered a_2 and b_3/v between 10^{-4} and 1, and b_4 between 10^{-5} and 1. The theoretical scan covers light masses from 5 to 60 GeV; the detector-level analysis covers only 15 to 60 GeV. Conditions on the phase transition and successful tunnelling select a subset of that parameter space. The resulting relation between the scalar mass and exotic Higgs branching fraction motivates the experimental sensitivity target.

At finite temperature, the relevant quantity is the free-energy density as a function of the doublet and singlet backgrounds. Schematically, it contains the tree potential, zero-temperature quantum corrections, and thermal contributions. A first-order transition requires competing minima with a barrier between them. The critical temperature marks degeneracy of their free energies, and transition completion requires bubbles of the new phase to nucleate[82].

The collider measures a zero-temperature decay rate. Its connection to the thermal history passes through the parameters of the scalar potential: the same interactions that enter g_{hss} also influence the potential away from the present vacuum. A projected bound can test the selected thermal scenarios when their predicted rate exceeds that bound.

A.2 Collision sample and detector response

A.2.1 Production kinematics and the mass grid

The benchmark consists of unpolarised e^+e^- collisions at $\sqrt{s} = 240$ GeV with an integrated luminosity of 5 ab^{-1}. Higgsstrahlung supplies the production process,

$$e^+e^- \to Zh, \qquad Z \to \ell^+\ell^-, \qquad h \to ss \to (b\bar{b})(b\bar{b}), \quad \ell = e, \mu. \tag{A.11}$$

The charged-lepton pair tags the associated Z boson. The remaining particles carry the Higgs decay, with two dijet resonances expected at the same mass. Known initial-state four-momentum also permits the Higgs system to be selected through the recoil against the leptons.

At Born level, two-body Zh production fixes the Higgs energy and momentum:

$$E_h = \frac{s + m_h^2 - m_Z^2}{2\sqrt{s}}, \qquad |\boldsymbol{p}_h| = \frac{\sqrt{\lambda_K(s, m_h^2, m_Z^2)}}{2\sqrt{s}}, \tag{A.12}$$

where $\lambda_K(x,y,z) = x^2+y^2+z^2-2xy-2xz-2yz$ is the kinematic triangle function, unrelated to the quartic coupling λ. The Higgs therefore has a boost in the collision frame even before its decay is considered. Beam radiation and reconstruction alter the ideal two-body kinematics.

In the Higgs rest frame, the two scalars have energy $m_h/2$ and momentum $m_h\beta_s/2$. Neglecting the bottom-quark mass for this kinematic illustration, their daughter energies in the same frame are

$$E^{(h)}_{b,\bar{b}} = \frac{m_h}{4}(1 \pm \beta_s \cos\vartheta^*), \qquad E^{(h)}_{\mathrm{min,max}} = \frac{m_h}{4}(1 \mp \beta_s), \tag{A.13}$$

with ϑ^* measured relative to the scalar flight direction in the scalar rest frame. The allowed interval is much wider for a light scalar. A 15 GeV hypothesis can produce a very soft daughter even though each scalar carries 62.5 GeV in the Higgs frame.

Table A.1 Illustrative parton kinematics in the Higgs rest frame for $m_h = 125$ GeV, neglecting the bottom-quark mass. Values follow directly from Eqs. (A.13) and (A.14); they are not reconstructed efficiencies.

m_s [GeV]	β_s	$E^{(h)}_{\mathrm{min}}$ [GeV]	$E^{(h)}_{\mathrm{max}}$ [GeV]	$\psi^{(h)}_{\mathrm{min}}$ [degrees]
15	0.971	0.91	61.59	27.8
30	0.877	3.84	58.66	57.4
45	0.694	9.56	52.94	92.1
60	0.280	22.50	40.00	147.5

The opening angle also changes. For massless daughters with equal energies in the Higgs frame, $m_s^2 = 2E_bE_{\bar{b}}(1-\cos\psi)$ gives the minimum opening angle

$$\psi^{(h)}_{\mathrm{min}} = 2\arcsin\left(\frac{2m_s}{m_h}\right). \tag{A.14}$$

Unequal daughter energies correspond to larger opening angles at the same scalar mass. This angle is defined in the Higgs frame. The additional Higgs boost, showering, and hadronisation change the observed jet geometry. The parton-level relations nevertheless explain why one classifier and one set of jet-resolution assumptions need not perform equally well across the full mass grid.

A.2.2 Generation, transport, and reconstruction

Ten signal hypotheses are simulated, with $m_s = 15, 20, \ldots, 60$ GeV. FeynRules supplies the model implementation. MadGraph5_aMC@NLO generates signal events at leading order, and WHIZARD generates the Standard Model background samples. PYTHIA 8 handles

showering and hadronisation in the simulation chain. Mokka, based on Geant4, transports particles through the reference detector, followed by digitisation and Arbor particle-flow reconstruction[81].

Particle flow uses charged-particle momenta from tracking together with calorimeter information for neutral particles. Associating tracks with calorimeter deposits prevents counting the same charged particle twice. In a dense jet, incorrect associations can alter both its reconstructed energy and its constituents. This is one reason to retain the complete detector chain when assessing four-jet masses, flavour scores, and the particle-multiplicity inputs to the classifier.

The detector concept combines silicon tracking and a time projection chamber with finely segmented electromagnetic and hadronic calorimeters. Vertex information supports heavy-flavour identification, while particle flow combines tracking and calorimeter measurements to reconstruct the visible final state. The design expectation quoted in the study is a jet-energy resolution of approximately 3–5% for jet energies between 20 and 100 GeV.

For the analysis, backgrounds are grouped by their relation to the Higgs decay. The Zh, $h \to b\bar{b}$ sample is treated separately because ordinary two-bottom-quark decays can resemble the four-jet signal after clustering. Other Higgs decay modes form a second category, and processes without a Higgs boson form a third. This grouping also defines the background-normalisation variations used in the statistical analysis.

The grouping also separates two classification problems. Non-Higgs events can often be rejected through the lepton tag and recoil requirement. Ordinary Higgs decays share the production kinematics of the signal and can survive those requirements. Their rejection depends more directly on the hadronic system: whether it resembles two equal-mass scalar decays, whether all four jets carry bottom flavour, and whether the reconstructed constituents support four distinct jets.

A.3 From reconstructed particles to scalar candidates

A.3.1 Lepton tag and recoil selection

Electrons and muons are identified with the LICH algorithm, which uses the tracking and high-granularity calorimeter response. The reference performance quoted in the study is 99.7% efficiency for electrons and 99.9% for muons, with misidentification rates below

0.07%.

Identifying a lepton and isolating it address different ambiguities. Identification tests whether the detector response is consistent with an electron or muon. Isolation tests whether substantial nearby activity suggests that the candidate belongs to a jet. Leptons from heavy-hadron decays can satisfy identification while failing isolation. Requiring two isolated candidates therefore helps associate the selected pair with the Z decay. The same-flavour and opposite-charge conditions then restrict the allowed dilepton combinations before any mass requirement is applied.

Lepton isolation uses the additional energy E_{cone} around the candidate inside an angular cone satisfying $\cos\theta_{\mathrm{cone}} > 0.98$. Expressing both energies numerically in GeV, the requirement is

$$\left(\frac{E_{\mathrm{cone}}}{\mathrm{GeV}}\right)^2 < 4\left(\frac{E_\ell}{\mathrm{GeV}}\right) + 12.2. \tag{A.15}$$

An accepted event contains an isolated opposite-sign pair of the same flavour. The dilepton invariant mass selects a Z candidate, and the recoil mass selects an accompanying Higgs candidate. In the nominal centre-of-mass frame,

$$M_{\mathrm{rec}}^2 = (\sqrt{s} - E_{\ell\ell})^2 - |\boldsymbol{p}_{\ell\ell}|^2 = s + M_{\ell\ell}^2 - 2\sqrt{s}E_{\ell\ell}. \tag{A.16}$$

The second form makes explicit how the lepton measurements constrain the unobserved recoil system before its jets are assigned.

The recoil construction follows from four-momentum conservation. With $P_{\mathrm{ini}} = (\sqrt{s}, \mathbf{0})$ and $P_{\ell\ell} = (E_{\ell\ell}, \boldsymbol{p}_{\ell\ell})$, the recoil four-vector is $P_{\mathrm{ini}} - P_{\ell\ell}$. Its invariant mass depends only on the measured leptons and the assumed initial state. A Higgs candidate can thus be selected before deciding which hadronic particles form either scalar candidate. This avoids putting the dijet pairing into the Higgs tag.

For small measurement changes at fixed nominal s, differentiation gives

$$\delta M_{\mathrm{rec}} = \frac{M_{\ell\ell}}{M_{\mathrm{rec}}}\delta M_{\ell\ell} - \frac{\sqrt{s}}{M_{\mathrm{rec}}}\delta E_{\ell\ell}. \tag{A.17}$$

The two variations are correlated because they are formed from the same leptons. Their covariance must be included when propagating a resolution; treating the terms as independent generally loses information. Radiation from the beams also changes the hard collision four-momentum relative to the nominal value used in the reconstruction. The recoil window is therefore wider than an ideal two-body delta function.

Table A.2 records the numerical preselection as tabulated in the source manuscript.[1] The polar-angle observable is denoted by $\theta_{\ell\ell}$ as in that manuscript, and ϕ denotes the angle between the selected lepton tracks.

The angular acceptance, mass window, and recoil window act on different features of the event. Angular requirements restrict poorly accepted geometries and suppress backgrounds concentrated along the beams. The dilepton window selects a pair compatible with the associated Z. The recoil window then asks whether the rest of the event has a mass compatible with a Higgs.

Table A.2 Tabulated preselection for the CEPC study. The electron and muon channels use different angular acceptance values. Numerical inputs from Ref.[81].

Observable	Requirement
Each lepton's transverse momentum	$p_{T,\ell} > 20$ GeV
Electron-channel polar-angle variable	$\lvert\cos\theta_{ee}\rvert < 0.71$
Muon-channel polar-angle variable	$\lvert\cos\theta_{\mu\mu}\rvert < 0.81$
Lepton-track angle	$\cos\phi > -0.74$
Dilepton invariant mass	$77.5 < M_{\ell\ell} < 104.5$ GeV
Recoil mass	$124 < M_{\mathrm{rec}} < 140$ GeV

A.3.2 Four-jet assignment and heavy flavour

After removing the selected leptons, the remaining particle-flow objects are clustered into exactly four jets. LCFIPlus provides Durham clustering together with vertex finding and flavour-tagging information. Exclusive clustering fixes the number of reconstructed jets for both signal and background. An ordinary $h \to b\bar{b}$ decay can therefore yield four reconstructed jets, so the jet count alone cannot distinguish the two Higgs decay hypotheses.

The Durham procedure compares pairs using an energy-weighted angular separation. Its usual distance has the form

$$y_{ij}^{\mathrm{D}} = \frac{2\min(E_i^2, E_j^2)(1-\cos\theta_{ij})}{Q^2}, \tag{A.18}$$

where Q supplies an event-energy normalisation. Combining the pair with the smallest distance and repeating the procedure gives a clustering sequence. The exclusive four-jet prescription stops at the requested multiplicity. The distance favours merging soft or nearby

[1] The source table specifies a lepton transverse-momentum threshold and $\cos\phi > -0.74$. Its accompanying prose instead describes an energy threshold and $\cos\phi > -0.93$. The tabulated requirements are used here.

objects, which explains its dependence on both the energy hierarchy and the angular structure.

There are three ways to partition four jets into two pairs. For each partition we form the two dijet masses and select the assignment with the smallest absolute difference. If its masses are m_a and m_b, the reconstructed scalar-mass estimator and its consistency variable are

$$\overline{m}_{bb} = \frac{m_a + m_b}{2}, \qquad \Delta m_{bb} = |m_a - m_b|. \tag{A.19}$$

The average estimates the common scalar mass, while the difference tests the equal-mass hypothesis. Detector response and incorrect pairings broaden the distribution. The simulated mass shapes in Figure A.1 are retained as the detector-level reference for this reconstruction.

For jets numbered 1 through 4, the partitions are $(12)(34)$, $(13)(24)$, and $(14)(23)$. Each partition uses every jet exactly once. Minimising the mass difference introduces no input value for m_s, so the same assignment rule can be applied throughout the mass scan. It also permits an incorrect partition to win if detector response makes its two masses more similar.

The average and difference retain different information. A common shift of both reconstructed masses changes $\overline{m}_{bb}$ while leaving their difference unchanged. A distortion affecting only one candidate changes both quantities. Their correlation with jet energies and decay angles is useful to the classifier.

For each jet i, LCFIPlus returns a bottom-flavour score L_i between zero and one, based on quantities including track impact parameters and secondary vertices. The event discriminator combines the four scores as

$$P_b = \prod_{i=1}^{4} L_i, \qquad P_{\neg b} = \prod_{i=1}^{4} (1 - L_i), \qquad f_b = \frac{P_b}{P_b + P_{\neg b}}, \qquad B_{\text{ineff}} = \ln(1 - f_b). \tag{A.20}$$

Events in which all four jets have large bottom-flavour scores tend towards more negative B_{ineff}. This continuous observable retains information that would be lost by requiring a fixed number of jets to pass a binary tag.

The algebra of this combination makes its behaviour explicit. Define the sum of the four single-jet log odds by

$$A_b = \sum_{i=1}^{4} \ln \frac{L_i}{1 - L_i}. \tag{A.21}$$

Then $f_b = (1 + \exp[-A_b])^{-1}$ and $B_{\text{ineff}} = -\ln(1 + \exp[A_b])$. The transformation is monotonic: larger combined bottom-flavour evidence gives a more negative value of B_{ineff}. A jet

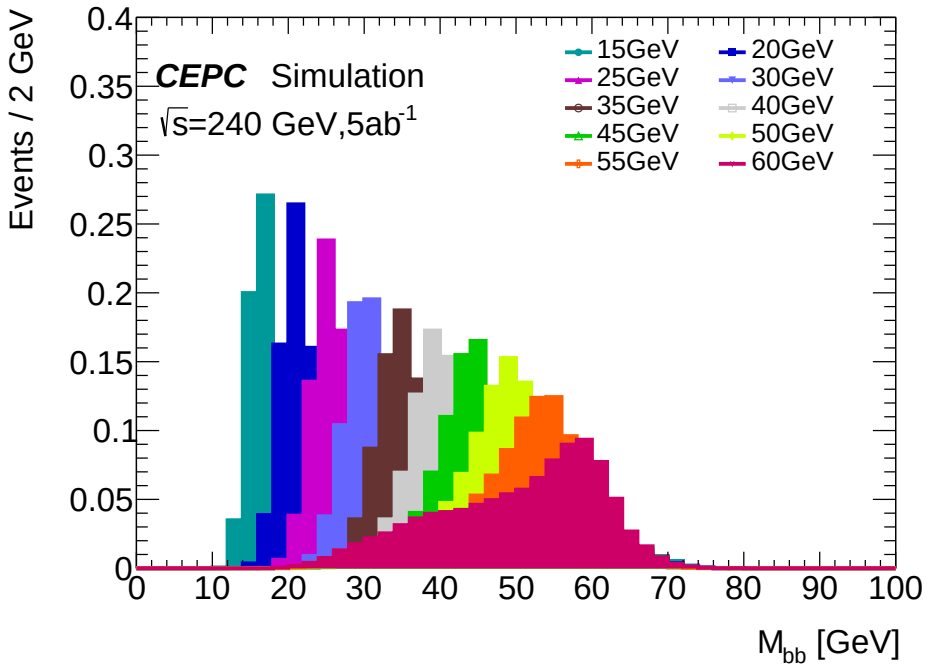


Figure A.1 Average dijet-mass estimator $\overline{m}_{bb}$ after pairing four jets into two scalar candidates in $e^+e^- \to Zh$, $Z \to e^+e^-, \mu^+\mu^-$, $h \to ss \to 4b$ simulation at 240 GeV. The ten histograms correspond to $m_s = 15, 20, \ldots, 60$ GeV and show the reconstructed mass spread. The source plot denotes this estimator by M_{bb}. Figure reproduced from Ref.[81].

with a small score can offset several jets with large scores through the log-odds sum. The event variable consequently tests the four-jet flavour pattern.

A.4 Classification across the scalar-mass grid

The preselection leaves Higgs and non-Higgs events whose kinematic distributions overlap the signal. A gradient-boosted decision tree combines flavour, mass, and angular information to distinguish them. The implementation uses LightGBM[83], with a separate classifier for each simulated scalar mass. This permits the classification boundary to follow the change in dijet kinematics across the grid.

A.4.1 Information carried by the inputs

The 24 inputs are grouped in Table A.3. Besides the reconstructed masses, they include the four individual jet energies and the recoil against the jet system. The constituent observables help identify events in which exclusive clustering has split a less complex hadronic state

into four jets. For these inputs, constituents must have energy above 0.4 GeV; selected distance measures use the seven constituents with the highest transverse momenta. The angular inputs describe the scalar candidates and their daughter jets. Only two jet-helicity quantities are needed because the second jet in each pair gives the same absolute helicity cosine.

The mass variables test several parts of the signal chain. The dilepton recoil mass tests associated Higgs production. The four-jet invariant mass tests whether the measured hadronic system is compatible with the Higgs decay, while the recoil against those jets tests consistency with the associated system. The average dijet mass is more directly sensitive to the light scalar. These observables share reconstructed particles and are therefore correlated. Energy lost from the hadronic system affects them differently from an error in the lepton measurement.

Constituent information complements those invariant masses. Forcing a two-jet event into four jets can produce the requested multiplicity without creating four independently populated hadronic systems. Particle counts and separation scales can expose that structure. The 0.4 GeV constituent threshold defines which particles contribute to these features, so their values depend on reconstruction as well as on the hard decay.

Decay angles supply further information about how momenta are distributed inside the two reconstructed candidates. A spin-zero particle has no preferred decay axis in its rest frame, but boosts, acceptance, and event selection distort the measured angular distributions. Background jets need not follow the same two-resonance geometry. Angular discrimination is assessed with the reconstructed distributions in Figure A.2.

Table A.3 The 24 classifier inputs, grouped by the reconstructed information they supply. Counts refer to individual scalar-valued inputs. Definitions follow Ref.[81].

Input group	Quantities	Count
Lepton tag	Transverse momenta of the two leptons; dilepton recoil mass	3
Jet system	Four jet energies; four-jet invariant mass; four-jet recoil mass	6
Scalar pairing	$\overline{m}_{bb}$ and Δm_{bb}	2
Constituents	Total particle multiplicity in the four jets; y_{12}, y_{23}, y_{34}, y_{45}, y_{56}, y_{67} distance measures	7
Decay geometry	Two boosted-angle cosines; opening-angle cosine of the scalar pair; two absolute jet-helicity cosines	5
Heavy flavour	B_{ineff}	1
Total		24

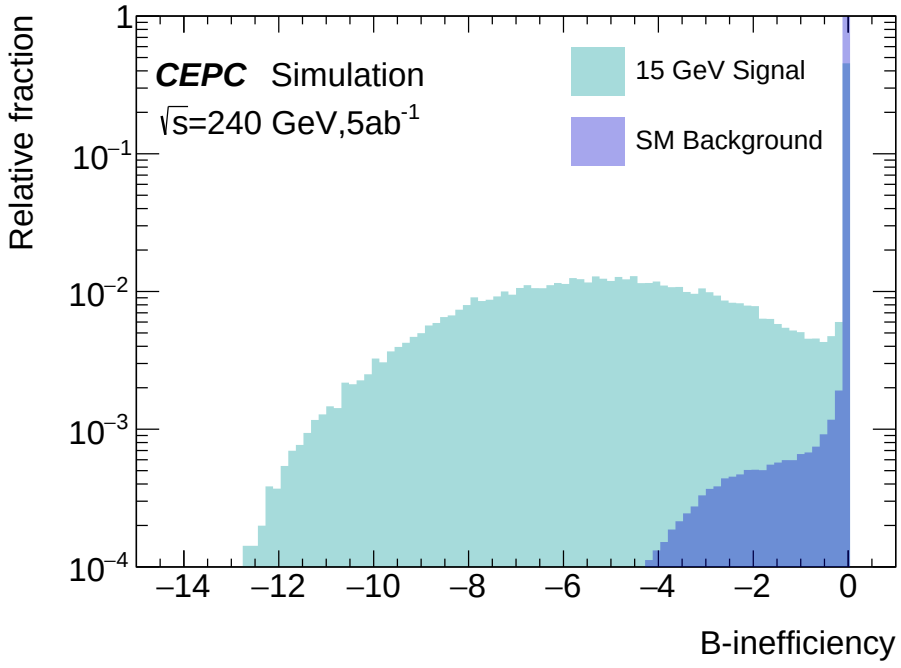


Figure A.2 Normalised classifier-input distributions for a 15 GeV scalar signal and Standard Model background at $\sqrt{s} = 240$ GeV. Figure reproduced from Ref.[81]. Combined bottom-flavour observable B_{ineff}, denoted b_{ineff} in the source plot.

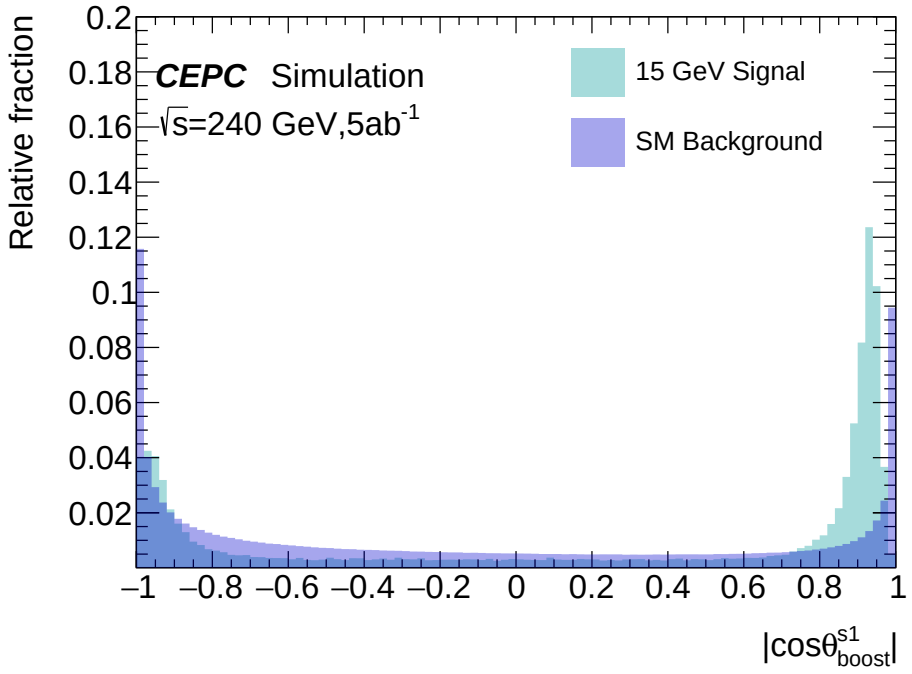


Figure A.2 (continued) Normalised classifier-input distributions for a 15 GeV scalar signal and Standard Model background at $\sqrt{s} = 240$ GeV. Figure reproduced from Ref.[81]. Boosted-angle cosine of one scalar candidate.

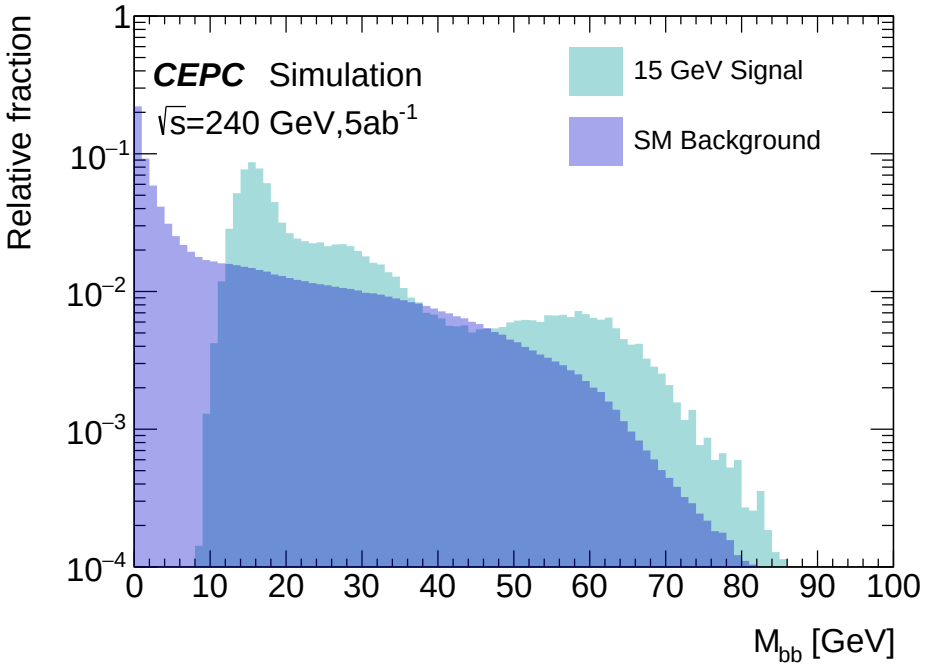


Figure A.2 (continued) Normalised classifier-input distributions for a 15 GeV scalar signal and Standard Model background at $\sqrt{s} = 240$ GeV. Figure reproduced from Ref.[81]. Average dijet mass $\overline{m}_{bb}$, denoted M_{bb} in the source plot.

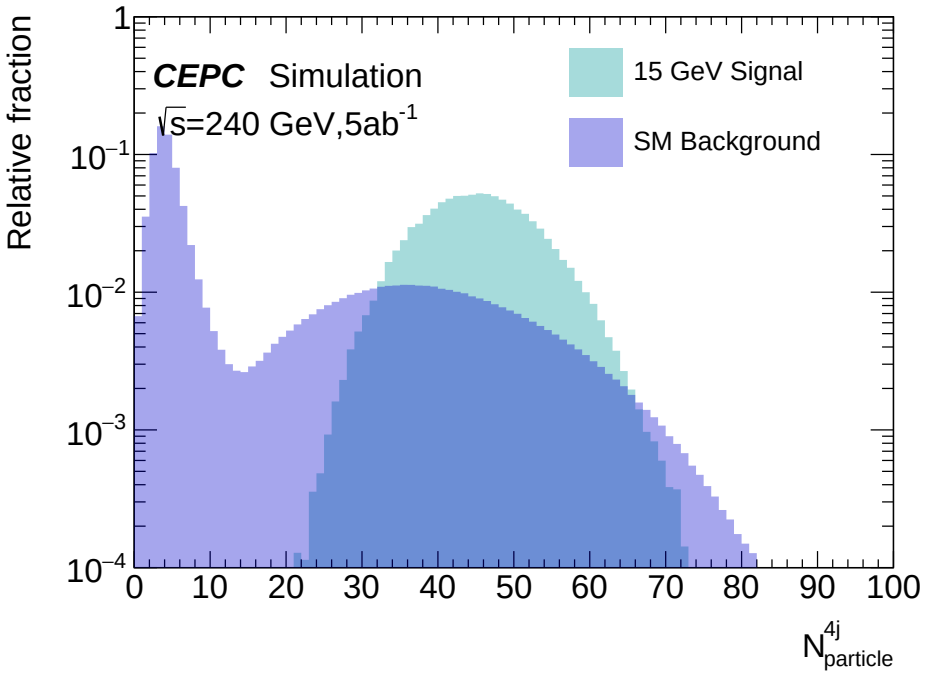


Figure A.2 (continued) Normalised classifier-input distributions for a 15 GeV scalar signal and Standard Model background at $\sqrt{s} = 240$ GeV. Figure reproduced from Ref.[81]. Particle multiplicity in the four jets.

Optuna tunes each model through 100 hyperparameter trials. The settings retained in the original analysis are listed in Table A.4. LightGBM grows trees by selecting leaves to split, so the allowed number of leaves and minimum leaf population control the complexity of the classifier. Data subsampling, feature subsampling, and regularisation are also varied. The reported area under the receiver-operating-characteristic curve is about 0.99 for each mass hypothesis. The limit calculation uses the score distribution with its event yields and uncertainties.

A.4.2 Boosting, model complexity, and the final score

A boosted tree classifier builds an additive score from successive trees. Writing the tree response as $T_t(\boldsymbol{x})$ for input vector $\boldsymbol{x}$, a general representation is

$$F(\boldsymbol{x}) = F_0 + \sum_{t=1}^{N_T} \nu_t T_t(\boldsymbol{x}). \tag{A.22}$$

Each tree partitions the input space through conditional decisions, so a mass requirement can depend on a flavour score or a jet energy. Such conditional boundaries explain how the classifier uses correlations without requiring a separate hand-designed cut for every combination of observables. The coefficients ν_t control the contribution of each tree. The number of trees is a separate training parameter from the leaf-count and regularisation settings.

The leaf count limits the number of terminal regions in a tree, while the minimum leaf population restricts how sparsely a region can be populated. Subsampling changes which events or features are available during training. Regularisation penalises overly complex responses. These controls are relevant because a leaf-wise growth rule can continue to refine a small region of input space if that region offers the largest loss reduction.

A.5 Uncertainties and statistical treatment

A.5.1 Rate variations and detector response

The dominant Zh, $h \to b\bar{b}$ background receives a 5% normalisation variation. The other background categories receive 100% variations. A flat 1% flavour-tagging uncertainty is included, motivated by the control-sample method discussed in the original study and enlarged to account for the four-jet final state. Luminosity and lepton-identification uncertainties are neglected in that projection.

Table A.4 Optimised classifier settings retained from Ref.[81]. N_L is the maximum leaf count, n_L the minimum leaf population, and f_D and f_F the data and feature fractions. The last two columns give the L1 and L2 regularisation strengths.

m_s [GeV]	N_L	n_L	f_D	f_F	λ_1	λ_2
15	11	52	0.462	0.732	2.10×10^{-2}	1.13×10^{-7}
20	250	23	0.568	0.621	1.02×10^{-1}	3.74×10^{-1}
25	13	59	0.706	0.695	9.24×10^{-6}	1.21×10^{-4}
30	3	11	0.521	0.533	2.54×10^{-8}	1.11
35	4	53	0.879	0.885	1.59×10^{-4}	2.00×10^{-5}
40	215	17	0.875	0.660	2.56×10^{-2}	3.02×10^{-4}
45	98	39	0.811	0.603	1.29×10^{-8}	5.41×10^{-5}
50	185	77	0.680	0.634	7.62×10^{-6}	6.78×10^{-5}
55	110	92	0.981	0.562	7.18×10^{-7}	4.65×10^{-1}
60	12	78	0.702	0.818	3.68×10^{-5}	5.19×10^{-4}

Soft-jet response is assessed using an e^+e^- k_t clustering implementation in FastJet. The comparison between reconstructed and truth jets measures one resolution contribution. Comparing FastJet with LCFIPlus supplies the second term, which accounts for differences between the clustering packages. Treating these two terms as independent, the study combines them according to

$$\delta^2_{\mathrm{JER}} = \sigma^2_{\mathrm{FastJet-truth}} + \sigma^2_{\mathrm{FastJet-LCFIPlus}}. \tag{A.23}$$

Figure A.3 shows the resulting energy dependence. The larger resolution at low energy is relevant to the reconstructed mass shapes and motivates using the detector simulation throughout the projection.

The raw classifier score is the final fit observable. Its distribution for the 15 GeV hypothesis is shown in Figure A.4. The background is concentrated at lower scores, while the signal extends into the bins with larger scores. A schematic form of the binned likelihood is

$$\mathcal{L}(\mu, \boldsymbol{\eta}) = \prod_{c,k} \mathrm{Pois}(n_{ck} \mid \mu s_{ck}(\boldsymbol{\eta}) + b_{ck}(\boldsymbol{\eta})) \prod_r C_r(\eta_r), \tag{A.24}$$

where c labels the electron or muon channel, k a score bin, and μ the signal strength relative to the reference signal template. The functions C_r constrain nuisance parameters for the systematic and finite-sample uncertainties. The source analysis uses Gaussian, log-normal, or Poisson constraints as appropriate. The product over score bins retains both rate and shape information.

The electron and muon information is combined in a profile-likelihood test, and expected 95% confidence-level upper bounds are obtained with the CL_s method[84-85]. After fitting,

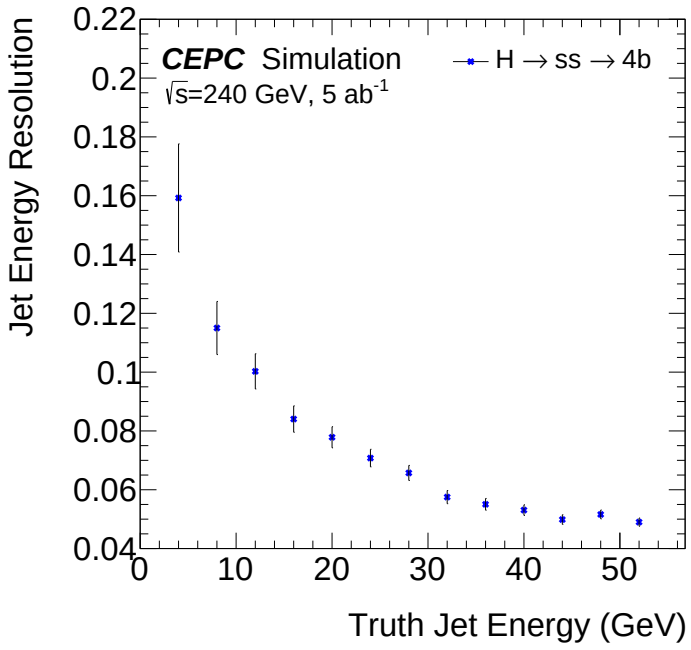


Figure A.3 Jet-energy-resolution estimate for $h \to ss \to 4b$ simulation at $\sqrt{s} = 240$ GeV. The horizontal coordinate is truth-jet energy; the resolution combines reconstructed-to-truth and FastJet-to-LCFIPlus comparison terms in quadrature. Bars reflect statistical uncertainty. The larger low-energy spread characterises the soft-jet response. Figure reproduced from Ref.[81].

the original analysis reports a small effect from the jet-resolution uncertainty and an effect near 1% from flavour tagging. The assumed background uncertainties have a larger effect, particularly at the lowest scalar mass.

A.5.2 Nuisance parameters and sensitivity

A rate uncertainty and a shape uncertainty enter a binned analysis in different ways. Multiplying a background template by one factor changes its total yield while preserving the relative bin populations. A change in jet response can instead move events between score bins, because jet energies enter both the explicit energy inputs and reconstructed masses. The total selected yield can change at the same time when events cross a preselection boundary.

For small nuisance variations around the nominal model, write the expected count in a bin schematically as

$$\nu_{ck}(\mu, \boldsymbol{\eta}) \simeq \mu s^0_{ck} + b^0_{ck} + \sum_r D_{ck,r}\eta_r. \tag{A.25}$$

The derivative $D_{ck,r}$ describes how that bin responds to a nuisance parameter near the chosen

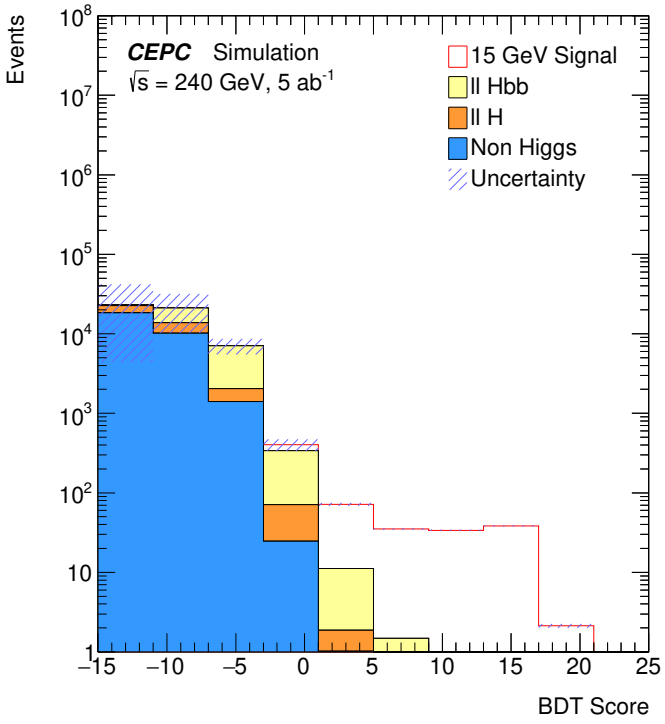


Figure A.4 Classifier-score distribution for the 15 GeV scalar hypothesis at $\sqrt{s} = 240$ GeV and 5 ab^{-1}, combining electron and muon channels before the fit. The signal template is compared with Zh, $h \to b\bar{b}$, other Higgs decays, and non-Higgs backgrounds; the band represents the background uncertainty. Larger scores favour the signal. Figure reproduced from Ref.[81].

expansion point. A common normalisation gives derivatives proportional to the template, while a shape variation can increase some bins and decrease others.

Profiling allows the nuisance parameters to move within their constraints as the signal strength is tested. Bins with little expected signal can help constrain a background component that also contributes to the signal-sensitive bins. If a nuisance variation resembles the signal shape, the fit has less power to distinguish the two. In a local quadratic approximation, this loss is expressed by

$$I_{\mu\mu}^{\text{prof}} = I_{\mu\mu} - I_{\mu\eta} I_{\eta\eta}^{-1} I_{\eta\mu}, \tag{A.26}$$

where I is the information matrix, including the nuisance constraints. The subtracted term reflects the freedom of nuisance parameters to absorb part of the signal-like change. This gives a statistical reason for keeping the full score distribution rather than only its highest-score bin.

A.5.3 Profile tests and expected upper bounds

For a specified scalar mass, the fit tests one nonnegative signal strength μ. Let $\widehat{\mu}$ and $\widehat{\boldsymbol{\eta}}$ maximise the likelihood globally, with $\widehat{\mu} \geq 0$, and let $\widehat{\boldsymbol{\eta}}_\mu$ maximise it at fixed μ. A one-sided test for an upper bound can be written as

$$q_\mu = \begin{cases} -2\ln \dfrac{\mathcal{L}(\mu, \widehat{\boldsymbol{\eta}}_\mu)}{\mathcal{L}(\widehat{\mu}, \widehat{\boldsymbol{\eta}})}, & 0 \leq \widehat{\mu} \leq \mu, \\ 0, & \widehat{\mu} > \mu. \end{cases} \tag{A.27}$$

Large values mean that the tested strength describes the data less well than the best fit. The second branch prevents a fluctuation towards more signal from counting as evidence against a smaller signal strength. The physical lower bound on $\widehat{\mu}$ matters when the unconstrained best fit would have a negative signal yield.

The distributions of this statistic under the tested signal-plus-background and background hypotheses define the confidence construction. In a continuous-statistic notation, let $p_{s+b} = P(q_\mu \geq q_{\mu,\mathrm{obs}} \mid \mu)$ and $p_b = P(q_\mu \geq q_{\mu,\mathrm{obs}} \mid 0)$. The modified frequentist ratio is

$$\mathrm{CL}_s = \frac{p_{s+b}}{p_b}. \tag{A.28}$$

An upper bound occurs where this ratio reaches 0.05. With these two probabilities defined using the same upper tail, the denominator is the background probability p_b. It reduces the tendency to exclude a signal for which the experiment has little separation from background. The source projection uses the CL_s construction with profile-likelihood methods[84-85].

An expected bound describes repeated hypothetical background-only experiments. The Asimov construction sets each bin count to its background expectation to estimate median sensitivity; fluctuations around background generate the expected bands. Asymptotic formulae approximate the required test-statistic distributions when their regularity and sample-size conditions are adequate[84]. Their application depends on the bin populations and nuisance constraints of the statistical model.

A.6 Projected reach and its interpretation

A.6.1 Yield normalisation and decay assumptions

The published normalisation for the result is

$$R_{ss} = \frac{\sigma(e^+e^- \to Zh)\ \mathrm{Br}(h \to ss)}{\sigma_{Zh}^{\mathrm{SM}}}. \tag{A.29}$$

For a Standard Model Higgsstrahlung rate, this ratio reduces to the exotic Higgs branching fraction, with the scalar decay prescription held fixed. Its connection to the selected four-bottom-quark yield can be written as

$$N_{\mathrm{sig}} = \mathcal{L}_{\mathrm{int}}\, \sigma_{Zh}^{\mathrm{SM}} R_{ss} [\mathrm{Br}(s \to b\bar{b})]^2 \sum_{\ell=e,\mu} \mathrm{Br}(Z \to \ell^+\ell^-)\epsilon_\ell(m_s). \tag{A.30}$$

Here ϵ_ℓ includes the selection and reconstruction response. The squared scalar branching fraction enters because both scalars must decay to bottom quarks. The bounds below retain the decay assumptions of the original signal model. A reinterpretation with a different scalar branching fraction or a displaced decay would require a corresponding change in the predicted rate or acceptance.

The fit strength and the reported ratio are related by the reference signal normalisation. If s_{ck} is constructed for a chosen value R_{ss}^{ref}, the limit on μ gives $R_{ss}^{95} = \mu_{95} R_{ss}^{\mathrm{ref}}$, with the reference yield set by Eq. (A.30).

In the minimal mixing model, the associated production cross section scales as c_α^2. Thus $R_{ss} = c_\alpha^2 \,\mathrm{Br}(h \to ss)$ at leading order under that coupling assumption. For the benchmark mixing, this production correction is of relative order 10^{-4}, so the ratio is numerically close to the exotic branching fraction.

For illustration, if only the scalar branching fraction changes from B_b^{ref} to B_b^{new} while masses, prompt lifetime, and conditional $b\bar{b}$ decay kinematics remain the same, the rate rescales as $(B_b^{\mathrm{new}}/B_b^{\mathrm{ref}})^2$. A bound on R_{ss} would then scale with the inverse of this factor. This relation does not apply unchanged if the new model alters the decay length, the final-state particles, or the distributions used by the classifier.

A.6.2 Mass dependence of the published result

Across the ten mass hypotheses, R_{ss}^{95} lies between 3.50×10^{-4} and 6.63×10^{-4}. The strongest point is at 20 GeV, while the weakest is at 50 GeV; the endpoint values alone would therefore understate the variation over the grid. Relative to the cut-and-count approach studied in the same work, the multivariate method improves the projected bounds by a mass-dependent factor of approximately two to four. Flavour information and correlations among the reconstructed observables contribute to this improvement.

The comparison in Figure A.5 shows why this channel was useful for testing light-singlet phase-transition scenarios. Under the specified model assumptions, the projected collider

Table A.5 Expected combined-channel upper bounds at 95% confidence level for 5 ab^{-1} at 240 GeV. The second column is 10^4 times the dimensionless ratio in Eq. (A.29). Values retained from Ref.[81].

m_s [GeV]	$10^4 R_{ss}^{95}$
15	4.07
20	3.50
25	5.45
30	4.49
35	5.34
40	5.50
45	6.32
50	6.63
55	6.09
60	4.58

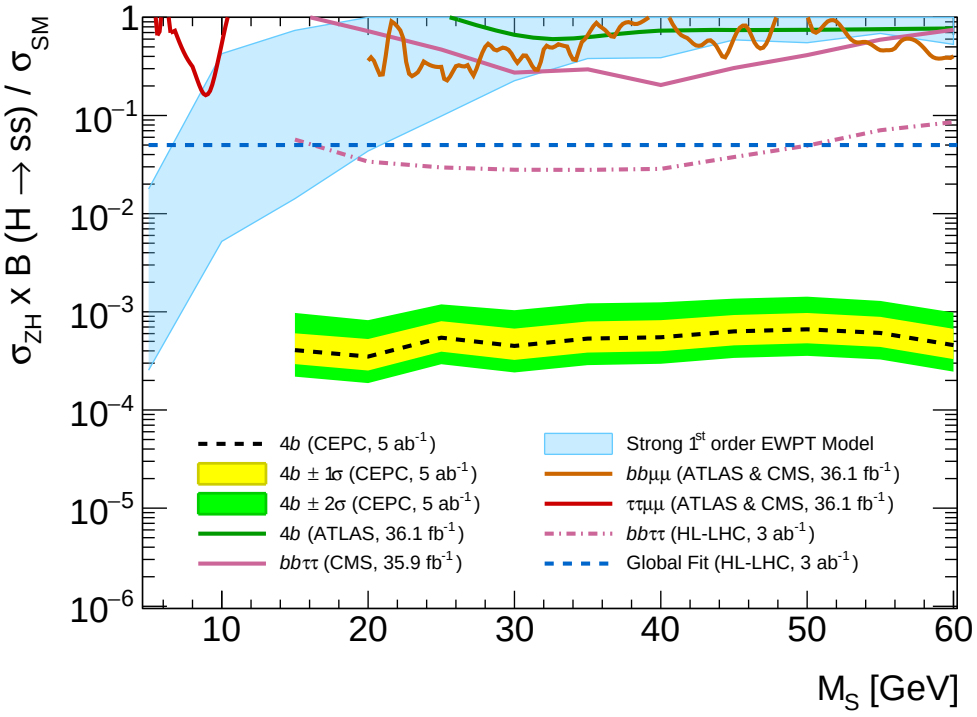


Figure A.5 Expected combined-channel 95% confidence-level upper bound on $R_{ss} = \sigma_{Zh}\,\mathrm{Br}(h \to ss)/\sigma_{Zh}^{\mathrm{SM}}$ versus scalar mass for 240 GeV CEPC collisions and 5 ab^{-1}. The dotted curve is the median bound; the bands show the one- and two-standard-deviation ranges. The shaded region is the strong first-order electroweak-transition scan of Ref.[82]. LHC results and HL-LHC projections are those used in the 2023 analysis. Figure reproduced from Ref.[81].

reach overlaps the parameter region associated with a strong first-order transition and successful tunnelling. The detector-level sensitivity is weaker than the idealised estimates discussed in Ref.[82]; reconstruction and experimental uncertainties affect how the theory target translates into an accessible decay rate.

A.6.3 Sensitivity scaling and limits on model parameters

A simple expansion illustrates why score information can improve a limit. If background expectations are known and the tested signal is small in each relevant bin, the background-only Fisher information is

$$I_{\mu\mu} = \sum_{c,k} \frac{s_{ck}^2}{b_{ck}}, \qquad \sigma_\mu \simeq I_{\mu\mu}^{-1/2}. \tag{A.31}$$

The formula follows from differentiating the Poisson log likelihood at $\mu = 0$. Bins contribute according to both their signal yield and their background contamination. In this limit, keeping bins with different signal-to-background ratios generally preserves more information than merging them. A cut-and-count analysis replaces the distribution by one accepted yield and cannot retain those differences. Nuisance parameters modify the estimate through profiling, as in Eq. (A.26).

When both s_{ck} and b_{ck} grow proportionally to exposure and the analysis remains dominated by background counting fluctuations, $I_{\mu\mu}$ grows with luminosity and the typical bound scales as $\mathcal{L}_{\rm int}^{-1/2}$. For an effectively background-free region, a fixed upper bound on the number of signal events instead leads to an inverse-luminosity rate bound. A background uncertainty that cannot be constrained by the data can weaken either improvement.

Finally, a branching-fraction constraint can be translated into the cubic coupling only after specifying the Higgs width and production model. With the decay assumptions of Eq. (A.9), let B_{ss} denote a bound on $\mathrm{Br}(h \to ss)$ at fixed mass. Combining the width and branching-fraction expressions gives

$$|g_{hss}|^2 \leq \frac{32\pi m_h}{\beta_s}\, c_\alpha^2 \Gamma_{\rm SM}(m_h) \frac{B_{ss}}{1 - B_{ss}}. \tag{A.32}$$

The measured quantity remains a rate; the right-hand side supplies the model-dependent conversion. Near threshold, β_s is small, so a given bound on the branching fraction permits a larger coupling. Additional Higgs decay modes would change the total-width relation, and

additional light-scalar decay modes would change the four-bottom-quark yield. Both effects must be included before comparing a different scalar theory with the tabulated limits.

List of Research Achievements

学术论文

[1] **Xuliang Zhu** et al. Prospective study of light dark matter search in positron beam mode with DarkSHINE experiment initiative. arXiv:2609.05900, 2026 (accepted by Frontiers of Physics for publication).

[2] **Xuliang Zhu**. DarkSHINE Simulation software framework. Nucl. Part. Phys. Proc. 346 (2024) 57. DOI: 10.1016/j.nuclphysbps.2024.07.003.

[3] Zhen Wang, **Xuliang Zhu**, et al. Probing electroweak phase transition at CEPC via exotic Higgs decays with $4b$ final states. Letters in High Energy Physics, 2023, 2023: 436. DOI: 10.31526/lhep.2023.436.

[4] Jing Chen, …, **Xu-Liang Zhu**, …, et al. Prospective study of light dark matter search with a newly proposed Dark SHINE experiment. Sci. China-Phys. Mech. Astron. 66(1): 211062 (2023). DOI: 10.1007/s11433-022-1983-8.

[5] Zhi-Yu Zhao, …, **Xu-Liang Zhu**, …, et al. Design of a LYSO crystal electromagnetic calorimeter for dark photon detection in the DarkSHINE experiment. Nucl. Sci. Tech. 36 (2025) 41. DOI: 10.1007/s41365-024-01618-8.

[6] Zhen Wang, …, **Xu-Liang Zhu**, …, et al. Design of hadronic calorimeter for DarkSHINE experiment. Nucl. Sci. Tech. 35 (2024) 148. DOI: 10.1007/s41365-024-01502-5.

[7] Georges Aad et al. (ATLAS Collaboration, including **Xuliang Zhu**). Search for Dark Matter Produced in Association with a Dark Higgs Boson in the $b\bar{b}$ Final State Using pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector. Phys. Rev. Lett. 134(12) (2025) 121801. DOI: 10.1103/PhysRevLett.134.121801.

[8] Georges Aad et al. (ATLAS Collaboration, including **Xuliang Zhu**). Search for triple Higgs boson production in the $6b$ final state using pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector. Phys. Rev. D 111(3) (2025) 032006. DOI: 10.1103/PhysRevD.111.032006.

[9] Xiang Chen, Xue-Chun Chen, Le He, et al. (including **Xu-Liang Zhu**). Simulation and reconstruction software for the DarkSHINE dark photon search experiment. Manuscript in preparation.

研究报告

[1] DarkSHINE Collaboration (including **Xu-Liang Zhu**). DarkSHINE Baseline Design Report: Physics Prospects and Detector Technologies, Version 2 (preliminary).