Light Dark Matter eXperiment (LDMX)

Torsten Åkesson,[1] Owen Colegrove,[2] Giulia Collura,[2] Valentina Dutta,[2]
Bertrand Echenard,[3] Joshua Hiltbrand,[4] David G. Hitlin,[3] Joseph Incandela,[2]
John Jaros,[5] Robert Johnson,[6] Gordan Krnjaic,[7] Jeremiah Mans,[4] Takashi Maruyama,[5]
Jeremy McCormick,[5] Omar Moreno,[5] Timothy Nelson,[5] Gavin Niendorf,[2] Reese Petersen,[4]
Ruth Pöttgen,[1] Philip Schuster,[5] Natalia Toro,[5] Nhan Tran,[7] and Andrew Whitbeck[7]

[1]*Lund University, Department of Physics, Box 118, 221 00 Lund, Sweden*
[2]*University of California at Santa Barbara, Santa Barbara, CA 93106, USA*
[3]*California Institute of Technology, Pasadena, CA 91125, USA*
[4]*University of Minnesota, Minneapolis, MN 55455, USA*
[5]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[6]*Santa Cruz Institute for Particle Physics,*
University of California at Santa Cruz, Santa Cruz, CA 95064, USA
[7]*Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*
(Dated: June 15, 2018 : To be posted on LDMX Confluence Page)

We present an initial design study for a small-scale accelerator experiment to discover sub-GeV dark matter, probing dark matter couplings to electrons well beyond that motivated by thermal freezeout, and over a large mass range. The experiment, referred to as the Light Dark Matter Experiment (LDMX), would also be sensitive to a range of visibly and invisibly decaying dark sector particles, thereby addressing many of the science drivers highlighted in the 2017 US Cosmic Visions New Ideas in Dark Matter Community Report. The experiment will measure missing momentum in low-current, high repetition rate electron-nuclear fixed-target collisions. The LDMX design makes use of a low-mass tracking system and spectrometer magnet to measure the momentum of incoming electrons with high purity, and to cleanly reconstruct electron recoils, thereby providing a measure of missing momentum. A high-speed, high-granularity, Si-W calorimeter with MIP sensitivity is used to reject potential high rate bremsstrahlung background at trigger level, and to work in tandem with a scintillator-based hadron calorimeter to veto rare photo-nuclear reactions. The design leverages new and existing calorimeter technology under development for the HL-LHC, as well as existing tracking technology and experience from the HPS experiment. In this paper, we present our initial design concept, detailed GEANT-based studies of detector performance, signal and background processes, and a preliminary analysis approach. Using these results, we show how a first phase of LDMX with 4×10^{14} electrons on target could test much of the parameter space for thermal scalar and Majorana fermion dark matter beyond the reach of existing and planned experiments.

Contents

I. INTRODUCTION

Discovering the particle nature of dark matter (DM) is perhaps the most pressing challenge facing elementary particle physics today. Among the simplest possibilities is one in which dark matter arose as a thermal relic from the hot early Universe, which only requires small non-gravitational interactions between dark and familiar matter, and is robustly viable over the MeV to TeV mass range. Testing the hypothesis that the dark matter abundance arises from weak W- and Z- boson mediated interactions has been the primary focus of direct and indirect detection experiments to date, which are most sensitive to dark matter particles with masses ranging from a few GeV to a TeV. However, the lower mass range of MeV to GeV, where the most stable forms of ordinary matter are found, has remained stubbornly difficult to explore with existing experiments.

In the last several years, powerful ideas to probe "light dark matter" (LDM) in the sub-GeV mass range have emerged from efforts to test the intriguing possibility that dark matter is part of a dark sector that is neutral under all Standard Model (SM) forces (see [1, 2] for recent reviews). As with WIMPs, an attractive sensitivity milestone is motivated by the requirement that thermal freeze-out reactions give rise to an appropriate abundance of dark matter. This casts a spotlight on dark matter interactions with electrons that is only a few orders of magnitude beyond existing accelerator based sensitivity. To reach this exciting goal, our aim is to use an electron beam to produce dark matter in fixed-target collisions, making use of missing energy and momentum to identify and measure dark matter reactions [3, 4]. Already, NA64 at CERN has carried out a first physics run for a fixed-target electron beam experiment using missing energy as the identifying signature [5, 6]. That experiment promises to reach a sub-GeV dark matter sensitivity surpassing all existing constraints by 2020 [1, 7], but will nonetheless fall short of the required sensitivity primarily due to luminosity limitations. At masses above O(500) MeV, Belle II's future missing mass measurements are expected to provide the required sensitivity, but will fall short at lower masses due to luminosity and background limitations.

The "Light Dark Matter eXperiment" (LDMX) described in this note is designed to meet the following science goals:

- Explore dark matter interactions with sensitivity that extends three orders of magnitude beyond NA64's 2020 reach. Combined with Belle II's future missing mass measurements, this would provide the sensitivity needed to test all simple scenarios of dark matter freeze-out into Standard Model final states below a few GeV. This goal was highlighted in [2], will provide LDMX with excellent discovery potential, and is the primary science driver for the experiment.

- Using missing momentum measurements, explore broad and important new territory for secluded dark matter models, milli-charge particles, invisibly decaying dark photons, axions, and dark higgs particles. By extension, explore significaint new territory for SIMP, ELDER, asymmetric, and freeze-in dark matter scenarios.

- Using LDMX as a short baseline beam dump (parasitically), provide sensitivity to displaced visibly decaying dark photons, axions, inelastic dark matter, dark higgs, and other long-lived dark sector particles.

LDMX would cover a uniquely broad range of the dark sector science highlighted in the US Cosmic Vision New Ideas in Dark Matter Community Report [2], with special emphasis on the simplest thermal sub-GeV dark matter scenarios. We believe that LDMX, along with an appropriate set of complementary experiments, would therefore provide the foundation for a successful light dark matter program in the US or abroad.

The technical design considerations for LDMX are as follows. An electron beam incident on a thin target can produce dark matter particles through a "dark bremsstrahlung" process, in which most of the incident electron's energy is typically carried away by the invisible dark matter. To search for this process, LDMX reconstructs the kinematics of each beam electron both up- and down-stream of the target using low-mass tracking detectors. The up-stream tracker tags the incoming beam electrons while the down-stream tracker selects the low-energy, moderate transverse-momentum recoils of the beam electrons. Calorimetry is then used to veto events with an energetic forward photon or any additional forward-recoiling charged particles or neutral hadrons. Because each electron passes through the detector, the experiment must contend with high event rates in the tracker and electromagnetic calorimeter. Therefore, LDMX requires low-mass tracking that provides high-purity tagging for incoming electrons and clean, efficient reconstruction of recoils in a high-rate environment. The calorimetry for LDMX must simultaneously be fast enough to support this high rate of background events, most of which are "easy" to reject based on their high electromagnetic energy deposition, and sensitive enough to reject rare but subtle processes where a hard bremsstrahlung photon undergoes a photo-nuclear reaction in the target or in the calorimeter itself. These simultaneous requirements call for a high-speed, high-granularity calorimeter with MIP sensitivity to identify photo-nuclear products, used in conjunction with a hadron calorimeter that experiences much lower event rates. As described in this paper, LDMX plans to meet these technical challenges by leveraging technology under development for the HL-LHC and experience from the Heavy Photon Search (HPS) experiment.

To achieve sufficient statistics, LDMX proposes to use a low-current ($\sim$pA) but high bunch-repetition (~ 40 MHz) electron beam with multi-GeV energy. A beam with 10^8 electrons/second on target and energy in the 4 to 16 GeV range can explore most of the sub-GeV dark matter parameter space, while remaining below threshold for production of neutrinos, which are an irreducible background. Three options for such a beam are currently under consideration – the proposed 4-8 GeV DASEL beamline at SLAC [8], the 11 GeV CEBAF beam at Jefferson Laboratory [9], or a proposed 3.5-16 GeV beam derived from slow SPS extraction at CERN [10]. For historical and practical reasons, our design study has focused on the use of DASEL, and is therefore based on 4 GeV incoming electrons. This is conservative, as beams with higher energy, in the 8-16 GeV range, alleviate most of the potential backgrounds that LDMX will need to contend with.

Our general strategy is to run the experiment in two phases. In Phase I, we target a modest total required luminosity of 0.8 pb^{-1} corresponding to 4×10^{14} tagged electrons on target. During this phase, we would produce strong physics results with only minor deviations from established detector technologies. The studies and results in this paper primarily focus on Phase I performance at a beam energy of 4 GeV, the beam energy of a commissioning run at the proposed SLAC DASEL facility. However, we will also discuss requirements and implications for Phase II of the experiment which aims for $\sim 10^{16}$ tagged electrons on target at a beam energy of 8 GeVor higher.

This paper is largely self-contained, and includes a discussion of the dark matter science goals for LDMX (Section II); a summary of the signal and background processes (Section III); a description of the detector design concept and ongoing developments (Section IV); a detailing of the physics simulation and basic detector performance (Section V); studies of the background rejection strategies (Section VI) and the corresponding signal efficiency (Section VII); the physics reach of LDMX for several light dark matter scenarios (Section VIII); a discussion of the Phase II requirements (Section IX); and conclusions (Section X).

While this paper will focus on missing momentum and energy signatures exclusively, a companion phenomenology paper [11] will explore applications of LDMX to a broader variety of dark matter and other dark sector scenarios, including displaced decay signatures. Work on the detailed

experimental aspects of displaced decay signagtures, among others, is left to future work.

II.   SCIENCE GOALS

The primary science driver for LDMX is to search for dark matter particles in the sub-GeV mass range — a region that is simultaneously well motivated by the thermal freeze-out hypothesis for the origin of dark matter and experimentally open and accessible territory. In simple models, a thermal origin for dark matter implies a *minimum* interaction strength between dark and ordinary matter (as does any thermal contact between dark and ordinary matter in the early Universe). For DM masses below 200 MeV, LDMX aims to *fully* explore the scalar thermal DM parameter space in its first phase (and a signficant part of the fermion DM parameter space), and then fully explore the fermion thermal DM parameter space in its second phase. Exploring light dark matter to this level of sensitivity is a science priority highlighted in recent community reports [1, 2], and resonates with the dark matter priorities emphasized in the P5 report [12]. This section provides a self-contained discussion of thermal dark matter, realizations of sub-GeV thermal dark matter in particular, and the precise sensitivity targets that this general hypothesis motivates. A more comprehensive discussion of light dark matter may be found in [1, 4].

Beyond thermal dark matter, the use of missing momentum measurements also has powerful applications to many other models, including secluded, SIMP, ELDER, asymmetric, and freeze-in dark matter scenarios, as well as milli-charge particles, invisibly decaying dark photons, axions, and dark higgs particles. These applications, as well as the use of displaced decay signatures in LDMX is described in [11].

A.   Thermal Dark Matter: A Target of Opportunity

Perhaps the simplest possibility for the origin of dark matter is that its abundance arises from non-gravitational interactions with Standard Model particles in the hot early Universe. This compelling scenario is largely model independent and only requires the DM-SM interaction rate exceed the Hubble expansion rate at some point in the early Universe. If such a condition is satisfied, dark matter automatically reaches thermal equilibrium with visible matter and a residual DM abundance is guaranteed. This mechanism for generating the DM abundance is attractive and important for many reasons.

First, this mechanism is simple and generically realized because equilibrium is hard to avoid even for tiny couplings between dark and visible matter. Therefore, most *discoverable* models of DM with non-gravitational interactions fall into this category. Second, this mechanism implies a minimum annihilation rate $\sim 10^{-26}$ cm$^3 s^{-1}$ in order to avoid producing an overabundance of dark matter at "freeze-out". This minimum annihilation rate defines a minimum cross section which must be experimentally probed to rule out dark matter of thermal origins. Finally, the allowed thermal DM mass range is well defined – see Fig. 1. The most natural mass range is comparable to that of Standard Model matter, so that DM annihilation into SM final states is generically possible and well-motivated.

B.   Light Thermal Dark Matter (LDM)

If DM is realized in the upper half of the thermal mass window $\sim$ GeV - 10 TeV, it can be a Weakly Interacting Massive Particle (WIMP) charged under the electroweak force. This has been the traditional focus of dark matter direct and indirect detection experiments, driven in part by the well known connection between WIMPs and supersymmetry (SUSY), whose DM candidates

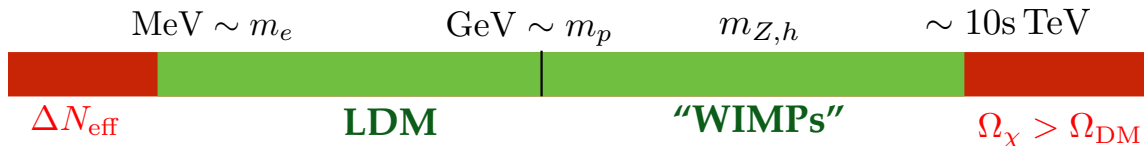| MeV $\sim m_e$	GeV $\sim m_p$	$m_{Z,h}$	$\sim$ 10s TeV
$\Delta N_{\rm eff}$	LDM	"WIMPs"	$\Omega_\chi > \Omega_{\rm DM}$

FIG. 1: The allowed mass range over which DM can thermalize with the SM in the early universe and yield the observed relic abundance via annihilation. For masses below $\lesssim$ MeV, thermal DM freezes out during BBN after neutrinos decouple from the SM, thereby transferring large amounts of entropy to SM particles, which modifies the effective number of relativistic degrees of freedom, $\Delta N_{\rm eff}$. Independently of this issue, for masses below $\sim$ 10 keV, DM is too hot to accommodate the observed matter power spectrum [13] and for masses above $\gtrsim$ 10 TeV, a perturbative annihilation rate cannot achieve the correct relic abundance in simple models [14].

realize this paradigm. However, powerful null results from direct and indirect detection experiments have largely ruled out the simplest WIMP scenarios by several orders of magnitude, and the remaining parameter space is largely cornered by upcoming experiments.

The lower half of the thermal mass window, $\sim$ MeV - GeV, has remained stubbornly difficult to test with traditional experiments designed to probe WIMPs and is not well explored. This is unfortunate because the sub-GeV mass range is well-motivated by "hidden sector" (or "dark sector") scenarios in which dark matter is simply a particle with its own forces and interactions, neutral under the Standard Model, but with sufficient coupling to visible matter that thermal equilibrium is achieved in the early Universe. This mass range is also independently motivated by asymmetric dark matter scenarios, in which dark matter carries a net particle number in analogy with the baryon asymmetry observed in visible matter. The particle physics community has highlighted these scenarios as among the most important to test, in the P5 report [12], the recent Dark Sectors 2016 community report [1], and the US Cosmic Visions New Ideas in Dark Matter community report [2]. For these reasons, and to help focus our design efforts, the primary science driver for LDMX is the exploration of dark matter interactions with electrons to a level of sensitivity needed to decisively test most predictive thermal dark matter scenarios over nearly the entire sub-GeV mass range. At the same time, many other dark matter and dark sector scenarios will also be explored, as discussed in [11].

C. A Benchmark Scenario for LDM

In the MeV-GeV mass range, viable models of LDM have the following properties:

- **Light Forces:** There have to be comparably light force carriers to mediate an efficient annihilation rate for thermal freeze-out (this follows from a simple generalization of the Lee-Weinberg bound [15]).

- **Neutrality:** Both the DM and the mediator must be singlets under the full SM gauge group; otherwise they would have been produced and detected at LEP or at hadron colliders.

These properties single out the hidden sector scenario highlighted in [1, 2, 12], which is the focus of considerable experimental activity. Thus, for the remainder of this note, we will use one of the simplest and most representative hidden sector models in the literature – a dark matter particle charged under a $U(1)$ gauge field (i.e. "dark QED"). Sensitivity to a variety of other new physics models, mediator particles, and dark matter will be described in a companion paper [11].

$$\sigma v \propto \alpha_D^2 \qquad\qquad \sigma v \propto \epsilon^2 \alpha_D$$
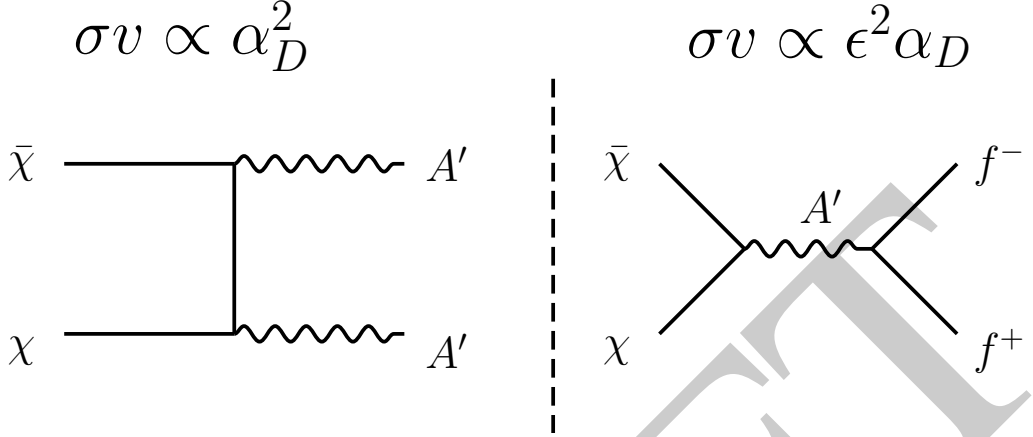

FIG. 2: Feynman diagrams for LDM with secluded annihilation (left) with $m_\chi > m_{A'}$ and direct annihilation (right) with $m_\chi < m_{A'}$. In the secluded regime, the dark photon decays visibly to kinematically accessible SM final states and motivates experimental searches for hidden forces (see [1]), but the DM annihilation cross section is independent of the A' coupling to visible matter. In the direct annihilation regime, the cross section for achieving the correct relic density depends on the parameter ϵ which couples the A' to charged SM particles, so there is a minimum value of this coupling for each choice of χ mass that realizes a thermal history in the early universe. These minimum values define predictive experimental targets for discovery or falsification (see Fig. 4).

We define the LDM particle to be χ and the $U(1)$ gauge boson (popularly called a "dark photon" mediator) A'. The general Lagrangian for this family of models contains

$$\mathscr{L} \supset -\frac{1}{4} F'^{\mu\nu} F'_{\mu\nu} + \frac{m_{A'}^2}{2} A'_\mu A'^\mu - A'_\mu (\epsilon e J_{\text{EM}}^\mu + g_D J_D^\mu), \tag{1}$$

where ϵ is the kinetic mixing parameter, $m_{A'}$ is the dark photon mass, and $J_{\text{EM}}^\mu \equiv \sum_f Q_f \bar{f} \gamma^\mu f$ is the SM electromagnetic current where f is an SM fermion with charge Q_f, $g_D \equiv \sqrt{4\pi\alpha_D}$ is the $U(1)_D$ coupling constant, and J_D is the dark matter current. Although each possible choice for χ has a different form for J_D, the relic density has the same dependence on the four model parameters $\{\epsilon, g_D, m_\chi, m_{A'}\}$ and can be captured in full generality with this setup.

This framework permits two qualitatively distinct annihilation scenarios depending on the A' and χ masses.

- **Secluded Annihilation:** For $m_{A'} < m_\chi$, DM annihilates predominantly into A' pairs as depicted on the left panel of Fig. 2. This annihilation rate is independent of the SM-A' coupling ϵ. The simplest version of this scenario is robustly constrained by CMB data [16], which rules out DM masses below $O(10)$ GeV for simple secluded annihilation models. More complex secluded models remain viable for low DM masses; these are potentially discoverable by LDMX, and are discussed in [11], but they do not provide a sharp parameter space target.

- **Direct Annihilation:** For $m_{A'} > m_\chi$, annihilation proceeds via $\chi\chi \to A'^* \to f f$ to SM fermions f through a virtual mediator. This scenario is quite predictive, because the SM-A' coupling ϵ must be large enough, and the A' mass small enough, in order to achieve the thermal relic cross-section. No robust constraint on this case can be extracted from CMB data.
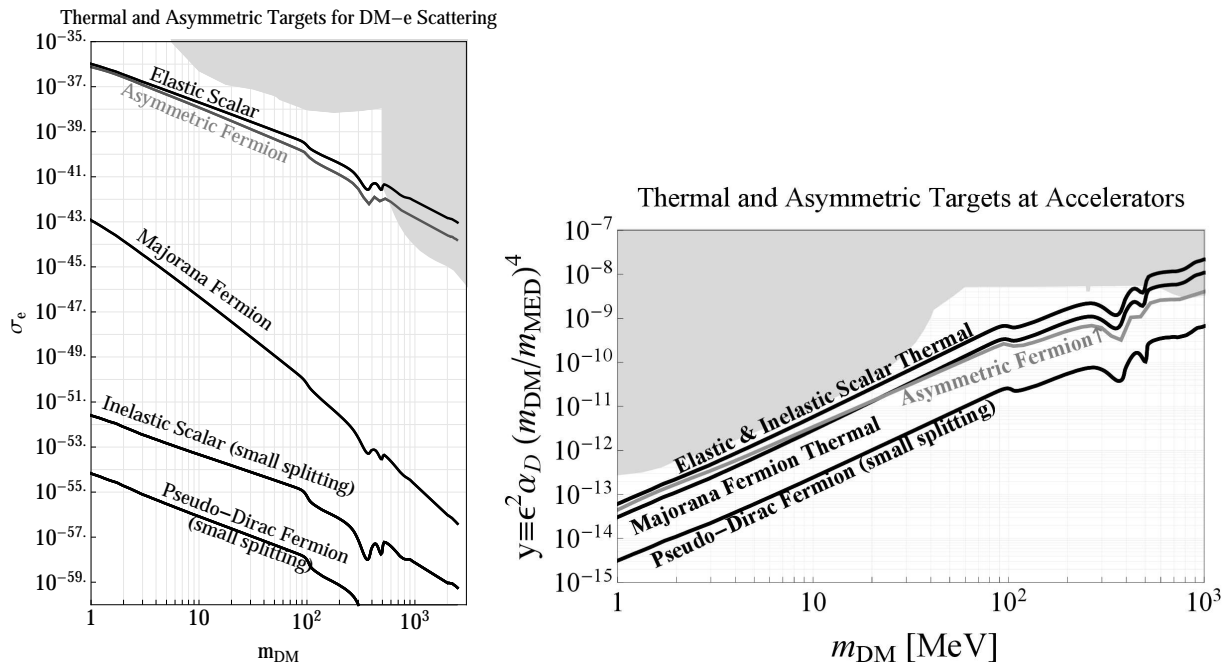
Thermal and Asymmetric Targets for DM−e Scattering

Thermal and Asymmetric Targets at Accelerators

FIG. 3: **Left:** Thermal targets for representative DM candidates plotted in terms of the σ_e the electron-recoil direct detection cross section. The target for each scenario differs considerably because for many scenarios σ_e contains spin or DM velocity suppression in the non-relativistic limit. **Right:** By contrast, the thermal targets presented in the y vs. m_χ plane reveals their intrinsic similarity. These targets are invariant for accelerator and collider searches, which involve DM production and/or detection in a relativistic setting for which the detection rates of each DM candidate are parametrically similar. Both plots are taken from [2] and also show a target for asymmetric fermion dark matter, a commonly discussed variation on the thermal-origin framework. This scenario and the resulting target are discussed further in [2].

Therefore, the observed DM abundance implies a *minimum* DM production rate at accelerators. The detailed phenomenology depends on the ratio $m_{A'}/m_\chi$. If $m_\chi < m_{A'} < 2m_\chi$, the mediator decay to DM is kinematically forbidden so that visible dark photon searches and DM searches like LDMX are both relevant. If instead $m_\chi < m_{A'}/2$, the mediator decays primarily to DM, making searches for DM production even more sensitive and absolutely essential. This case represents most of the parameter space for direct annihilation, so we focus on it from now on.

To comprehensively study direct annihilation models, we follow conventions in the literature (see [1]), and introduce the dimensionless interaction strength y as

$$\sigma v(\chi\chi \to A'^* \to ff) \propto \epsilon^2 \alpha_D \frac{m_\chi^2}{m_{A'}^4} = \frac{y}{m_\chi^2} \quad , \quad y \equiv \epsilon^2 \alpha_D \left(\frac{m_\chi}{m_{A'}}\right)^4 \tag{2}$$

This is a convenient variable for quantifying sensitivity because for each choice of m_χ there is a unique value of y compatible with thermal freeze-out independently of the individual values of α_D, ϵ and $m_\chi/m_{A'}$. Reaching experimental sensitivity to this benchmark for masses between MeV – GeV would provide nearly decisive coverage of this class of models [51].

The viable choices for χ (which vary according to spin and type of mass) whose relic density arises from direct annihilation can be simply enumerated
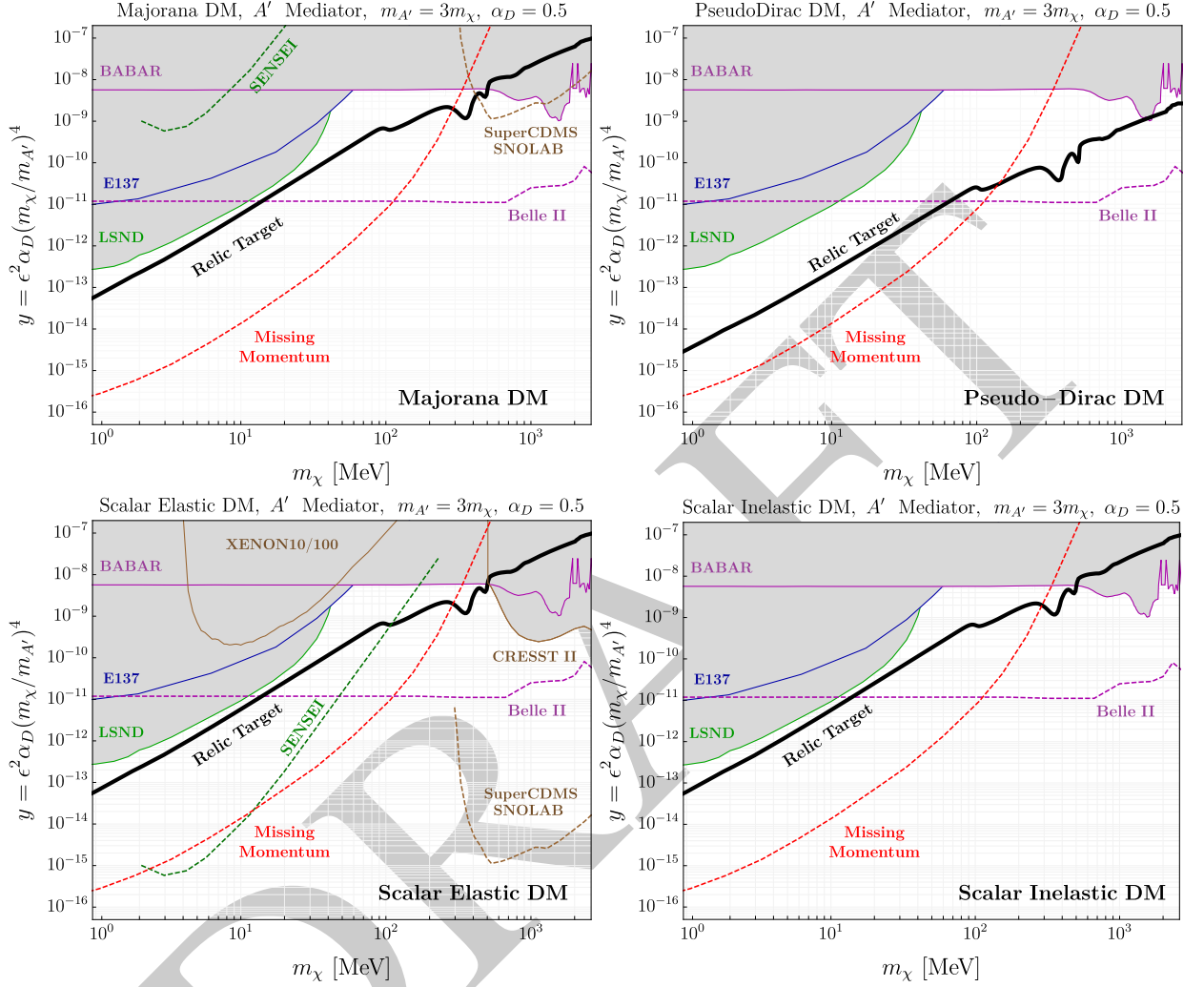
FIG. 4: The parameter space for LDM and future experimental projections in the y vs. m_χ plane plotted against the thermal relic targets for representative scalar and fermion DM candidates coupled to a dark photon A' – see text for a discussion. The red dashed curve represents the ultimate reach of an LDMX-style missing momentum experiment.

- **Majorana Dark Matter:** In this scenario, χ is a Majorana fermion. It therefore coupled through an axial-vector current

$$J_D^\mu = \overline{\chi}\gamma^\mu\gamma^5\chi, \tag{3}$$

which introduces velocity dependence in nonrelativistic scattering prcesses $\sigma_{\rm scat} \propto v^2$ and suppresses signals which involve DM interactions at low velocities. The thermal target and parameter space for this model are presented in the upper left panel of Fig. 4.

- **"Pseudo" Dirac Dark Matter:** Finally, χ can be a Dirac fermion. If the mass term for χ are entirely $U(1)_D$ preserving, then the model is already constrained by CMB data, unless it is asymmetric. If the mass terms for χ include $U(1)_D$ breaking (by analogy to the $SU(2)_W$ breaking mass terms of particles in the Standard Model), then χ splits into two Majorana

fermions (in the mass basis) which couple off-diagonally to the A' through

$$J_D^\mu = \overline{\chi}_1 \gamma^\mu \chi_2 + \overline{\chi}_2 \gamma^\mu \chi_1 \quad \text{(mass basis)}. \tag{4}$$

as well as a small diagonal axial current coupling to the A'. The thermal target and parameter space for this model are presented in the upper right panel of Fig. 4.

- **Scalar Elastic:** In this scenario, χ is a complex scalar particle with $U(1)_D$ preserving mass terms, and current

$$J_D^\mu \equiv i(\chi^* \partial^\mu \chi - \chi \partial^\mu \chi^*). \tag{5}$$

The annihilation cross section for this model is p-wave suppressed, so $\sigma v(\chi\chi^* \to f\bar{f}) \propto v^2$ and therefore requires a slightly larger coupling to achieve freeze out relative to other scenarios. This model also yields elastic signatures at direct detection experiments, so it can be probed with multiple complementary techniques. The thermal target and parameter space for this model are presented in the lower left panel of Fig. 4.

- **Scalar Inelastic Dark Matter:** In this scenario, χ is a complex scalar particle with $U(1)_D$ breaking mass terms (by analogy to the $SU(2)_W$ breaking mass terms of particles in the Standard Model). Therefore, χ couples to A' inelastically and must transition to a slightly heavier state in order to scatter through the current

$$J_D^\mu = i(\chi_1^* \partial^\mu \chi_2 - \chi_2^* \partial^\mu \chi_1), \tag{6}$$

which typically suppresses direct detection signals even for small mass differences between $\chi_{1,2}$. Thus, this model can only be probed robustly at accelerators. The thermal target and parameter space for this model are presented in the lower right panel of Fig. 4.

For each of these models, there is a slightly different y vs. m_χ thermal target in Fig. 4, but they all share the same parametric dependence on the product $\alpha_D \epsilon^2 (m_\chi/m_{A'})^4$. Furthermore, all of the models on this list are safe from bounds on late-time energy injection around $T \sim$ eV from Planck measurements of the CMB temperature anisotropies [17].

All existing constraints, collected in the y vs m_χ parameter space, are depicted in Fig. 4 alongside projections for various experimental techniques (see [2] for a discussion). Also shown are the thermal targets for each of the LDM candidates listed above, which are largely invariant in this parameter space regardless of the values of ϵ, α_D, and $m_\chi/m_{A'}$ separately. Direct detection constraints are also naturally expressed as functions of y and m_χ, but accelerator-based constraints are not — for example, collider production bounds depend on $m_{A'}$ and ϵ. However, within the predictive framework of direct DM annihilation, both α_D and $m_\chi/m_{A'}$ are bounded from above, accelerator-based constraints typically become *stronger* as either α_D or $m_\chi/m_{A'}$ is decreased (specifically, lowering α_D expands the constrained regions downward in Figure 4 while lowering $m_\chi/m_{A'}$ moves them downward and to the left). Therefore, we plot conservative versions of these constraints with $\alpha_D = 0.5$ and $m_\chi/m_{A'} = 1/3$ near their upper limits to reveal gaps in the DM parameter space consistent with the assumption of direct annihilation [52]. For more details about this parameter space see [4].

So long as $m_{A'} > 2m_\chi$, the dominant DM production process in LDMX is direct production of the mediator A' followed by decay to χ particles. This occurs in most of the parameter space, and therefore we have focused on this case in our design studies. In the region $m_\chi < m_{A'} < 2m_\chi$, DM is instead produced in LDMX through a virtual A' with comparable sensitivity (see [4]). For
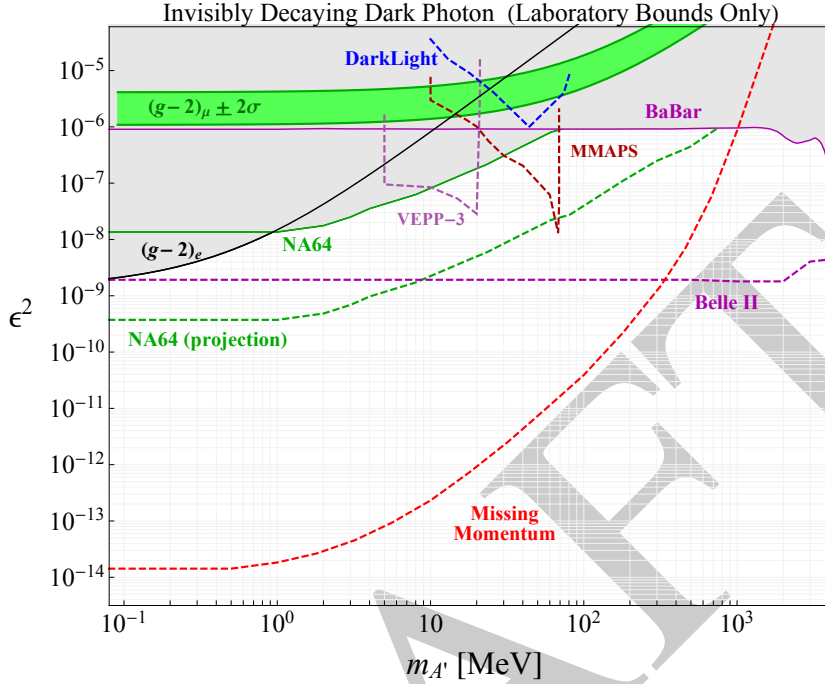
FIG. 5: The parameter space for an invisibly decaying dark photon A' in the ϵ^2 vs $m_{A'}$ plane. Also shown are various experimental constraints from [2].

A' masses that are significantly larger than a few GeV, DM production is dominated by virtual A' mediated reactions with contact operator behavior.

Independent of the details of χ, LDMX is sensitive to invisible decay channels of mediators that couple to electrons, such as an A'. Fig. 5 shows the available parameter space for an invisibly decaying A' in the ϵ^2 vs. $m_{A'}$ plane alongside projections for future experiments; this figure makes no assumptions about dark matter as long as the A' decays invisibly with unit branching fraction.

Although our emphasis in this section has been on various models of DM with direct annihilation through a dark photon mediator A', the missing momentum technique can probe multiple other mediator scenarios with equally powerful sensitivity to the corresponding theoretical targets. For example, both dark and visible matter could both be directly charged under a new $U(1)$ group which gauges an anomaly-free combination of SM quantum numbers (e.g. baryon minus lepton number). Such new forces can also mediate DM direct annihilation to SM particles with thermal targets analogous to those presented in Fig. 4. Furthermore, missing momentum techniques can also probe strongly interacting dark sectors [18], milli-charged particles, minimal dark photons, minimal $U(1)$ gauge bosons, axion like particles, and light new leptophilic scalars particles [11].

D. Searching for Light Dark Matter Production

As discussed above, the primary science goal for LDMX is to search for the process depicted in Fig. 6, wherein a DM particle-antiparticle pair is radiated off a beam electron as it scatters off a target nucleus. The DM production is mediated by a kinetically mixed dark photon – see Eq. (1). Depending on the dark photon and DM masses, the leading contribution to this process may be *decay* of a dark photon into a DM particle-antiparticle pair ($m_{A'} > 2m_\chi$) or pair-production through a virtual dark photon. In either case, a constraint on the DM particle production rate can be
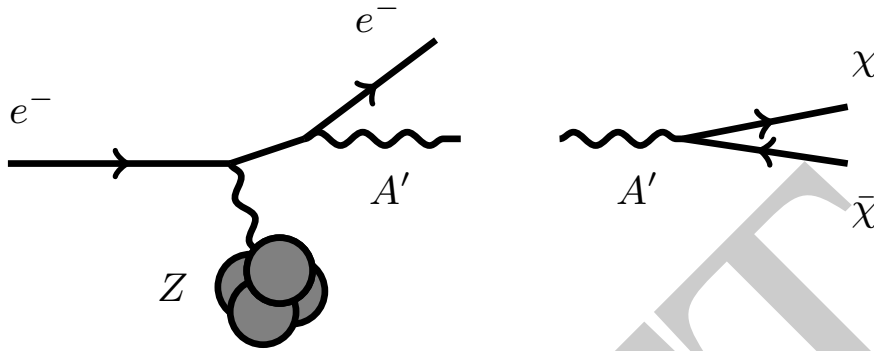
FIG. 6: Feynman diagram for radiation of an LDM particle-antiparticle pair off a beam electron as it scatters off a target nucleus. The DM production is mediated by a kinetically mixed dark photon (see (1)). Measuring this (and similar) reactions is the primary science goal of LDMX.

used to infer a robust bound on the interaction strength y, which can in turn be directly compared to the targets from thermal freeze-out shown in Figure 4. This note will focus, for concreteness, on the former case, though the final-state kinematics is extremely similar for either on- or off-shell dark photons and the same search strategies apply to both.

LDMX will search for this process using the *energy-angle kinematics* of the recoiling electron, or "missing momentum" approach. As discussed in detail in Section III, this kinematics is distinctive, with the recoiling electron typically carrying a small fraction of the beam energy (the rest is carried by the DM and hence invisible) and receiving an appreciable transverse kick from the DM production process. Thus, the experimental signature for the signal comprises three basic features: (i) a reconstructed recoiling electron with energy substantially less than beam energy but also (ii) detectable, with measurable transverse momentum, and (iii) an absence of any other activity in the final state.

This search strategy has distinct advantages over other approaches that have been used to detect DM production:

- **Missing mass** (as in BABAR, PADME, MMAPS*, and VEPP-3* – (* indicates proposed experiments) relies on full reconstruction of *all* recoiling particles, is only practical in e^+e^- collisions, and requires a much lower luminosity, greatly reducing production yield and hence sensitivity.

- **DM re-scattering** in a detector downstream of the production point (as in E137, Mini-Boone, and BDX*) can use very intense beams of either protons or electrons, but the low probability of DM scattering weakens sensitivity relative to what is possible in a kinematic search – whereas the kinematic signals of DM production scale as the square of the weak DM coupling, re-scattering signals scale as the fourth power. As a result, even the most aggressive proposals with intense beams fall short of the anticipated LDMX reach.

- **Missing energy** (as in NA64), reconstructing only the energy (not the angle or 3-momentum), is closely related to the missing-momentum approach but with fewer kinematic handles to reject SM backgrounds and measure veto inefficiencies *in situ*. In addition, missing energy experiments lack the ability to distinguish final-state electrons from one or
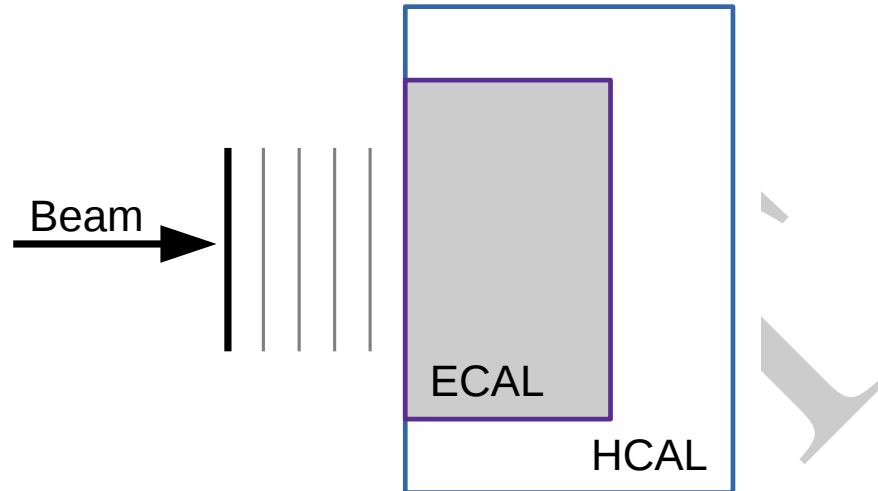
FIG. 7: Conceptual drawing of the LDMX experiment, showing the beam impacting on a thin tungsten target, the recoil tracker, the electromagnetic calorimeter, and hadron calorimeter.

more photons, introducing irreducible neutrino backgrounds to high-rate missing energy experiments [19]).

However, reaching the full potential of this technique places demanding constraints on the experiment and beamline supporting it. A high repetition rate of electrons is required ($\sim$ 50M e/sec on target for phase 1 and as much as $\sim$ 1G e/sec on target for phase 2), and so also a fast detector that can individually resolve the energies and angles of electrons incident on the detector, while simultaneously rejecting a variety of potential background processes varying in rate over many orders of magnitude. A conceptual cartoon diagram of the proposed experimental design is shown in Fig. 7, showing the alignment of the beam, the thin target, a tracker for the recoil electrons, and the required electromagnetic and hadron calorimeters to confirm the missing momentum signature. This cartoon will be helpful to the reader for understanding the discussion of signal and potential background reactions in Section III. The remainder of this note, from Section IV onward, describes the design in greater detail.
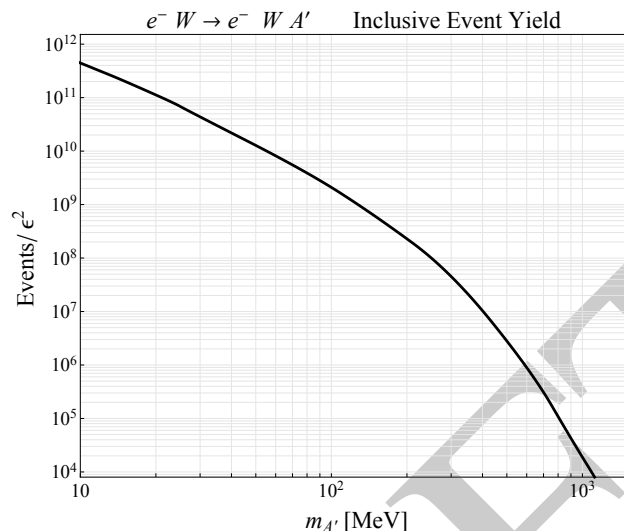
FIG. 8: Inclusive cross-section normalized to $\epsilon = 1$ for dark matter pair production off a tungsten target. The figure is adapted from [3], but utilizes signal samples for a 4 GeV electron beam and normalized to the phase 1 LDMX luminosity of 4×10^{14} EOT and target thickness $0.1X_0$.

III. SIGNAL AND BACKGROUND PROCESSES

A. Dark Matter Signal Production

This section briefly summarizes cross-sections and kinematic features of the DM production process in the case of an on-shell dark photon. More detail can be found in [3, 19]. The DM production cross-section σ_{DM} scales as the square of the kinetic mixing parameter ϵ, therefore, the quantity σ_{DM}/ϵ^2 and the kinematic distributions of the recoiling electron depend only on particle masses. Throughout this section, we assume the common benchmark value $m_{A'} = 3m_\chi$.

The normalized DM production cross-section on a tungsten target, as a function of DM mass, is shown in Fig. 8. For low masses, the production cross-section scales as $1/m_{A'}^2$; for higher masses, it is further suppressed by the steeply falling form-factor for scattering off the nucleus with high enough momentum exchange to produce an on-shell mediator.

The kinematics of the electron in DM production also depend on the A' mass. In general, as long as the A' or $\chi\bar{\chi}$ pair that is produced is heavy relative to the electron mass, the differential cross-section for DM production is peaked in the phase space where the DM carries away the majority of the beam energy and the electron carries relatively little. This is, of course, the exact opposite of the structure for the kinematics of the vast majority of bremsstrahlung events (with the tiny remainder constituting the main LDMX backgrounds, as discussed in the next section), and the primary kinematic handle for a missing momentum search. This qualitative behavior is accentuated for larger χ and A' masses, and less dramatic for lower masses comparable to m_e, as shown in the top left panel of Figure 9. The selection efficiency of a requirement $E_e < E_{cut}$ is shown on the bottom left panel as a function of E_{cut} (the fraction of events accepted saturates below 100% for some mass points because events with backward-going recoil electrons are not included for this plot). A nominal value $E_{cut} = 0.3E_{beam} = 1.2$ GeV (for 4 GeV beam energy) is assumed as a signal selection throughout this note. While this cut has not yet been optimized, it does keep 75-90% of signal events over the full range of masses considered. Reducing it by a factor of 2 would only slightly degrade signal efficiency. In the remainder of this section we also

impose a cut $E_e > 50$ MeV motivated by the degradation of tracking acceptance at lower energies. This too could be relaxed in future studies, but is included now to be conservative.

The top right panel of Figure 9 shows the electron p_T distribution for events that pass this $E_e < 1.2$ GeV cut for a range of DM and A' masses. These are to be contrasted with the sharply falling p_T distribution from bremsstrahlung, which (even after accounting for multiple scattering in a 10% r.l. target) falls off as $1/p_T^3$ for $p_T \gtrsim 4$ MeV. As we will see below, bremsstrahlung-originated events dominate the background, so this kinematic difference is a powerful one.

The optimal strategy for using angular/p_T information in a DM search will depend on the ultimate rejection power of the experiment's visible-particle veto, and varies depending on the DM mass of interest. Our approach to designing LDMX is to require that **no** kinematic information be needed to mitigate potential backgrounds that descend from bremsstrahlung photons. Instead, our goal is to fully measure the kinematics of the recoiling electron in as unbiased a manner as possible. In this way, any candidate signal events can be cross-checked against the expected features of dark matter production and its mass measured, thereby strengthening our discriminating power in the case of a discovery. If however a substantial number of events pass the energy selection and overall detector veto, either due to unexpected background reactions or poorer than planned performance, then the recoil electron p_T measurement can be used either to *reject* these events or to measure the bremsstrahlung component. Requiring a *minimum* recoil electron p_T of 20–50 MeV offers substantial kinematic rejection of bremsstrahlung background, but substantially degrades the experiment's acceptance for A' masses below ~ 100 MeV (i.e. DM masses below ~ 30 MeVin our plots). At lower masses, a 2D selection using both energy and angle information, or a multi-component signal+background fit to the 2D p_T-energy distribution may be the more effective use of this measurement.

Another very important kinematic feature is the angular/transverse separation between the recoiling electron and the direction of the beam. The larger this separation, the better the spatial separation between the recoil electron's shower and other final-state particles that one would like to veto. Figure 10 (left) shows the angle of the recoiling electron at the target for events passing the energy cut, while Figure 10 (right) shows, for the same events, the transverse separation of the recoiling electron from the beam spot at the face of the ECAL for the detector geometry described in Section IV.

B.   Background Processes

This section summarizes and briefly describes the important background processes for a dark photon signal. This is an introduction to the main challenges presented by each background, and later in Sec. VI, detailed study of the rejection of the backgrounds will be presented.

As an orientation, the highest rate process is the situation when the electron does not interact in the target and showers in the calorimetry. If the electron does undergo bremsstrahlung then the next set of processes are related to the potential fate of the photon. In rarer instances, the photon can interact without undergoing a typical electromagnetic shower, and present unique detector challenges. A similar process can happen for electrons undergoing a nuclear interaction. Finally, we discuss the rates of neutrino backgrounds.
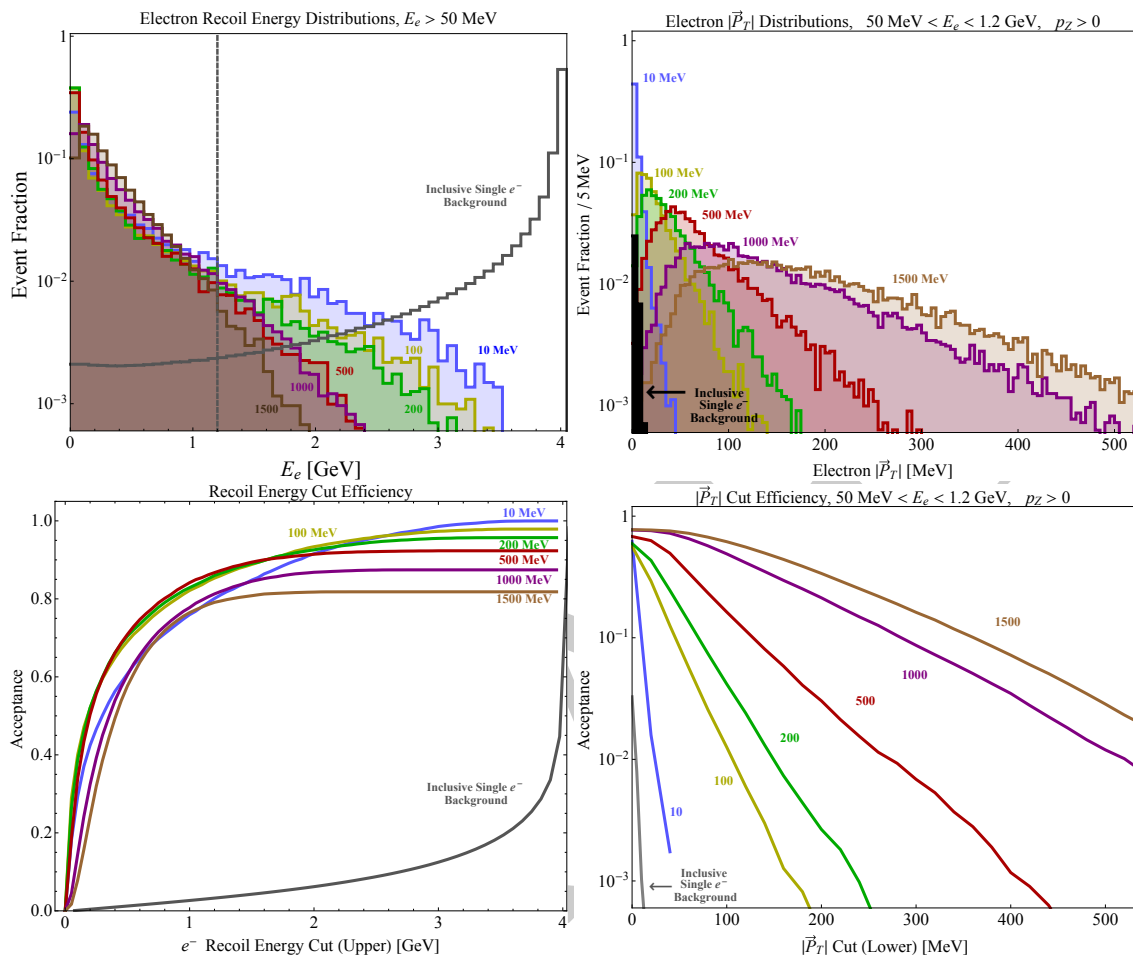
FIG. 9: Top: Electron energy (left) and p_T (right) spectra for DM pair radiation process, at various dark matter masses. Bottom Left: Selection efficiency for energy cut $E_e < E_{cut}$, as a function of E_{cut}, on inclusive signal events, The nominal cut is $E_{cut} = 0.3E_{beam}$. Bottom Right: Selection efficiency for p_T cut $p_{T,e} > p_{T,cut}$, as a function of $p_{T,cut}$, on events with 50 MeV $< E_e < E_{cut}$. In all panels, the numbers next to each curve indicate A' mass. Also included in each plot is the corresponding inclusive single electron background distribution.

1. Incident low-energy particles/beam impurities.

At the level of of 10^{16} incident electrons, a large number of electrons with low energies - consistent with signal recoils - are certain to intersect the target even for the most stringent requirement on the purity of the incoming beam. Such electrons, reconstructed correctly by the detectors downstream of the target, cannot be distinguished from signal.

For this reason, LDMX employs a tagging tracker to measure the trajectories of incoming electrons to ensure that each recoil corresponds to a clean 4 GeV beam electron entering the apparatus on the expected trajectory to veto any apparent signal recoils originating with beam impurities.
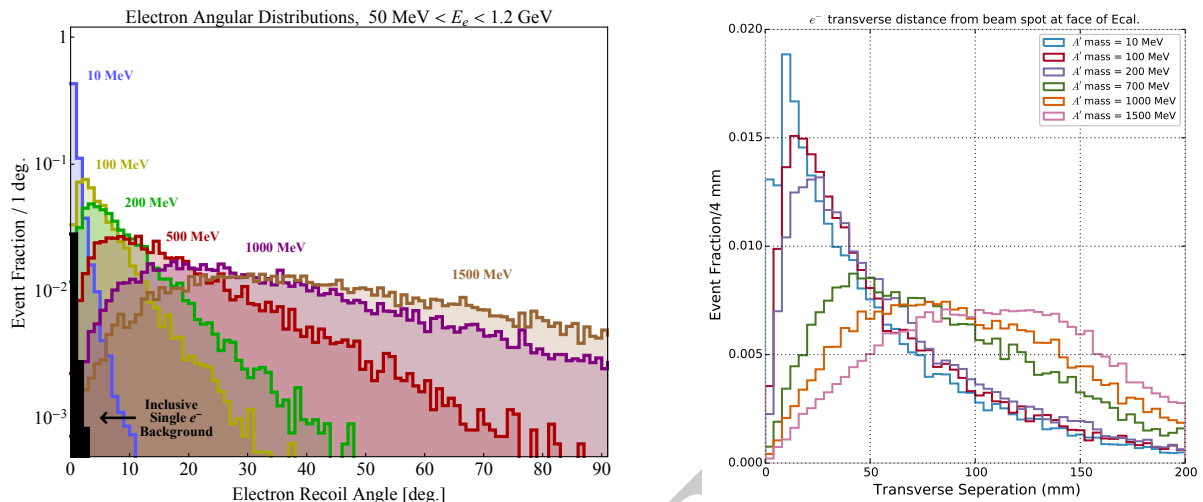
FIG. 10: Left: Angular distribution of signal events passing the energy cut $E_e < 1.2$ GeV. The numbers next to each curve indicate A' mass. Right: Transverse separation of the recoiling electron from the beam spot at the face of the ECAL for the detector geometry described in Section IV.

2. Electrons that do not interact in the target

These electrons experience some straggling in the trackers and target, but do not lose appreciable energy. These events feature a hard track through both trackers and typically include a high-energy (≈ 4 GeV) shower in the ECal. Occasionally, such events may have lower energy in the ECal due to electronuclear or photonuclear interactions occurring during the shower development in ECal.

3. Hard bremsstrahlung

These occur in the target (or, less frequently, in a tracking layer), such that the electron track through the recoil tracker falls below $E_{cut} = 1.2$ GeV, with a relative rate of 3×10^{-2} per incident electron. These events have two showers in the ECAL, with combined shower energy ≈ 4 GeV, separated by $1 - 2$ cm or more depending on the electron energy. In rejecting these backgrounds, it becomes important to resolve the electron from the photon and measure the photon energy with high resolution. In this context, a photon that is very poorly measured can give rise to a fake dark photon signal.

The discussion above assumes that the bremsstrahlung photon does not interact until it hits the ECal. Of course, the photon may convert in the target (or, less frequently, in a tracking layer). This background topology occurs at a relative rate of 1.5×10^{-3} per incident electron. It is generally easier to detect than the non-interacting bremsstrahlung: in addition to the energy deposition in the ECal, the e^+e^- pair from the conversion produces one or two additional tracks in the recoil tracker, which can be vetoed. The same considerations apply to the ($\sim 1.5 \times 10^{-4}$ per incident electron) background of trident reactions $e^- \rightarrow e^- e^+ e^-$ in the target where at least one of the outgoing e^- has energy below $E_{cut} = 1.2$ GeV.

A more challenging variation of hard bremsstrahlung is when, rather than producing a typical electromagnetic shower, the bremsstrahlung photon undergoes a rare process such as a photonuclear reaction or conversion to $\mu^+\mu^-$ (or $\pi^+\pi^-$). These processes typically deposit less detectable
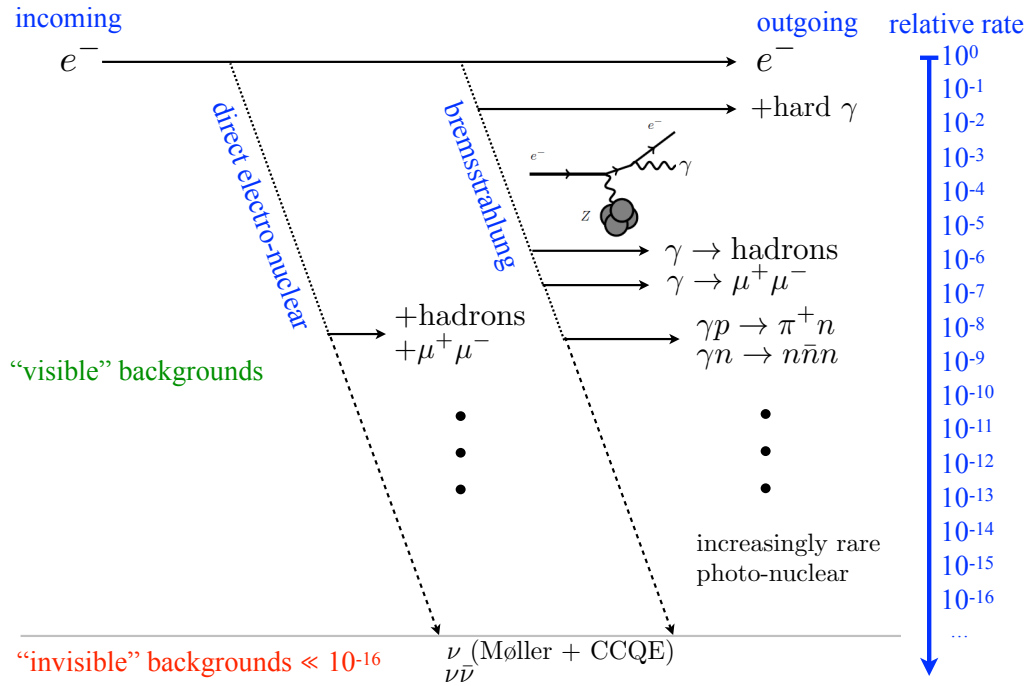
FIG. 11: Flow of important (veto design driving) potential background processes and their raw rates relative to the number of beam electrons incident on the target.

energy in the ECal, leading to a significant undermeasurement of the photon energy. We now discuss these photon processes which represent a challenging class of rarer backgrounds. These additional photon interactions can occur in the target or the calorimeter, with each case presenting unique challenges.

Hard bremsstrahlung + photonuclear reaction in the target or ECal. These are events where the hard bremsstrahlung photon does not undergo conversion but instead undergoes a photonuclear reaction in the target area or one of the first layers of the ECal (typically in a tungsten layer). The photonuclear cross-section is roughly 1000x smaller than the conversion cross-section, so these events will occur with overall relative rate of 1.7×10^{-5} per incident electron.

The photonuclear processes initiated by $2.8 - 4$ GeV photons can result in a wide range of final states. When pions are produced in the interaction and escape the nucleus, they typically give rise to either "tracks" ($\pi^{\pm}$) or substantial energy deposition (π^0) in the ECal. In many cases, a large number of low-energy protons and neutrons are liberated from a heavy nucleus; some of the protons may deposit energy in the ECal, and some of the neutrons can be detected by the HCal. Two rare but important classes of these events are characterized by only two to three $\mathcal{O}(\text{GeV})$, moderate-angle neutrons escaping from the nucleus, or a single forward ~ 3 GeV neutron. These events must be rejected by the HCal.

The same photonuclear reactions can also occur in the target. These events have 50x lower rate (relative to reactions in the ECAL) because of the thin target material, and tracking provides

additional handles to reject them, but they are also more difficult to detect calorimetrically because the effective angular coverage of the HCal is reduced.

Photon conversion to muons Analogous to the usual conversion reaction, photons can instead convert to a pair of muons with a rate suppressed by m_e^2/m_μ^2. Such events occur at a relative rate of 9.1×10^{-7} per incident electron. Muon events leave a qualitatively different detector signature than other background processes, namely one or two "tracks" passing through the calorimeter. Such processes can occur in the target, tracking system, or the photoconversion process can occur in the calorimeters. To ensure that wide-angle muon pair production is well modeled, we have included both coherent conversion $\gamma Z \to \mu^+\mu^- Z$ off a high-Z nucleus and the rarer incoherent conversion off a single nucleon. In addition, muon pairs can be produced in the absence of a preceding hard bremsstrahlung by trident reactions, illustrated in Fig. 12.
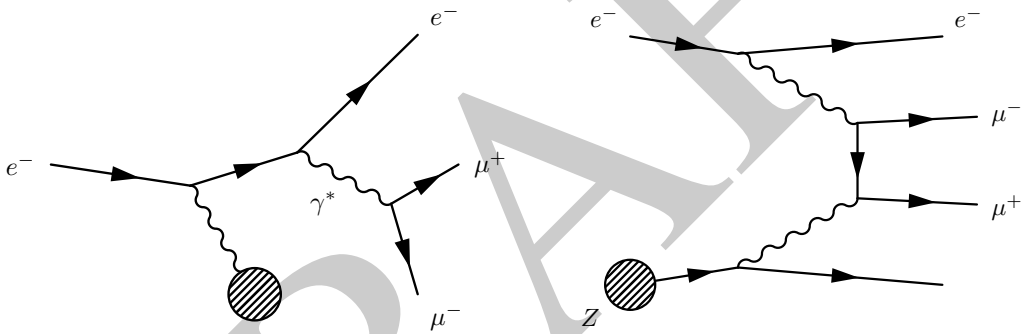

FIG. 12: Reducible potential background reactions, photoconversion (left) and Bethe-Heitler (right), with pair-produced muons in the final state.

4. Electronuclear interactions in target.

Electronuclear interactions present a similar composition to photonuclear reactions but occur at a lower rate and have a broader p_T distribution. Figure 13 shows the expected distribution of the recoil electron transverse momentum distribution for photonuclear and electronuclear events. Because electronuclear reactions in the ECal have the additional handle of the tracking system which can reject full energy electrons, the primary concern becomes electronuclear reactions that occur in the target area. However, because they are kinematically similar to the in-target photonuclear background, the same background rejection strategy is used for both interactions.

5. Neutrino backgrounds.

The backgrounds discussed above are all "instrumental" in that they rely on mismeasurement of energy escaping in visible particles (e.g. hadrons or muons) to fake a missing energy signal. Neutrino production, by contrast, transfers energy from the incident electron to particles that are (essentially) undetectable at LDMX. The rates for these processes are extremely small.

The dominant neutrino-production process, charged-current electron-nuclear scattering, has an inclusive cross-section of ~ 8 fb/GeV/nucleon leading to $O(5)$ such events expected in phase 1 of
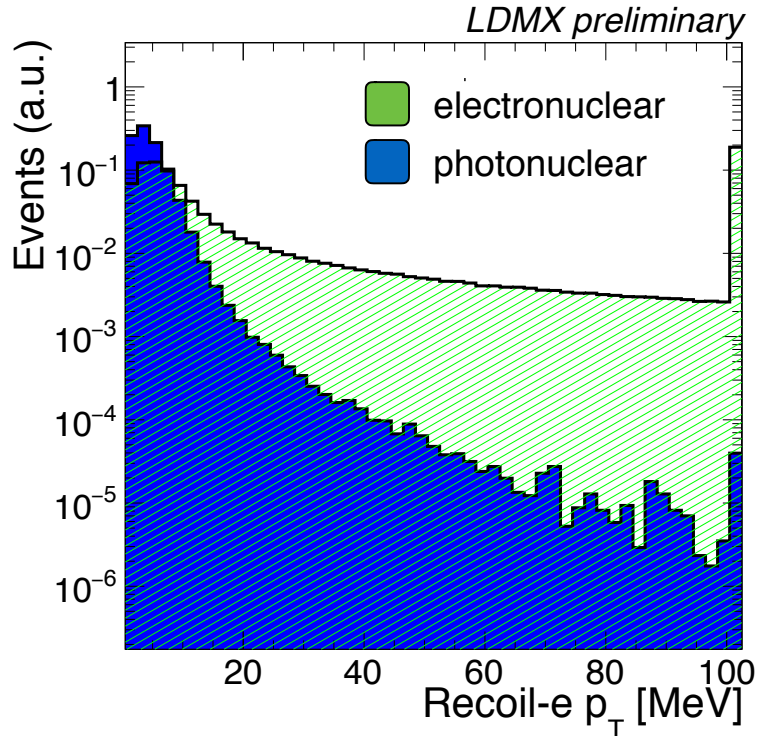
FIG. 13: Distributions of recoil transverse momentum for photonuclear events originating in the ECAL and electronuclear events originating in the target. The last bin shows the integrated yield for $p_{\mathrm{T}} > 100$ MeV.

LDMX. This is not a background for LDMX because there is no recoiling electron in the final state. However, rarer processes and/or multiple interactions can mimic the LDMX signal. Although the rates are small, it is worth elaborating on these background processes' yield to ensure that they can be neglected and/or measured. In particular, the latter two kinds of processes are reducible but may need to be considered in scaling up LDMX by 1-2 orders of magnitude in luminosity.

Neutrino backgrounds can be divided into three classes depending on the origin of the apparent recoil electron (from the neutrino-production process, from a second interaction in the target, or as a fake):

Neutrino trident processes $eN \to e\nu\bar{\nu}N$ are an *irreducible physics background* — the final state has both a low-energy recoil electron (with broad p_T distribution) and significant missing energy carried by invisible particles. Indeed, neutrino trident events are a direct counterpart of our signal process, mediated by electroweak bosons rather than a new force carrier. However, their rate in LDMX or even the most ambitious future upgrades is negligible. Parametrically, this reaction is suppressed relative to hard bremsstrahlung by $\frac{\alpha_2}{\pi}\frac{m_e^2 M_{\max}^2}{M_Z^4} \sim 10^{-18}$ where $M_{\max}^2 \sim E_{beam}/R_{nuc}$ is the pair mass scale at which the nuclear form factor strongly suppresses the trident process. For LDMX at 4 GeV, a neutrino trident cross-section of 20 ab was found using MadGraph/MadEvent, leading to a yield of 10^{-5} such events originating from the target in phase 1. While this background is clearly irrelevant for the currently proposed experiment, it does set an in-principle floor to the sensitivity of the missing momentum search program.

Multiple interactions in target + charged-current scattering Supposing that the charged-current scattering produces no visible final-state particles, a background to the LDMX signal could still arise from the combination of two processes occurring within the target — for example, production of a Moller electron at $O(100)$ MeV followed by the charged-current scattering of the incident electron, leading to a "low-energy e + high-energy ν" final state. Another similar final state arises from (1) soft bremsstrahlung by the incident electron, (2) asymmetric conversion of the resulting photon, with low-energy positron absorbed in the target and high-energy electron escaping, and (3) charged-current scattering of the hard incident electron. These processes are both expected to lead to $O(0.01)$ background events at LDMX.

Inelastic charged-current scattering (fake electron) The charged-current scattering cross-section for multi-GeV electrons (or neutrinos) is dominated by resonant excitation of a nucleon and deep inelastic scattering, which can lead to charged pions in the final state. There is some region of phase space where these pions, if they fake an electron, would mimic the LDMX signal. It is difficult to estimate this rate quantitatively. However, since it represents only a fraction of the final-state phase space for inclusive CC scattering, the yield should be no more than $O(1)$ event and likely smaller. Furthermore, since electrons and pions can be distinguished calorimetrically, this background is expected to be negligible.

C. Experimental Strategy

This section summarizes the high level design considerations for LDMX. We have two goals – to measure the distinguishing properties of dark matter production (see Section III A), and to reject potential backgrounds for this process (see section III B).

To measure the signal process, we need to precisely measure the momentum of both the incoming and recoiling electron. The relative transverse momentum provides a measure of the mass of the produced dark matter, and the overall relative momentum provides a measure (to a high degree) of its direction. Thus, we want to use a thin target and low material tracking (minimize the impact of multiple scattering) in a magnetic field to measure the signal process. This is illustrated by the left cartoon in Fig. 14.

Having measured the momentum change across the target, our main goal is then to capture the fate of bremsstrahlung photons and non-interacting electrons. The fate of such photons is illustrated in the background flow chart shown in Fig. 11. A typical candidate background reaction is illustrated by the right cartoon in Fig. 14.

Both non-interacting electrons, and electrons that undergo appreciable bremsstrahlung in the target give rise to EM showers (totalling the beam energy) in the ECAL system, but these are easily rejected with good shower energy reconstruction resolution. Therefore, these events can be readily rejected at trigger level, thereby limiting the event rate of the experiment to a manageable O(10) kHz or lower rate.

At the level of roughly $\sim 2 \times 10^{-5}$ per incident electron, photo-nuclear (PN) processes occur, and at roughly $\sim 10^{-6}$ muon conversion processes occur. The ECAL will continue to reject the majority of PN events due to energy deposition from π^0 production, but good MIP sensitivity is required to reject the small subset of events that contain no energetic π^0, but rather energetic $\pi^\pm$, p, or photon conversions to muon pairs. In the system described in this paper, this can in practice provide another factor of $\sim 10^3$ rejection, depending on the type of final state.

In practice, the most difficult potential background for the ECAL to reject is one in which the hard bremsstrahlung photon converts all of its energy into a single (or few body) neutral hadronic
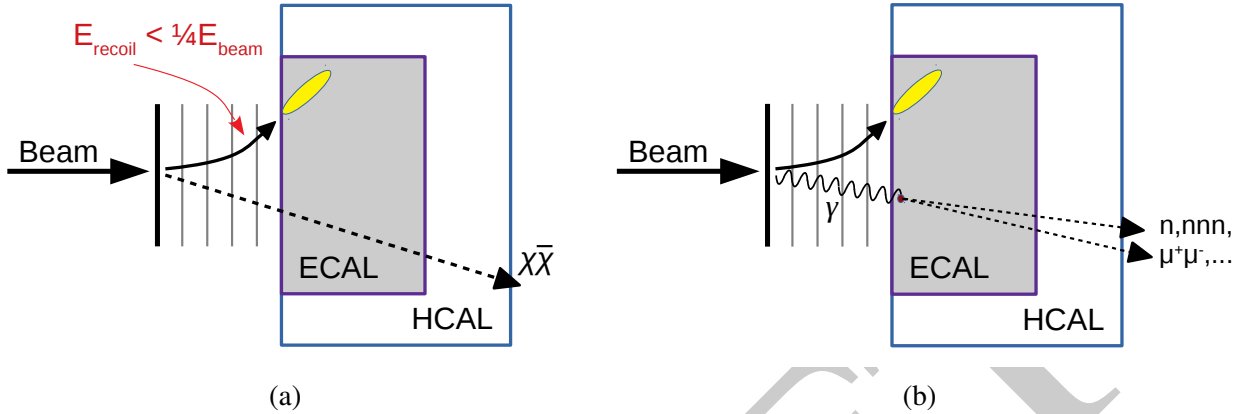
FIG. 14: Conceptual schematic of a signal process (a) and dominant background (b) processes.

final state. This occurs at a relative rate of $\sim 10^{-3}$ per incident hard photo-nuclear reaction (on W), but these usually have a hard charged pion or proton in the final state. Thus, the region of phase space where the MIP is soft and invisible posses the most threat of producing a background, and this is expected at the $\sim 4 \times 10^{-4}$ per hard photo-nuclear interaction (on W). Per incident 4 GeV electron on Tungsten absorber, this corresponds to $\sim 10^{-8}$ in relative rate. For a benchmark of 1×10^{14} electrons on target, we would face up to $\sim 10^{6}$ events with a single hard forward neutron and very little else in the ECAL (other than the recoil electron). This drives the performance requirement of the hadronic veto – we require better than 10^{-6} neutron rejection inefficiency in the few GeV energy range. In practice, an HCAL veto meeting this requirement is also sufficiently sensitive to muons to veto the remainder of the photon conversions to muon pairs (and by extension, pion pairs). Moreover, this level of inefficiency provides a great deal of redundancy against potential failures of the ECAL veto with respect to photo-nuclear, electro-nuclear, or MIP conversion events.

IV. DETECTOR CONCEPT

As described in Sec. III A, the signature of DM production used by LDMX involves (i) substantial energy loss by the electron (e.g. recoil with $\lesssim 30\%$ of incident energy), (ii) a potentially large transverse momentum kick, and (iii) the absence of any additional visible final-state particles that could carry away energy lost by the electron. In the first phase of LDMX we will search for this signature in a sample of $4 \cdot 10^{14}$ incident electrons, delivered onto a 10% radiation-length ($0.1\ X_0$) target with a 46 MHz bunch spacing and average charge of one e^- per bunch. The target thickness has not been precisely optimized, but represents a reasonable compromise between event rate and transverse momentum resolution over a wide range of dark matter mass. Performing the measurement of each kinematic component (i and ii above) on every incident electron requires a tracker and electromagnetic calorimetry (ECal) downstream of the target, with a sensitive area that extends over the nominal beam axis. Placing the tracker in a weak magnetic field allows us to use some of the most striking of these events in which a soft, wide-angle recoil electron doesn't penetrate into the ECal. In addition, stray low-energy particles from beam halo must be rejected with very high efficiency, as these could mimic the DM signal. This motivates another tracker with stronger B-field upstream of the target that precisely measures the momentum and trajectory of the incoming beam electrons. While the energy and angle resolution requirements for LDMX are modest, the experiment poses two main challenges. First, every incident electron passes through the trackers and showers in the ECal; the detectors placed directly in the beam line must therefore contend with high radiation doses and a large event rate. In order to mitigate this issue, the beamspot for the experiment must be quite large - of order 10-100 cm^2. Second, the requirement of an event veto for any additional visible final-state particles to take advantage of (iii), that's robust enough to handle a variety of rare backgrounds is not only an important physics performance driver for the tracking and ECal, but also calls for the addition of a dedicated hadron veto system surrounding the sides and back of the ECal.

These four detector systems – the "tagging tracker" upstream of the target, "recoil tracker" downstream, a radiation hard forward ECal, and the hadronic veto system, form the majority of the LDMX experimental concept. To keep the detector compact and the field in the ECal minimal, we place the tagging tracker inside the bore of a dipole magnet and the recoil tracker in its fringe field. The layout for LDMX is illustrated in Figures 15 and 16.

The tracking proposed for LDMX borrows heavily from the Silicon Vertex Tracker (SVT) of the HPS experiment [20], a fixed target search for visibly decaying dark photons using the CEBAF beam at the Thomas Jefferson National Accelerator Facility (JLab). The HPS SVT was designed to provide high-purity, high-precision tracking for low-momentum ($\lesssim 3$ GeV) electrons at high occupancies (up to 4 Mhz/mm^2) with the nearly-continuous (2 ns bunch spacing) CEBAF beam. The SVT meets these requirements with low-mass construction ($0.7\%\ X_0$ per 3d hit) and excellent hit-time resolution (~ 2 ns). Given very similar requirements, the solutions developed for HPS are also optimal for LDMX.

Similarly, a natural solution to the granularity, timing, and radiation hardness required of the ECal is to exploit the silicon calorimetry designed for the CMS Phase-II high granularity calorimeter (HGC) upgrade in the forward region of the detector.[21] The Tungsten-Silicon sampling calorimeter is an ideal technology with very good energy resolution for electromagnetic showers, potential for MIP tracking, and the ability to withstand high radiation levels from non-interacting electrons. In addition, having been designed for 25 ns bunch spacing with up to 200 proton-proton collisions per bunch crossing in the High-Luminosity phase of the CERN LHC, the HGC has rate capabilities that are ideal for the LDMX ECal. The primary physics trigger will also require a
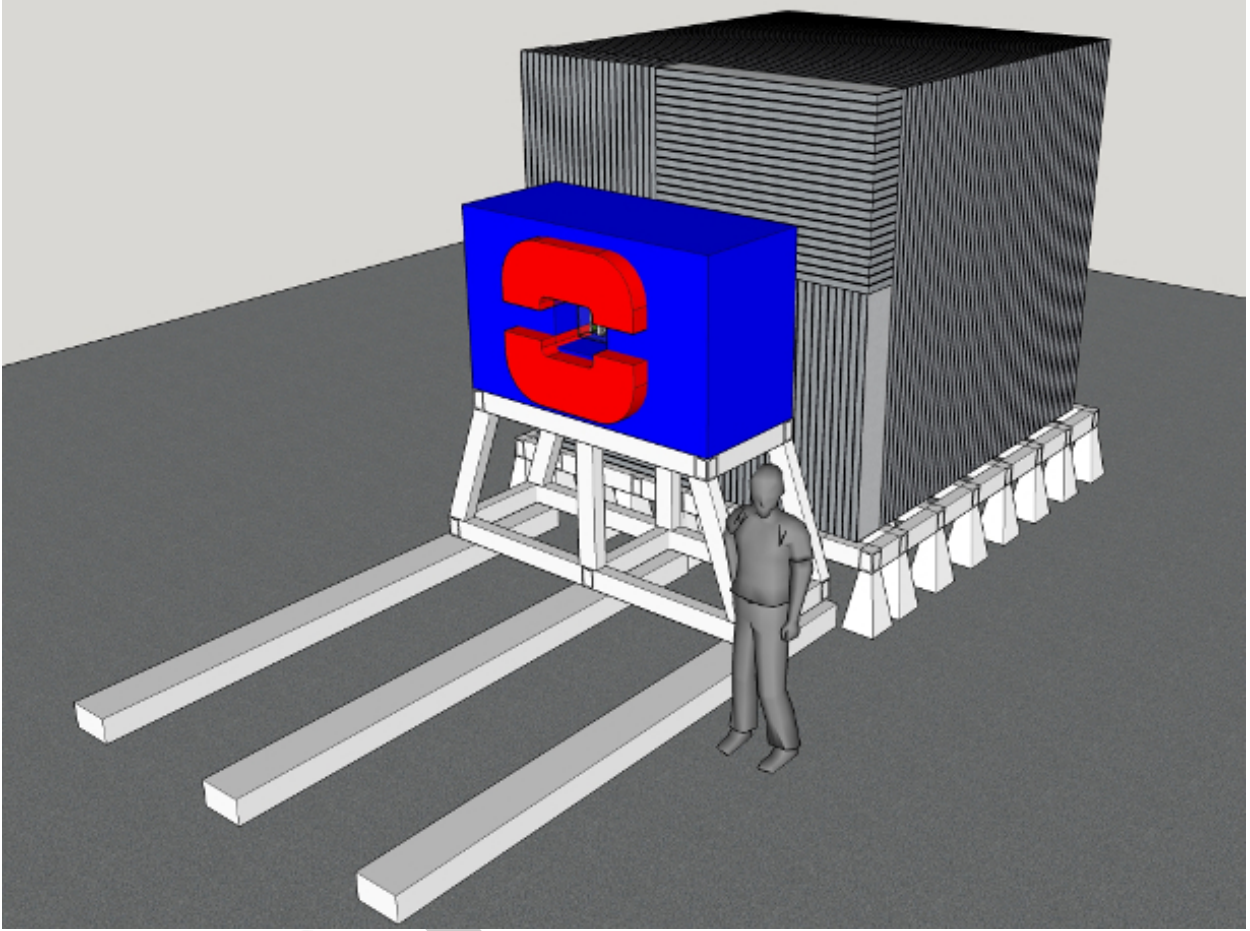
FIG. 15: An overview of the LDMX detector showing the full detector apparatus with a person for scale.

positive signal in a scintillator pad overlaying the target that's coincident with *low* (or no) energy deposition in the ECal in order to avoid triggering on empty buckets. Finally, the hadronic veto is designed to be a Steel-Plastic Scintillator sampling calorimeter (HCal) with high sampling fraction for extremely high neutron detection efficiency and good angular coverage for large-angle background processes. Like the ECal, the HCal could take advantage of technology being developed for other experiments, such as high speed SiPM detector readout designed for the current CMS hadronic calorimeter or the Mu2e experiment.

In the following sub-sections, we describe the LDMX beamline and detector sub-systems in more detail.

A. Beamline and Spectrometer Magnet

The LDMX beamline consists of a large-diameter beampipe terminating in a thin vacuum window immediately upstream of an analyzing magnet inside of which the tagging and recoil trackers are installed. The analyzing magnet is a common 18D36 dipole magnet with a 14-inch vertical gap and operated at a central field of 1.5 T. The magnet is rotated by approximately 100 mrad about the vertical axis with respect to the upstream beamline so that as the incoming 4 GeV beam is deflected by the field, it follows the desired trajectory to the target. In particular, the incoming
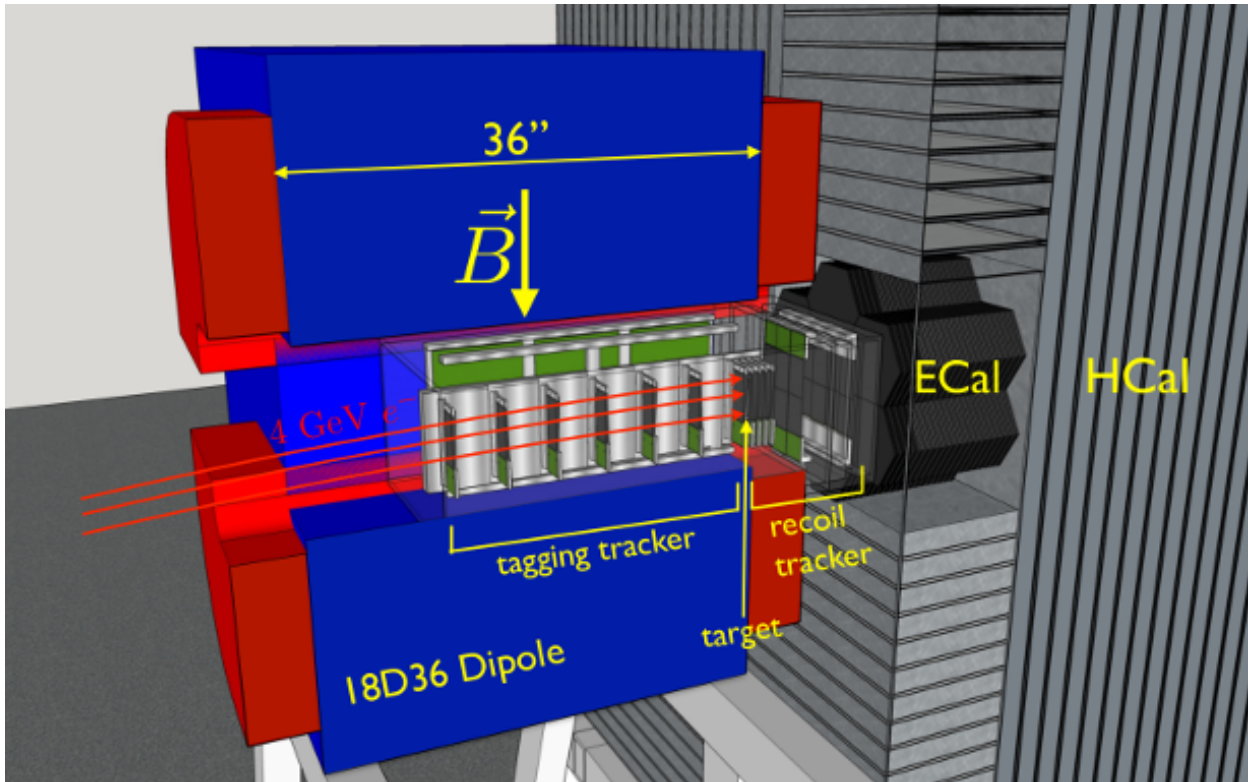
FIG. 16: A cutaway overview of the LDMX detector showing, from left to right, the trackers and target inside the spectrometer dipole, the forward ECal, and the HCal. Although not yet studied, the Side HCal may be extended forward (transparent region surrounding the ECal) to provide better coverage for wide-angle bremsstrahlung and neutral hadrons originating from photo-nuclear reactions in the target and the front of the ECal.

beam arrives at normal incidence to, and centered on, the target, which is laterally centered in the magnet bore at $z = 40$ cm relative to the center of the magnet. Although the magnet gap differs, this arrangement is very similar to that employed by the HPS experiment at JLab.

A number of 18D36 magnets, not currently in use, are available at SLAC, along with the steel required to adjust the magnet gap as may be required to suit our purposes. These include a magnet that is already assembled with the 14-inch gap planned for LDMX. It was tested to 1.0 T in 1978, at which point 199 kW of power was dissipated. Based on the current capacity of the other similar magnets with smaller gaps, it is expected that this magnet can be operated at 1.5 T, resulting in a power dissipation of approximately 450 kW and requiring an approximately 55 gpm flow of cooling water. If this magnet proves to be suitable for LDMX, it will be split, cleaned up, and reassembled before testing and carefully mapping the field in the tracking volumes. Although the final location of LDMX has not been determined, the downstream end of the SLAC ESA beamline would permit construction of a large HCal and could be accommodate LDMX operation along with other experiments.

B. Overview of Tracking and Target Systems

Although the tagging and recoil trackers function as two distinct systems, they use common technologies and share the same support structures and data acquisition hardware. In particular, the first four layers of the recoil tracker are identical to the layers of the tagging tracker and share a common support and cooling structure, as shown in Figure 17. The key element of this
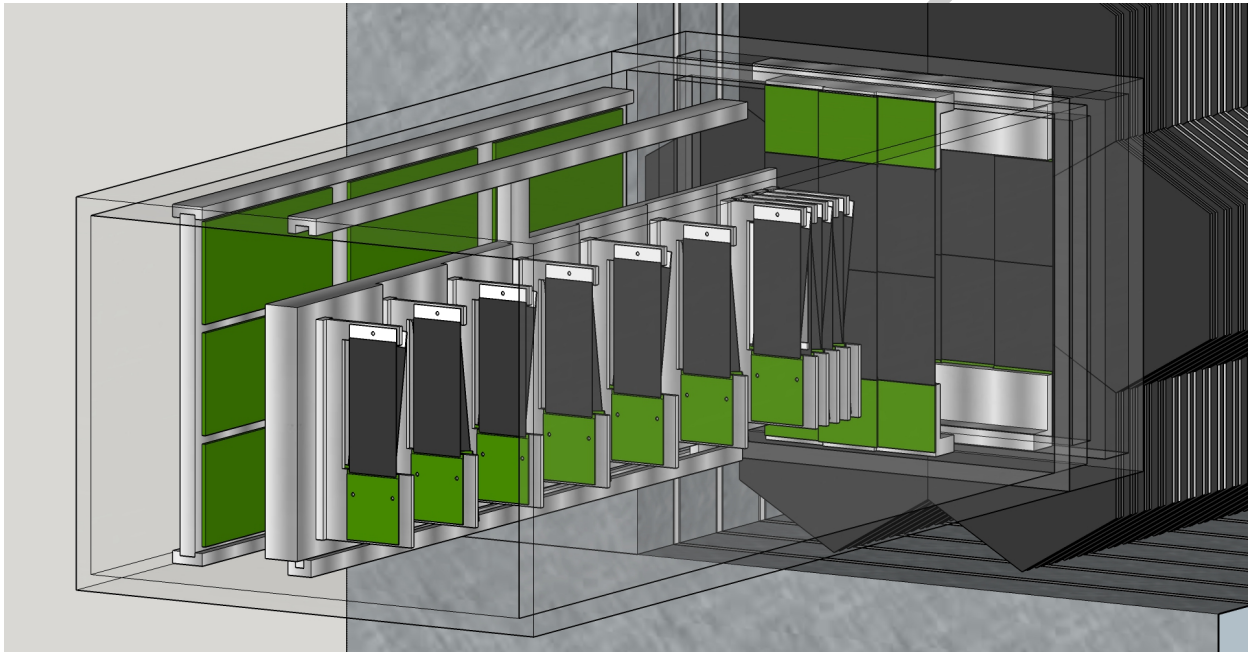

FIG. 17: An overview of the tracking systems and target inside the LDMX magnet.

upstream support structure is a vertically-oriented aluminum plate onto which the stereo modules are mounted. To provide cooling, a copper tube through which coolant flows is pressed into a machined groove in the plate. Starting from the upstream end of the magnet, the plate slides into precision kinematic mounts in a support box that is aligned and locked in place within the magnet bore. Another similar plate slides into the support box on the positron side and hosts the Front End Boards (FEBs) that distribute power and control signals from the DAQ and also digitize raw data from the modules for transfer to the external DAQ. The last two layers of the recoil tracker, being much larger, are supported on another structure: a cooled support ring onto which the single-sided, axial-only modules are mounted. This support ring is installed from the downstream end of the magnet, engaging precision kinematic mounts in the support box for precise alignment to the upstream stereo modules. The cooling lines for all three cooled structures—the upstream and downstream tracker supports and the FEB support—are routed to a cooling manifold at the upstream end of the magnet which, in turn, connects to a cooling feedthrough with dielectric breaks on the outside of an environmental enclosure which shields the detector from light and RF and maintains an environment of dry gas.

Overall, this design is similar to that of the HPS tracker, although with some important simplifications. First, because the radiation dose in the LDMX tracking system is modest, cooling is only needed to remove heat from the readout electronics. It is not necessary to keep the silicon itself cold. Therefore, cooling water that is close to room temperature can be used and there are no significant issues of differential thermal expansion to be concerned with. Second, the LDMX de-

tector is in no danger from the nominal LCLS-2 beam, so it does not need to be remotely movable, in contrast to HPS. Finally and most significantly, the LDMX detector does not need to operate inside the beam vacuum as is the case for HPS. This greatly simplifies many elements of the design, the material selection, and the construction techniques.

It is the target interposed between the last layer of the tagging tracker and the first layer of the recoil tracker that determines the different functional requirements of the tagging and recoil tracking systems. The target is a 350 micron tungsten sheet, comprising 10% of a radiation length ($0.1\ X_o$). This choice of thickness provides a good balance between signal rate and transverse momentum transfer due to multiple scattering, which limits the potential utility of transverse momentum as a signal discriminator. The tungsten sheet is glued to a stack of two 2 mm planes of PVT scintillator that enables a fast count of the incoming electrons in each bunch as required to select the appropriate threshold employed by the ECal trigger as discussed in the trigger description. The scintillator-tungsten sandwich is mounted in an aluminum frame that is inserted into the upstream tracker support plate from the positron side to simplify the process of replacing or swapping the target, as shown in Figure 18. With the scintillator mounted on the downstream face
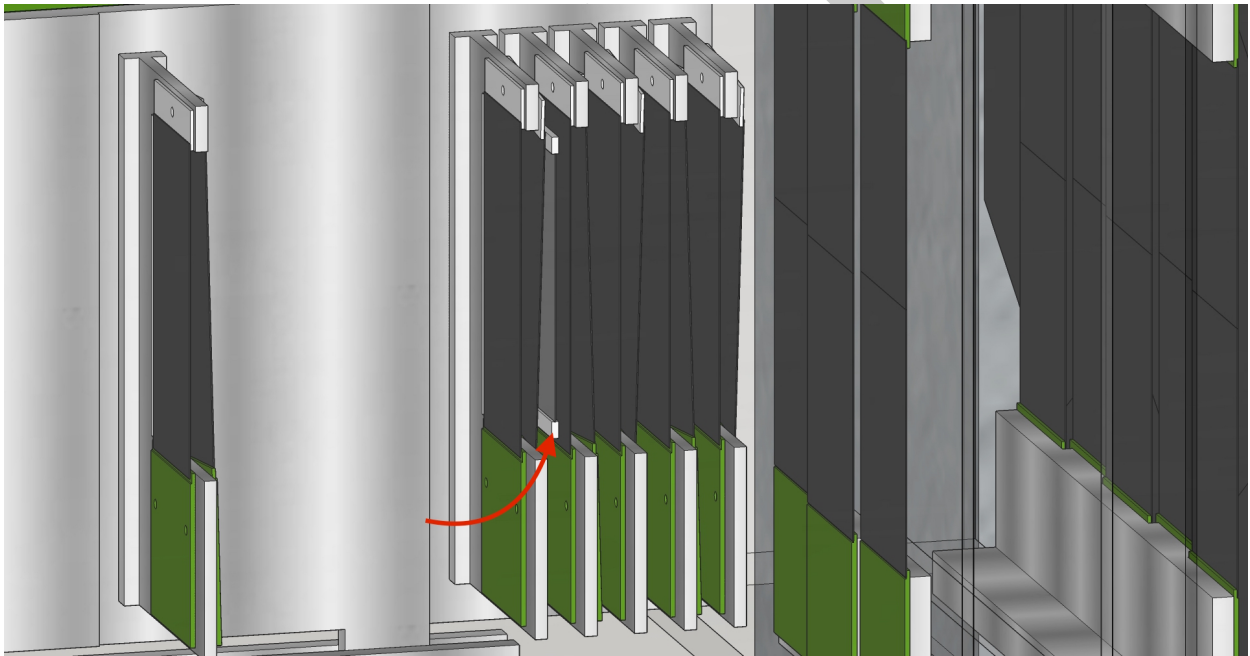
FIG. 18: The LDMX target, indicated by the red arrow, installed between the last tagging tracker layer and first recoil tracker layer.

of the target, it can also be used to help identify events in which a bremsstrahlung photon has undergone a photonuclear reaction in the target. The structure and readout of the scintillator are discussed below as part of the trigger system.

1. Tagging Tracker

The tagging tracker is designed to unambiguously identify incoming beam electrons that have the correct energy, as well as precisely measure their momentum, direction, and impact position at the target. The key elements of the design are determined by this goal. First, the long, narrow

layout of the tagging tracker accepts only beam electrons that have roughly the momentum and trajectory expected for the incoming beam. This eliminates off-energy beam electrons. Second the layers have low mass and are spaced far apart in a significant magnetic field to ensure a precise measurement of momentum. Finally, a large number of layers ensures the redundancy that's required for high-purity pattern recognition.

The layout and resolution of the tagging tracker are summarized in Table I. It consists of six

TABLE I: The layout and resolution of the tagging tracker.

Layer	1	2	3	4	5	6	7
z-position, relative to target (mm)	-607.5	-507.5	-407.5	-307.5	-207.5	-107.5	-7.5
Stereo Angle (mrad)	-100	100	-100	100	-100	100	-100
Bend plane (horizontal) resolution (μm)	$\sim$6	$\sim$6	$\sim$6	$\sim$6	$\sim$6	$\sim$6	$\sim$6
Non-bend (vertical) resolution (μm)	$\sim$60	$\sim$60	$\sim$60	$\sim$60	$\sim$60	$\sim$60	$\sim$60

double-sided, silicon microstrip modules arranged at 10 cm intervals upstream of the target, with the first module centered at $z = -7.5$ mm relative to the target. The modules are positioned laterally within the magnet bore such that they are centered along the path of incoming 4 GeV beam electrons. Each module places a pair of 4 cm $\times$ 10 cm sensors back to back: one sensor with vertically oriented strips for the best momentum resolution, and the other at ±100 mrad stereo angle to improve pattern recognition and provide three-dimensional tracking. The sensors are standard p^+-in-n type silicon microstrip sensors identical to those used for HPS. These sensors have 30(60) μm sense(readout) pitch to provide excellent spatial resolution at high S/N ratios. They are operable to at least 350 V bias for radiation tolerance.

The sensors are read out with CMS APV25 ASICs operated in multi-peak mode, which provides reconstruction of hit times with a resolution of approximately 2 ns. At very high occupancies, six-sample readout can be used to distinguish hits that overlap in time down to 50 ns, which limits the readout rate, and therefore the trigger rate, to approximately 50 kHz. However, at the low hit occupancies anticipated in LDMX, three-sample readout may suffice, enabling a maximum trigger rate approaching 100 kHz. These sensors are mounted on standard FR4 hybrid circuit boards, which provide the power conditioning and I^{2}C control for the APV25, as well as a thermal path to the cooled support structure.

The sensors and hybrids are assembled into half-modules, which are the smallest non-reconfigurable units of the tagger tracker. Each half module consists of a single sensor that's glued to the hybrid at one end with conductive epoxy to provide bias voltage and support. A thin ceramic plate is glued to the other end of the sensor for support. A pair of half modules is attached to either side of an aluminum module support with screws to form a module. Finally these double-sided modules are attached to the upstream support plate described in Section IV B to position them along the beamline and provide cooling for the hybrids as shown in Figure 19. This support arrangement places only silicon in the tracking volume and no material on the side of the tagger towards which electrons bend in order to avoid secondaries that would be generated by degraded beam electrons hitting material close to the tracking planes.

2. Recoil Tracker

The recoil tracker is designed to identify low-momentum (50 MeV to 1.2 GeV) recoil electrons and precisely determine their momentum, direction, and impact position at the target. In addition,
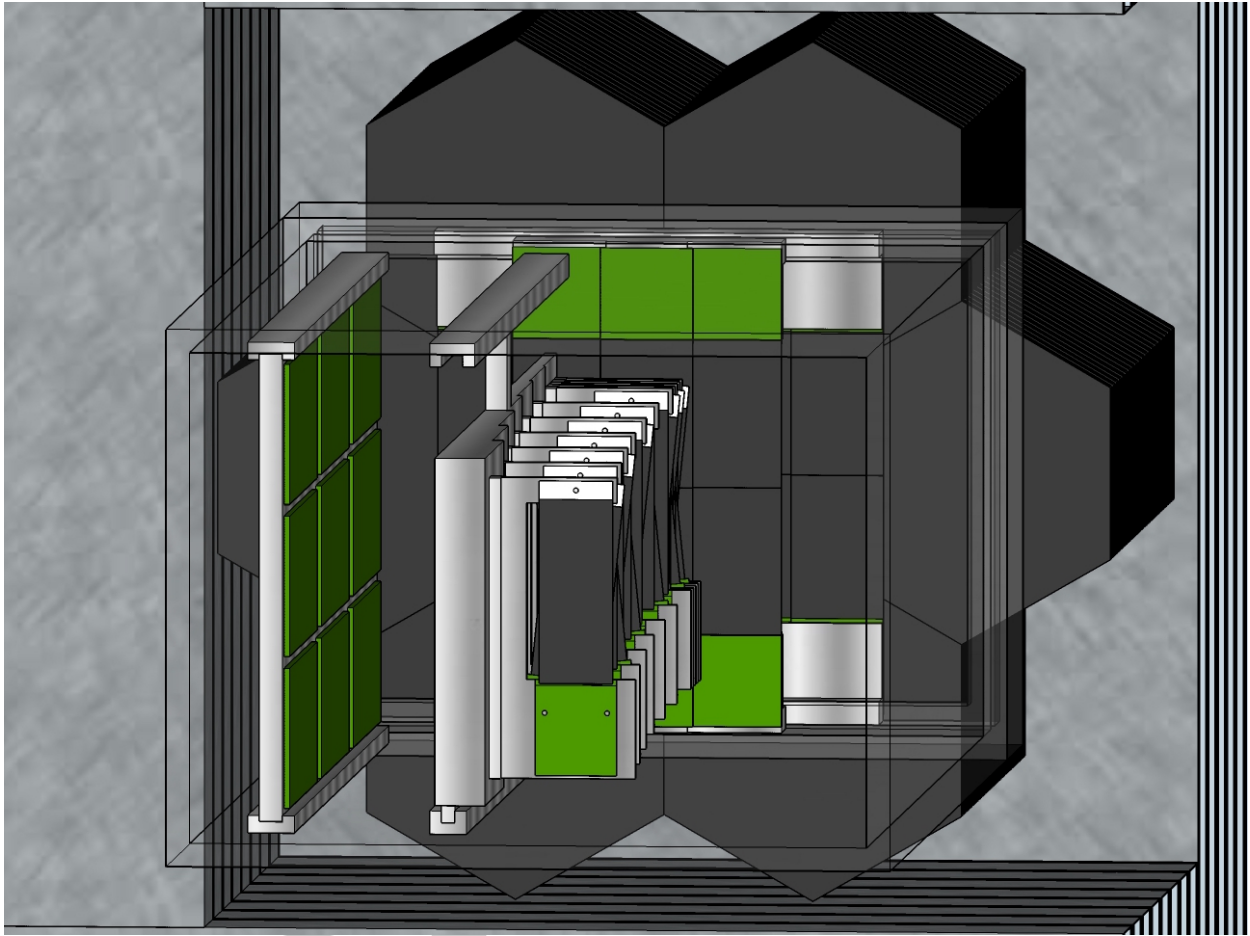
FIG. 19: A beams' eye view of the tagging and recoil trackers.

it must work together with the calorimeters to correctly distinguish low-momentum signal recoils from scattered beam electrons and multi-particle backgrounds. The key elements of the design are determined by this goal. First, the recoil tracker is placed at the end of the magnet in the beginning of the fringe field to optimize tracking for particles up to two orders of magnitude softer than the beam-energy electrons measured by the tagging tracker. Second, the recoil tracker is wide for good acceptance in angle and momentum and is longitudinally compact to minimize the distance from the target to the calorimeters to maintain good angular coverage. Finally, the recoil tracker provides 3-d tracking near the target for measurement of both direction and impact parameter with good resolution, but emphasizes low mass density over the longest possible lever arm further downstream to deliver the best possible momentum resolution. This design delivers good momentum resolution for both multiple-scattering limited, low-momentum tracks and beam energy electrons that travel along a nearly straight path in the fringe field.

The layout and resolution of the recoil tracker are summarized in Table II. It consists of four stereo layers located immediately downstream of the target and two axial layers at larger intervals in front of the ECal. The stereo layers are double-sided modules of silicon microstrips arranged at 15 mm intervals downstream of the target, with the first module centered at $z = +7.5$ mm relative to the target. These modules are laterally centered on the target and the center of the magnet bore and are identical to the modules of the tagger tracker that are mounted upstream on the same support plate.

TABLE II: The layout and resolution of the recoil tracker.

Layer	1	2	3	4	5	6
z-position, relative to target (mm)	+7.5	+22.5	+37.5	+52.5	+90	+180
Stereo Angle (mrad)	100	-100	100	-100	-	-
Bend plane (horizontal) resolution (μm)	≈6	≈6	≈6	≈6	≈6	≈6
Non-bend (vertical) resolution (μm)	≈60	≈60	≈60	≈60	-	-

The thinner axial-only layers, at $z = +90$ mm and $z = +180$ mm, are mounted on a separate support structure at the downstream end of the magnet bore as described in Section IV B. They have a somewhat different module design as shown in Figure 20. Each consists of a pair of sensors,
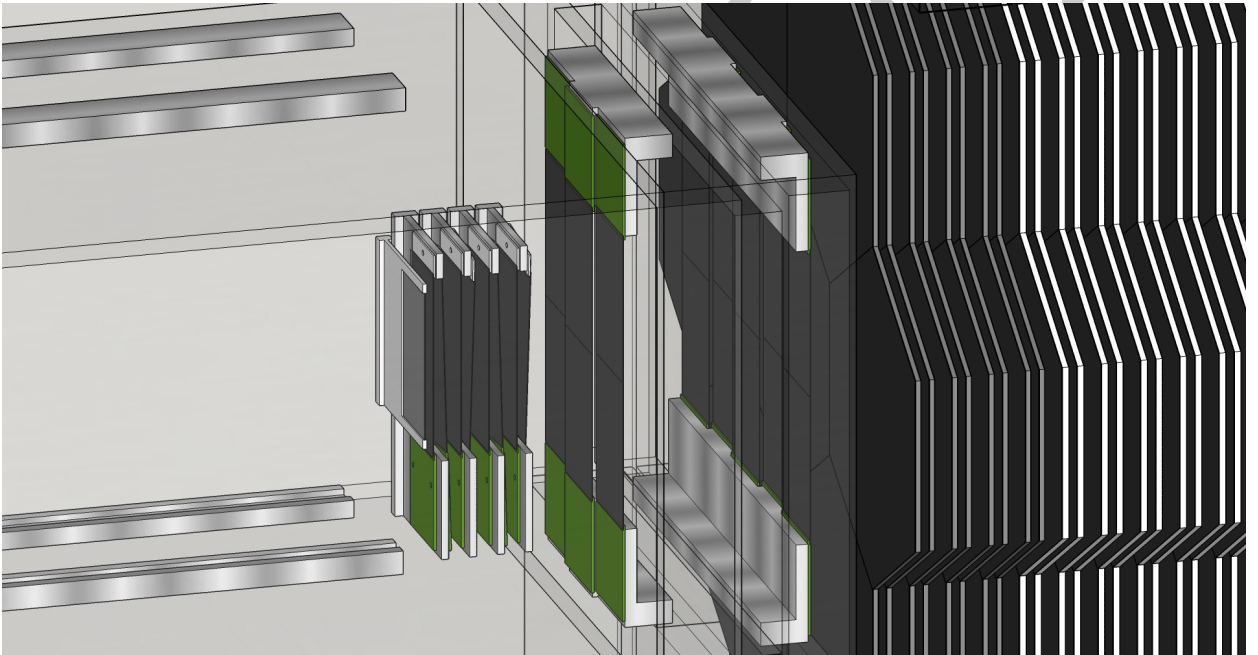

FIG. 20: The target and recoil tracker.

glued end-to-end, with an APV25-based FR4 hybrid circuit board at each end of this structure to read out the two sensors. The sensors are standard p^+- in - n silicon microstrip sensors, but are somewhat shorter and wider than those used in the stereo modules and therefore require six APV25 chips to read out each sensor instead of five. The modules are supported at both ends by screw attachment of the hybrids to castellated support blocks attached to the cooled support structure.

C. Forward Electromagnetic Calorimeter

A high granularity, Si-W sampling calorimeter will be used for the ECal of LDMX. Its components and design draws on the High Granularity Calorimeter (HGC) for the forward calorimeter upgrade of the CMS experiment for the HL-LHC [21]. In particular, the hexagonal sensors, front-end readout electronics and front-end trigger architecture will be the same. This is facilitated by the fact that the the sampling time for the CMS HGC is comparable to that planned for LDMX.

An ECal that is based on silicon pad sensors is well-suited to the LDMX experiment whose

principal task for Phase I, with $4 \cdot 10^{14}$ electrons on target (EOT), will be the identification of photons and electrons with very high efficiency and very good energy resolution. The ECal will also need to play a role in the identification of hadrons. However, a hadronic veto calorimeter, described below, will be used to identify hadrons from photonuclear and electronuclear reactions that pass through the ECal without interactions that produce any identifiable signals. The high granularity of the ECal provides the ability to track charged hadrons that cross multiple silicon layers and to identify isolated charged hadrons that range out in a single layer, depositing charge compatible with what is expected for several minimum ionizing particles (MIP).

For LDMX Phase II, we envision operation at higher energy and higher average number of electrons per 20 ns sampling time. While Phase I is the main topic of this note, we mention Phase II in the context of the ECal because high granularity is needed even more in that higher rate environment and because a silicon based calorimeter has the radiation tolerance needed to sustain the higher fluences.

With a reasonably good transverse spread of beam electrons, the granularity of the apparatus can be used to help identify overlapping showers. The central stack of hexagons would use the highest granularity available. The highest granularity being considered for CMS modules is ∼432 pads of area 0.52 cm^2 on a sensor made from an 8" wafer. The sensor would be read out by 6 HGCROC front end ASICs on a printed circuit board (PCB) that completely covers the sensor. The HGCROC front end readout chip comprises 78 channels (72 reading out standard cells, 2 reading out calibration cells, and 4 channels not connected to any sensor cells for common-mode noise estimation) and is designed in a radiation-hard 130nm CMOS technology. It may be possible to double the pad count to 864, with pads having area 0.26 cm^2, if it is determined that this would bring a significant improvement in performance for operating with multiple electrons per 20 ns sample time. Note that independent of the sensor layout, LDMX would use 500 or 700 μm thick silicon as compared to a maximum of 320 μm thick silicon used by CMS in order to get better energy resolution and enhance the potential to detect hadrons. For sensors with the same granularity as CMS sensors, this could be done without the need for new masks to be designed and fabricated.

As noted above, silicon is chosen to tolerate the fairly substantial radiation doses expected to be encountered in LDMX Phase II, and even in the central stack of the ECal in Phase I. Simulations of the radiation environment performed with FLUKA [22] indicate that, while electromagnetic interactions are the dominant source of irradiation within the LDMX environment, a substantial hadronic component is also present. In all operational scenarios under consideration, a diffuse beam would be used to spread the electrons over a cross-sectional area at the target of at least 16 cm^2.

The effective 1 MeV neutron-equivalent fluence at shower max for a 4 GeV electron beam spread to an area of 32 cm^2 is shown in Fig. 21. For 10^{14} EOT, the fluence is seen to peak at around 3×10^{13} $n/$cm^2. This can be linearly projected to $2.4 - 6 \times 10^{14}$ $n/$cm^2 for $4 - 10 \times 10^{14}$ EOT with a 16 cm^2 beamspot. This is certainly high enough to necessitate changes in the biasing of the silicon and require the silicon to be operated colder than room temperature. It is, however, substantially lower than the worst case expectation for the CMS HGC. The CMS experiment has performed radiation studies for 300, 200 and 120 μm sensors exposed to maximum fluences of roughly 9×10^{14}, 4×10^{15} and 1.5×10^{16} 1 MeV equivalent neutrons per cm^2, respectively. The full depletion voltage of the sensor, V_d, is proportional to the square of the thickness, which can lead to a reduced operational lifetime for thicker sensors in a high radiation environment. In the 120 μm case, the noise is much higher for irradiated sensors, reducing the signal-to-noise ratio to about 2. For the fluences expected in LDMX, the CMS studies suggest that a 432 channel sensor

with thickness of 300 μm would have a signal-to-noise ratio of about 11 for $5 \cdot 10^{14}$ n/cm^2. The radiation environment of the LDMX experiment may, however, be low enough to allow the use of 500 μm thick sensors, which would boost the signal by 67%. Moreover, with 500 μm thick sensors there would be 10 % improvement in the energy resolution of 4 GeV electrons over that obtained with 300 μm sensors. A drawback of thicker sensors is that they increase the space between absorbers, thereby increasing the Molière radius. In any case, the minimum signal-to-noise ratio expected at the end of operation of LDMX, even for 300 μm sensors, would be adequate to identify MIPs.

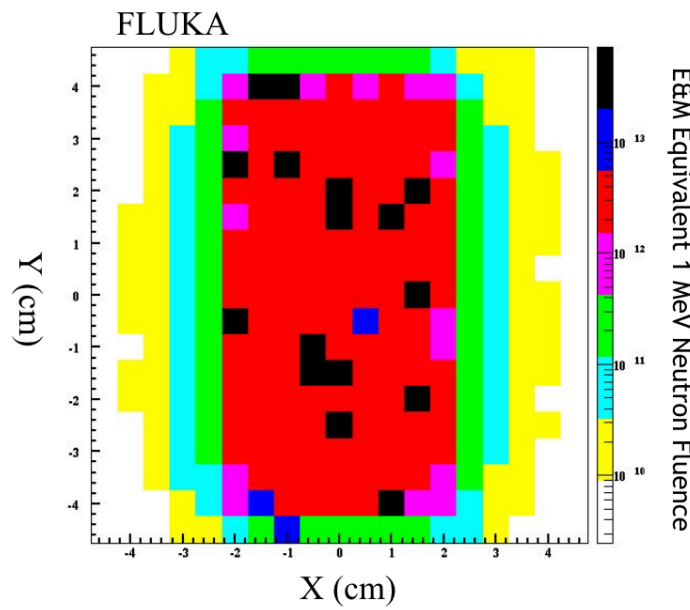
FIG. 21: Effective 1 MeV neutron equivalent fluence at shower max after 10^{14} electrons on target with a beamspot area of 32 cm^2. The maximum effective fluence for this condition is found to be of order 3×10^{13} n/cm^2. Scaling to a total fluence of 4×10^{14} electrons with a beamspot of 16 cm^2, the expected fluence would be 2.4×10^{14} n/cm^2.

Module components and the first completed CMS prototype HGC, a nominal 6-inch, 128 channel module, are shown in Fig. 22. The current design of the ECal used in the studies presented in this note makes use of 32 Silicon layers, each comprised of a central module surrounded by a ring of 6 modules. The hexagonal sensors are to be produced on '8-inch' wafers, yielding a span of 170 mm between parallel sides. The basic ECal readout unit is a hexagonal module containing a single hexagonal sensor. The module has a metal baseplate that serves as electromagnetic absorber, and so has been chosen to be predominantly tungsten. It must also transport heat from the module electronics and sensor to a cooling/support plane. In order to enhance thermal transport without introducing an unacceptably large mis-match in thermal expansion coefficients between the silicon and the baseplate, CMS has chosen an alloy of 25% Cu and 75% W. In LDMX we have considered a pure W baseplate for the module, which however could be affixed to a Cu cooling/support plane similar to the CMS design.

Note that in the module shown in Fig. 22, the corners of the sensor and PCB have been removed to give access to mounting holes on the baseplate that are used to position and attach the modules
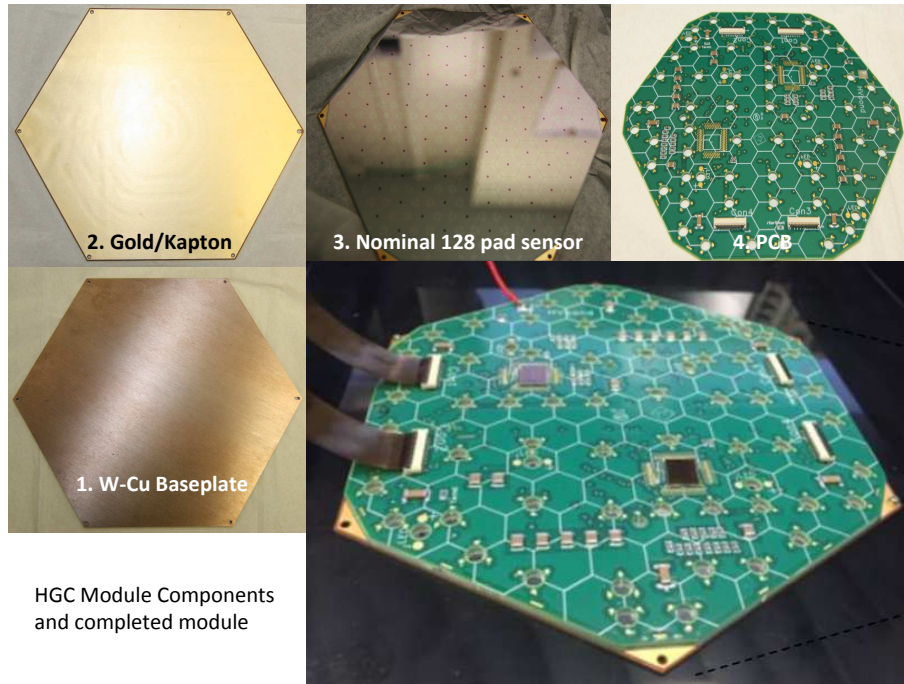
FIG. 22: Prototype CMS HGC 6 inch, 128 channel module with (1) a W/Cu (75%/25%) baseplate, to which (2) a thin Gold-on-Kapton sheet is glued for insulation and to enable biasing of (3) a Si sensor glued directly on top of it. Finally, (4) a PCB is glued on the sensor. Connections to each set of 3 adjacent pads on the sensor are made by wirebonds through 4 mm diameter holes.

precisely on the cooling plate. In CMS the cooling surface is a large 5 mm–thick planar Cu sheet with embedded CO_2 cooling lines, where the Cu contributes to the absorber material budget. For LDMX we are planning to use a simpler, liquid cooling system, since the dark currents and thermal load are lower in the less severe radiation environment. We have begun to study ECal performance with a very thin and lightweight cooling layer, made up of small diameter, thin-walled stainless steel cooling lines embedded in carbon fiber. A lightweight cooling plane could be useful in the upstream layers of the ECal where we find that it is beneficial for the rejection of some relatively rare types of photo-nuclear interactions to start with a thin absorber, increasing the thickness to a nominal value near 1 X_0 over some 5 layers or so.

We first simulated ECal and carried out performance studies for a system with 42 layers of silicon for which the first 10 layers have absorbers of $0.8X_0$ and the remaining 32 layers having absorbers of $1.0X_0$. We chose this geometry to allow us to understand shower containment and energy resolution very well, and we have found that at least for a 4 GeV electron beam, we can manage with fewer sensing layers. We have therefore begun studies and estimated costs for a 32 layer, 40 X_0 device. This would have 224 installed silicon modules. The active region of the ECal would then be roughly 30 cm in depth and about 51 cm wide.

For the electronics, we plan to use the CMS HGCROC front-end chip and PCB design. The CMS development schedule is such that the final components will be produced in substantial quantities in time for the construction of the LDMX detector. The HGCROC, for instance, is expected to be ready for mass-production by the end of 2019. The PCB should be finalized even earlier, based on prototype, pin-compatible front-end chips. For the DAQ and trigger, each plane of modules will be connected to a motherboard housing trigger and readout/control FPGAs as

discussed in Section IV F.

The cooling planes would be stacked in an enclosed and thermally insulated box through which dry nitrogen gas or dry air would flow. The support planes with modules attached would be slotted into a box-like frame designed to allow for some level of reconfiguration of absorber and sensing layers. A simple approach with some of these same characteristics has been used for the CMS test-beam studies of the first HGC modules at FNAL and CERN. The box and support structure could be relatively simple and constructed from hard aluminum or stainless steel. Feed-throughs would be needed for the data and trigger fibers, the power and bias cables, and the cooling and gas lines. Fortunately, the density of services is relatively low in each plane. With appropriate care and attention in assembly and installation, we believe it will be possible to achieve a very good thermal and gas enclosure.

D. Hadronic Veto System

As the forward electromagnetic calorimeter contains most, but not all, of the energy of electromagnetic showers, and does not efficiently detect neutral hadrons or muons, a high efficiency hadronic veto calorimeter (HCal) system is an important component of LDMX. The main functions of the hadronic calorimeter are to detect neutral hadrons, mainly neutrons produced in photonuclear reactions in the ECal or the target with very high efficiency, to be sensitive to electromagnetic showers that escape the ECal, and to detect minimum-ionizing particles (MIPs) such as muons. A scintillator-based sampling calorimeter with a large number of nuclear interaction lengths of steel absorber meets these requirements.

Based on our studies of the backgrounds from hadronic processes, the hadronic veto system must identify neutral hadrons in the energy range from approximately 100 MeV to several GeV with high efficiency. The most problematic events typically contain either a single high energy neutral hadron, or multiple lower energy neutral hadrons. The required efficiency for lower energy neutrons can be achieved with absorber plate sampling thicknesses in the range of 10% to 30% of a strong interaction lengths (λ_A). The sampling fluctuation contribution to the energy resolution of the HCal varies inversely as the square root of the absorber thickness. In order to reduce the probability of a single high energy forward-going neutron to escape without interacting to the required negligible level, a total HCal depth of approximately 16 λ_A of the primary steel absorber is required. We also surround the ECal with a Side HCal in order to intercept neutral hadrons produced at large polar angles.

As the HCal rates are low, transverse segmentation is not a very demanding requirement for a veto, although some complementary measurements, such as studying hadronic systems in electronuclear scattering reactions, require reasonable angular resolution of the hadronic showers. A design with minimal cracks and/or dead spaces is important. We have, accordingly, compared both tile and bar realizations. A bar geometry requires many fewer channels than a tile geometry, and at the shower multiplicities we will face, does not suffer from a significant occupancy problem. We have therefore adopted a bar geometry, in an (x, y) configuration, with ambiguity resolution accomplished by means of timing correlations for signals arriving at the two ends of each bar.

The HCal dimensions and absorber segmentation are being optimized with detailed GEANT4 simulations of both single particle response (see Sec. V G) and various background processes (see Sec. VI C 3). We have explored several geometries with different steel absorber thicknesses, ranging from 2mm to 100mm. Our initial background studies are based on the geometry depicted in Fig. 15. The Main (Side) HCal size was conservatively set to 3m × 3m, enabling us to explore smaller transverse dimensions as well. The Main HCal is 13 λ_A in depth with absorber layers

50 mm thick (0.3 λ_A), which corresponds to a depth of roughly 3 m. The Side HCal design is similar to that of the Main HCal, but with finer sampling ($\sim 0.07\ \lambda_A$, and is $\sim\ 3\lambda_A$ in depth).

The final HCal design will be optimized as we improve our understanding of background processes. For example, the transverse size will likely be less than 3 m, we may vary the sampling fraction with depth, with thinner absorber in the front of the HCal for improved energy resolution, and the total HCal depth can reduced or extended in a modular way based on studies such as those of the photonuclear modeling presented in Sec. V D.

Enabling HCal technologies   The HCal scintillator design is based on that of the Mu2e Cosmic Ray Veto (CRV) system. The scintillator is in the form of doped polystyrene bars, of the type developed for the Mu2e CRV. Extruded polystyrene bars have been produced at Fermilab for several decades. The Mu2e bars are 20 mm thick $\times$ 50 mm wide, co-extruded with an integrated TiO_2 reflector. The extrusion also includes through holes into which a wavelength-shifting fiber is inserted (see Fig. 24 and Fig. 25). The Mu2e system uses two 1.4 mm fibers per bar. The scintillator response to minimum ionizing particles (MIPs) has been measured for 120 GeV protons on a 20 mm thick extruded polystyrene bar, read out with a 1.4 mm Kuraray scintillating fiber and a 2 mm$\times$ 2 mm Hamamatsu SiPM [23]. On average, each MIP at normal incidence deposited 3.9 MeV in a bar, corresponding to a response of 50 photo-electrons/fiber at each bar end (Fig. 23). Based on these measurements, we expect that a 15 mm $\times$ 50 mm bar with a single fiber will yield 75 photoelectrons per fiber per bar (both ends), which should provide an adequate signal for LDMX. In the Side HCal, due to space constraints, the readout will be single-ended. We have already produced prototype bars with a single hole, and we will assume a single fiber per bar in the following discussion. The transverse spatial resolution measured by Mu2e is 15 mm, while the resolution along a 2 m bar, determined by the relative timing of the signals at either end of the bar, has been measured to be 150mm [23]. Slightly longer bars should offer similar performance, adequate for ambiguity resolution in the crossed x, y configuration. We plan to attach the scintillator bars to the steel absorber, and to support this structure by an external frame. In this manner, there is ample access to the sides, top and bottom of the HCal for signal, SiPM bias, and monitoring cabling.

Two existing readout electronics systems providing adequate performance have been identified. The merits of each are currently being explored; we briefly describe them below. The existence of multiple options demonstrates that an HCal readout system is technically feasible. Given the basic dimensions of the HCal, the readout system must be able to service signals from approximately 5,000 wavelength-shifting fibers.

The first system is that already developed for the Mu2e CRV. This could either be adopted as a more-or-less "turn-key" system or with modifications (such as a single fiber per bar) to reduce the number of readout channels. The basic unit is made of four bars glued together to form a 200 mm "quad-bar", containing four individual fibers, which are read out by SiPMs at each end of the bar, mounted on a modified CRV Counter Mother Board (CMB) (see below). Fig. 25 shows a corner of the HCal, detailing the quad-bar and fiber structure. Fig. 26 shows a CRV Counter Mother Board (CMB) mounted on each end of two bars, to readout the four fibers with 2mm$\times$2mm Hamamatsu SiPMs. For LDMX with a single fiber per bar, the CMB would be expanded in length to encompass four bars. The CMB provides bias to the SiPMs, has a temperature monitor, and two flasher LEDs for independent monitoring and calibration of each 50 mm bar. In the event of a SiPM failure, the Counter Mother Boards, which hold the four SiPMs for each quad-bar, can be replaced by removing two mounting screws. The four SiPM signals are transmitted to a Front End Board (FEB) on four shielded twisted pairs via an HDMI-2 cable.
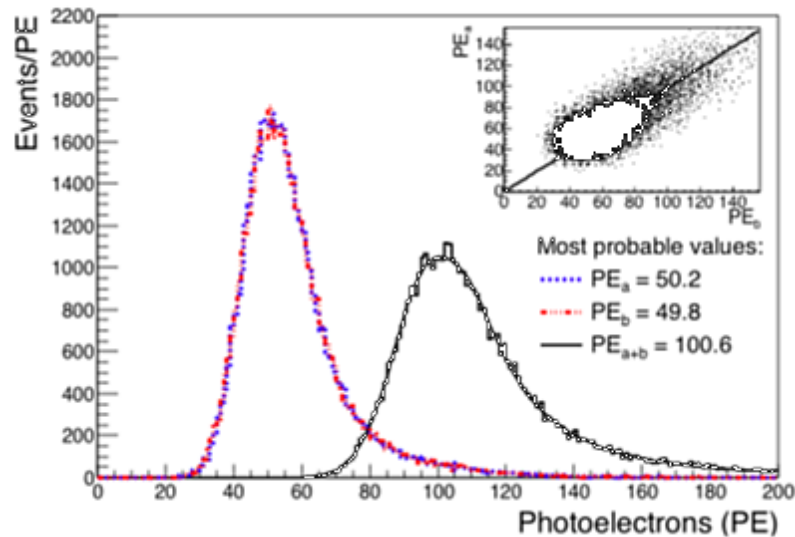
FIG. 23: Measured photoelectron yield for 120 GeV protons on a 2 cm thick extruded polystyrene bar, read out with a 1.4 mm Kuraray scintillating fiber by a 2 mm × 2 mm Hamamatsu SiPM. The beam was 1 m from the end of the bar.
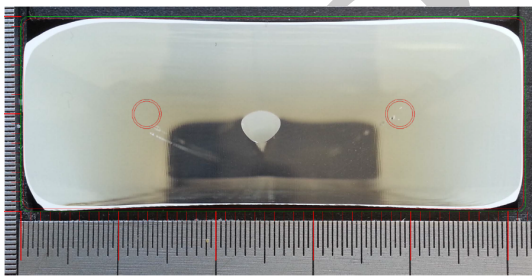

FIG. 24: Photograph of the end of a 20 mm×50 mm extruded polystyrene bar, co-extruded with a TiO_2 diffuse reflecting layer and containing a single hole for a wavelength-shifting fiber. We expect to use a 15 mm × 50 mm bar for the HCal.
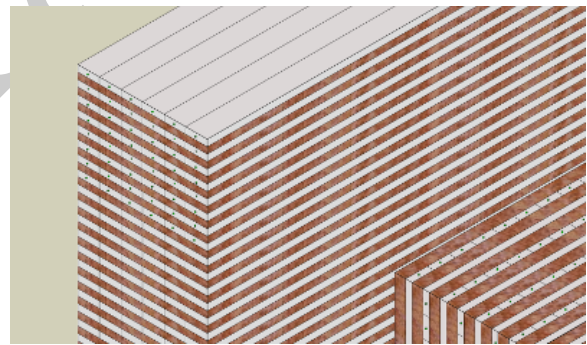

FIG. 25: Detail of the front corner of the HCal, showing the 15mm × 50mm bars, each containing a single wavelength-shifting fiber.

A Front End Board (FEB) services 16 CMBs, digitizing a total of 64 SiPM signals. The Readout Controller (ROC) chassis, which receives the signals from 24 FEBs, also provides the 48 volt bias to the SiPMs and the power to the FEBs, all over a CAT 6 cable. The readout of the bars in the Side HCal is similar to that described for the Main HCal. As the energy of wide angle hadronic showers is lower, the sampling is reduced to 12.5 mm steel and the scintillator bar thickness is also reduced to 15 mm. This necessitates a design with a reduced thickness of the CMB. The rest of the FEB to ROC readout chain is unchanged.

The second readout system is based on the CMS hadronic calorimeter system. In contrast to the Mu2e system, the CMS readout electronics system is a fiber plant scheme where fibers are taken from the scintillator to a centralized SiPM location called a readout box (RBX). The RBX
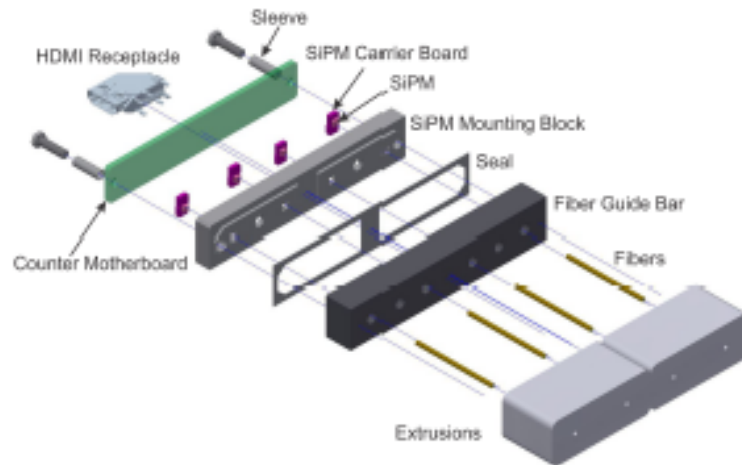
FIG. 26: Exploded view of the Mu2e CRV Counter Mother Board (CMB), SiPM holder and fiber guide structure that's attached by two screws to the end of each 100 mm wide scintillator di-bar. The CMB provides a light-tight seal, holds the fiber ends in registration against the four 2 mm×2 mm SiPMs, and has an HDMI receptacle to bring four channels to the Front End Boards (FEBs). The CMB also has a temperature monitor and two calibration LEDs. For an LDMX with one fiber per bar, the CMB would be modified to service four bars.
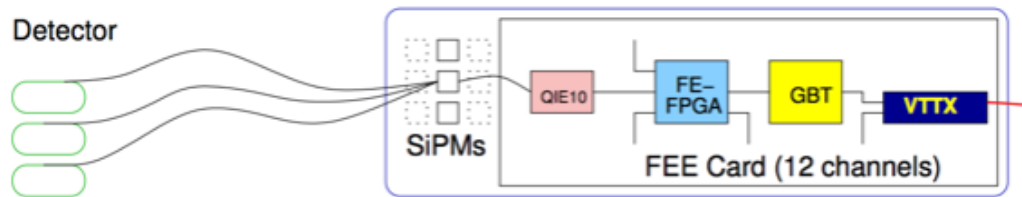

FIG. 27: The CMS HCal electronics readout chain

is described in more detail below. Signals from wavelength-shifting fibers in the scintillating bars are transported to the RBX via clear fiber cables. The CMS system can optically gang up to 6 clear fibers onto a single large area SiPM thus reducing the channel count and effective segmentation. The light transmission efficiency of the wavelength-shifting fiber-to-clear-fiber combination is approximately 75%.

The clear fibers transport the signal from the scintillating fibers to a readout module (RM). An RM consists of an optical decoder unit (ODU) which organizes the clear fibers for ganging and is installed directly onto the SiPM mounting unit. There are 64 SiPMs in a single RM. The SiPM signals are then sent to the QIE board which includes a QIE11 digitizer ASIC that digitizes the SiPM signal, which is then sent to the backend electronics via the CERN VTTX transceiver. Four RMs are contained in one RBX. A schematic of the front-end electronics readout chain is given in Fig. 27. The current CMS HCal system is designed for triggering at a 40 MHz rate.

E. Trigger System

The LDMX trigger system is designed to reduce the typical beam particle arrival rate of 46 MHz to a rate of 5 kHz for storage and analysis. The selection is performed by a combination of dedicated hardware and software running on general-purpose computers.

The first stage trigger is implemented in hardware and allows the selection of both candidate events for dark matter production and important samples for calibration and detector performance monitoring. The overall trigger management is provided by a trigger manager board, which receives inputs from the various triggering subsystems including the ECal and HCal. The latency requirements of the trigger calculation are set by the tracker readout ASIC to 2 μs.

The primary physics trigger is based on the ECal and is designed to select events with energy deposition significantly lower than the full beam energy. The ECal HGCROC calculates the total energy in 2×2 fundamental cells for every 46 MHz bucket. The energy information is transferred by digital data link to the periphery of the calorimeter, where sums are made over larger regions and transferred by optical link to the trigger electronics. The total energy is then used to select the events. The details of the electronics for the ECal trigger are discussed along with the DAQ for the ECal below.

The use of a missing-energy calorimeter trigger requires information on the expected energy, which varies with the number of incoming beam electrons in each accelerator bucket. The trigger scintillator system is designed to provide this information to the trigger calculation. It is important that this system makes an accurate count of the number of electrons. If the actual number of incoming electrons is systematically higher than the number reported by the trigger scintillators, the missing energy will appear to be small (or likely negative), resulting in all events being vetoed. Conversely, if the actual number of incoming electrons is lower than the number reported by the trigger scintillators, the missing energy will appear to be large, resulting in most events being triggered.
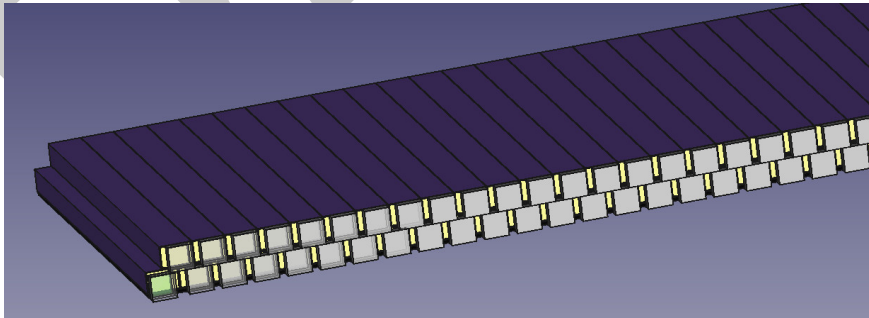

FIG. 28: Drawing of the concept for the target and trigger scintillator system, showing the relationship of the SiPMs to the scintillator strips. Each scintillator strip is wrapped in ESR foil.

The trigger scintillator system is constructed as two planes of thirty $4 \text{ cm} \times 3 \text{ mm} \times 2 \text{ mm}$ strips of scintillator each wrapped with 3M ESR film, which provides a very high reflectivity. Each strip is connected to a SiPM photodetector. The structure is shown in Fig. 28. The strips are used to count the number of incoming electrons. The natural fluctuations in light production and collection in the scintillator imply that it is difficult to identify the number of electrons passing through a given strip. Instead, the logic is simply based on the number of strips above threshold. The strips in a given row have a gap between them set by the ESR film and construction tolerances

to $\approx 200\,\mu$m. The second layer of strips is offset from the first by 1.5 mm so that the gaps in the two layers do not align.

The count of incoming electrons can be lower than the actual number either due to multiple electrons hitting a single strip or by an electron slipping through a dead area. The probability of such an issue occurring in a single layer with twenty-five illuminated 3 mm-width strips increases with the number of incoming electrons, such that it increases from a 7% inefficiency for one initial electron to 40% inefficiency for four initial electrons. The trigger counts the number of incoming electrons using algorithms which combine the information from the two different layers, to improve the efficiency from a single layer. Two different algorithms have been tested. One algorithm simply takes the larger of the number of strips hit in either layer as the number of incoming electrons. The other algorithm attempts to identify patterns which are consistent with multiple electrons per strip by comparing the data in the two layers. The simpler algorithm improves the inefficiency for four electrons to 20%, while the more-complex algorithm provides an additional enhancement to 15%.

The trigger scintillator is read out using an array of 2×2 mm SiPMs connected to QIE11 electronics developed for the CMS HCal upgrade. The SiPMs are coupled directly to the edge of the scintillator strip. The SiPMs and QIE11 readout cards will be located in the environmental enclosure for the tracker along with the tracker front end electronics. The SiPMs will be cooled in common with the tracker elements to achieve very low thermal noise levels in the trigger scintillator readout. The readout electronics will continuously digitize the SiPM signal, providing an integrated charge as well as time-of-arrival measurement for the pulse with an LSB of 500 ps. Both amplitude and timing information can be provided to the trigger, allowing correction of the calorimeter amplitude for time walk effects already at trigger level.

F.   DAQ

The off-detector trigger and readout electronics are assumed to share a single ATCA crate, using the Reconfigurable Cluster Element (RCE) generic computational building block developed by SLAC. The RCE is based on the Xilinx Zynq 7 System-On-Chip technology, running a dual core 1 GHz ARM processor with 1 GB of DDR3 memory tightly integrated with on-chip programmable logic (FPGA). The RCE blocks are integrated into Data Processing Modules (DPM), each of which hosts two RCE blocks. The DPMs receive data from rear transition modules (RTM), which are customized to adapt to the specific portion of the experiment that is being read out. Up to four DPMs can be mounted on a single Cluster-On-Board (COB) ATCA blade, handling up to 96 optical data links. The modular ATCA design permits LDMX to re-use architecture and functionality from other DAQ systems whose components are similar to those used by LDMX.

Overall DAQ Architecture   One of the DPMs functions as the trigger interface, which receives trigger signals from the optical fiber module on the RTM, distributes clock and trigger signals, and handles communication with the software trigger supervisor. Specific interfaces to beamline and infrastructure services are provided through a custom RTM. Final DAQ is carried out through the 10 Gbps Ethernet readout path available separately from each COB. The COBs in a single ATCA crate are also interconnected by 10 Gbps readout and contain ARM processors that can be used to execute high-level trigger algorithms.

To establish a conservative design for the DAQ, we consider a target DAQ bandwidth of 25 kHz (trigger rejection factor of 5×10^{-4}). This provides a factor of five safety compared with the combined expectation for the performance of the hardware trigger and for the data volume in
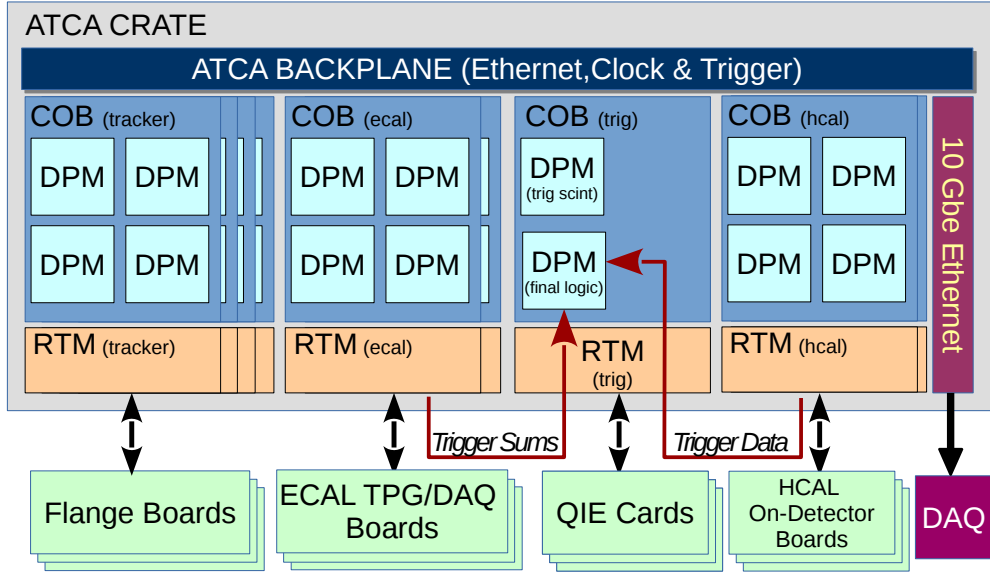
FIG. 29: Schematic block diagram of the RCE-based DAQ.

triggered events. The breakdown of data volumes by subsystem is shown in Table III and sums to an estimated total of 3.1 kilobytes per event, for a total DAQ bandwidth of 77.5 MBps. This bandwidth is within the capability of the 10 Gbps Ethernet bandwidth with a comfortable safety factor of ≈ 10. The 10 Gbps ethernet connections from the COBs will be connected to one or more data storage and data-quality-monitoring PCs. To estimate the potential total data volume, we assume there is no high-level trigger, in which case the full sample of 2×10^{11} events would require 500 TB of storage.

TABLE III: Estimated DAQ data volume per channel.

Subsystem	kbytes/event	Estimate Notes
Trigger	0.5	Includes event accounting information (32B), trigger sums from each module of the ECal (210 B), trigger-counter readout (144B), and readout of trigger information from the HCal (128B).
Tracker	1.0	Each raw hit requires 20B for id and data, and we expect 50 hits/event, including noise hits.
ECal	1.2	Each hit with TDC is 10B including 4B for data and 6B for addressing and overhead, and each (low-amplitude) hit without TDC is 8B. For $\mu = 1$, we expect an average of 125 hits. We assume are all TDC hits for the estimate.
HCal	0.4	Each channel's readout is 20B for id and data, and we budget for 20 channels in a typical event.
Total	3.1	

Tracking DAQ   The data acquisition system for the tracking detectors supports the continuous 46 MHz readout and processing of signals from the silicon-strip sensors. It also selects and transfers those events that were identified by the trigger system to the back-end DAQ for further event processing at rates approaching 50 kHz. The system adopted is an evolution of the DAQ

used for the HPS SVT and consists of APV25 readout ASICs hosted on hybrid circuit boards integrated into the tracking modules, a set of Front End Boards (FEBs) inside the magnet bore for power distribution and signal digitization, flange boards that transmit power and data through the wall of the SVT environmental enclosure, and the remotely located ATCA-based data acquisition hardware and power supplies.

The APV25 ASIC, initially developed for the Compact Muon Solenoid silicon tracker at the Large Hadron Collider at CERN, was chosen because it provides excellent signal to noise, analog output for optimal spatial resolution, and signal pulse-shape sampling capability for good hit time resolution. Each hybrid board has five or six analog output lines (one for each of the APV25 ASICs) that are sent to an FEB using LVDS signals over about 50 cm of twisted-pair magnet wire. At the front-end readout board, a preamplifier scales the APV25 differential current output to match the range of a 14-bit Analog to Digital Converter (ADC). Each front-end board services four hybrid boards. The ADC operates at the system clock frequency of 46 MHz. The digitized output from the front-end board is sent through compact 8-pair mini-SAS cables to flanges on the SVT enclosure to connect to the external DAQ, which resides many meters away. The front-end readout board houses an FPGA and buffers signals to allow for control of the distribution of clock, trigger and I2C communication with the APV25 ASICs. To further simplify the services and minimize cabling that enter through the flanges, it contains linear regulators to distribute and regulate three low voltage power lines to each of the APV25 ASICs, as well as passing through the high voltage bias. Figure 30 shows a schematic layout of this part of the readout chain.
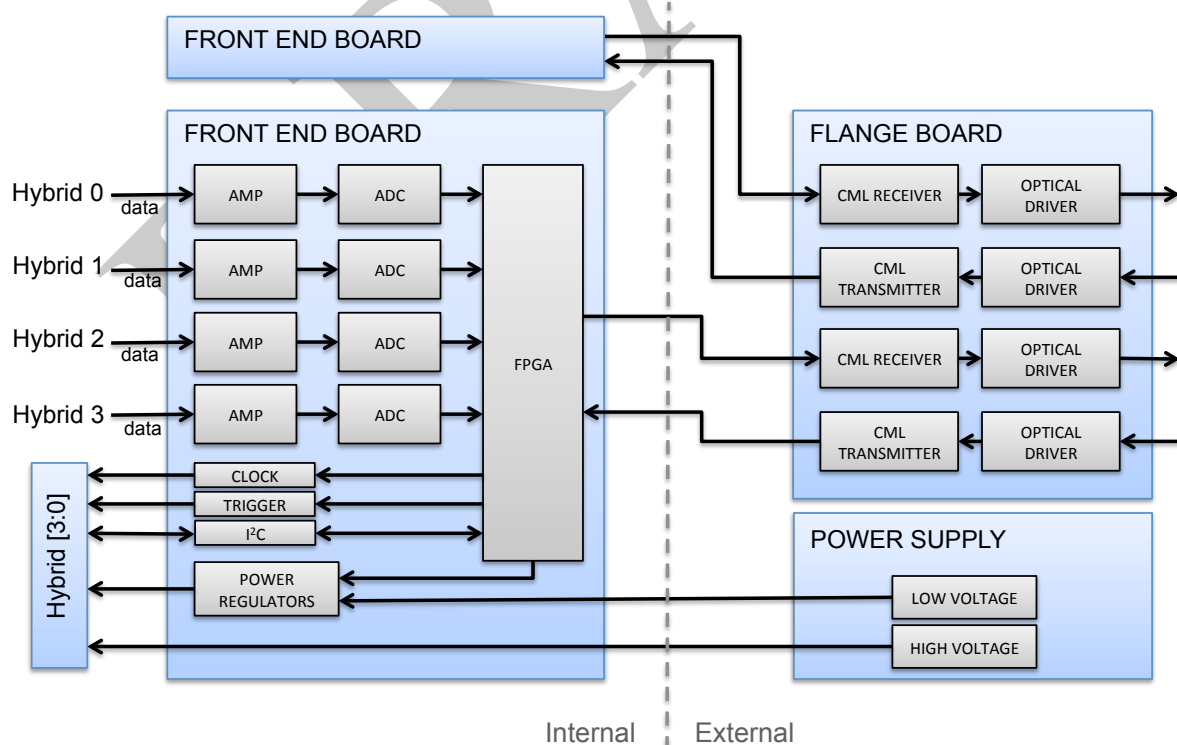

FIG. 30: Schematic view of the downstream part of the tracker DAQ.

The digitized signals are converted to optical signals just outside the flanges on custom-built flange boards. Each flange board houses optical drivers to handle the electrical-optical conversion and to transmit the optical signals over fibers to the ATCA crate. The flange board also interfaces

the low and high voltage power transmission from the Wiener MPOD power supplies to the front-end boards located inside the SVT enclosure.

The optical signals from four hybrid boards on one third of a flange board are received at one of four sections of the Rear Transition Module (RTM) board in the ATCA crate. Each section of the RTM connects to one of four DPMs. In order to minimize the complexity of the system inside the SVT enclosure, all signal processing is done at the DPM level. Each data DPM receives the digitized signals from the RTM, applies thresholds for data reduction and organizes the sample data into Ethernet frames. Four COBs are sufficient to handle all of the hybrid boards of the trackers. The maximum readout rate of the tracking system is approximately 50 kHz, limited by the APV25 readout rate for six-sample, multi-peak readout.

ECal DAQ   As described above, signals in the Si/W ECal are digitized in the HGCROC and transferred on 1.28 Gbps DC-coupled copper data links to the periphery of the calorimeter. At the edge of the detector, the copper links are received by FPGAs and processed before transmission on optical links off the detector. The FPGAs also provide the fast-control and slow-control interfaces to the off-detector electronics. The architecture of the trigger primitive and readout electronics of the ECal is shown in Fig. 31.
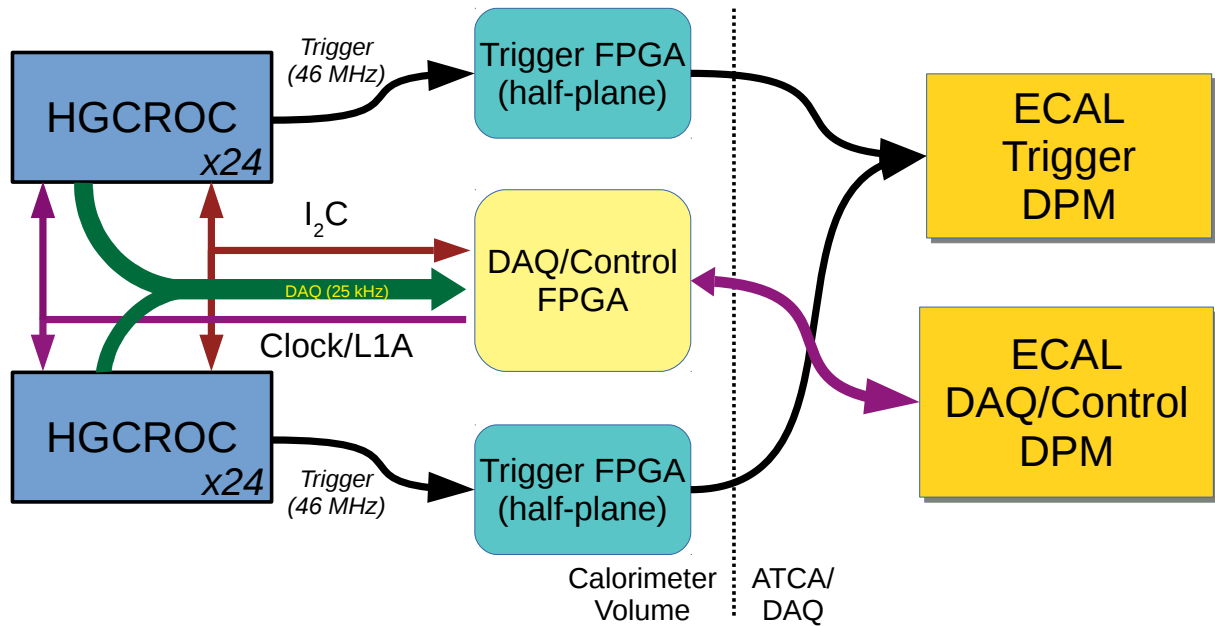

FIG. 31: Schematic view of the trigger and readout for the ECal.

For trigger purposes, groups of nine cells are summed inside the HGCROC to form a trigger cell and the energy for each trigger cell is transferred on the electrical links to an FPGA at the periphery of the ECal. Two links are required for the trigger readout of each HGCROC. An additional copper data link is used for readout of triggered event data. Each module requires six HGCROC chips, resulting in a total of eighteen high-speed links for each module.

The ECal layout includes seven modules on each layer, with the central core module surrounded by six "peripheral" modules. All the modules of a layer are connected to a motherboard PCB using stacking connectors. This PCB also houses three data processing FPGAs, two for trigger and one for DAQ and controls.

For trigger purposes, the plane is divided into two "half-planes" each containing 21 HGCROC chips. The trigger data for a half-plane is transported to a single FPGA on the periphery where summing and other initial local processing is carried out. The trigger data is then transferred by optical link to the ATCA off-detector card where the final trigger calculation is performed. A total of 68 optical links is required for the trigger data, so that the full calculation can be performed in a single COB card.

The 42 DAQ links from the HGCROC on a plane are connected to the DAQ FPGA on the motherboard. The DAQ FPGA is also responsible for controls and uses a bi-directional data link to communicate with the electronics in the ATCA crate, where a second COB will be used for this purpose.

Trigger and HCal Scintillator DAQ   The target scintillators are read out using SiPM phototransducers and QIE11 ASICs developed as part of the CMS HCal Phase 1 upgrade [24]. The QIE11 is a deadtimeless charge-integrating ADC with an internal TDC capable of 500 ps timing resolution. The data from the QIE11 is read out via an Igloo2 FPGA onto a pair of digital optical data links that operate asynchronously at 5 Gbps. The readout of the QIE cards is continous. For every 46 MHz bunch, the data links will digitize and transmit on the fiber the complete data for every channel. Each fiber carries the data for five channels, so the full trigger scintillator DAQ will require twelve fibers. In the ATCA crates, the trigger scintillator readout and control will occupy a single DPM on one COB, which could be shared with another portion of the DAQ if required.

For the hadron calorimeter, we estimate that 1.25 trigger/readout fibers will be required for each layer of the calorimeter in the case of either electronics design under consideration. In all cases, the data bandwidth is dominated by that required for triggering information. Depending on the number of layers in the HCal, either one or two COBs will be required for the HCal readout, where the second COB could be shared with the trigger scintillator readout and the final trigger decision DPM.

G.   Computing

During the design stage, LDMX will be focused on the generation of Phase I Monte Carlo samples. The Phase I MC samples will be on the scale of 300 TB and will take ~30M CPU hours to generate. During Phase I running, LDMX is expected to collect 550 TB of raw data, all of which will require long-term tape storage. Reconstruction of a fraction of the raw data on a daily basis is crucial to monitoring the online performance of the detector and will go a long way in mitigating any issues that may arise. Doing so in a timely manner will require 4k CPU hours per day and 200 dedicated nodes. Reconstruction of the full raw dataset will require 6M CPU hours and 5 PB of disk space. For Phase II running these numbers are expected to increase by a factor of 2-10.

V. SIMULATION AND PERFORMANCE

A. Simulation Framework

The propagation of particles through the detector along with their interactions with material is simulated using the object-oriented LIGT (LDMX Interface to GEANT4 Toolkit) framework. LIGT wraps a custom version of the GEANT4 [25] toolkit (10.02.p03) and adds several services including dedicated process and cross-section biasing capabilities, an event model, ROOT [26] based persistence, GDML [27] based geometry system as well as an analysis/reconstruction pipeline. The framework was designed with flexibility and ease of use in mind, allowing for the quick study of rare physics processes using a wide range of detector concepts. The structure of the simulation is described in Figure 32.
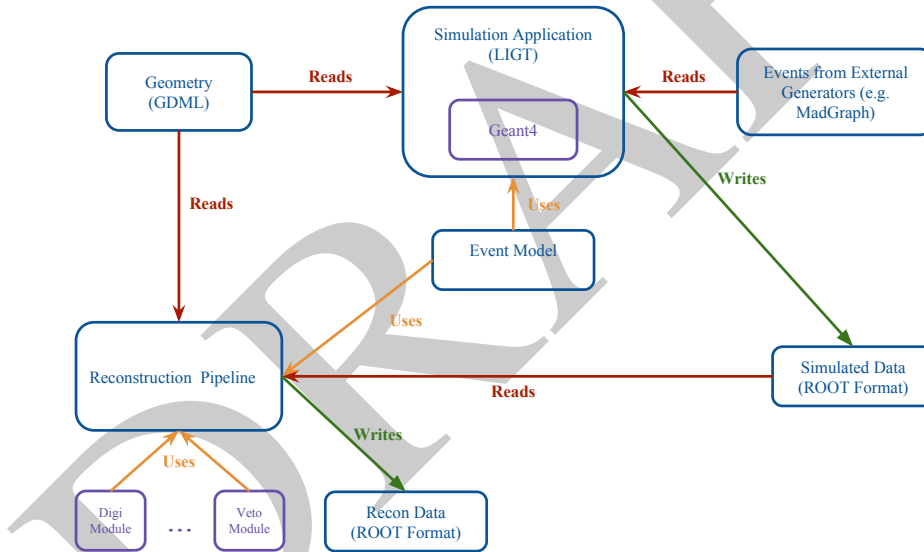

FIG. 32: Diagram showing the structure of the LIGT simulation and reconstruction application.

Event generation is based mainly on sources provided by GEANT4 using the physics list FTFP_BERT augmented by the list G4GammaConversionToMuons . FTFP_BERT contains all standard electromagnetic processes and makes use of the Bertini cascade to model hadronic interactions. By default, the Bertini cascade model is used by GEANT4 for energies below 3.5 GeV. In order to ensure that the Bertini model covered our phase space, the threshold was increase to 10 GeV through a direct modification to GEANT4. In addition to FTFP_BERT, the G4GammaConversionToMuons is used in order to add the additional reaction $\gamma \rightarrow \mu\mu$ to the list of physics processes.

In addition to GEANT4 sources, LIGT can also read events written in the Les Houches Event (LHE) format [28] used by the event generator MadGraph/MadEvent [29]. An XML parser is used to read the event blocks containing the information characterizing each particle. The four-momentum and PDG ID of each of the particles is used in the creation of GEANT4 primary particle objects originating from a common vertex. By default, the vertex is set to (0, 0, 0) but can be uniformly spread over a volume. These particles are then used by the tracking simulation.

A description of the geometry is provided to both GEANT4 and the reconstruction application at runtime using the XML based Geometry Description Markup Language (GDML). Using GDML, all detectors and their corresponding materials can be defined using a simple XML tree structure. Extensions to the GDML format also allow specifying which detectors are "sensitive" as well as the definition of additional detector properties such as unique identifiers and visualization parameters. Furthermore, the same detector description is used to define geometries that can be used by the "parallel worlds" toolkit provided by GEANT4. This is commonly used to define scoring planes which aren't desired to be part of the main geometry. A rendering of the detector being used in the simulation is shown in Figure 33.
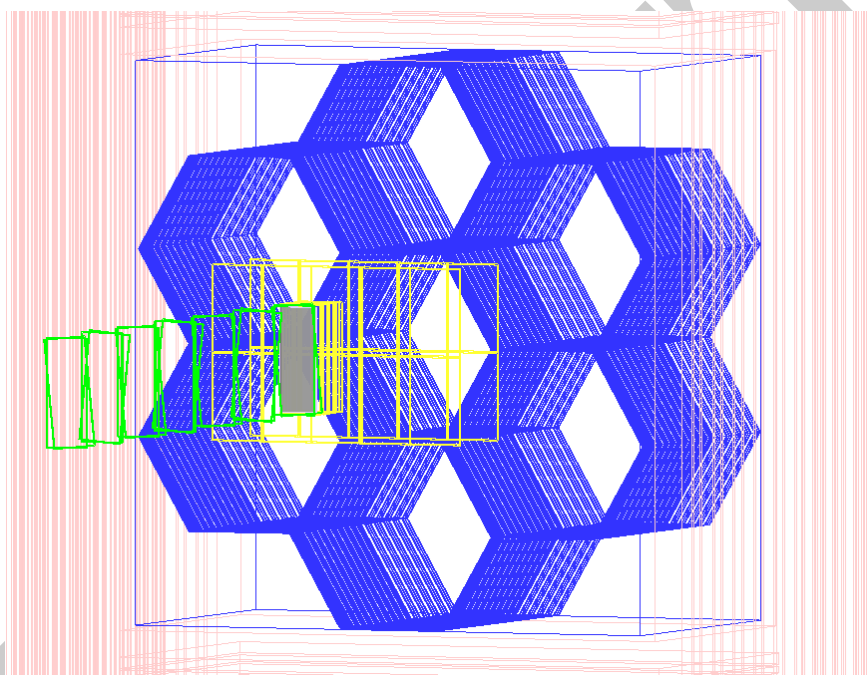

FIG. 33: Rendering of the LDMX detector showing the Tagger (green) and Recoil (yellow) trackers, target (grey), Ecal (blue) and Hcal (pink) as it appears in the simulation and reconstruction application. The trigger pads and magnet are also included but are not shown.

To add additional flexibility to GEANT4, all user actions (e.g. event, stepping) are instantiated via an abstract factory. This allows their initialization to occur at runtime which, in turn, makes the simulation configurable. An important use case for such a design is in the simulation of rare physics processes. A series of "filter" modules implementing the GEANT4 stepping action interface are used to stop the tracking of events that don't satisfy specific criteria. For example, in the case of an electron incident on a target, a filter can be used to select events where the electron bremsstrahlungs and falls below a certain momentum threshold. Another filter can then be use to select event where the bremsstrahlung photon undergoes a photonuclear reaction. Several sets of these modules are used to quickly and efficiently simulate several rare processes including photo/electronuclear.

When a subdetector volume is declared sensitive, if a particle enters it and takes a "step" (i.e. interacts), information associated with that interaction (e.g. energy loss, position) is generated. Subdetector specific processors are then used to convert the information associated with a step into a GEANT4 hit which, in turn, is stored in a container. Once registered with GEANT4, hit containers can be accessed by the rest of the application. At the end of each event, all available

GEANT4 hits are converted to LDMX event model objects and persisted to files using the ROOT data format.

The ROOT based output is processed using a configurable modular reconstruction pipeline with each stage of the pipeline instantiated via an abstract factory. A stage is used to either process the event model hit objects read from the input file or to process physics objects generated by another stage. If a stage generates a collection of physics objects (e.g. tracks), it is made available to all subsequent modules in the stage. Currently, the reconstruction pipeline includes stages that digitize all event model hits, adds noise, performs basic tracking finding and applies various vetoes optimized to reject previously described backgrounds. Once all modules have been executed, the resulting collection of objects is persisted to files using the ROOT data format.

B.   Dark Matter Signal Generation

The new physics of dark matter production is modeled using an external event generator based on MadGraph/MadEvent4. Here we describe this generator, its validation, and the interface with Geant4.

MadGraph is an automated tool for calculating the tree-level amplitudes for arbitrary physics processes, which allows users to define Feynman rules for new physics models; MadEvent is a Monte Carlo event generator based on MadGraph. MadGraph/MadEvent4 (MG/ME) was designed for the study of high-energy collider reactions, but minor modifications to the code (e.g. introducing non-zero masses for incident particles and for the electron, and an electromagnetic form factor as described in [30] for the nucleus) allow for its application to fixed-target processes. These modifications and a new-physics model that introduces a dark photon with arbitrary mass and kinetic mixing ϵ with the photon has been used for the APEX test run [31] and HPS experiment [20]. For LDMX, we have added LDM particles (either fermions or scalars) that couple to the dark photon with an arbitrary interaction strength g_D. This allows us to simulate the signal process of DM particle pair-production via either decay of an on-shell A' or off-shell A' exchange. This report focuses on the on-shell production process, though the kinematics of the two are very similar.

Within MG/ME, we generate events for the DM production process $e^-W \to e^-WA', A' \to \chi\bar{\chi}$ where W represents the tungsten nucleus, χ represents the dark matter particle and $\bar{\chi}$ its antiparticle. Events are generated assuming a 4 GeV incident electron and tungsten nucleus at rest as the initial state. MG/ME computes a Monte Carlo approximation of the inclusive cross-section for this process, and generates a sample of unweighted events in the LHE format. The inclusive cross-section computed by MadGraph is stable within 1% and is consistent within $\sim 30\%$ with independent calculation of the cross-section in the Weizsacker-Williams (WW) approximation from [32]. The deviations from the WW inclusive cross-section are largest at high and low masses, and compatible with the size of errors expected in the WW approximation.

Simulating the detector response of these starts by smearing the vertex position of the recoiling electron and W nucleus uniformly over the thickness of the target and a transverse region spanning ± 1 cm in the x direction and ± 2 cm in the y direction about the nominal center of the target. Both the recoiling electron and nucleus are then tracked through the detector. Although the events also include the A', it is currently ignored by the tracking simulation. For the signal efficiency studies described in Section VII, events are generated assuming different A' masses.

C.   Background Sample Generation

A standard set of background samples has been used for the background rejection studies in Section VI and signal efficiency studies in Section VII. Unilke the signal events, most of these are generated directly in Geant4 10.02.p03, with modifications to the photonuclear, electronuclear, and photon-conversion processes described below to achieve the accuracy needed for LDMX studies.

As discussed in Section III B, the leading backgrounds for LDMX arise from either a rare interaction of a hard bremsstrahlung photon or (more infrequently by a factor of $O(\alpha/\pi)$) a rare hard interaction of the primary electron. It is neither necessary nor feasible to generate Monte Carlo events with the full Phase I statistics of incident electrons. Rather, we use a combination of bremsstrahlung preselection and the Geant4 occurence biasing toolkit to simulate the smaller number of events expected for each of these rare interaction types.

Hard Bremsstrahlung Filter Pre-Selection   Many of the important background processes identified in Section III B begin with a hard bremsstrahlung reaction in the target, where the resulting high-energy photon subsequently undergoes a rare interaction in the target, recoil tracker, or ECal. Events surviving the tracking and ECal selections are those where the electron recoiling from the bremsstrahlung reaction has energy below 1.2 GeV, which we call "hard brem" events. For a 10% X_0 target, "hard brem" events comprise only about 3% of all electron interactions. To simulate them efficiently, we implement a Geant4 filter that propagates the primary electron through the target, putting all secondaries on a waiting stack to minimize the computational expense of simulating non-selected events. To maintain an filter efficient for signals within detector resolutions, events where the electron leaves the target with an energy exceeding 1.5 GeV, does not undergo bremsstrahlung, or stops in the target are killed after minimal simulation time. For surviving events — those where the electron undergoes at least one bremsstrahlung reaction and leaves the target with at most 1.5 GeV energy — the electron track is placed in a waiting stack until other interactions have been simulated. This allows additional selections to be applied on the bremsstrahlung photon and its subsequent interactions, before incurring the computational cost of simulating an electron shower.

Cross-Section Biasing for Photonuclear Processes   The rare photon interactions relevant for LDMX have cross-sections $10^{-3} - 10^{-5}$ times the conversion cross-section. Thus, even with the above optimizations, further techniques to enhance the simulated rate of the rare process are required.

Starting in version 10.0, Geant4 has introduced an occurence biasing toolkit that facilitates simulating particle interactions with a biased interaction law, e.g. an enhanced cross-section for a rare process such as photonuclear interaction or muon conversion. Each event is assigned a weight to account for the effect of cross-section biasing on both the occurence probability of the biased interactions (a factor of $\sigma_{\mathrm{biased}}/\sigma_{\mathrm{physical}}$) and the survival probability of the interacting particle z (a factor of $e^{z/\ell_{\mathrm{physical}} - z/\ell_{\mathrm{biased}}}$, where ℓ is the particle's interaction length).

For LDMX ECal photonuclear studies, we adopt the following biasing prescription:

$$\sigma_{\mathrm{PN}}^{\mathrm{biased}} = B\sigma_{\mathrm{PN}}^{\mathrm{physical}} \qquad \sigma_{\mathrm{conv}}^{\mathrm{biased}} = \sigma_{\mathrm{conv}}^{\mathrm{physical}} - (B-1)\sigma_{\mathrm{PN}}^{\mathrm{physical}}. \tag{7}$$

This biasing scheme guarantees that the *total* interaction cross-section, and therefore the photon's non-interaction probability, is unaffected by the biasing. Thus all photonuclear interactions have a uniform weight $1/B$, and their distribution through the detector volume matches the expected

physical distribution. For photonuclear reactions, a biasing factor $B_{\text{ECal-PN}} = 450$ is chosen so that the biased conversion cross-section is positive in all detector materials, while bringing the biased cross-section in Tungsten within an order of magnitude of the physical conversion cross-section for efficient simulation. To obtain an exclusive photonuclear sample corresponding to N_e electrons on target, we generate events corresponding to N_e/B_{PN} electrons on target with the hard bremsstrahlung filter, cross-section biasing, and a second filter that *requires* a hard photon to undergo a photonuclear reaction. To generate MC statistics corresponding to $N_{\text{EOT}} = 1.2 \times 10^{14}$ electrons on target, we generate biased MC events for 2.6×10^{11} incident electrons, corresponding to a sample of 2.7×10^9 photonuclear events.

For the generation of photonuclear events originating from the target, the biasing description described above is also used with a biasing factor $B_{\text{target-PN}} = 450$. By generating events for 1.3×10^{11} incident electrons, we can model the target-area PN background corresponding to $N_{\text{EOT}} = 2.3 \times 10^{14}$ electrons on target.

Cross-Section Biasing for Electronuclear Processes In generating target-area electronuclear events, event yields are lower than the photonuclear counterpart and accurate modeling of the z distribution is less critical. Therefore, to increase generation efficiency, we adopt the simpler biasing scheme

$$\sigma_{\text{EN}}^{\text{biased}} = B_{\text{target-EN}} \sigma_{\text{EN}}^{\text{physical}} \tag{8}$$

with a biasing factor $B_{target-EN} = 5 \times 10^4$. By generating events for 5.0×10^9 incident electrons, we can model the target-area EN background corresponding to $N_{EOT} = 3.8 \times 10^{14}$ electrons on target.

D. Background Modeling

a. Photonuclear and Electronuclear Modeling As discussed in Sec. III C, background processes of particular concern for LDMX are those where the vast majority of the energy of the incident photon is carried by a small number of particles. It is therefore important to understand not just the overall photonuclear rate, but the yields and kinematics of exclusive final states that are particularly difficult to veto.

This motivates using a photonuclear model based on nuclear interaction data at energies near the $3 - 4$ GeV energy scale crucial for LDMX. At this energy, the interactions of secondaries in a large nucleus can also be significant; these are typically modeled using an intranuclear cascade.

These considerations strongly motivate using the Bertini Cascade model in GEANT4 [33]. This model is based on measured cross-sections for photon-hadron, hadron-hadron, and photon-dihadron reactions within the nucleus, and includes 2-to-9-body final states as well as kaon production. Final-state phase space distributions for $2 \rightarrow 2$ processes are taken from data at low energies, with extrapolations or parametrizations at multi-GeV energies, while multi-body phase space distributions are modeled as uniform. The Bertini Cascade model has been broadly validated and is the default model for photonuclear reactions in Geant with < 3.5 GeV incident photons (with higher-energy reactions modeled using a high-energy string model) and for electronuclear reactions with virtual photon energies < 10 GeV.

In GEANT4, the Bertini cascade model is the default model for photonuclear reactions initiated by < 3.5 GeV incident photons (with higher-energy reactions modeled using a high-energy string model). The model is believed to be accurate up to ~ 10 GeV incident energies [34], and indeed is used in GEANT4 for this purpose for other primary particles. Given the importance for LDMX of

modeling the full physics of the photonuclear final state properly, all simulations have been done using a modified FTFP_BERT physics list, where photonuclear processes with < 10 GeV photons are always modeled using the Bertini cascade.

Electronuclear interactions are closely related to photonuclear reactions, since they proceed at these energies through a virtual photon. They are simulated in Geant4 using the equivalent-photon approximation, with Bertini cascade as the default generator for equivalent photon energies below 10 GeV. In the following discussion, we focus on photonuclear reactions for concreteness but the same considerations (for both physics and simulation) apply to electronuclear processes.

Rare but Important Final States Notwithstanding the validation of the Bertini model, LDMX's sensitivity depends on the rates of sub-dominant processes in small corners of phase space. These have not received as much scrutiny in the validation of the Bertini model as inclusive distributions, and we have found that the inputs to the Bertini model that govern their rates are, in several cases, significantly overestimated relative to data and/or physical considerations. Therefore, we have investigated in detail three classes of event that are design drivers for LDMX: single forward neutrons (which drive the depth needed for the HCal), moderate-angle neutron pairs (which drive the width needed for the back HCal), and hard backscattered hadrons in the "cumulative" region (which set a sensitivity floor, since they are quite difficult to reject).

We have found all three of these event types to be significantly over-populated by the Bertini model, relative to data-driven expectations, though some exclusive processes that contribute are under-populated. We have identified the origin of these discrepancies and they all appear straightforward to correct by means of minor corrections to the input data and/or algorithms used in the Bertini cascade. Discussions with the Geant4 authors and code modifications to this effect are underway.

Comparisons of the default GEANT4 Bertini Cascade output to relevant data are elaborated in Appendix A, as are the origins for these discrepancies in the Bertini Cascade model. Here, we briefly summarize our findings and the implementation of *post hoc* corrections to the photonuclear model in analyses presented here (presently, several of these corrections are applied only to the small post-veto samples considered in Secs. VI C 3 and **??**).

In the case of single- and di-neutron final states, Geant4 vastly overestimates (by a factor of 1000) the rate of two-body dinucleon dissociation by 3–4 GeV photons. This leads to an overestimate of the inclusive single- and di-neutron event yields by factors of 100 and 500 respectively. Although the code modifications required to remedy this are straightforward, improving the accuracy of this simulation requires re-generating the entire event sample with new nuclear cross-section models, and is not currently feasible. Therefore, we have not attempted to correct any studies for this effect, except that when working with sufficiently small post-veto samples, we have down-weighted single- and di-neutron events by factors of 100 and 500 respectively.

The over-population of the "cumulative" region is illustrated, for example, in Figure 34 — the momentum distribution of protons backscattered beyond $100°$ in 5 GeV e-Pb scattering, as extracted from CLAS data [35] is compatible with the theoretically expected exponential distribution, while an unphysical tail is evident in the Bertini Cascade model. . This tail has been traced to an unphysical model of the transitions between regions of the nucleus with different densities — highly energetic hadrons at grazing incidence to the boundary between regions are taken to reflect off this boundary, when they should escape the nucleus in a more physical model. These events have been removed from our sample by a two-fold approach: First, in inclusive studies, events with hard nucleons at $\theta > 100°$ and $w \equiv (p + T)/2(\sqrt{1.25} - 0.5\cos\theta) > 1400$ MeV are downweighted to a uniform exponential, removing the "ankle" in the inclusive distribution and

FIG. 34: Distribution of proton momenta at $\theta_p > 100°$ in 5 GeV electron scattering off ^{208}Pb. The black histogram shows the measured yield in a skim of CLAS eg2 data [35], for events with the electron in the CLAS acceptance. The blue histogram shows the results of a GEANT4 simulation before final-state down-weighting.
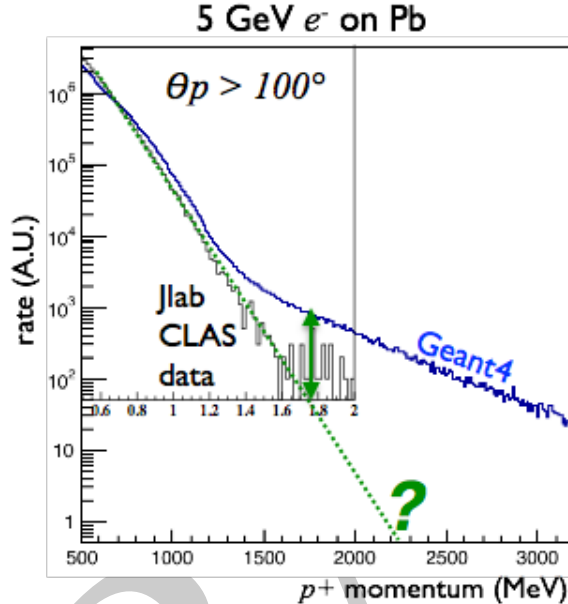
softening this feature in the more backwards phase space (this procedure was adopted before the origin of the tail in Fig. 34 and related distributions was known; the form of the weighting function w is motivated by the phenomenological model of [36]). This weighting does not remove all of the unphysical tail events — for example those with backscattered pions or wide-angle neutrons are not weighted by this procedure. The following step isn't yet reflected in tables! In the analysis of smaller post-HCal-veto samples in Secs. VI C 3 and **??**, we disregard the w-dependent event weights, but instead have resimulated all events that survived the vetoes from their seeds, using a modified GEANT4 version that eliminates the unphysical reflections from the outset. Essentially all of these events, once resimulated, involve forward rather than backward-going hadrons and are easily vetoed by the HCal. These events are, therefore, not included in the event counts.

We also caution the reader that the single- and di-neutron final state yields (and potentially multi-neutron final states, which are likely dominated by cascades of related reactions) are over-estimated in the samples used in this whitepaper. It is expected that the inefficiency of the ECal veto and the size required of the HCal veto to achieve low background will both be lower when more realistic samples are used.

b.   Modeling of Processes with Muons   The dominant source of muon pairs in LDMX is photon conversion to $\mu^+\mu^-$ in coherent scattering off a nucleus (in the target or ECal). This process is included in GEANT4, but was found to have an unphysical form factor. In addition, approximations to the phase-space distribution of outgoing muons were found to be inaccurate on the tails. LDMX dimuon samples were therefore generated with a modified version of the G4PhotonConversionToMuons class, which uses the full dimuon phase space distribution from [37] eqn (2.3) (assuming elastic recoil of the nucleus and keeping only the W_2 component of the result).

To illustrate the effect of the modified photon conversion process, we compare kinematics of

photon conversion to muons in GEANT4, for a monochromatic 4 GeV photon beam incident on a thin target, to a matrix element calculation using MadGraph and MadEvent. In Fig. 35 (left), we show the momentum transfer q of the photon conversion process where $Q = p_{\mu+} + p_{\mu-} - p_\gamma$ and p denotes the three-momentum of the incoming photon and outgoing muon. Different colored lines are shown for the three different models of the coherent process: default GEANT, our custom version of GEANT with improved form factor and phase space integration, and the expectation from MadGraph. At high momentum transfer, Q, the muons tend to decay at higher angle in the lab frame (with less lorentz boost) thus making their detection more challenging. This can be seen in Fig. 35 (right) which illustrates the polar angle of the muon.

We also show the expectation from the incoherent process (also simulated using MadGraph and MadEvent in magenta which has an overall smaller cross-section than the coherent process and therefore, its normalization is plotted relative to the coherent process. The incoherent-conversion process dominates at large Q, where the nuclear form factor suppresses the coherent process. It can contribute an $O(1)$ correction to the phase space involving wide-angle muons.
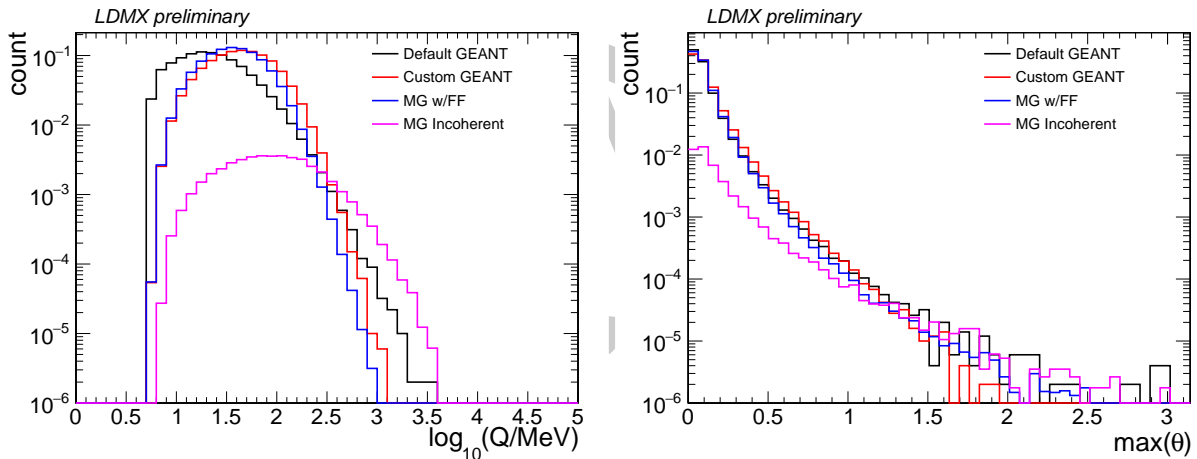

FIG. 35: Kinematics of background coherent photoconversion to muons using default GEANT, custom GEANT, MadGraph with a form factor as given in [37], and the incoherent process as well. On the left is q, the momentum transfer and on the right is the maximum polar angle of the two muons.

E. Tracking System Simulation and Performance

1. Tagging Tracker

Two methods are used to benchmark the performance of the tagging tracker and better understand the constraints it provides in preventing mis-tags. First, the momentum and impact parameter resolutions at the target are determined using an analytic model of the tagging tracker that includes the effect of intrinsic resolutions and multiple scattering in the tracker planes. Second, full simulation is used to confirm these resolutions and understand reconstruction efficiencies and susceptibility to background.

For incoming 4 GeV electrons, the presence of seven three-dimensional measurements of the trajectories ensures near perfect tracking efficiency and a very low rate of fake tracks. The ability

of the tagging tracker to reject beam impurities is discussed further in Section VI. The analytic model finds a longitudinal momentum resolution of approximately 1% and the corresponding full simulation results show good general agreement, as shown in Fig. 36. The transverse momentum resolutions are found to be 1.0 MeV and 1.4 MeV in the horizontal and vertical directions, respectively, which are small compared to the 4 MeV smearing in transverse momentum from multiple scattering in the $10\%X_0$ target. Meanwhile, the impact parameter resolution for 4 GeV electrons is expected to be approximately 7 μm (48 μm) in the horizontal(vertical) direction. Again, the full simulation shows good general agreement, as shown in Fig. 37. These results indicate that tight requirements can be made in both the energy and trajectory at the target, which respectively serve to clearly identify signal events and prevent accidental mis-association between incoming particles and tracks reconstructed in the recoil tracker.
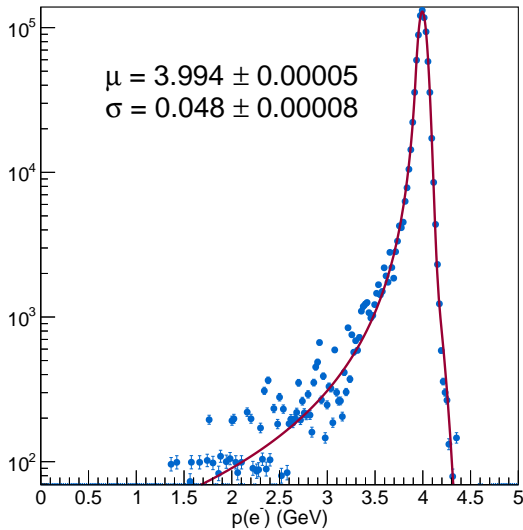

FIG. 36: The longitudinal momentum reconstructed by the tagging tracker for a sample of 4 GeV beam electrons. Excellent momentum resolution allows tight selection against any off-energy component of the beam.
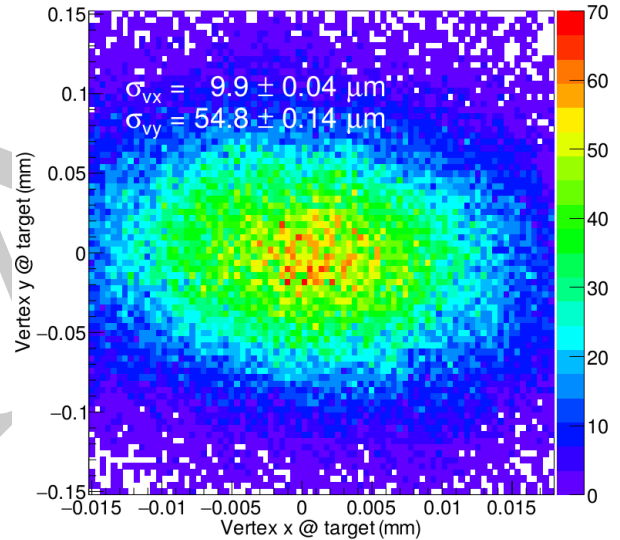

FIG. 37: The $x - y$ position at the target reconstructed by the tagging tracker for a sample of 4 GeV beam electrons. Excellent impact parameter resolution provides a strong constraint on a matching recoil track.

In summary, the design of the tagging tracker appears robust enough to provide unambiguous tagging of incoming electrons with the nominal beam energy for Phase I of LDMX. While the system can clearly deal with higher occupancies, further study will be required to find the beam intensity limits for Phase II.

2. Recoil Tracker Performance

The recoil tracker must have a large acceptance for recoiling electrons characteristic of signal events with good resolution for transverse momentum and impact position at the target, both of which are critical for unambiguously associating those recoils with incoming electrons identified by the tagging tracker. While good reconstruction efficiency for signal recoils is important, it is even more important to have good efficiency for charged tracks over the largest possible acceptance

to help the calorimeter veto background events with additional charged particles in the final state. In addition, the tracker should have sufficient momentum resolution that it can assist the ECal in identifying events in which an incoming beam electron passes through the target and tracker without significant energy loss. Finally, the recoil tracker can help identify events in which a bremsstrahlung photon generated after the end of the tagging tracker undergoes a photonuclear reaction in the target or tracker material. As with the tagging tracker, an analytic model of the recoil tracker is used to estimate momentum and impact parameter resolutions, while full simulation and reconstruction are used to confirm those resolutions, to estimate signal acceptances and tracking efficiencies, and to test the ability of the recoil tracker to reject backgrounds. Resolutions will be discussed further here while acceptance and efficiency for signal events will be discussed in Section VII and background rejection will be discussed in Section VI.

Measurement of the position and transverse recoil momentum at the target requires precise determination of the track angle at the target. For the transverse momentum, a good curvature measurement is also required in order to set the overall momentum scale. At least two 3-d measurements directly downstream of the target are needed to determine the recoil angle, and at least one additional bend-plane measurement is needed for curvature. For low-momentum tracks, it is sufficient for the third measurement to be in the first four closely-spaced layers, but for high momentum tracks that are nearly straight, hits in both of the downstream axial layers are necessary for good momentum resolution.

The ability to distinguish signal from background using the recoil transverse momentum is obviously limited by the multiple scattering in the target, where multiple scattering in a 10% X_0 target results in a 4 MeV smearing in transverse momentum. Using the analytic model of the recoil tracker, the material budget and single-hit resolutions were designed so that the transverse momentum resolution is limited by multiple scattering in the target over the momentum range for signal recoils. This has been verified in full simulation, as shown in Fig. 38. Meanwhile, the impact parameter resolution, shown in Fig. 39, strongly constrains the phase space for mis-reconstructed tracks to point to the same location in the target as the incoming track reconstructed in the tagging tracker.

F.   ECAL Simulation and Performace

The simulation of the ECal uses the GEANT4 framework coupled with an electronics simulation that converts the energy deposits in the silicon into simulated digital hits. Noise is introduced at the digitization stage using a realistic model based on tests performed with the CMS HGCROC. The noise RMS depends on the properties of the readout pad and is linearly dependent on its capacitance, which is calculated to be 27.56 pF. The RMS is estimated in this model to be approximately 1506 electrons. Under the assumptions that a MIP produces about 33000 electrons in ~ 500 μm of silicon, and a response of 0.13 MeV per MIP, this translates to a noise RMS of approximately 0.006 MeV. A readout threshold of 4 times the noise RMS is found to be reasonably conservative, resulting in sufficiently few noise hits being recorded without adversely affecting the probability of reading out genuine hits.

A simulated interaction of a 4 GeV electron in LDMX is shown in Fig. 40. A typical 4 GeV electron event results in O(800) MIPs being observed in the calorimeter cells.

The ECal is responsible for measuring electrons and photons with extremely high efficiency and good energy resolution. This includes inclusive, primarily electromagnetic, showers, but also rarer photon and electron interactions such as photon conversion to a muon pair or hadronic, photonuclear processes. The power and versatility of the ECal to detect this wide range of signatures
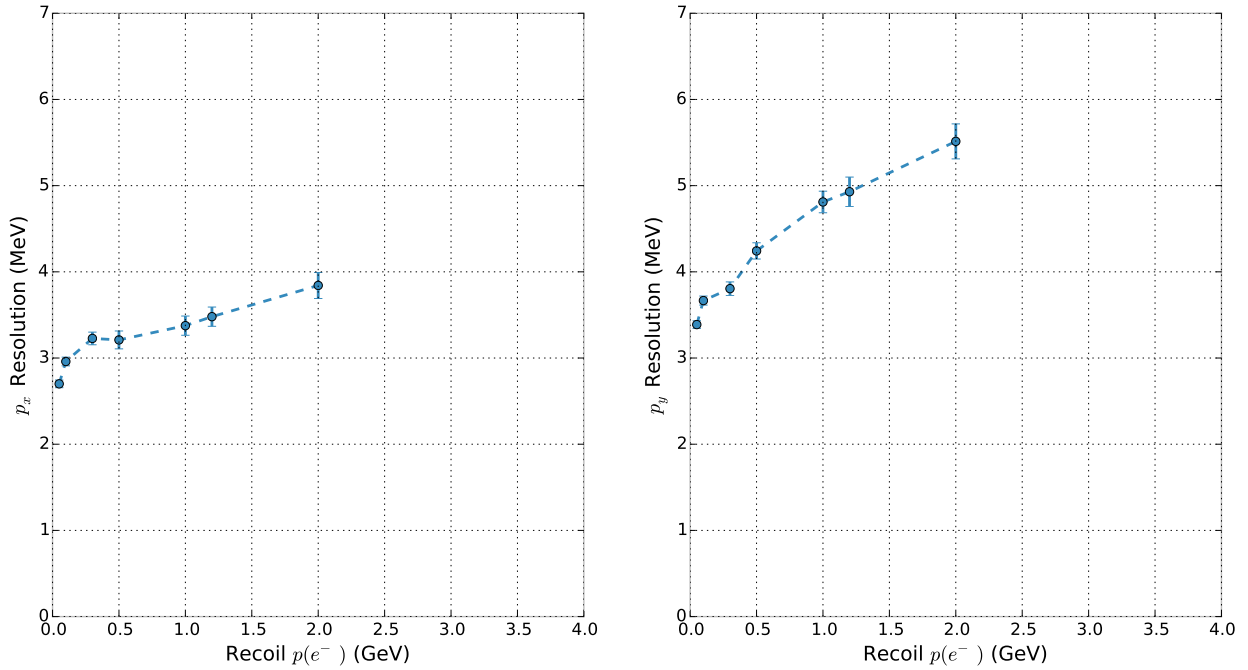
FIG. 38: Resolutions on the components of the momentum transverse to the target for signal recoils. Simulated recoils are assumed to originate at a random depth in the target so that the average resolution is less than the 4 MeV smearing from multiple scattering in the full target thickness. Only the vertical component of the recoil momentum shows significant degradation due to detector resolution at the highest recoil momenta considered.
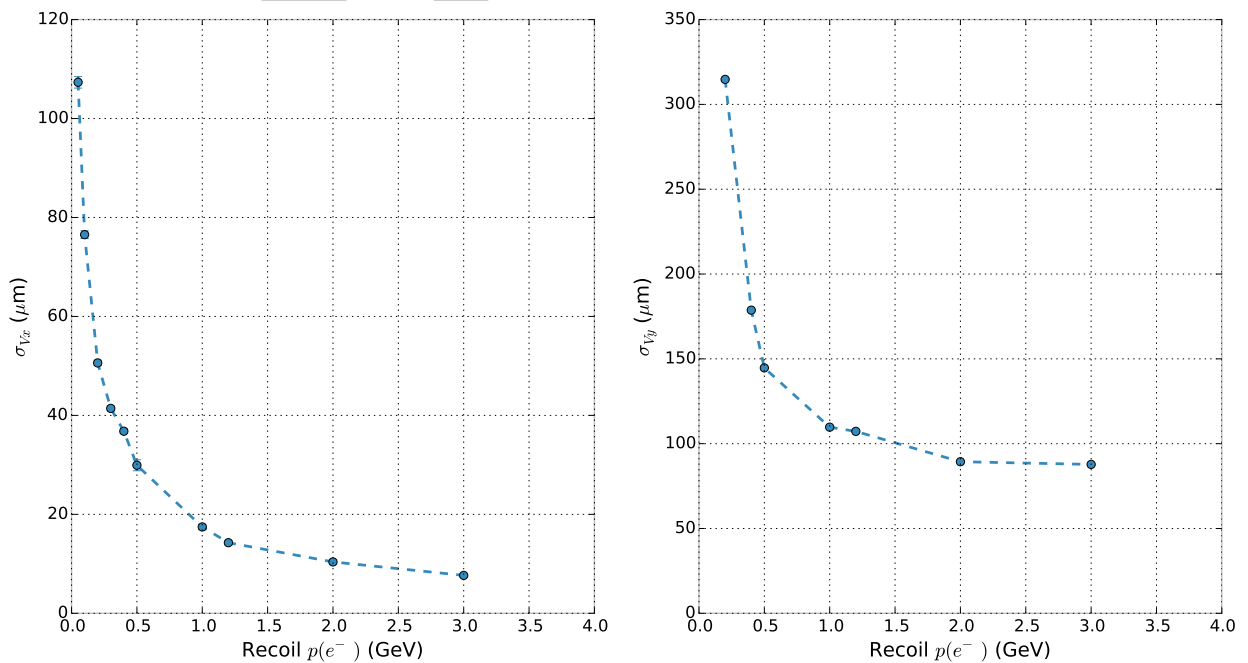

FIG. 39: Recoil tracker impact parameter resolutions. Requiring a common impact position at the target with both the tagging and recoil tracker strongly selects against associating mis-reconstructed recoils with a tagged incoming electron.
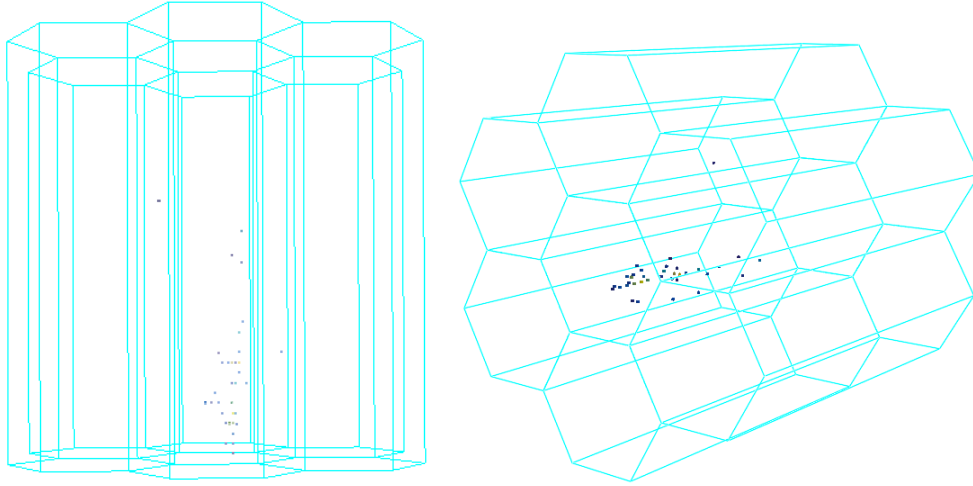
FIG. 40: An event display for a GEANT4 simulation of a 4 GeV electron in the LDMX detector.

is a strength of the proposed detector technology.

a. Inclusive shower Reconstructed electromagnetic shower energy is the most powerful discriminant between signal and the most common, purely electromagnetic bremsstrahlung background. As shown in Fig. 9, applying an upper bound on the energy deposited in the ECal is sufficient to reduce this background to negligible levels.

b. Photonuclear interactions Rarer background processes can cause small, irregular, or negligible energy depositions in the ECal, possibly leading to a fake signal signature. In particular, photonuclear processes, in which the photon interaction results in the production of hadrons in the ECal, may have this effect.

The optimization strategy discussed later in this section primarily focuses on events in which the recoil electron will hit the ECal. However, events in which the recoil electron is emitted at large angles and misses the ECal represent a very interesting category with a distinct experimental signature. For signal events in this category, very little energy is deposited in the ECal, with most of the energy coming from pure noise and soft photons that the recoil electron may have produced as it approached the ECal. In contrast, for photonuclear processes the photon can deposit a substantial amount of energy in the ECal. Event displays for a signal event with $m_{A'} = 0.1$ GeV (left) and a photonuclear event (right) in which the recoil electron misses the ECal are shown in Fig. 41 to illustrate this difference. The distinction between the two classes of events, with the recoil electron missing or hitting the ECal, respectively, is illustrated in Fig. 42, which shows in each case distributions of the total energy deposited in the ECal, computed as the energy sum of all hits above the readout threshold.

The definition of the ECal fiducial region used to determine the classification of events in this scheme is illustrated in Fig. 43. Tracking information is used to determine where the recoil electron is expected to hit the ECal face. If that position is within 5 mm of any ECal cell center, shown in red and white in Fig. 43, the recoil electron is considered to be inside the fiducial volume. Smaller radii were empirically found to be too small due to the presence of dead spots where electrons could hit the fiducial region but still be misclassified as missing the ECal.

The recoil electron is found to miss the ECal in a significant fraction of signal events. However
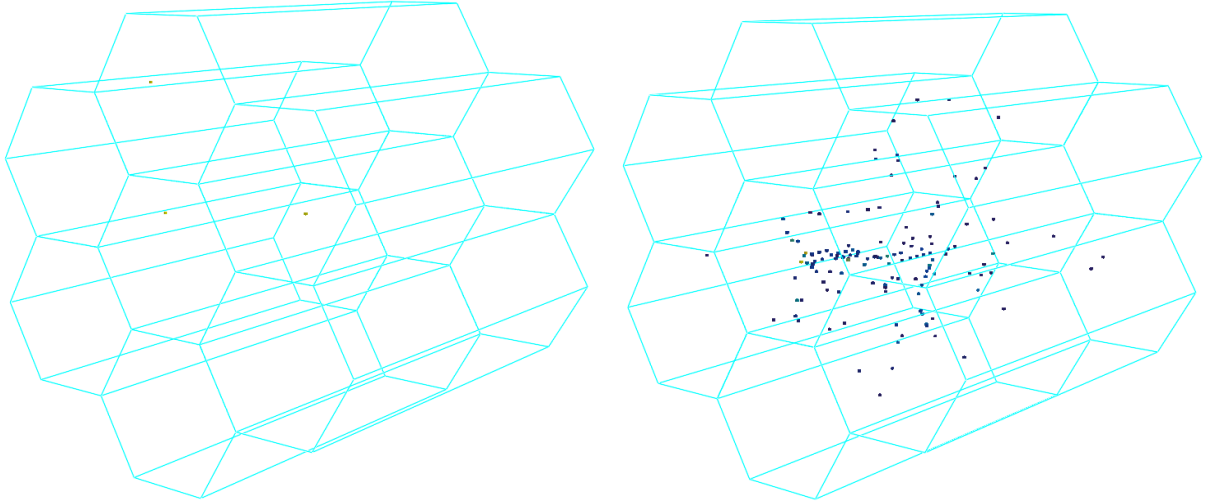
FIG. 41: Event display for a $M_A = 0.1$ GeV signal event (left) and a photonuclear event (right) where in both cases the recoil electron misses the ECal.
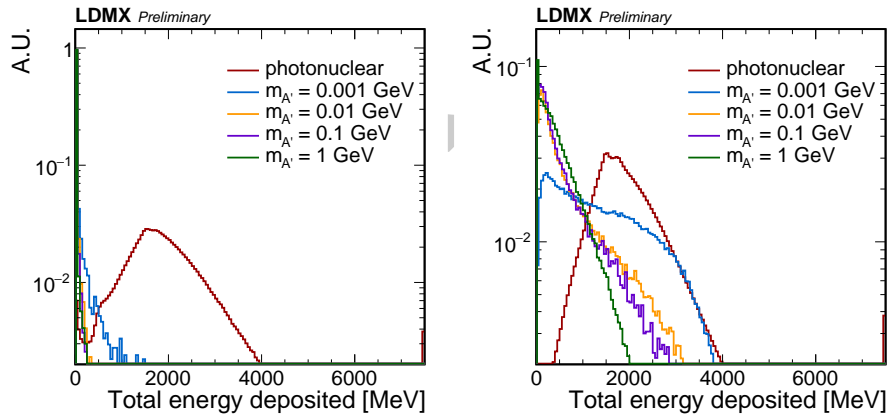

FIG. 42: Total energy deposited in the ECal for events in which the recoil electron misses the ECal (left) or is within the ECal (right), for the photonuclear background and for signal processes with four different mediator masses. All distributions are normalized to unit area for comparison. For all distributions shown in this subsection, the last bin includes the overflow.

this fraction is reduced when we require the presence of a recoil electron identified from tracking information, which is required to have energy below 1.2 GeV. The fraction of events where the recoil electron missed the ECal for signal processes with four different mediator masses and for photonuclear processes are shown in Table IV for events in two categories: all events, events passing the trigger, and events passing the trigger that additionally have a reconstructed recoil electron with energy below 1.2 GeV. One apparent trend in the table is that the probability of
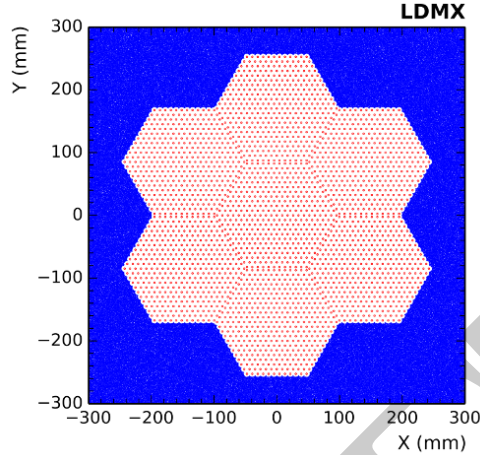
FIG. 43: Two million points outside of the fiducial region were generated from a uniform distribution and are shown in blue. The hex cell centers are shown in red, with any areas that are white or red corresponding to the fiducial region.

the recoil electron missing the ECal decreases as mediator mass decreases. The likelihood that the recoil electron misses the ECal is largely determined by the kinematics of the mediator. As the mediator mass decreases the recoil electron is likely to carry more energy and be deflected at smaller angles, so the recoil electron is less likely to miss the ECal.

TABLE IV: Fraction of events after sequential selection requirements in which the recoil electron misses the ECal for signal processes of four different mediator masses and for photonuclear processes.

Selection	1.0 GeV A'	0.1 GeV A'	0.01 GeV A'	0.001 GeV A'	Photonuclear
Trigger	0.38	0.36	0.25	0.07	0.07
Recoil energy < 1.2 GeV	0.12	0.08	0.03	0.01	0.01

Several variables, constructed from information read out from the ECal, have been found to be powerful in distinguishing photonuclear processes from signal in the ECal. The most powerful discriminating variables are related to the energy distributions in the ECal. The photon and recoil electron both produce hits in the ECal in photonuclear interactions. This motivates the use of the variables listed below that capture the increased energy and number of hits expected in photonuclear events.

- Number of readout hits: the total number of hits in the ECal above the readout threshold.

- Total energy deposited: the energy sum of all hits above the readout threshold.

- Total tight isolated energy deposited: the energy sum of all isolated hits in the inner ring. A hit is isolated if none of its neighbor cells are above the minimum readout threshold, and the inner ring is defined as the cells immediately neighboring the one containing the shower centroid, the energy weighted average x,y position of all hits in an event.

- Highest energy in a single cell: the highest amount of energy deposited in a single cell.

Distributions of these variables for photonuclear and signal processes are shown in Fig. 44. Only events in which the recoil electron is within the ECal fiducial region that pass the trigger are plotted. We define a series of cuts on this set of variables, designed to preserve very high signal efficiency while rejecting a substantial fraction of photonuclear events. Table V lists the signal and background efficiencies corresponding to these cuts, applied in sequence.
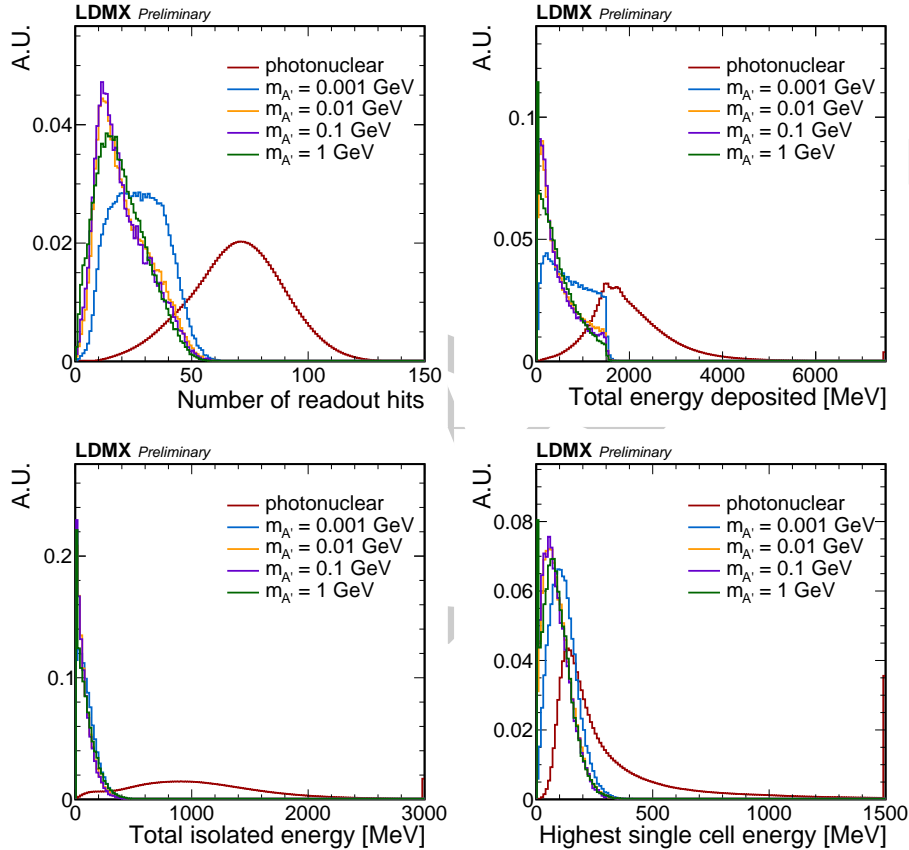

FIG. 44: Distributions of quantities related to the energy deposited in the ECal for photonuclear and signal processes in events passing the trigger in which the recoil electron is within the ECal fiducial region. From top left to bottom right: number of readout hits, total energy deposited, total tight isolated energy deposited, and highest energy in a single cell. All distributions are normalized to unit area.

Further discriminating power between signal and photonuclear processes can be obtained by exploiting information about the shower profile in the longitudinal and transverse directions within the ECal. In the case of these variables, the separation between signal and background distributions is not sufficient to allow us to place a single cut with high signal efficiency as well as high background rejection. However, the shapes of the signal and background distributions are sufficiently distinct to motivate the use of these variables in a multivariate analysis based on a boosted decision tree (BDT), in combination with the energy-related variables described above. The quantities related to the longitudinal and transverse shower profiles that are included in the BDT are described and motivated below.

TABLE V: Signal and background efficiencies corresponding to cuts applied on energy-related ECal observables, following the trigger requirement. The selection requirement listed on each row is applied in sequence to those listed in previous rows. Only events in which the recoil electron is within the ECal fiducial region are considered.

Selection	1.0 GeV A'	0.1 GeV A'	0.01 GeV A'	0.001 GeV A'	Photonuclear
Number of readout hits < 60	1.00	1.00	1.00	0.999	0.311
Total energy deposited < 1500 MeV	0.997	0.994	0.994	0.984	0.208
Total isolated energy < 400 MeV	0.990	0.991	0.990	0.979	0.094
Highest single cell energy < 300 MeV	0.986	0.988	0.986	0.972	0.090

1. Longitudinal shower distributions: In photonuclear interactions, a substantial amount of energy is deposited in the back layers of the ECal, leading to a deeper longitudinal profile for photonuclear events, while the recoil electron produces a shallow string of hits in the front of the ECal, which combined with the photon depositing energy in the back of the ECal, leads to a broader longitudinal profile. The variables listed below characterize the effects of this broader, deeper longitudinal profile.

   - Deepest layer hit: the layer number of the layer closest to the back HCAL that has a readout hit. The ECal layers are assigned integers in the range [0,32], where the 0th layer is closest to the tracker and the 32nd layer is closest to the back HCAL.

   - Average layer hit: the energy weighted average of the layer numbers corresponding to all readout hits in an event, where each hit has its layer number weighted individually by the energy of the hit.

   - Standard deviation of layers hit: the energy weighted standard deviation of layer numbers for all hits in an event. Similar to the average layer hit, each individual hit has its layer number weighted by the energy of the hit.

2. Transverse shower distributions: In photonuclear interactions, the shower resulting from the photon is found to also have a much wider transverse profile than the recoil electron. The separation of the photon and recoil electron from the magnetic field also contributes to the wider transverse profile of photonuclear events. The variables listed below characterize the broad transverse profile of photonuclear events.

   - Transverse RMS: a two dimensional, energy weighted RMS centered on the shower centroid.

   - X and Y standard deviations: the energy weighted standard deviations for the x and y positions of all hits in an event. As before, the x and y positions of each hit are individually weighted by the energy of the hit.

The BDT is trained against photonuclear events; the signal sample used for training corresponds to a mixture of events simulated with four different mediator masses (0.001 GeV, 0.01 GeV, 0.1 GeV, and 1 GeV). Figure 45 shows the distribution of the BDT discriminator value for signal and background events after requiring that they pass the trigger and have a recoil electron that is within the ECal. The ROC curves showing the signal efficiencies for different mediator masses as a function of the background efficiency corresponding to different BDT thresholds are plotted in Fig. 46. A BDT threshold of 0.94 corresponds to a rejection of 99% of photonuclear events in

the fiducial sample, for signal efficiencies ranging between 70 and $\sim 80\%$. By applying a more stringent BDT cut, we can achieve background efficiencies at the sub-percent level, while still retaining reasonably high signal efficiencies.
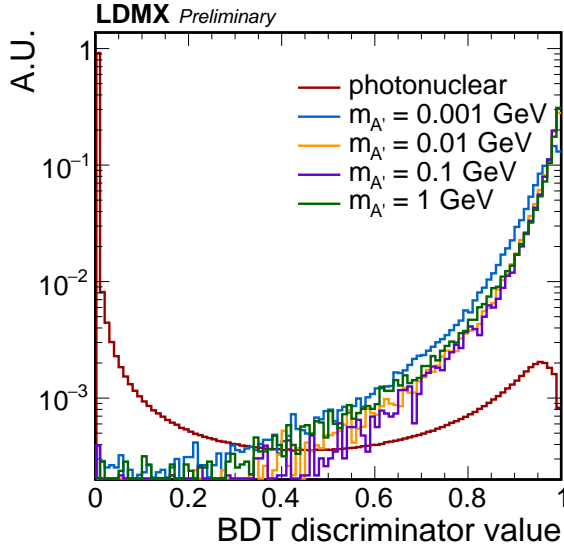

FIG. 45: Distributions of the ECal BDT discriminator value for signal and photonuclear events passing the trigger in which the recoil electron is within the ECal fiducial region. All distributions are normalized to unit area.

Muons and minimum ionizing particles   Another important signature to distinguish in the ECal are those of minimum ionizing particles (MIPs). Long tracks from MIPs can form in the ECal due to the high transverse and longitudinal segmentation of the detector. Efficiently identifying such MIP tracks can be a strong handle for vetoing backgrounds and is complementary to vetoes from the HCal system. In order to understand the energy range of muons which potentially do not leave a complementary handle in the HCal, we perform a simple study to understand the MIP energy loss of the muon in the ECal. This is shown in Fig. 47 where the typical muon loses roughly 500 MeV of energy, on average, in the ECal. Therefore muons in that energy range would rely solely on the ECal information for rejection.

Further, muons present another unique case but much rarer scenario. In particular, if a muon decays in flight in the ECal, thus not leaving a veto in the HCal, it can leave a challenging signature of $\mu^- \to e^- + \nu_e + \bar{\nu}_\mu$. Here the muon will lose most of its energy and decay softly in the ECal. However, since the muons are pair-produced in the ECal, we find that this rarer background is subdominant to other backgrounds.

G.   HCAL Simulation and Performace

The HCAL performance is studied using simulations based on the GEANT4 framework, combined with a model of the scintillator and readout responses to convert the energy deposited inside
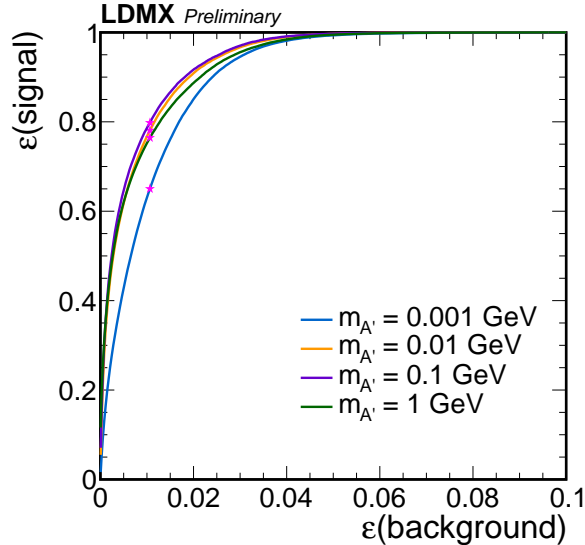
FIG. 46: ROC curves for the ECal BDT evaluated for signal and photonuclear events passing the trigger in which the recoil electron is within the ECal fiducial region.

the active medium into photoelectrons (PEs). As described previously (see section IV D), the scintillator layers are longitudinally segmented in an alternate (x, y) configuration to allow ambiguity resolution. In the current simulation, each bar is 20 mm thick and 50 mm wide.

Single neutron penetration studies to optimize the HCal depth and transverse dimensions have been done with steel absorber thickness between 2mm and 100 mm. The transverse dimensions were 3×3 meters, with smaller sizes studied using dimensional cuts. The bulk of the full background studies have been done using a 50 mm plates and a total HCal depth of 13 λ_A.

In both cases, the scintillator response was simulated by generating a Poisson-distributed number of photo-electrons based on the energy deposited in each scintillating bar, and adding the noise. An early estimate of 10 PE / MeV of energy deposited in the bar was used; future simulations will be based on the photo-electron yield measured by the Mu2e experiment and the updated bar geometry (see section IV D). These values includes quenching effects on the scintillation light yields (Birks' law), parametrized from earlier measurements of plastic scintillators [38]. The noise is generated from a Poisson process with a mean of 0.004 PE / bar. We consider an event to be vetoed if *any* bar contains at least 3 PEs.

In the most challenging cases, the hadronic veto system must detect a few neutrons having energies ranging from 100 MeV to a several GeV. To guide the design of the detector, we perform studies to characterize the detection efficiency of a single, normally incident neutron. The veto inefficiency as a function of the HCal depth is shown in Fig. 48 for different incident neutron energies and absorber thicknesses. As expected, thinner absorbers are less ineffective at vetoing energetic neutrons, while the veto capabilities of thicker absorbers decreases for low-energy neutrons as they are fully absorbed in the steel plates. We also study a very deep sampling hadronic calorimeter with 50 mm thick absorber to understand the rate of single energetic (few GeV) neutrons punching through the HCal as a function of depth. The results are displayed in Fig. 49. The results are displayed in Fig. 49. After a few layers, the inefficiency decreases approxmiately by an
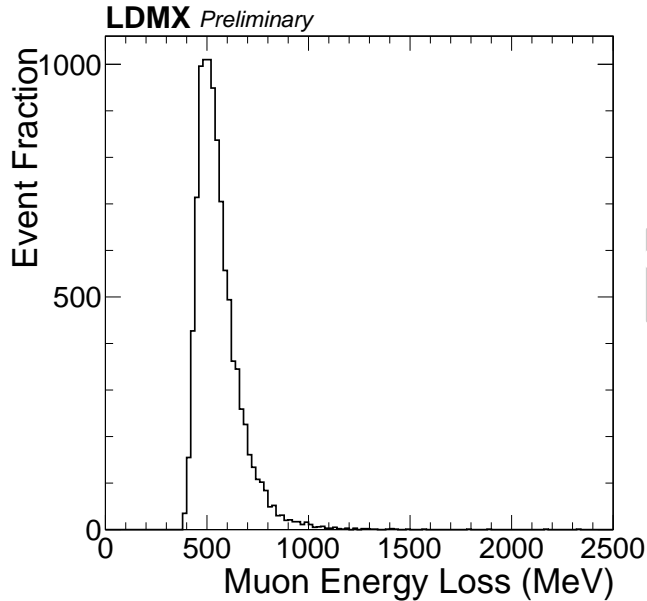
FIG. 47: Energy loss of muons with incident energy of 2.5 GeV onto the ECal system.

order of magnitude per 0.6m of HCal length. A depth of ~ 3m should be sufficient to reach the required single neutron inefficiency (see Sec. V D).

The Main HCal transverse dimensions are studied using a sample of photo-nuclear events corresponding to 10^{14} EOT (see Sec. V C). We select events containing a single reconstructed track, a value of the BDT output compatible with the signal hypothesis, and no veto hit found in the Side HCal. These PN events typically contain a few energetic neutrons emitted forward and must be vetoed by the Main HCal. For each event, we determine the position of the veto hit minimizing the transverse dimension. The distribution of the corresponding transverse and longitudinal positions is displayed in Fig. 50. In Fig. 50, the longitudinal size L is defined as $L = 2 \times \min(\max(|x|, |y|))$ where x and y are the coordinates of each hit in the HCal. While these preliminary results point towards a transverse dimension of the order of 2.5m-3m, additional studies are required to improve our understanding of the rate and kinematics of photo-nuclear interaction at moderate to high angles. For more discussion, see Sec. V D.

The energy resolution is estimated using the RMS of the sum all PEs inside the HCal (normalized to the mean value). The resolution as a function of the incident particle energy is displayed in Fig. 51 for electrons, pions and neutrons. The data are well approximated by a function of the form $a \oplus b/\sqrt{E}$. The energy resolution for 1 GeV neutrons is around 45% for 20 mm thick absorber, increasing to $\sim 70\%$ for a thickness of 50mm.
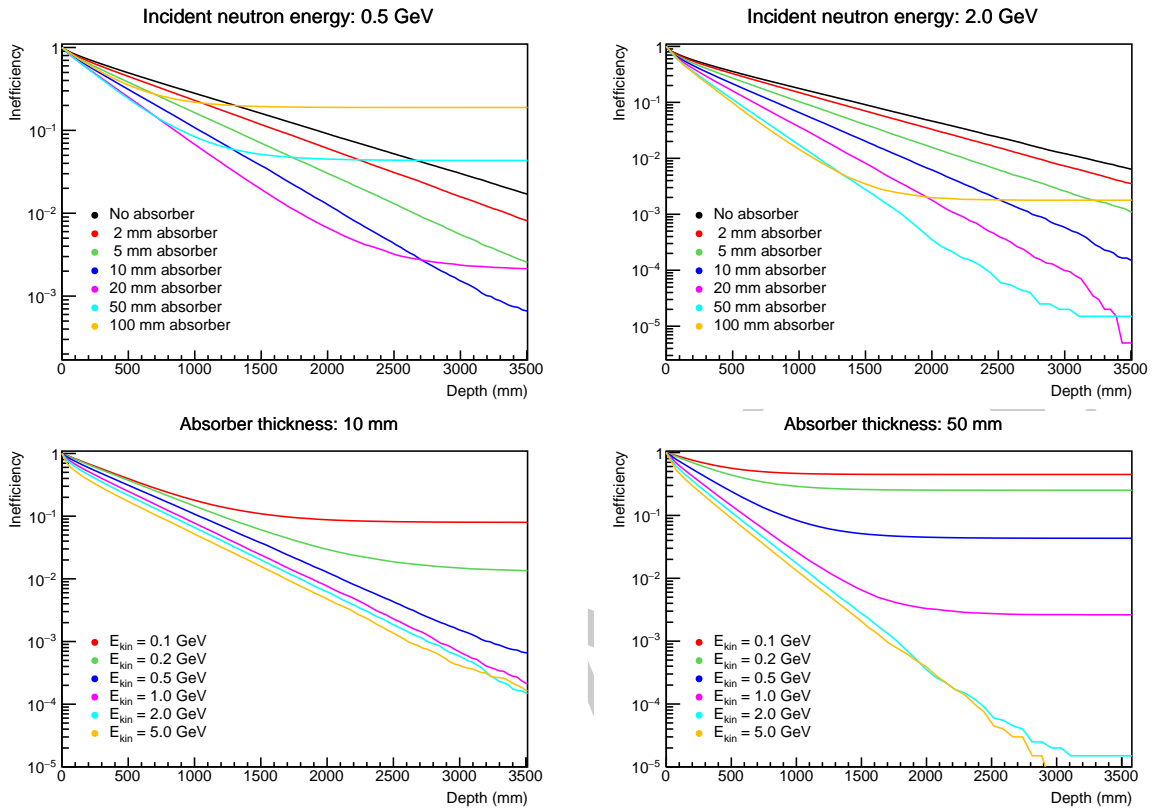
FIG. 48: Top: Single neutron veto inefficiency as a function of the sampling fraction for (left) 500 MeV and (right) 2 GeV incident neutrons. Bottom: Single neutron veto inefficiency as a function of the incident neutron energy for (left) 10 mm and (right) 50 mm absorber thickness.
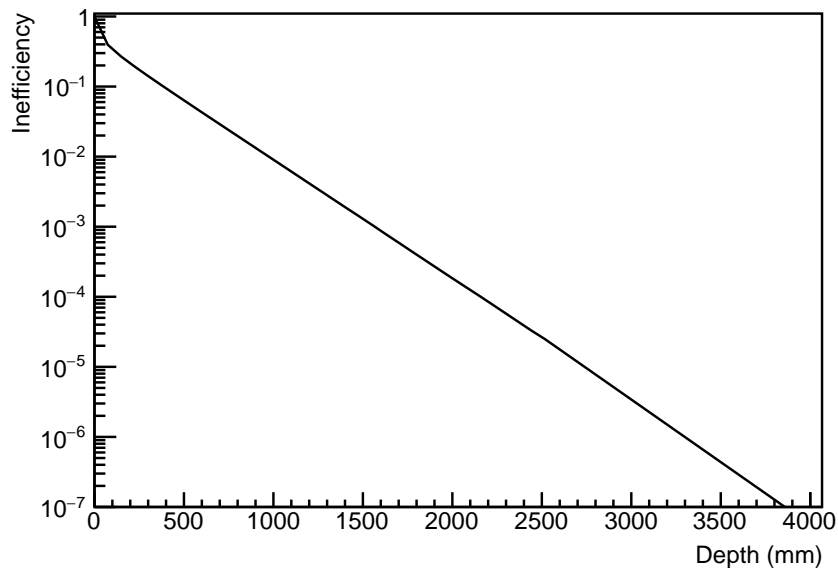
FIG. 49: Single neutron inefficiency as a function of depth of the HCal for a neutron of incident energy of 3 GeV.
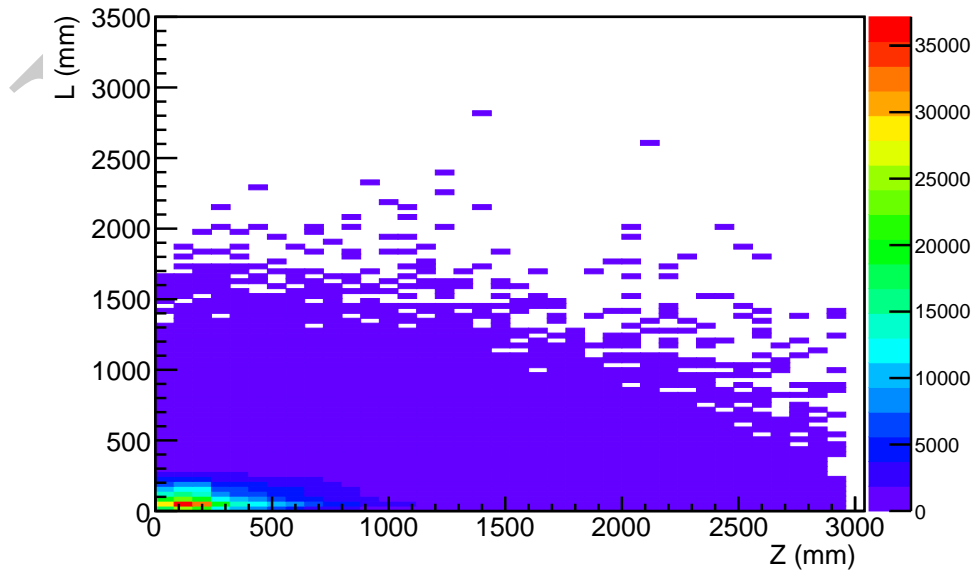
FIG. 50: The transverse (L) and longitudinal position (Z) of the veto hit minimizing the transverse HCal dimension. A HCal size of 3m $\times$ 3m$\times$ 3m is selected as a starting point to study the required dimensions.

FIG. 51: The energy resolution as a function of the incident particle energy for (top left) electrons, (top right) pions and (bottom) neutrons. The data are fit with a function of the form $a \oplus b/\sqrt{E}$.

TABLE VI: Draft trigger menu for LDMX, showing the primary contributions to the trigger budget for a $\mu_e = 1.0$ 46 MHz beam rate, with no corrections for any overlaps.

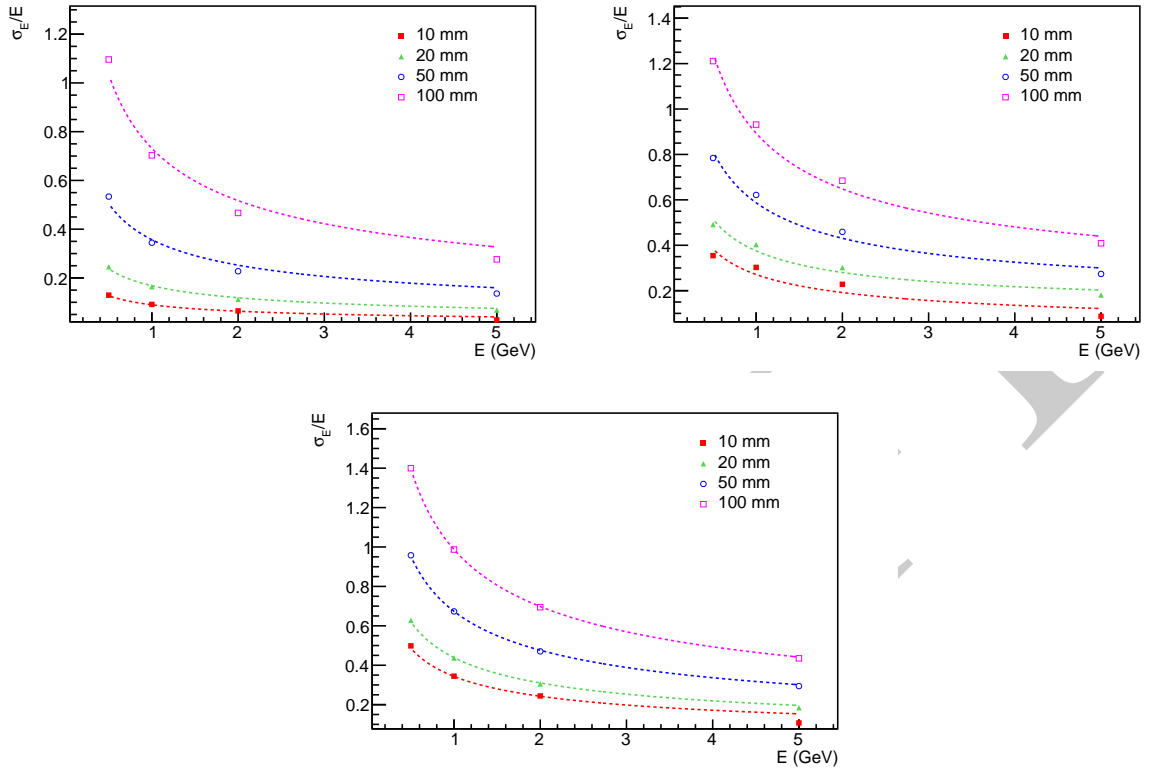| Trigger	Prescale factor	Rate (Hz)
Physics Trigger	1	*4000*
Background-Measurement Triggers		*500*
ECAL Missing-Energy > 1 GeV	5000	100
HCAL hit > 2 MIP	1000	100
HCAL hit > 20 MIP	1	100
HCAL MIP track		200
Detector-Monitoring Triggers		*500*
Zero-bias (trigger scintillator ignored)	4.6×10^5	100
Beam-arrival (trigger scintillator)	1.5×10^5	300
Empty-detector (trigger scintillator veto)		100
Total Trigger Budget		5000

H. Trigger System Performance

As described above, the primary physics trigger is based on the total energy observed in the calorimeter, combined with a requirement to identify the number of incoming electrons in the trigger scintillators. The performance of this trigger for signal is discussed in detail in Section VII A.

Besides the primary physics trigger, the LDMX trigger system will also allow the selection of events for calibration, alignment, and background studies. Each event will be marked with the set of triggers that fired. The trigger will include input from the scintillator calorimeter to allow selection of events with hadrons or muons. Pre-scaled samples of events will also be acquired with a zero-bias requirement, using only a requirement on energy deposition in the trigger scintillators, and with a veto on energy in the trigger scintillator.

An initial draft trigger menu is shown in Table VI. The table is built using a target trigger rate of 5 kHz, which provides a safety factor of ten compared with the limits expected from the tracker DAQ, which provides the strongest limit on trigger rate. As result, the trigger and DAQ systems have sufficient capacity to add additional monitoring triggers, to increase the bandwidth assigned to these triggers, or to define new physics triggers, e.g. triggers appropriate for nuclear physics studies. The signal trigger bandwidth can also be increased if required, for example for a larger average number of incoming beam electrons per bucket.

VI.   BACKGROUND REJECTION STUDIES

A.   Rejection of Beam Backgrounds

An incoming charged particle within the signal recoil acceptance that is wrongly reconstructed as a good incoming beam electron in the tagging tracker would be an irreducible background. However, incoming 4 GeV beam electrons have a precisely-defined trajectory through the tagging tracker which - together with the design of the tracker - makes the likelihood of such reconstruction errors vanishingly small.

In order for an incoming low momentum particle to fake a beam-energy electron in the tagger, a number of conditions must simultaneously be met. First, the incoming particle must reach the first layer of the tagging tracker or it will not intersect any other material before hitting the magnet. Because the first layer of the tagger is well inside the field region of the magnet, most low momentum particles are swept away, and only particles within a small region of energy-angle phase space will reach the first tagging layer. Second, that particle must undergo a hard scatter in the first layer, of order 100 mrad, but within a window approximately 1 mrad wide both vertically and horizontally, that puts it on the correct trajectory to make a hit in the correct location in the next layer. Finally, the particle must repeat a similarly unlikely scatter 5 more times in each subsequent layer so that it makes on-trajectory hits in all the correct locations in all of the layers of the tagging tracker before reaching the target and passing into the recoil tracker.
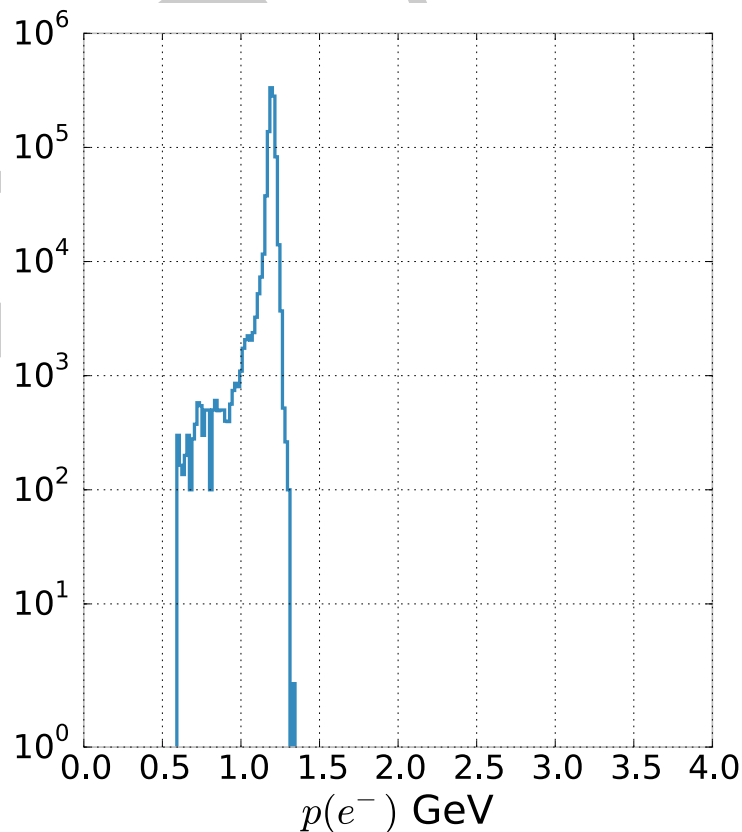

FIG. 52: Reconstructed total momentum for a sample of 1.2 GeV beam electrons in the tagging tracker. Of 10^6 incident particles, none is reconstructed with momentum greater than 1.4 GeV.

Taking into consideration the design of the tagging tracker, these requirements place a very heavy penalty on any off-energy component in the beam. First, incoming particles with momentum less than approximately 500 MeV will not hit the first layer of the tagger unless they are significantly off-trajectory as well. Furthermore, even at 500 MeV, scatters of more than $10°$ are needed in order for the incoming particle to appear to be on the correct trajectory. It is clear then that the most challenging scenario is large contamination with incoming charged particles at the top of the momentum range for signal recoils, nominally 1.2 GeV. Such particles have the highest likelihood of reaching the first layer of the tagging tracker without being bent away by the magnetic field and require much smaller scatters and/or track reconstruction errors to result in fake tags. In order for a 1.2 GeV particle to make a trajectory through the tracker consistent with a 4 GeV track, six successive scatters of approximately 50 milliradians must occur, each suppressing the rate by several orders of magnitude. It is therefore expected that even relatively impure beam will not present a significant background to the experiment.

As a preliminary test of the robustness of the tagging tracker against mis-tagging off-energy incoming electrons, a sample of simulated 1.2 GeV electrons on a trajectory that allows them to pass through all seven layers of the tagging tracker was used to estimate how often they could be mis-reconstructed as full-energy beam electrons. The results, including noise and pattern recognition effects, are shown in Figure 52 and confirm that such particles cannot be mistaken for 4 GeV electrons at a level of less than one part in 10^6, clearly limited here by simulation statistics.

B. Rejection of Non-Interacting Electrons

The LDMX target is only a small fraction of a radiation length, so the leading order process for an incoming 4 GeV electron is to pass through the target and both trackers without any significant energy loss and shower electromagnetically in the ECal. Such events can constitute a background if these electrons are measured by the recoil tracker and ECal as having energy consistent with signal, less than 1.2 GeV. Therefore, the combination of the recoil tracker and ECal must be able to reject 4 GeV electrons as recoils within the energy range for signal at the level of 4×10^{14} and 10^{16} for Phase I and Phase II respectively.

As will be shown in Section VI C 1, the resolution of the ECal alone is sufficient to suppress backgrounds from 4 GeV electromagnetic deposits so that they will not be reconstructed as having less than 1.2 GeV to the level of less than 1 event in 10^{15}. The recoil tracker, although optimized for low-energy recoils, adds similar, if not better, discriminating power for such events. As can be seen from Figure 53, the resolution of the recoil tracker for 4 GeV electrons that pass through the target and trackers without significant energy loss provides at least five more orders of magnitude of rejection of this background, where much larger statistics would be required to find the point where fake tracks due to noise and pattern recognition errors could contribute. Obviously, in combination with an energy measurement from the ECal this background will be insignificant.

C. Rejection of Backgrounds with a Hard Bremsstrahlung Photon

A class of background processes which presents varying challenges on the detector are the case when the incoming electron undergoes a hard bremsstrahlung in the target and loses a majority of its energy. Then, based on how the photon interacts, a signal-like signature can be mimicked. For example, if the electron has less than 1.2 GeV of energy and the photon leaves no detectable signature in the detector, this would be a signal-like signature.
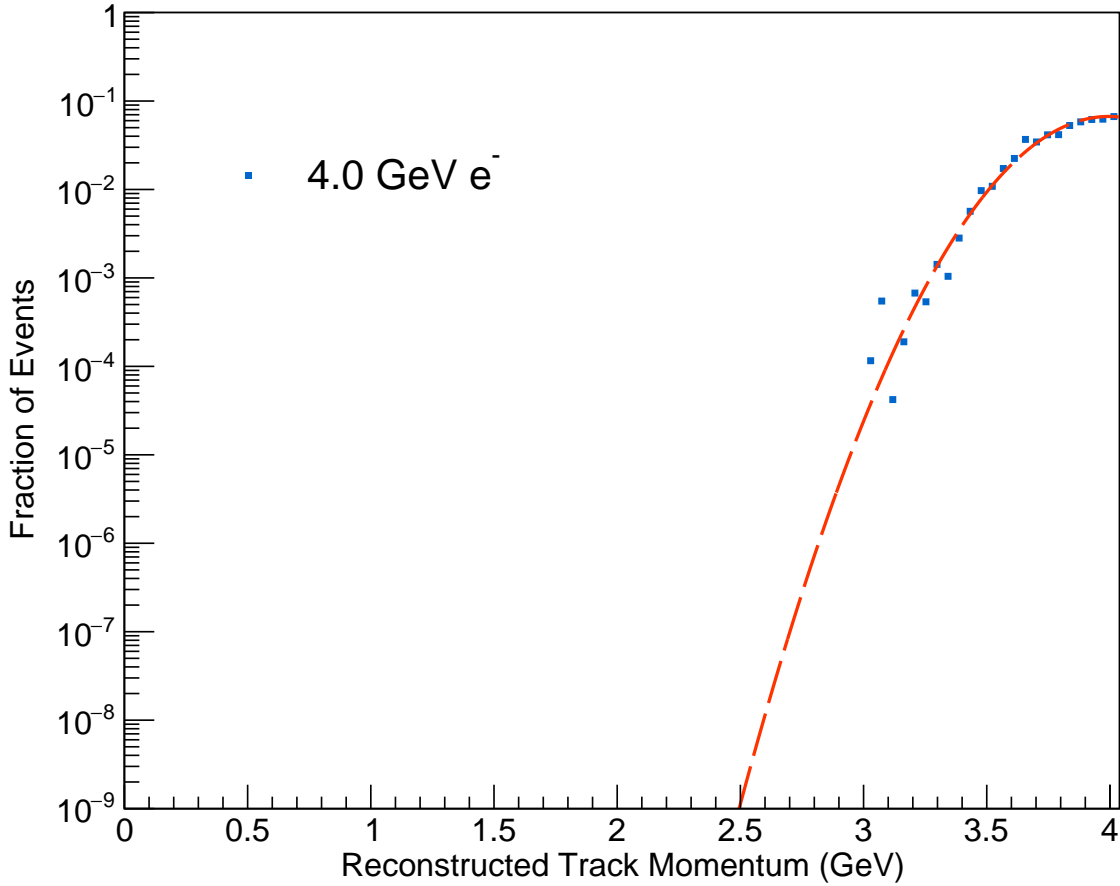
FIG. 53: The reconstructed recoil track total momentum for a sample of 4 GeV beam electrons that experience very little energy loss through the target and trackers. We find that the tracker is able to reject non-interacting electrons at level of at least 10^{-5}, limited only by Monte Carlo statistics for this study.

The basic detector signature for the nominal dark matter candidate is an incident 4 GeV electron, measured by the tagging tracker, which interacts in the target and leaves a less than 1.2 GeV electron track in the recoil tracker matched to a shower in the ECal and no other detector signals. This would pass the trigger requirement discussed in Sec. V H. In the following subsections, we study the cases where the photon from hard bremsstrahlung could mimic this signal-like signature. Specifically, we explore the following scenarios in detail:

- The photon undergoes a normal EM shower in the ECal but sampling calorimeter fluctuations result in a significant mismeasurement of the photon

- The photon converts, through a number of different processes, into a $\mu^+\mu^-$ pair

- The photon undergoes a photonuclear or electronuclear interaction resulting in a final state with hadrons

The samples used to to study these processes are detailed in Tab. VII.

Simulated Sample	Number of events	EoT equivalent
Target $\gamma^* \to \mu\mu$	3.3×10^7	1.1×10^{15}
ECal $\gamma^* \to \mu\mu$	7.3×10^8	6.6×10^{14}
Target photonuclear	5.2×10^7	1×10^{14}
ECal photonuclear	2.0×10^9	1.17×10^{14}
Target electronuclear	3.2×10^7	3.8×10^{14}

TABLE VII: Dedicated simulation samples for background investigation.

1. Photons that Shower Electromagnetically

As the ECal is a sampling calorimeter, the natural electromagnetic shower fluctuations can produce showers which have a low charge deposited in the silicon detector compared with the total energy of the shower. For such an event to appear as a signal event, it would be necessary for the shower in the calorimeter to not only match the recoil electron in energy, but for the shower to appear in the calorimeter at a position consistent with the low-momentum electron's bending in the dipole field. The most-challenging region of phase space to reject is therefore where the momentum of the electron is near the upper acceptance limit for the analysis – at 1.2 GeV.

Accordingly, we studied the behavior of the ECal for events where a recoil electron of 1.2 GeV is produced at the target along with a 2.8 GeV photon. The electron and photon originated at the same position in the target and were propagated collinearly into the ECal. We disabled all photonuclear, electronuclear, and muon-conversion processes in GEANT4 to allow a factorized study of this background.

Based on the extrapolated fit, the expected number of events of this type to yield less than 1500 MeV in the ECal–assuming 4×10^{14} EoT–is $\sim 10^{-2}$ events. Here we have also taken into account the fact that only $\sim 6 \times 10^{-2}$ of 4 GeV electrons on target result in a recoil electron with less than 1.2 GeV energy. Thus, pure electromagnetic shower fluctuations are not expected to be a significant source of background for 4×10^{14} EoT or even upwards of 10^{15} EoT.

2. Photons Undergoing a Muon Conversion

Another challenging background for the LDMX detector which is qualitatively different from many of the other backgrounds is those which involve the production of muons. An overview of the dominant photoconversion and trident processes which produce muons is given in Sec. III B. Here we consider strategies to reject these muon background produced both in the target and in the ECal. The primary sample of muons produced through photoconversion is simulated in GEANT4 with the cross section scaled up by a large factor as described in Sec. V C. The kinematics are simulated with the modified differential cross-section based on an improved nuclear form-factor described in Sec. V D.

The main rejection handles for muon backgrounds, depending on where they are produced, are:

- Trigger information: The recoiling electron should be lower than some threshold such that the event passes the nominal trigger.

- HCal information: This is the primary rejection handle for muon backgrounds as muons typically travel through the tracking and ECal systems and leave a MIP track through the HCal with many signals which would be vetoed by the baseline HCal selection. The most
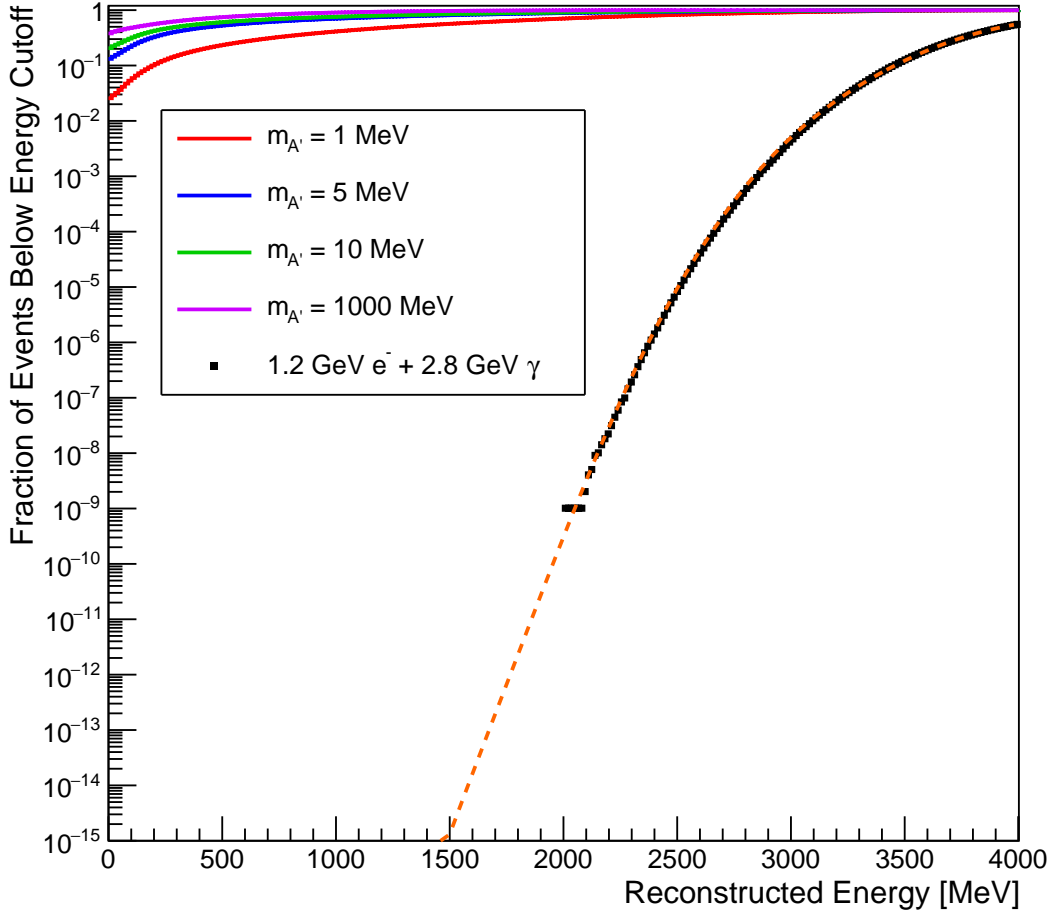
FIG. 54: The ECal reconstructed-energy distribution for 1.2 GeV electron plus 2.8 GeV photon events and signal events for various A' masses are plotted over a range extending to 4 GeV. By extrapolating the fit of the electron-plus-photon events, we find that the fraction of these events with a measured energy of less than 1500 MeV is on the order of 10^{-15}.

challenging scenarios to veto are often when the muon decays in the ECal producing neutrinos.

- Tracking information: If the muons are produced in the target, the recoil tracker system maybe identify more than one track. In this case, the event should be vetoed.

- ECal information: In the event that tracking and HCal systems do not veto the event, then the muons have often decayed in the ECal. The muons often to leave a MIP track in the ECal before decaying or ranging out and this, in addition to a large energy deposit, is the primary handle for rejecting muons in the ECal.

Because of the distinction between muon events produced in the target or ECal, we will discuss them separately.

a. Muons from the target For events where muons are produced in the target, a sample equivalent to 1.1×10^{15} EoT is produced. For a sample of 1.1×10^{15} EoT, we expect the electron to undergo a hard bremstrahlung at a rate of 0.0329 and of those resulting photons, we expect a

muon pair conversion rate of 9.1×10^{-7}. Therefore we generate 3.3×10^7 events where muons are produced in the target and produced a sample with this dedicated process.

Of the sample of dedicated muon events produced in the target, using the hadronic veto described in Sec. V G, we find only 9 muon pair events remain which have not been vetoed by the HCal. Of those remaining events, they are generically characterized by one hard forward muon, one soft muon, and one soft electron. Then the high energy muon decays in the ECal volume, the soft muon and electrons range out or decay in the ECal as well. Events of this type are easily vetoed by the tracking system which can individually identify the 3 charged particle tracks. In addition to the tracking information, the ECal information can provide complementary information to further veto this event.

To illustrate the rejection power of muon pairs produced in photoconversion in the target region, we select those events which pass the trigger selection. Then, of those events, we characterize the event by how many tracks are reconstructed and the number of hits found in the HCal system. This is plotted in Fig. 55 where on the x-axis is the number of HCal hits and the y-axis is the number of reconstructed tracks. We see that though there are several events which are not vetoed by the HCal, but most are vetoed by the tracking requirement. There is also the opposite scenario where the event is not vetoed by the tracking requirement but is vetoed by the HCal veto. This typically occurs when the recoiling electron and one of the muons is very soft and does not result in reconstructed tracks. However, the remaining hard muon is identified by the HCal veto.

The signal region is where there are *exactly* zero HCal hits and 1 reconstructed track. In Fig. 55, we find that there are 2 events which remain after the reconstructed track and HCal hit selection. Of those 2 events, neither event passes the ECal BDT which is described in Sec. V F. Therefore, we find no background events are contained in the signal region. This is summarized in Table. VIII.

b. Muons produced in the ECal We also study the case where the muon pair is produced from photoconversion in the ECal. Simulation is performed using GEANT4, again with the modified modeling of the photon conversion to a muon pair. In order to simulate a high statistics sample of muon pair background events, we perform final-state reweighting of this process as described in Sec. V C. We generate a sample that corresponds to $2 \times 10^{1}4$ EoT.

In the case where photon conversion occurs in the ECal, we no longer have a handle from the recoil tracking system. To illustrate the rejection power of muon pairs produced in the ECal, we select those events which pass the trigger selection. Then we characterize the event based on the number of hits in the HCal system. This is shown in Fig. 56. We find that a $\mathcal{O}(10)$ events remain after the HCal veto. These are typically events where one soft muon ranges out in the ECal and one hard muon decays into an electron and neutrinos. Therefore, we can use further ECal information to reject these remaining events which survive the HCal veto. As a simple means of veto, we use the same BDT that is used in to identify photonuclear events and these events are easily vetoed by the ECal BDT. For example, the event with the highest BDT score is shown in Fig. 57. Beyond the usual photonuclear shower handles, there is a clear muon MIP track in the ECal which could be identified easily with dedicated algorithms. In future work, we note that we can design an even more efficient ECal veto for muon MIPs in the ECal. The results for the rejection of muons originating in the ECal are also presented in Table VIII.

3. Photons Undergoing a Hard Photonuclear Interaction

A background process that is difficult to distinguish from signal is the case where a hard photon undergoes a photonuclear reaction in either the target or in the ECal absorber layer resulting in the conversion of the photon's energy into outgoing hadrons. The resulting final states can vary greatly

Cutflow and event yields

	ECal $\mu\mu$	**Target-area** $\mu\mu$
EoT equivalent	1.1×10^{15}	6.6×10^{14}
Events passing trigger	2.14×10^{7}	1.50×10^{8}
Passing HCal veto	36	169
Passing N_{track} veto	2	169
Passing ECal BDT veto	0	1

TABLE VIII: The estimated levels of dimuon background events after successive background rejection cuts.

because both nuclear disintegration and/or hadron production can occur. Since neutrons do not lose energy to ionization, and protons follow a very steep ionization curve, the detector response also greatly varies. Vetoing of such challenging processes requires all subdetector systems to work in complementarity.

Several handles have been studied in order to understand how to best veto these events. For example, as shown in Fig. 44, photonuclear interactions have distinctly different ECal shower signatures when compared to more common electromagnetic processes and serve as a powerful discriminator that can be used to veto these events. In processes involving neutrons that don't
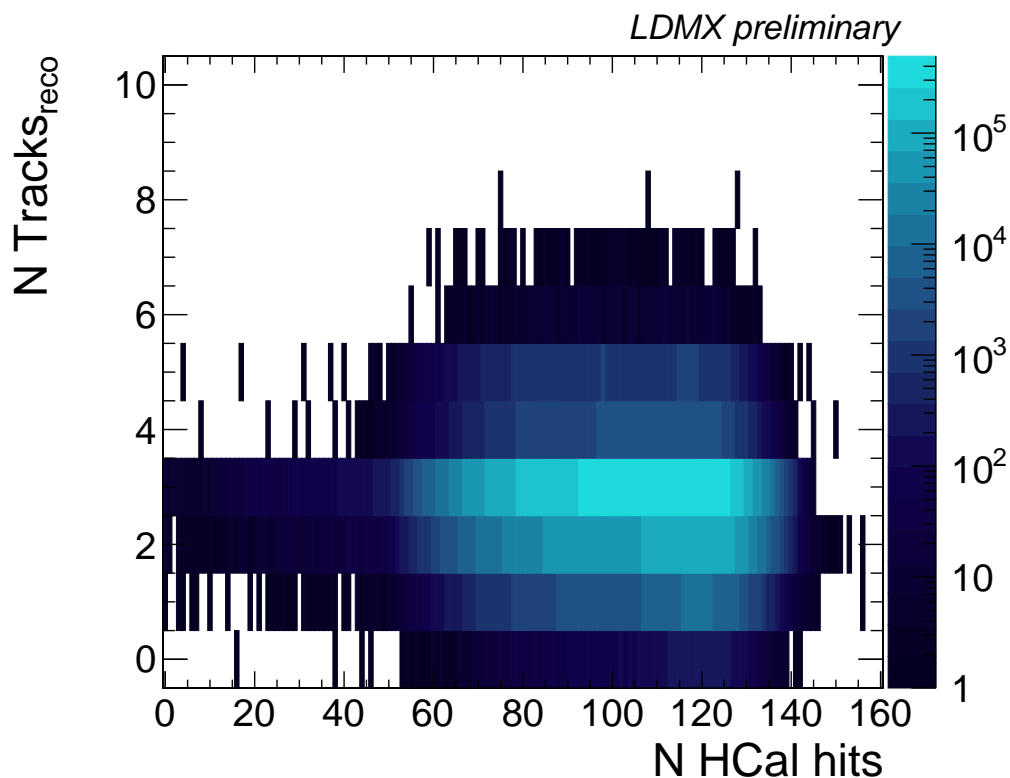

FIG. 55: Number of HCal hits versus reconstructed tracks for muon pair conversion background.
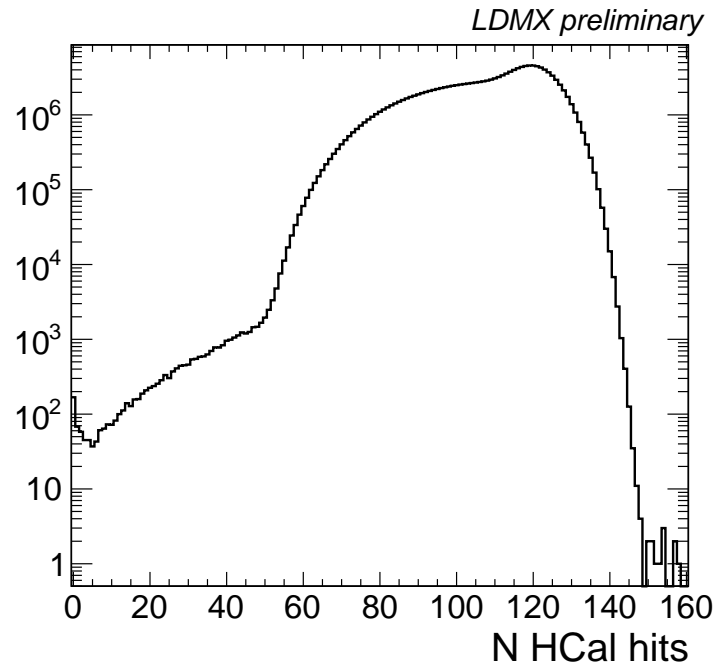
FIG. 56: Number of HCal hits for muon pairs produced in the ECal

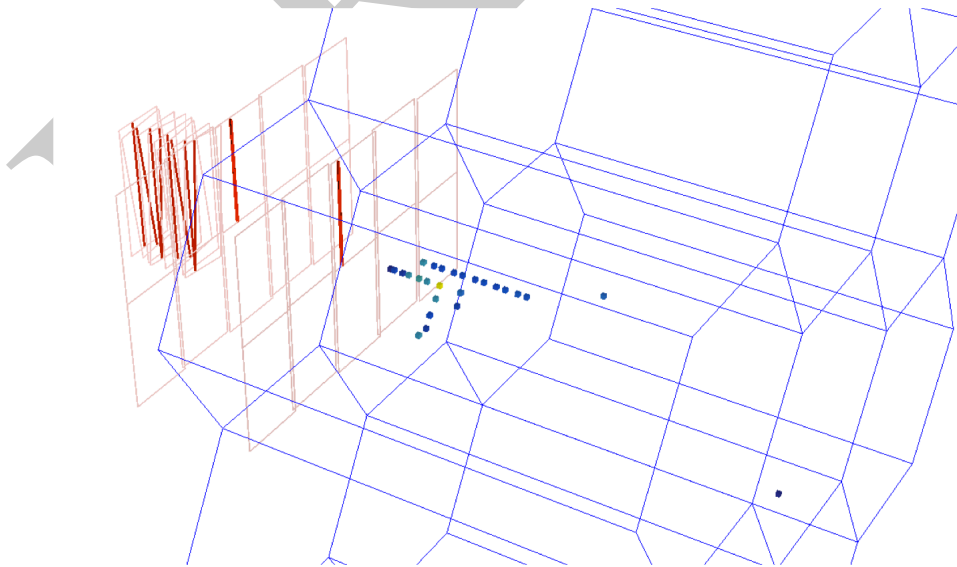FIG. 57: Event display of an $\gamma^* \to \mu\mu$ interaction in the ECal which does not pass the HCal veto.

interact in the ECal, the HCal is relied on to identify them. Finally, the recoil tracking system can be used to identify events with additional charged hadrons produced in these photonuclear pro-

Cutflow and event yields

	ECal PN	Target-area PN
EoT equivalent	1.17×10^{14}	1×10^{14}
$E_{recoil} < 1.5$ GeV, trigger requirement	2.7×10^8	2.2×10^7
ECal Shower-Profile BDT	2×10^6	8.2×10^5
HCal Max PE < 3	< 79	< 94
Single track with $p < 1.2$ GeV	<62	< 81
Recoil activity Cut	<31	N/A
ECal activity cut	<26	N/A
Tagger tracker activity cut	N/A	< 11

TABLE IX: The estimated levels of background photonuclear events after successive background rejection cuts. Note, these quoted background levels do not take into account the GEANT4 photonuclear modeling fixes. These numbers are expected to be reduced by orders of magnitude once the corrections are included. These corrections will be included in the next version of the draft.

cesses, either forward going if the photonuclear process occurs in the target or via back-scattering particles if the photonuclear interaction occurs in the ECal.

With this in mind, the photonuclear background rejection strategy is roughly broken into three main tasks:

1. shower profile rejection using the ECal

2. hadronic tails with a high efficiency, large angular coverage HCal veto

3. high track and hit multiplicity veto using the trackers

Development of the background rejection strategy was done using a sample of events where a hard bremsstrahlung occurs in target and the resulting photon undergoes a photonuclear reaction. As shown in Tab. VII, an ECal photonuclear (target-area photonuclear) sample of 2.0×10^9 (5.2×10^7) events was used which is what is expected assuming 1.17×10^{14} (1×10^{14}) EoT. The details of the sample generation are discussed in Sec. V C. Only the subset of events that pass the trigger requirements as described in Sec. V H are considered. A summary of the final rejection strategy is given in Table IX, the details of which are discussed below.

a. Shower profile rejection   As detailed in Sec. V F, there is a distinct difference in the ECal signatures between dark matter signal-like processes and photonuclear interactions. The variables with the strongest discrimination power are exploited to distinguish photonuclear reactions from normal electromagnetic showers in the ECal. Specifically, the information from these variables are combined using a gradient boosted decision tree (BDT) in order to further reject photonuclear reactions while maintaining a high signal efficiency.

Based on the BDT performance metrics discussed in Sec. V F, a cut of .94 on the BDT discriminant was determine to best balance the rejection of photonuclear background and the preservation of signal. Applying this requirement, roughly 98.7% and 96.0% of ECal photonuclear and target-area photonuclear background, respectively, is rejected all while preserving 65-85% of the signal, depending on the mediator mass. Those events that remain typically fall into one of two categories:

- photonuclear events that result in energetic backwards going pions

- photonuclear events that produce energetic neutrons

A subset of the first class of events can be rejected by making use of activity in the trackers while it is the task of the HCal to efficiently veto the second class of events.

b. Hadronic tails   A subset of events that the ECal has difficulty vetoing are typically characterized by neutral hadrons which escape the ECal without leaving any significant energy deposition. The HCal is left to detect these events with high efficiency.

As shown in Fig. 58, a strong discriminator between signal and photonuclear background is the layer with the maximum number of photoelectrons in the HCal. Typically, the recoil electron in a signal event will enter the ECal and shower with very little energy escaping into the HCal. As discussed in Sec. V G, noise will contribute on average 1 PE per event. Given that signal events will typical see very little activity in the HCal, most events are expected to have a max PE of 1. In contrast, photonuclear backgrounds will leave multiple hits in the HCal resulting in multiple PE's per layer. This is shown in Fig. 58 which overlays the maximum PE per event observed in signal and photonuclear background events. With this in mind, requiring an event to have a max PE < 3, will reject 99.99% of the remaining background while still preserving 99% of the remaining signal.
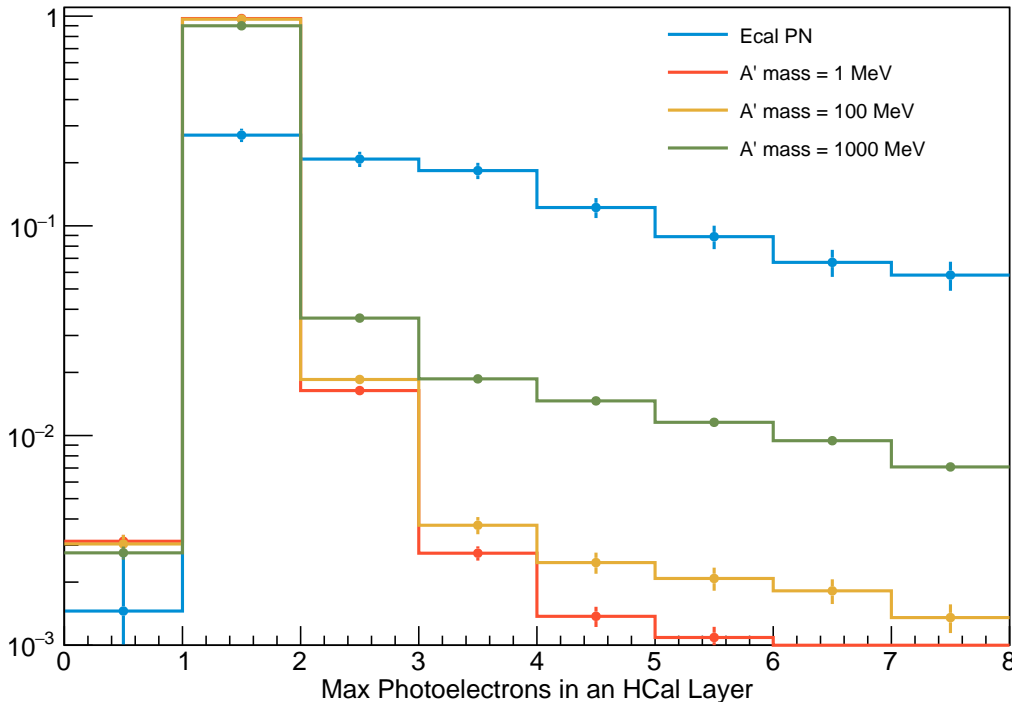

FIG. 58: The maximum photoelectrons in layer of the HCal per event observed for 1, 100 and 1000 MeV signal, target-area and ECal photonuclear events. Requiring the max PE < 3 is found to reject 99.99% of background events that are failed to be rejected by the ECal BDT while preserving 99% of the signal.

c. Tracker Veto   After exhausting all calorimeter handles, there still remains a subset of events with hard forward or backward going charged hadrons that can be rejected by the trackers. As shown in Fig. 59, a significant number of these remaining events have multiple reconstructed tracks in the Recoil tracker. Furthermore, as discussed in Sec III A, only signal events where the reconstructed momentum of the recoiling e^- is below 1.2 GeV are considered to encompass the

signal region. Therefore, simply requiring that the Recoil tracker has only a single reconstructed track with a momentum < 1.2 GeV vetoes an additional 25% and 23% of the remaining target and ECal photonuclear events, respectively.
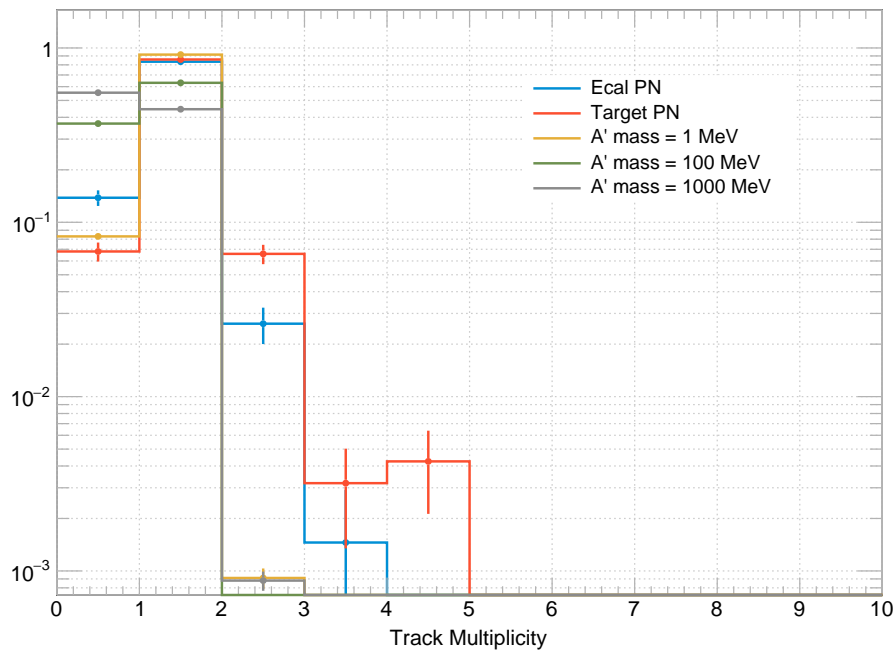

FIG. 59: The number of reconstructed tracks in ECal (blue) and target (orange) photonuclear events after the application of the calorimeter vetoes. For comparison, the number of reconstructed tracks for a 1 (yellow), 100 (green), 1000 (grey) MeV A' signal are also shown. After applying the calorimeter vetoes, a subset of events remains that contain hard forward or backward going charged hadrons that leave additional tracks in the recoil tracker.

For events that see a photonuclear reaction occur in the target, the Tagger tracker can be used to further reject these events. Given that the noise occupancy per tracker sensor is expected to be at the 10^{-4} level, the last layer of the Tagger tracker will typically only see a single hit from the incident beam electron. For photonuclear events, the backward going charged hadrons will leave multiple hits in the last layer. This can be seen in Fig. 60 which shows the hit multiplicity in the last Tagger tracker layer excluding the hit due to the incident beam electron. Requiring that no hits other than that due to the incident beam electron are observed in the last layer of the Tagger tracker, all but two of the remaining events are rejected while still preserving 99% of the remaining signal.

For ECal photonuclear events, the last layer of the Recoil tracker can be used in the same way the last layer of the Tagging tracker was used. It is expected that hits observed in the last layer of the Recoil tracker will have a corresponding track. However, this wont be the case in ECal photonuclear events because the backwards going charged hadrons are produced at angles such that they only hit the last layer of the recoil tracker. This can be seen in Fig 61, which shows the hit multiplicity in the last Recoil layer after excluding hits associated with tracks. Requiring that there are no hits other than those associated with tracks in the last layer of the recoil tracker rejects an additional 53% of the remaining events while preserving 96% to 97% of the remaining signal.

FIG. 60: The number of hits in the last layer of the Tagging tracker, excluding the contribution from the incident beam electron, for target-area (blue) photonuclear events after the application of calorimeter and recoil tracker vetoes. For comparison, 1 (orange), 100 (yellow), 1000 (green) MeV A' signal are also shown.

FIG. 61: The number of hits in the last layer of the Recoil tracker, excluding the contribution from hits associated with tracks, for ECal (blue) photonuclear events after the application of calorimeter and recoil tracker vetoes. For comparison, 1 (orange), 100 (yellow), 1000 (green) MeV A' signal are also shown.

Cutflow and event yields

	Target-area EN
EoT equivalent	3.8×10^{14}
$E_{recoil} < 1.5$ GeV, trigger requirement	1.4×10^{7}
ECal Shower-Profile BDT	6.2×10^{5}
HCal Max PE < 3	< 62
Single track with $p < 1.2$ GeV	< 38
Tagger tracker activity cut	< 6

TABLE X: The estimated levels of background electronuclear events after successive background rejection cuts. Note, these quoted background levels do not take into account the GEANT4 electronuclear modeling fixes. These numbers are expected to be reduced by orders of magnitude once the corrections are included. These corrections will be included in the next version of the draft.

D. Rejection of Electronuclear Interaction Backgrounds

Electronuclear interactions are a background component which are similar to the photonuclear interactions. Because we can track the momentum of an electron with the beam energy, we are less concerned with electronuclear interactions which occur in the ECal and primarily with those which occur in the target. Electronuclear interactions in the target are similar to the photonuclear interactions in the target, though they have a soft electron recoil momentum spectrum and a lower overall cross-section. The rejection strategy is again similar to the photonuclear interactions.

We simulated a sample of electronuclear interactions in the target corresponding to 3.8×10^{14} EoT. Of those events which pass the trigger criteria, we determine which events pass both the ECal BDT and HCal max PE selection. This is shown in Fig. 62. On the y-axis is the the BDT discriminant value and the x-axis is the maximum number of photoelectrons. The signal region is defined by events where the maximum PE is less than 3 and the BDT discriminant value is greater than 0.94. There are approximately 62 events which are not vetoed by the HCal and ECal selections. An additional $\sim$40% of events are rejected by requiring that the recoil tracker has only a single reconstructed track with a momentum below 1.2 GeV . For those events that pass the recoil tracker veto, we plot the hit multiplicity in the last layer of the tagger tracker. As shown in Fig. 63, requiring that no other hits other than that due to the incident beam electron are observed in the last layer of the Tagger tracker, rejects all but 6 events while still preserving 99% of the remaining signal.
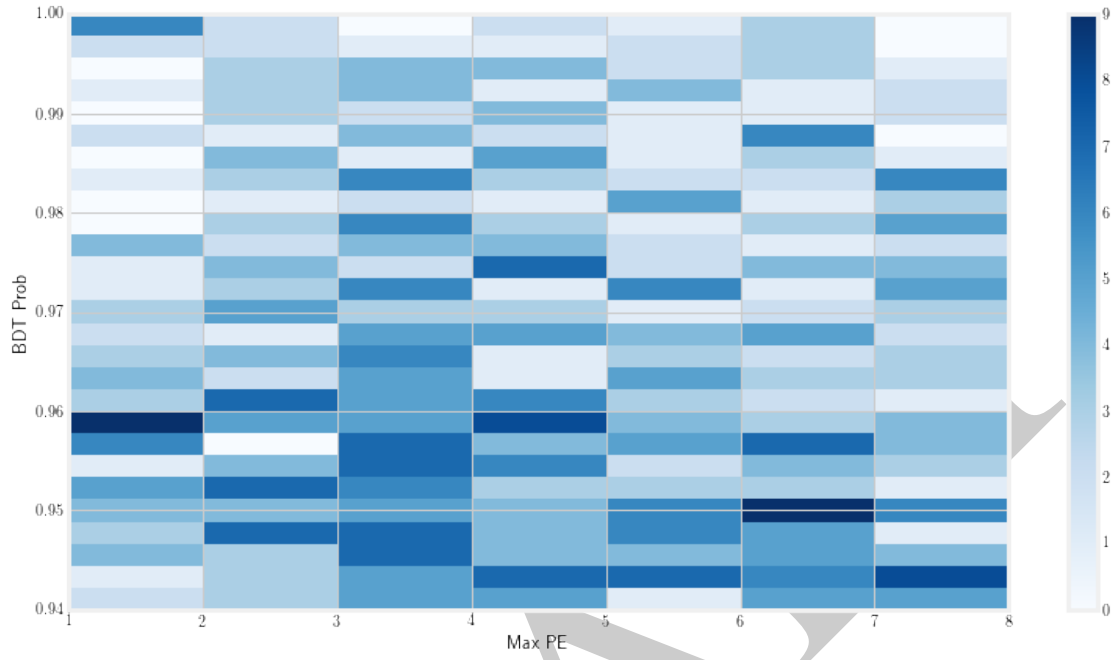
FIG. 62: The maximum number of photoelectrons versus the ECal BDT score for electronuclear events passing the trigger.
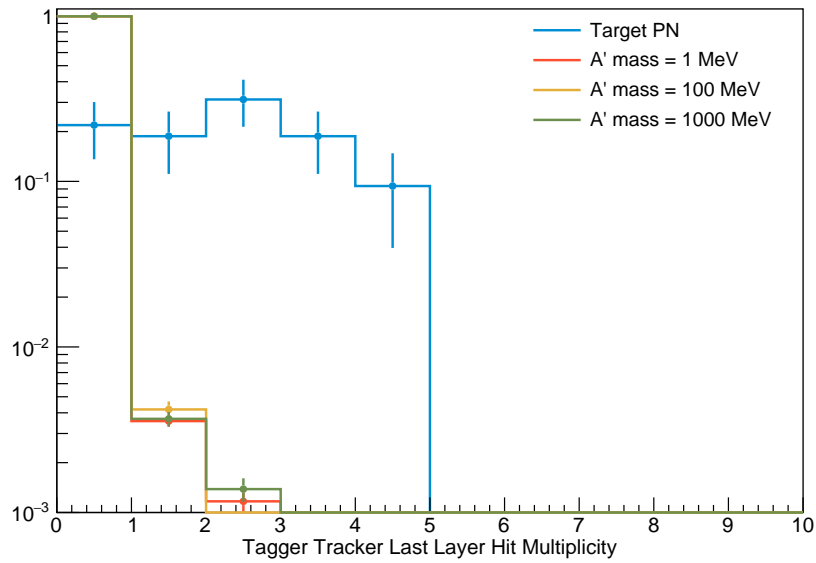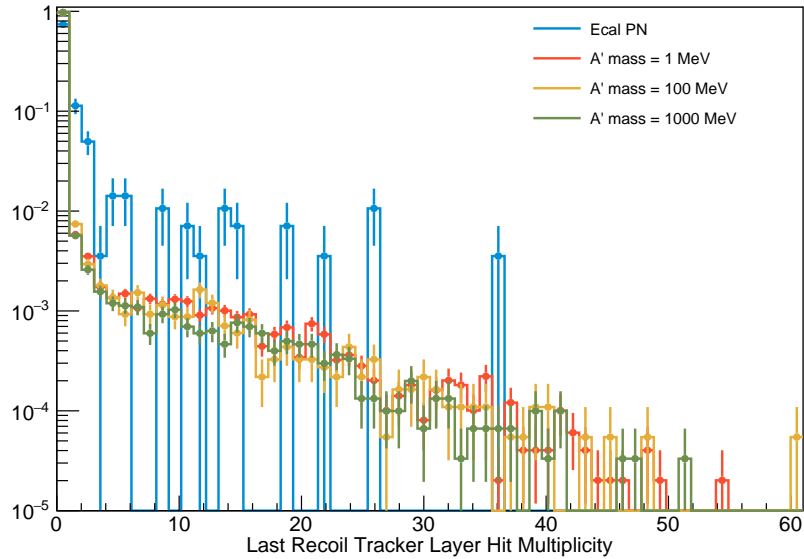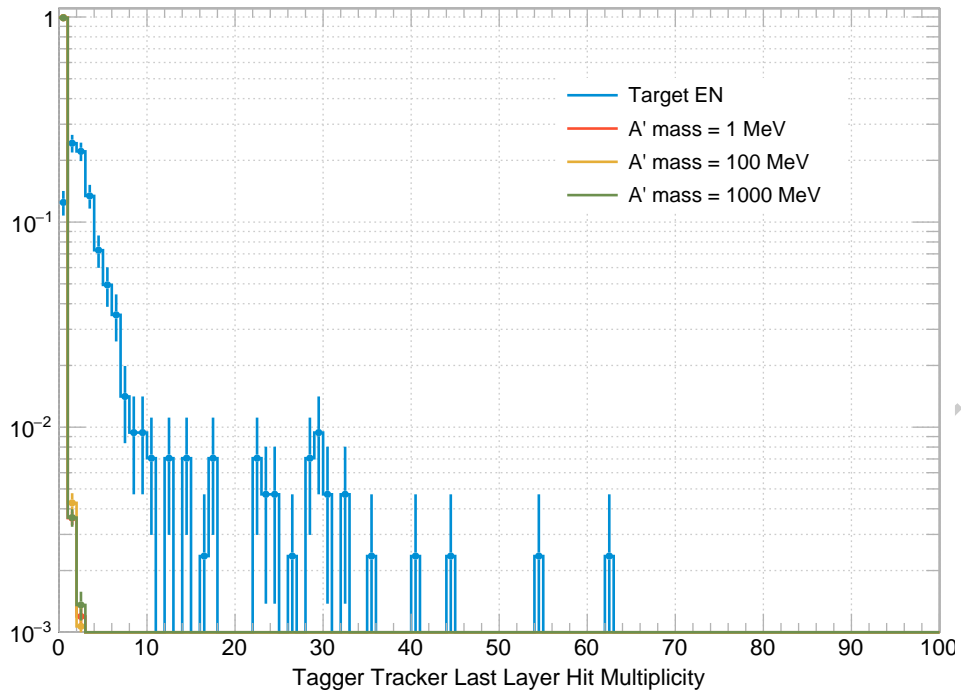
FIG. 63: The number of hits in the last layer of the Tagging tracker, excluding the contribution from the incident beam electron, for target-area (blue) electronuclear events after the application of calorimeter and recoil tracker vetoes. For comparison, 1 (orange), 100 (yellow), 1000 (green) MeV A' signal are also shown.

VII. SIGNAL EFFICIENCY STUDIES

A. Trigger Efficiency

As described above, the primary physics trigger is based on the missing energy in the event, based on the total energy observed in the calorimeter combined with the incoming energy as determined by the the number of incoming electrons in the trigger scintillators. While the average number of incoming beam electrons is 1, the number in any given bunch crossing will vary from zero to ten or more. For each number of incoming beam electrons n, a different threshold is set to reject events consistent with n full-energy electrons impacting the calorimeter while retaining events with $n - 1$ full-energy electrons and possibly one low-energy electron impacting the calorimeter. These thresholds are tuned to maximize the efficiency for signal, while keeping the calorimeter trigger bandwidth within its targeted allocation of 4 kHz. The resolution of the calorimeter implies that the rate for a fixed missing-energy threshold will increase with increasing numbers of incoming electrons.

Figure 64 shows the simulated performance of the primary physics trigger for a single incoming beam electron for signal and background.

The performance for the current optimization of the trigger threshold table for $\mu = 1$ is shown in Table XI. The signal inefficiency column takes into account that, for example, there are four opportunities for a signal event to be produced in an $n_{\text{beam}} = 4$ event.
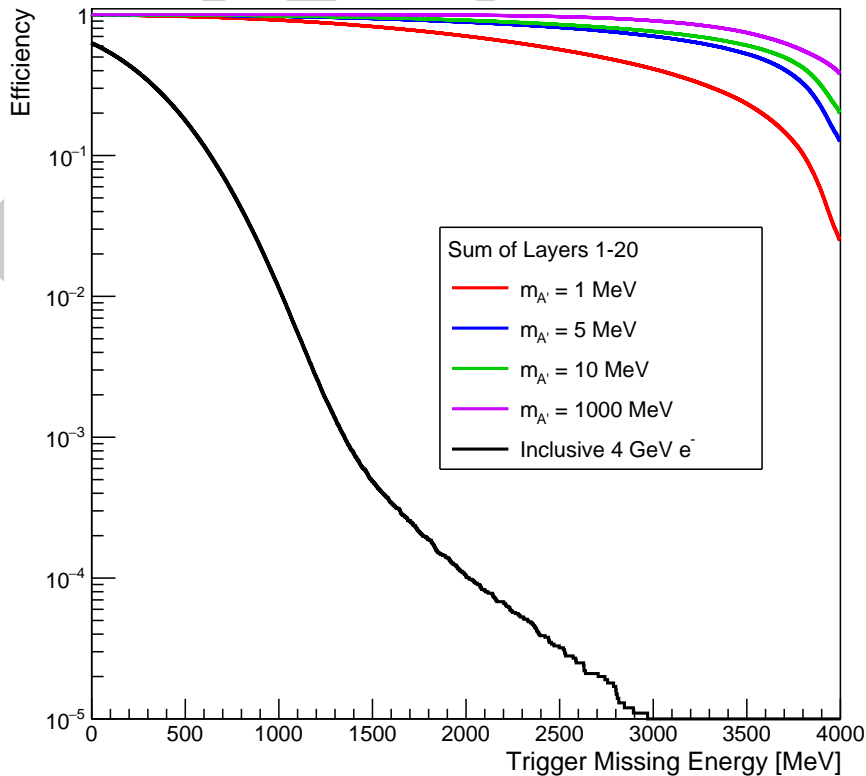

FIG. 64: Performance of the primary physics trigger for LDMX for a single incoming beam electron. The efficiency for signal electrons for different A' masses and the trigger rate for all backgrounds induced by beam electrons are shown as a function of the trigger missing-energy threshold.

TABLE XI: Trigger thresholds, rates, and total inefficiency contribution as function of number of incoming beam electrons for an average number of one electron per bunch.

$n_{\rm beam}$	Fraction of Bunches (Signal)	Trigger Scintillator Efficiency	Missing Energy Threshold [GeV]	Calorimeter Trigger Efficiency	Rate [Hz]	Signal Inefficiency
1	36.8% (36.8%)	100%	2.50	99.2%	588	0.3%
2	18.4% (36.8%)	97.4%	2.35	98.0%	1937	1.7%
3	6.1% (18.4%)	92.4%	2.70	91.6%	1238	2.8%
4	1.5% (6.1%)	84.3%	3.20	77.2%	268	1.6%
Total					4000	8.8%

B. Recoil Acceptance and Efficiency

As discussed in Sec. V E 2, the recoil tracker must not only have good acceptance for signal recoils but also for charged tracks over the largest possible acceptance. Since high-momentum signal recoils will nearly always pass through all six layers, the acceptance near the top of the energy range for signal recoils is near unity, only reduced by the small single-hit inefficiency in the last two layers. However, at low momentum a large number of tracks can escape detection. Therefore, in order to estimate the signal acceptance using simulated signal events, we apply both "loose" and "tight" track requirements. A loose track requires that the recoiling electrons leave hits in the first two 3-d layers which is sufficient for pattern recognition and angle estimation. For those events where the best vertex and p_t resolution is desired, the recoiling electrons are required to leave hits in at least two of the 3-d layers and at least four hits total. The resulting acceptance for the recoil tracker as a function of mediator mass using the loose (orange) and tight (green) track requirements is shown in Fig. 65 while the average of the track finding efficiencies for accepted tracks in those samples as a function of momentum is shown in Fig. 66.

The recoil tracker will play a leading role in rejecting beam, photonuclear, electronuclear and muon conversion backgrounds that originate in the target. Assuming the average number of incident beam electrons is 1, most signal events will see a single track in the recoil tracker. Furthermore, as discussed in Sec. III A, only events where the recoil electron has a momentum below 1.2 GeV are considered. Therefore, requiring an event to contain a single track in the recoil tracker with reconstructed momentum < 1.2 GeV, aides in the rejection of background while maintaining a high signal efficiency. The impact of these cuts on the signal efficiency using both the loose () and tight track requirement is shown in Fig. 67.

C. Signal Efficiency of Calorimeter Vetoes

As detailed in Sec. IV C, the difference in hit multiplicity and energy deposited in the ECal as well as the transverse and longitudinal shower shapes are combined into a BDT and used to discriminate between signal and photonuclear background. Based on the performance of the BDT, a threshold of .94 was found to balance the rejection of background versus the preservation of signal. The impact of applying such a requirement on signal events is shown in Fig. 68 as a function of mediator mass.

A typical signal event will see the recoil electron enter the ECal and shower with very little
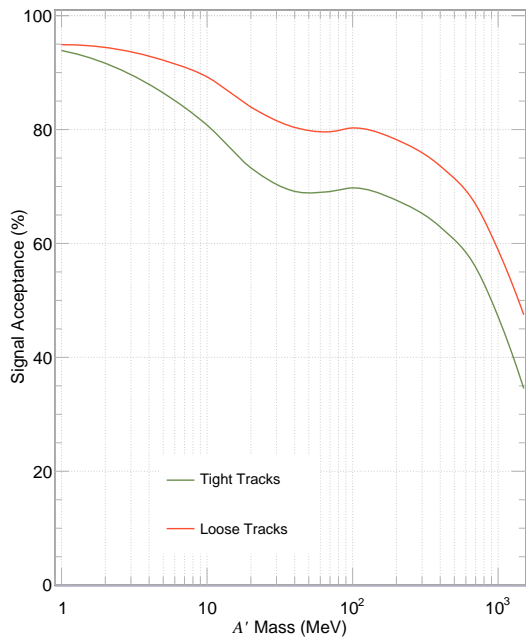
FIG. 65: The recoil tracker acceptance to signal recoils as a function of A' mass. The acceptance drops with A' mass as the polar angle of the recoils increases.
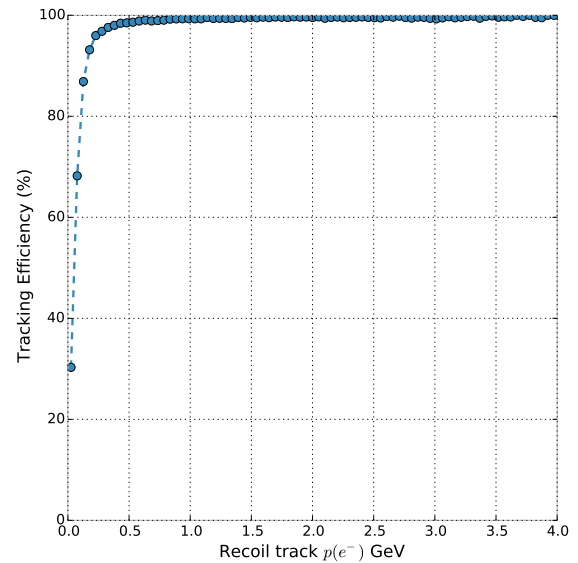
FIG. 66: The tracking efficiency for signal recoils as a function of the momentum for electrons that pass the acceptance criteria described in the text. Since data from all A' mass points are used, these efficiencies are averaged with respect to different recoil polar angle distributions for each sample.
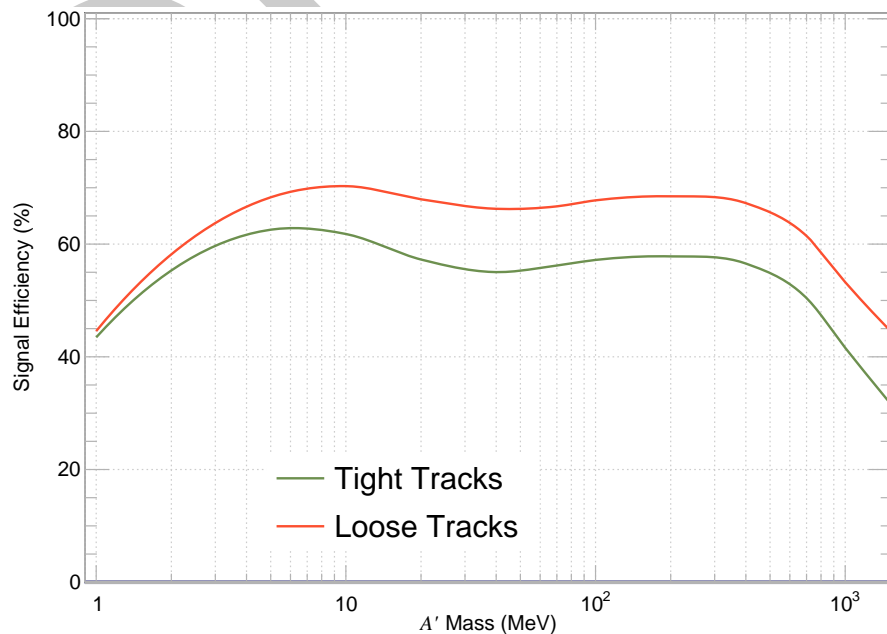
FIG. 67: The signal efficiency as a function of mediator mass after requiring an event to have only a single track below 1.2 GeV. The signal efficiency assuming a loose track requirement is shown in orange while the efficiency after imposing a tight track requirement is in green.

energy escaping into the HCal. In fact, assuming the average number of incident beam electrons is 1, most signal events will see a maximum PE in the HCal of 1 PE due to the expected noise contribution (see Sec. V G). In contrast, backgrounds will leave multiple hits in the HCal resulting in multiple PE's per layer. This is shown in Fig. 69, which shows the signal efficiency for different A' masses along the photonuclear background efficiency as a function of maximum PE. By requiring that the maximum PE observed in an event is < 3, the signal efficiency is maximized while also rejecting a large portion of the background. The impact of this cut on the signal is shown in Fig. 70. The drop in efficiency at high mass, is a result of signal recoils entering the HCal and showering, resulting in multiple PE's. Future iterations of the veto will mitigate the inefficiency by excluding the HCal region where the shower occurs when determining the maximum PE.

D. Signal Efficiency Summary

The signal efficiency after applying background rejection cuts discussed above is shown in Fig. 71 and summarized for three A' mediators in Tab. XII. After all cuts, the signal efficiency is approximately 50%. This efficiency is used in the calculation of the expected sensitivity described in the next section.
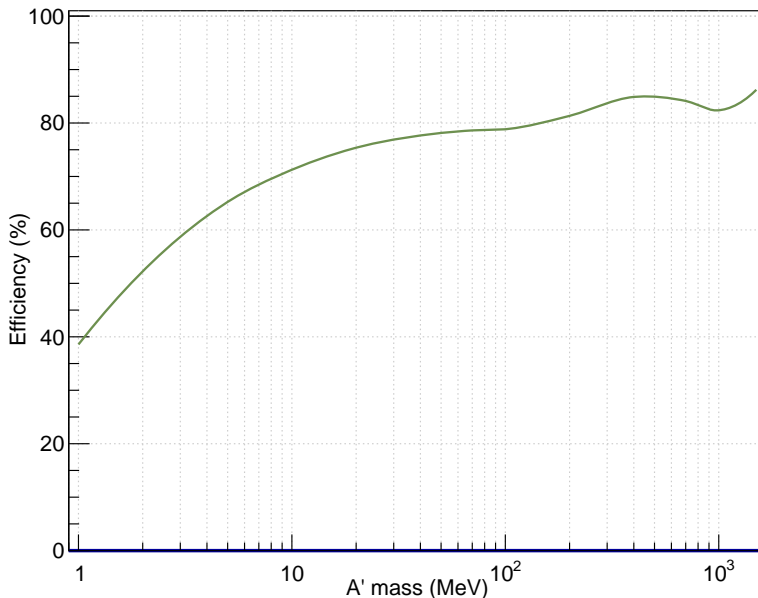
FIG. 68: The signal efficiency as a function of mediator mass using a BDT threshold of .94.
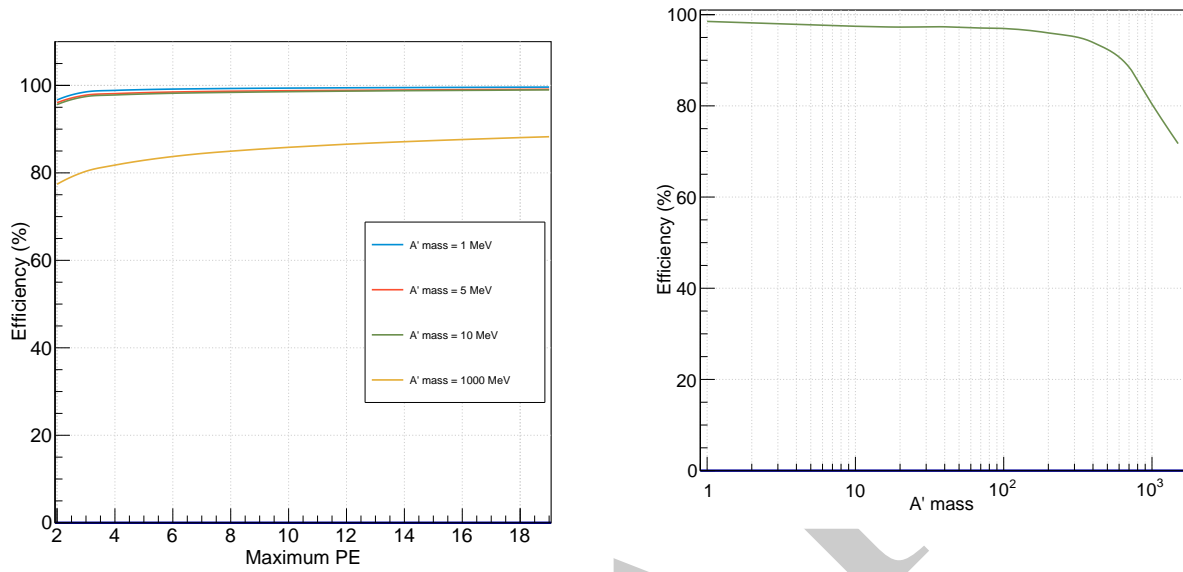
FIG. 69: The signal efficiency as a function of mediator mass as a function of maximum PE in the HCal.

FIG. 70: The signal efficiency as a function of mediator mass after requiring a maximum PE of < 3. The drop in efficiency at high mass is a result of signal recoils entering the HCal and showering resulting in multiple PE's.
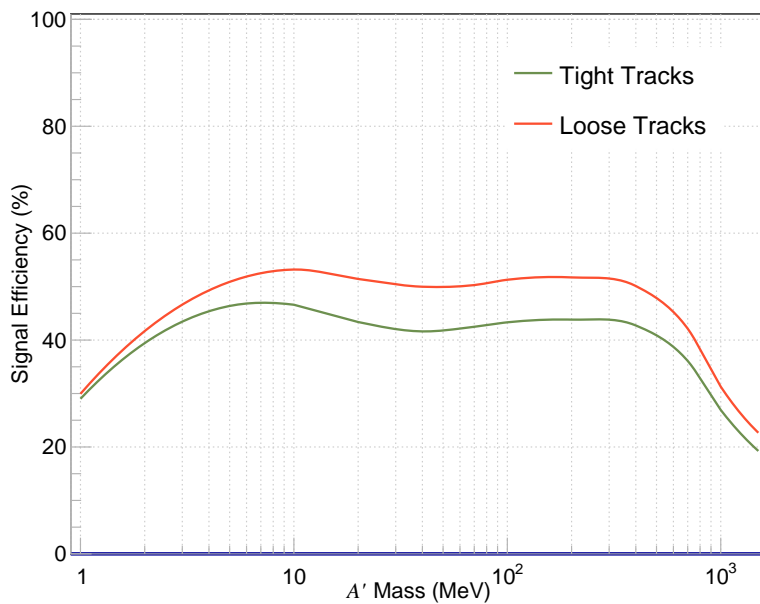
FIG. 71: The signal efficiency after applying all background rejection cuts.

Signal Cutflow

	1 MeV	100 MeV	1000 MeV
	ϵ (%)		
Trigger requirement	57	91	97
ECal Shower-Profile BDT	38	79	82
HCal Max PE < 3	37	76	63
Single track with $p < 1.2$ GeV	31	56	34
Recoil activity Cut	30	54	33
ECal activity cut	29	53	32
Tagger tracker activity cut	29	51	31

TABLE XII: Signal efficiency for 1, 100 and 1000 MeV after successive background rejection cuts.

VIII. PHASE I PROJECTED SENSITIVITY

Combining the preceding discussions of background rejection, signal efficiency, and dark matter production cross-sections, we can estimate the projected LDMX sensitivity for a 4×10^{14} EOT run. As mentioned in Sec. **??**, significant gains in sensitivity can be realized by exploiting the recoil electron pt shape with a larger sample size. For the low rates expected from simulation studies, only mild gains are possible. However, in the event that a significant deviation from the expected background rate is observed, the recoil electron pt spectrum will be a powerful tool for characterizing events and can even be used to constrain the mass of the mediator.

With an expected background of 0.5 events, the expected recoil electron p_T distribution for dark matter signal and backgrounds are shown in Fig. 72. The signal distributions are normalized to the cross section expected from the thermal relic target for the elastic scalar model. Based only on the total expected yields, we obtain the 90% C.L. sensitivity illustrated by the thick blue line in Figure 73. (We note that loosening the track requirements is likely possible for a zero-background search, which enhances the high-mass signal acceptance by 30–50% above what is shown).
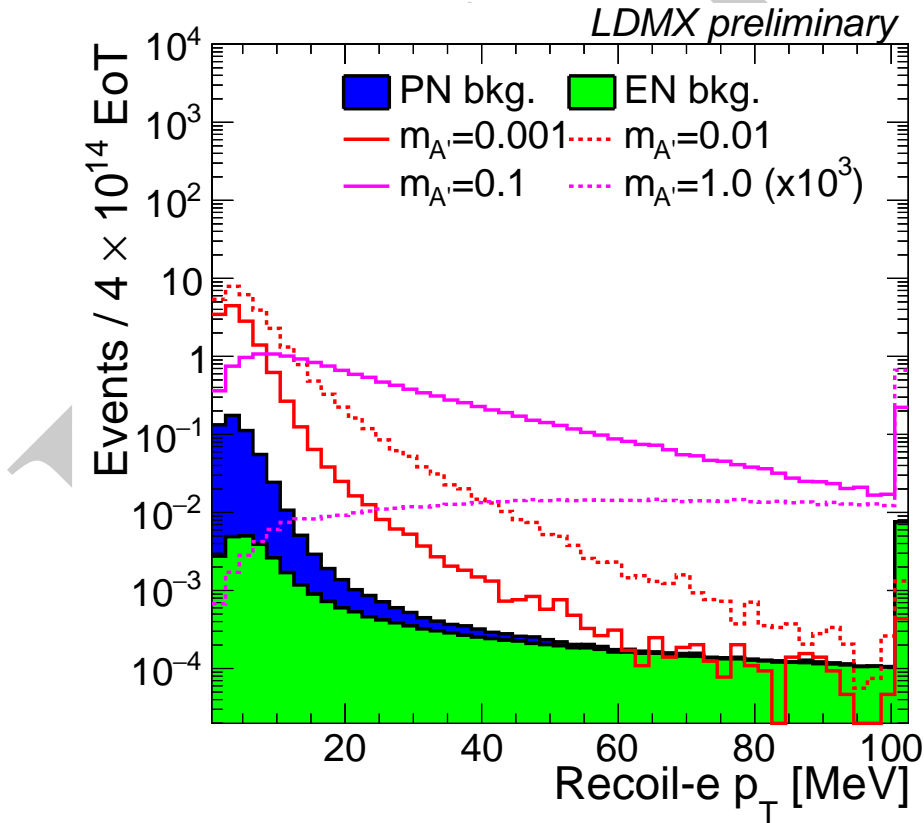
FIG. 72: Distribution of recoil electron transverse momentum for expected backgrounds, solid histograms, and dark matter signals with various mediator masses after all analysis selections. Signal yields are scaled to the elastic scalar thermal relic target.

We expect to reject background to the < 1 event level without any selection on the electron p_T (pending final numbers from GEANT4 resimulation!); however, if unexpectedly larger backgrounds are encountered electron p_T serves as a crucial discriminator and, over much of the mass range of interest, allows us to recover the low-background sensitivity. This is illustrated
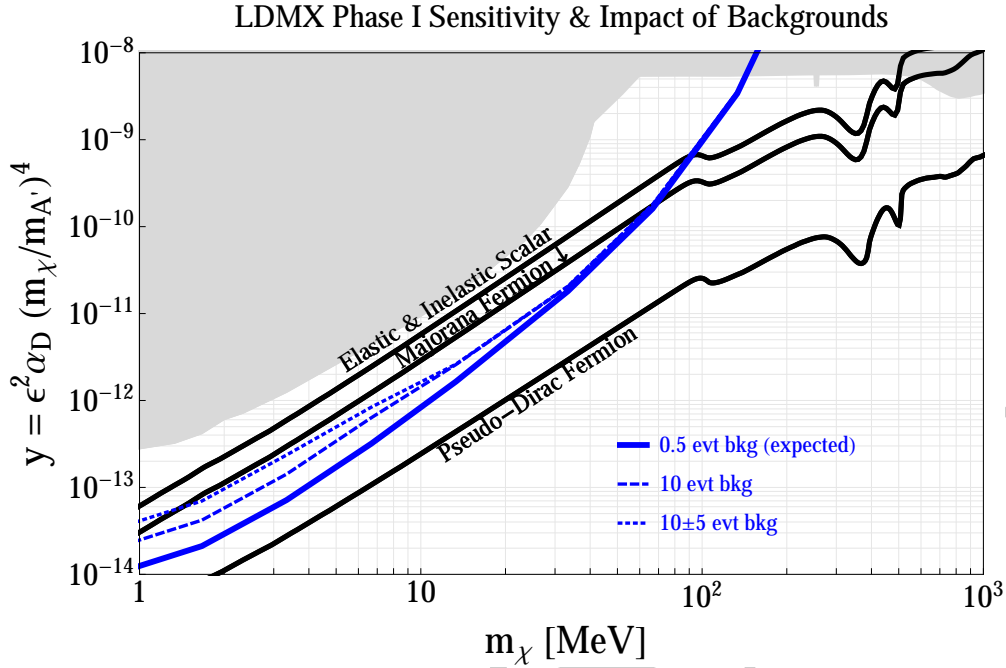
FIG. 73: Projected sensitivity for a 4×10^{14} EOT LDMX run (solid blue curve) and thermal relic targets. The dashed and dotted curves illustrate the robustness of this search to unexpected backgrounds at the 10-event level, with the p_T distribution expected for bremsstrahlung events. In this case, a mass-dependent optimized p_T cut can be used to reduce the background level, recovering nearly the same sensitivity at high dark matter masses. The dotted line further assumes, pessimistically, that the background can only be normalized to within 50% using veto sidebands as control regions.

by the dashed line in Fig. 73, where we have assumed a total background of 10 events with a bremsstrahlung-like electron p_T distribution (accounting for the electrons' scattering in the target). For dark matter masses above ~ 5 MeV, the p_T distribution of signal is substantially broader than the background $p_T$ distribution. A simple p_T cut, optimized separately for each candidate mass, can be used to recover excellent sensitivity to dark matter in the presence of backgrounds. The dotted line further illustrates the effect of systematic uncertainties in the background rate (for the sake of illustration, a 50% uncertainty is shown).

The following plots indicate several alternate interpretations of the LDMX sensitivity. The projected LDMX sensitivity can be interpreted as an *upper bound* on the coupling α_D of light dark matter to the mediator, as a function of mass, if we assume that the product of couplings $\alpha_D \epsilon^2$ lies precisely at the thermal relic line. This interpretation of the expected LDMX sensitivity is shown in Figure 74 for both Majorana and pseudo-Dirac dark matter. These figures also show several complementary direct and indirect constraints: direct searches for dark matter production at BaBar and LSND exclude the coupling ranges shaded in brown (in the case of pseudo-Dirac dark matter, the best existing constraints at low mass come from LSND and NA64, but are below the range of the plot). Demanding that the $U(1)_D$ coupling remain perturbative up to 1 TeV ("TeV perturbativity") leads to a constraint that the running coupling at m_χ must be below the "TeV perturbativity" line. Finally, demanding that the DM self-interaction cross-section lie below $\sigma/m = 1$ cm^2/g excludes the light-gray shaded region (a thin gray line illustrates the weaker constraint $\sigma/m < 10$ cm^2/g).

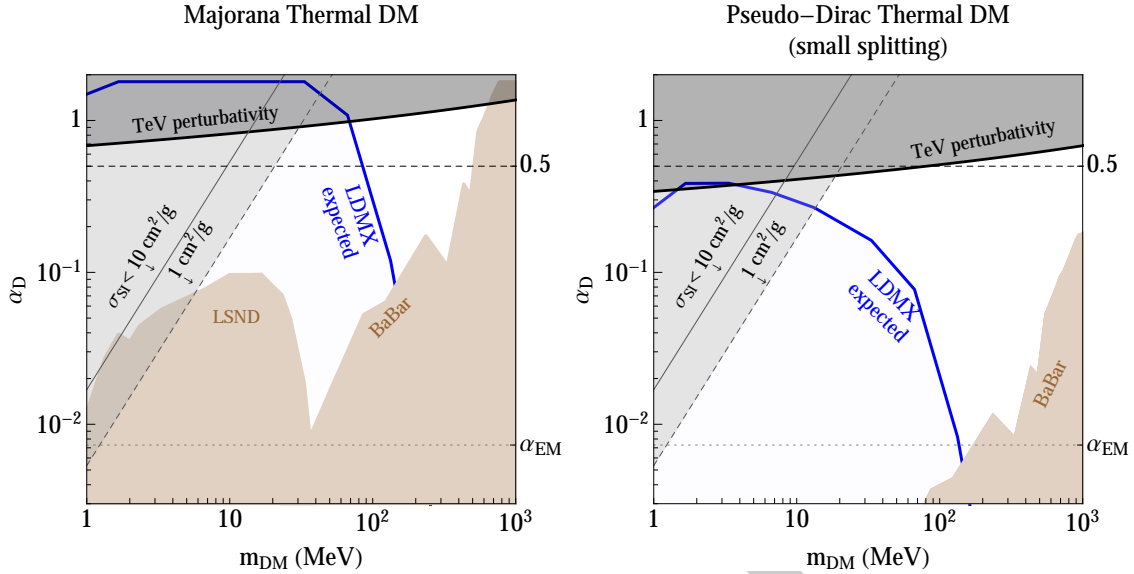Finally, we can map the LDMX sensitivity into an equivalent upper limit on the DM-nuclear

FIG. 74: Projected LDMX lower bounds on α_D vs. m_χ for thermal relic Majorana (left) and pseudo-Dirac (right) fermions. The region below the blue curves will be explored by LDMX phase I. The upper limits on α_D from TeV-scale perturbativity and from dark matter self-interaction bounds of 1 (10) cm^2/g are shown in dark and light gray, respectively. Direct searches lead to *lower bounds* because their yield scales as y/α_D (missing momentum/energy/mass) or y^2/α_D (beam dumps) while the thermal target depends only on y. The most significant constraints from existing accelerator-based experiments — a BaBar missing mass search [39] and a reinterpretation of the LSND neutral-current $\nu - e$ scattering measurement as a bound on light DM [40] — are indicated by the filled brown region extending from the bottom of the plot (for pseudo-Dirac DM, these and the recent NA64 missing energy search [6] are also the most sensitive, but their constraints are outside the range of the plot). LDMX expected sensitivity is indicated by a blue line, with light blue shading and assumes a $4\,10^{14}$ EOT run with 0.5 event background. The direct search constraints are abitrarily cut off at $\alpha_D = 1.8$ for visual clarity.

or DM-electron scattering cross-section vs. mass planes, under the assumption of a specific DM model. In the case of Majorana dark matter, this mapping is straightforward — direct detection for target $T = $ electron or nucleon is $\sigma_{\chi T} = 32\pi \frac{\alpha\alpha_D\epsilon^2\mu_{chiT}^2}{m_A^4}\left(\frac{\mu_{\chi T}v_{rel}}{m_\chi}\right)^2$, where $\mu_{\chi T}$ is the reduced map. In this case, the LDMX sensitivity (again, assuming a mass ratio $m_{A'}/m_\chi = 3$ and $\alpha_D = 0.5$) is shown in Figure 75. The effective cross-section sensitivity *improves* as α_D is decreased; as the mass ratio is increased the LDMX sensitivity region moves down (greater sensitivity) and to the left (lower masses). The effective cross-section sensitivity is model-dependent — for elastic scalar dark matter, LDMX sensitivity maps to higher cross-sections than those shown in the plot, while for inelastic models (scalar or fermion) LDMX sensitivity maps to lower cross-sections.
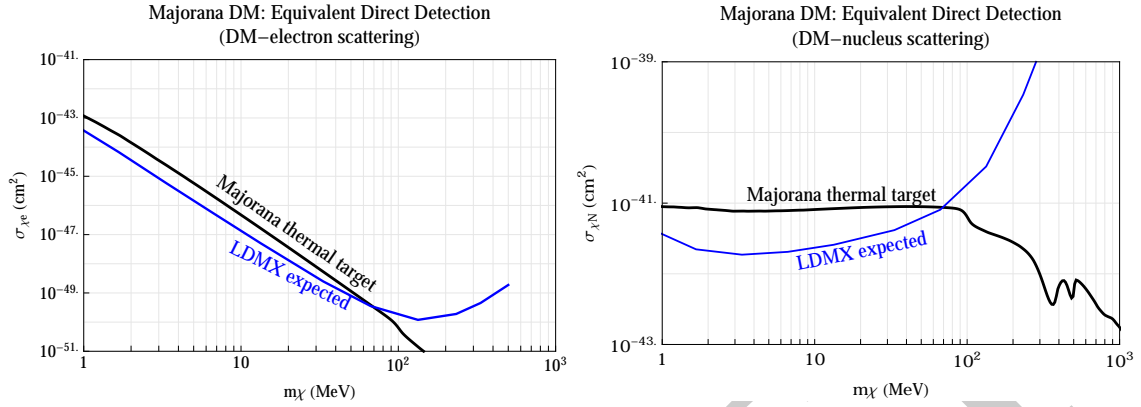
FIG. 75: Projected equivalent upper limits on DM-electron and DM-nucleon scattering, relevant for comparison with direct detection, under the assumption of Majorana dark matter (and hence a q^2 form factor in scattering). The constraints shown assume $\alpha_D = 0.5$ and $m_{A'}/m_\chi = 3$, and generally improve when these parameters are varied over their viable ranges for the thermal model. The Majorana thermal target for this scenario is also shown. This reinterpretation is model-dependent; among the thermal DM models discussed here, the equivalent cross-section sensitivity is greater for pseudo-Dirac and inelastic scalar models, and weaker for scalar elastically scattering DM.

IX.   PHASE II: EXTENDING THE REACH OF LDMX

In the preceding sections of this note we have demonstrated the potential for LDMX to achieve sensitivity to dark photons decaying to dark matter in the mass range below 1 GeV (see [11] for other applications). This was used as a performance driver for our intitial design study. We assumed 4×10^{14} electrons incident upon a $0.1 X_o$ tungsten target at an average rate of one 4 GeV electron per 20 ns. This could be achieved in ∼1 year of operation with a live time of ∼ 10^7 seconds. With this Phase I of the experiment, LDMX would be able to probe much of the thermal freeze-out parameter space for scalar and Majorana dark matter. In this section we discuss ways to extend the reach of LDMX to the challenging pseudo-Dirac fermion target and beyond wherever possible. We take two approaches. First, we consider what might realistically be achieved by means of small changes to the Phase I configuration. This approach allows us to estimate the impacts of the changes by extrapolating from our previous results. In the second approach, we consider more substantial changes to the experiment that could potentially provide significant enhancements, but whose impacts cannot be easily extrapolated from our Phase I studies. Taken together, we believe these two approaches bracket a reasonable range of possibilities and demonstrate the potential for LDMX to uniquely provide coverage for a wide array of dark matter models and other dark sector physics in the critically important sub-GeV mass range. Of course, a more detailed study is required to fully quantify LDMX Phase II performance and this will be the subject of a future note.

A.   Extrapolating from Phase I.

For the thermal-relic scenarios we are considering, the production of dark photons and the final state kinematics of both signal and background events, vary with mass. With this in mind, and in anticipation of other mass related considerations, we break the full mass range of interest into four regions of m_χ defined as follows: $(0.01, 20]$, $(20, 80]$, $(80, 150]$ and $(150, 300]$ MeV, corresponding to $0.03 < m_{A'} \leq 900$ MeV for $m_{A'}/m_\chi = 3$. We will show that the pseudo-Dirac fermion target may be reached for the first three mass ranges by means of simple changes to the Phase I experiment, while for the highest mass range, more significant changes may be required.

The small changes we consider are: (i) increased thickness of the target from $0.1 X_o$ to values ranging from $0.15 X_o$ to $0.4 X_o$, (ii) increased electron beam energy of 8 GeV as currently being planned for the LCLS-II high energy upgrade (LCLS-II HE), and (iii) up to 2 electrons on average per 20 ns. For the two highest mass ranges we also consider (iv) increasing the beam energy to 16 GeV, as would be possible in a CERN implementation of LDMX, and (v) changing the target material from tungsten to aluminum. For the highest mass range we consider (vi) a multiplicity of 5 electrons per 20 ns.

The impact of (i) is straightforward to estimate since the hard bremstrahlung rate is linearly proportional to the target thickness. A potential drawback of a thicker target is increased multiple scattering of the recoil electron. This adds uncertainty to the recoil electron p_t - an important quantity that can help discriminate signal from background, should that be necessary, or provide an estimate of an A' mass associated with candidate signal events. However, since the p_t generated by multiple scattering goes as the square root of target thickness, the change in resolution for small changes in thickness is modest.

We can clearly increase the target thickness in the two highest mass ranges where the recoil electron p_t for signal events is significantly greater than the level of dispersion introduced by multiple scattering in even a $0.4 X_o$ target. For the lower two mass ranges, we consider only

smaller changes to target thickness which have modest impact on p_t resolution.

The impact of (ii) and (iv), raising the beam energy, is also relatively straightforward to estimate since it can directly impact the dark photon (and dark matter) production cross section. The increase is negligible in the lowest mass range and becomes increasingly significant at higher masses, as illustrated in Figure 76.
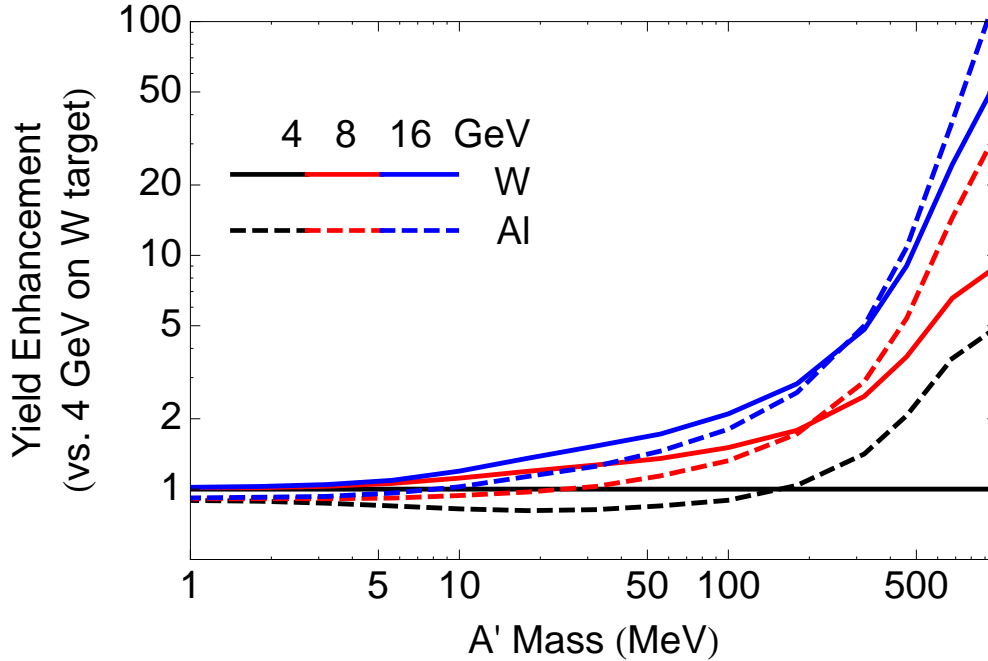

FIG. 76: Both beam energy and target material affect the dark photon production cross-section, with especially large effects at high masses. This figure illustrates how increasing the LDMX beam energy to 8 or 16 GeV, and/or switching from Tungsten to Aluminum targets (at 10% X_0 in each case), impacts the signal production cross-section for different dark photon masses. In each case, we assume the kinematic selection $E_{recoil} < 0.3 E_{beam}$ as was used in 4 GeV studies (this is probably too conservative for higher-energy beams).

One consequence of (iii), doubling the mean number of electrons on target per 20 ns sampling time, is a somewhat more challenging environment for triggering and reconstruction, which should not present any show-stoppers, particularly since the beam spread could also be doubled. Our detailed Phase I studies assumed a beam distribution of $\Delta x \times \Delta y = 2 \times 4\,\text{cm}^2$ at the target. In the DASEL implementation of LDMX this can be extended to $2 \times 8\,\text{cm}^2$, which could compensate for a doubling of the average number of electrons. Nevertheless, it is not as straightforward to extrapolate the impact of this change. The most common signal-like event type would be one in which there is a potential signal, as defined for the case of one electron on target, but now accompanied by another beam electron that showers in the calorimeter. The final state would contain an electron at beam energy, a soft recoil electron and noise in the case of signal, or noise and the remnants of a photonuclear interaction in the case of background. The relatively energetic beam electron would almost always generate an electromagnetic shower in the ECAL, and this shower could overlap with the signals induced by the photon coming from the interaction of a second electron with the target. If the photon has undergone a photonuclear interaction yielding hadrons that do not produce a large signal in the ECAL, then the potential to miss it is increased.

This means that one must either reject all events in which a beam electron in the ECAL is near to where a hard photon is expected to be, or the HCAL must be relied upon to veto such events (which looks feasible given the performance of the HCAL in this note). It follows that increasing the average number of electrons per sample period is likely to impact effective acceptance. For exactly two electrons randomly distributed over a 2×4 cm^2 target area, initial studies indicate that this kind of overlap can occur as often as $\sim 15\%$ of the time for the detector used in our Phase I study. As noted earlier, we can increase the target area to 2×8 cm^2 for which the overlap rate should be cut in half. For simplicity we will assume that this has an associated $\sim 10\%$ reduction in acceptance due to pileup in the ECAL. There is also the potential for pileup to create signal loss from the HCAL. Punch through from the EM shower of the beam electron could induce a veto in the HCAL. For the case of 2 electrons on target, this produces $\sim 1.7~(3.0)\%$ loss of signal for a veto on total energy in the HCAL above a threshold of 8 PE for 4 (8) GeV electrons.

The impact of (v), changing the target from W to a lower Z material such as Al, is an increase in the production cross section for dark photons that we have estimated to be about a factor of at least 2 for masses above about 100 MeV. As the mass increases, the minim momentum transfer $q_{min} = \frac{m_{A'}^2}{2E_{beam}}$ inceases, and as soon as this momentum transfer is comparable to the size of the nucleus, coherence in the scattering degrades resulting in a cross section suppression. For smaller (lower Z) nuclei, the onset of this behavior is pushed to higher mass.

The impact of (vi), increasing the electron multiplicity to an average of 7 electrons per 20 ns, is only considered for a CERN implementation where the beam can be spread over $\Delta x \times \Delta y = 30 \times 4$ cm^2. This is discussed in more detail below.

We now present some potential scenarios for each of the four mass ranges as summarized in Table XIII. For the lowest range, $0.01 \leq m_\chi < 20$ MeV, the signal production rate is relatively high and the pseudo-Dirac fermion target can be reached with a factor of ~ 3 increase in the luminosity relative to Phase I. This could be achieved with ~ 3 years of additional running or one could reduce the required run time by increasing the thickness of the target and/or doubling the mean number of electrons on target. There is no real gain from increased beam energy. Several options for achieving the desired increase in less time are shown in Table XIII.

For $20 \leq m_\chi < 80$ MeV the pseudo-Dirac fermion target can be reached with a factor of ~ 6 increase in luminosity relative to Phase-I. There is a small impact from increased beam energy on production of dark photons. We can also increase the W target thickness slightly to 0.15 X_o and/or double the number of electrons on target.

The $80 \leq m_\chi < 150$ MeV mass region is more challenging, with two orders of magnitude increase in sensitivity relative to Phase I needed to reach the pseudo-Dirac fermion target. Fortunately, in this mass range we gain a factor of ~ 4 in the A' production cross section by switching to an 8 GeV beam. Another factor of 4 is obtained by using a 0.4 X_o W target. This leaves us to find a factor of ~ 8, which can be achieved in several ways. One could double the number of electrons on target and run for ~ 4 years. Alternatively the tungsten target could be replaced by Aluminum to gain a factor of two, cutting the run time to 2 years. Shifting to the CERN implementation of LDMX, one could run at 16 GeV to gain a factor of 4 relative to 8 GeV. With a Tungsten target, one could then cut back to an average of one electron per sample period and run 2 years, or with an Al target one could bring the run time down to ~ 1 year.

The $150 \leq m_\chi < 300$ MeV mass region is the most challenging, with a factor of roughly 3×10^3 increase in sensitivity needed to reach the pseudo-Dirac fermion target at $m_\chi = 300$ MeV. This level cannot be reached even in a run time of 4 years with 16 GeV electrons, a $0.4 X_o$ target, and doubling the average number of electrons per sample period. However, with an Aluminum target, the sensitivity could be reached with pileup of 5 electrons on target. In this case one could

make use of the large beam spread of $32 \times 4\text{cm}^2$ that's possible with the CERN setup. The average electron density at the target would be about one-third that of the Phase I experiment but it would nevertheless be a significantly more complicated environment, where it is harder to maintain high signal efficiencies and high background veto efficiencies. For instance, the rate at which punch through would induce an HCAL veto for events with more than 8 PE is 8 (14) % for 4 (8) GeV beams.

Table XIII lists all of the example scenarios discussed above. Taken at face value, the scenarios imply that the lowest two mass regions could be probed to the level of the pseudo-Dirac fermion target in 2 years of running with a $0.15X_o$ W target, and an average of one 8 GeV electron per 20 ns sample period. The third highest mass range could be probed to the pseudo-Dirac fermion target with an 8 GeV electron beam in 2 years of running with a $0.4X_o$ Al target and an average of 2 electrons per sample period, or 1 year of running with a 16 GeV beam, a $0.4X_o$ Al target, and an average of one electron per 20 ns.

Assuming one could achieve efficient discrimination of signal and noise with 2 electrons per sample period, the first 3 mass ranges can be covered with LDMX at DASEL in as little as $\sim 3-4$ years of running with the upgraded LCLS2-HE accelerator. For $m_\chi = 300$ MeV, one could cover substantial new ground during these same 3 years of running, albeit without reaching the pseudo-Dirac fermion relic target. Belle II is projected to provide good coverage in the highest mass range but it would be prudent to consider how LDMX might also cover this region. This may be achievable with 16 GeV electrons at an average multiplicity of 5 electrons on an Al target, spread over $32 \times 4\text{cm}^2$. This would require $\sim$5 years of running.

TABLE XIII: Extrapolating to Phase II. The "Factor needed" is the increase in luminosity relative to Phase I that is needed to reach the pseudo-Dirac fermion target for masses in the range indicated. The "Factor achieved" is the approximate increase for a combination of changes to the Phase I experiment as indicated. The scenario denoted by "*" is reflected in Figure **??**

Mass Range [MeV]	Factor needed	E_e [GeV]	E_e Factor	Target $[X_o]$ $Mat'l$	Target Factor	μ_e	Years running	Factor achieved
$0.01 \leq M_\chi < 20$	3	-	1	0.2 W	2	1	1.5	
		-	1	0.1 W	1	2	1.5	$\sim$3
		-	1	0.15 W	1.5	2	1	
$20 \leq M_\chi < 80$	6	8	1.5	0.1 W	1	2	1.5	
		8	1.5	0.15 W	1.5	1	2	$\sim$6
		4	1	0.15 W	1.5	2	2	
$80 \leq M_\chi < 150$	10^2	8	4	0.4 W	4	2	4	
		* 8	4	0.4 Al	8	2	2	$\sim 10^2$
		16	16	0.4 W	4	1	2	
		16	16	0.4 Al	8	1	1	
$150 \leq M_\chi < 300$	3×10^3	8	4	0.4 Al	8	2	4	$\sim 2 \times 10^2$
		16	16	0.4 W	4	2	4	$\sim 5 \times 10^2$
		16	16	0.4 Al	8	5	5	$\sim 3 \times 10^3$

B. Further possibilities for improvement

The preceding discussion highlighted the fact that there are a number of fairly straightforward ways to enhance the performance of LDMX. The one exception is increasing the electron multiplicity per 20 ns sample. We mentioned that one way to deal with higher multiplicities is to spread the beam over a larger target area. The LCLS-II configuration would allow a doubling of the target

area from $\Delta x \times \Delta y = 2 \times 4$ cm^2 to 2×8 cm^2. At CERN one would have the ability to distribute the electron beam over an area as large as $\Delta x \times \Delta y = 30 \times 4$ cm^2. A broader beam in the transverse plane would certainly help. However, as alluded to above, this may not be enough to make it possible to make efficient use of the higher electron multiplicities that would allow LDMX to extend to higher masses. It is therefore worthwhile to consider other changes that would help to mitigate the effects of pileup. In this section we present several detector improvements that can reasonably be expected to be beneficial. Qualitative arguments are presented for the extent to which they may improve performance. A quantitative analysis of these possibilities will be the subject of a future note.

The ability to manage higher pileup could be improved by a number of changes to the Phase I concept. Larger beam spreads were discussed above. These would require a somewhat bigger or wider tracker, target and ECAL in the most extreme cases. It could also be of interest to spread the beam in time. In all of the accelerator scenarios under consideration, a 5 ns bunch structure is envisioned, so that 4 bunches would be contained in each 20 ns sampling period. For the cases where the average electron multiplicity is two or more, one could imagine placing half of the mean number of electrons in each of the second and third bunches. The tracker and calorimeters will have the capability to distinguish a 5 ns difference in signal times, thereby adding another dimension to help resolve overlapping objects. More mitigation of pileup effects can be achieved by increasing the granularity of the detector. The ECAL for instance could conceivably have modules with 2-3 times the number of pads. Because of the relatively low occupancy of the HCal, further studies are required to understand how much additional granularity may be needed; the time resolution envisioned of approximately 1 ns may be sufficient to suppress pileup for the HCal.

C.   Phase II Projected Sensitivity

The sensitivity of one representative future scenario, denoted by the "*" line in Table **??**, is compared to the sensitivity of "Phase 1" described throughout this whitepaper, in Figure 77.
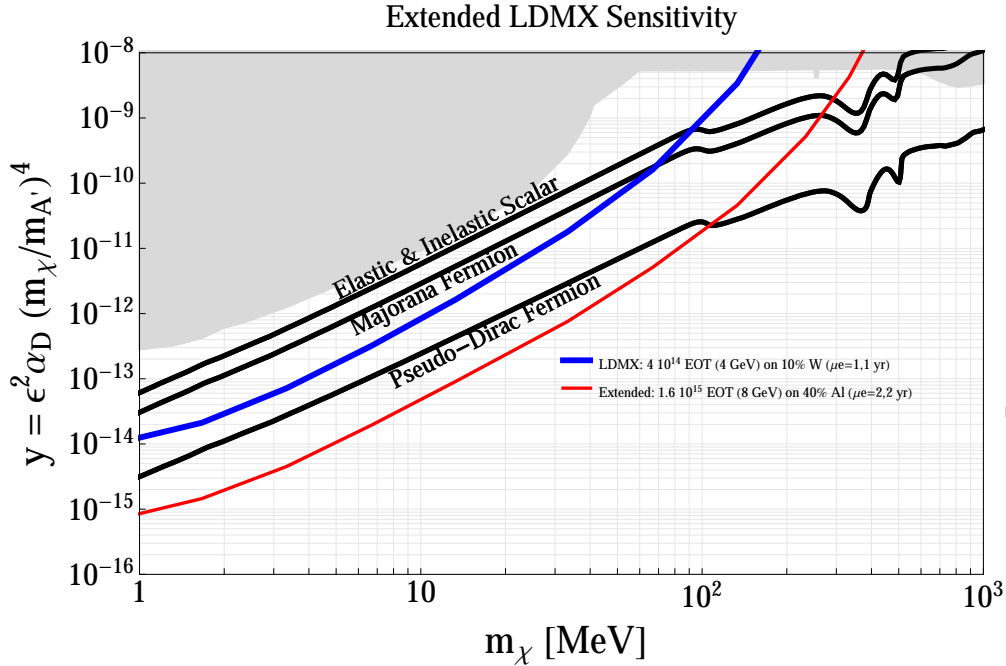
Extended LDMX Sensitivity

FIG. 77: The blue line is the sensitivity of the "Phase I" LDMX discussed throughout this whitepaper, with expected 0.5 background events. A scaling estimate of the sensitivity of the scenario denoted by the "*" line in Table **??** is illustrated by the red line. We have again assumed low background, which appears consistent with the expected reductions in both yield and rejection power at higher energies.

Appendices

Appendix A   PHOTONUCLEAR MODELING: FEW-NEUTRON FINAL STATES AND HARD BACKSCATTERS

In Section V D we highlighted three classes of event that are design drivers for LDMX, and which were found to be mismodeled in the default GEANT4 Bertini cascade model: single forward neutrons, moderate-angle neutron pairs, and hard backscattered hadrons in the "cumulative" region.

Here, we describe in more detail our findings about these classes of events, the data on which we base our estimates for their yields, and the path forward to improve modeling of these final states and phase-space regions in GEANT4.

a   Forward Single Neutrons and Neutron Pairs   The rates of single neutrons and neutron pairs are two design drivers for the HCal. When events have only a single forward neutron (accompanied only by hadrons at energies $< 100\,\text{MeV}$ and/or angles $> 90\,\text{deg}$), non-interaction of that neutron represents a single-point failure of the veto system. Requiring that these be rejected to the < 1 event level determines a lower bound on the depth of the detector. While single neutrons typically go forward due to energy conservation, pairs of neutrons are dominantly produced at nontrivial opening angle, so that their non-interaction probability is determined by the depth *and* transverse size of the HCal.

Neutron pairs can originate from a primary photon reaction on a dineutron $\gamma\,(nn) \rightarrow n\,n$, or from a sequence of reactions (e.g. $\gamma N \rightarrow \pi n, \quad \pi n \rightarrow n\pi$, where the pion backscatters at

low momentum in the 2nd reaction, effectively "converting" into a hard neutron) A single hard neutron can arise from a highly asymmetric 2-body final state — for example for $\gamma(nn) \to nn$ [$\gamma n \to n\pi^0$] with a 3 GeV incident photon, the phase space with $|cos\theta_{cm}| > 0.7$ [$cos\theta_{cm} < -0.77$] leads to a hard forward neutron with the other secondary backscattered. They can also arise from a smaller fraction of phase space for 3-body final states. The rates for these processes *per inclusive photonuclear interaction* for 3 GeV photons impinging on ^{184}W — according to estimates from data and to the default Bertini cascade model — are summarized in Tables **??**. As shown, the rates for these processes in Geant appear to be overestimated by 2 to 3 orders of magnitude! The origins for these discrepancies in the code have been identified, and work is underway on remedying them in collaboration with the Geant collaboration. In the following, we summarize the data used to extract our physical estimates as well as the origin for the unphysical event yields in the Geant4 code.

Primary Process	N_T/nucleus	$\sigma_T^{inclusive}$	Secondary process	Probability	Dineutron Yield
$\gamma(nn) \to \pi n$	300 *(3900)*	7 *(200)* nb	–	–	$1 \cdot 10^{-4}$ *(0.16)*
$\gamma n \to \pi^0 n$	110	0.9 μb	$\pi^0 n \to n\pi^0$	~ 0.01	$\sim 1 \cdot 10^{-4}$
$\gamma p \to \pi^+ n$	110	2.2 μb	$\pi^+ n \to n\pi^+$	~ 0.01	$\sim 1 \cdot 10^{-4}$
Total					$\lesssim \mathbf{3 \cdot 10^{-4}}$ *(0.16)*

TABLE XIV: Estimated yield of dineutron events *as a fraction of inclusive photonuclear reactions* for the exclusive processes discussed in the text, in the benchmark case of a 3 GeV photon impinging on ^{184}W. Where we have found discrepancies between the Bertini Cascade model and our calculations, the rate in the Bertini Cascade model is given in parentheses. Crucially, the yield for two-neutron final states from dineutron dissociation is overestimated in Geant4 by a factor of at least ~ 500, due to several factors discussed in the text. The cross-section of 7 nb is obtained by numerically integrating and interpolating deuteron dissociation cross-section data from [41]; if the dineutron dissociation cross-section is comparable to diproton dissociation yields measured in [42] a further suppression by 1/20 may be expected.

The yields of the two-step processes should be modeled accurately in the Bertini Cascade Model (detailed investigation is ongoing); rates given here are simply a rough estimate. To estimate the "probability" for the secondary reaction to convert a pion into a soft neutron is estimated as the ratio of the exclusive cross-section for elastic back-scattering ($cos\theta_{cm} < -0.75$) to the total pion-nucleon cross-section, at a reference energy of 1.5 GeV. These estimates suggest that cascades of single-nucleon processes have a higher rate than primary dineutron dissociation, and that the total yield of dineutron final states is overestimated by a factor of ~ 500. Note that neither our estimates nor the GEANT4 rates listed above account for the survival probability of the final-state neutrons (i.e. the probability that it exits the nucleus before scattering into a multi-particle final state), which will mildly suppress the dineutron rate relative to the above.

The rates for the relevant few-body photon-nucleon processes — inclusively and in the single-neutron phase space — are readily extracted from data. Single-pion photoproduction cross-sections off protons and neutrons and their angular distributions have been measured in e.g. [44, 45] at multi-GeV photon energies; the backscatter kinematics exhibits a universal form of $k^3 d\sigma/du$ for $E_\gamma > 4$GeV [43]. Neglecting the $< 20\%$ effect of shadowing for 3–4 GeV photons in heavy nuclei [46], these can simply be multiplied by A to estimate the per-nucleus production cross-section.

A similar approach is possible for two-body dinucleon dissociation, albeit with greater uncertainties in both the dissociation cross-section and effective number of dinucleons in a large nucleus.

Process	N_T/nucleus	$\sigma_T^{inclusive}$	kin. requirement	frac. in $1n$ kinematics	Yield in $1n$ kinematics
$\gamma(nn) \to n\,n$	300 *(3900)*	7 *(200)* nb	$\|q\| > 0.7$	0.9 *(0.3)*	$1 \cdot 10^{-4}$ *(0.05)*
$\gamma(pn) \to p\,n$	400 *(2600)*	7 *(200)* nb	$q < -0.7$	0.3 *(0.15)*	$3 \cdot 10^{-5}$ *(0.015)*
$\gamma p \to \pi^+ n$	74	2 μb	$q < -0.77$	0.02 *(0)*	$1\,10^{-4}$ *(0)*
$\gamma n \to \pi^0 n$	110	1 μb	$q < -0.77$	0.04 *(0)*	$2\,10^{-4}$ *(0)*
Total					$\sim \mathbf{4\,10^{-4}}$ *(0.06)*

TABLE XV: Estimated yield of effective single-neutron final states (with any additional neutrons produced backwards in the lab-frame) for the exclusive processes discussed in the text, in the benchmark case of a 3 GeV photon impinging on ^{184}W. Where we have found discrepancies between the Bertini Cascade model and our calculations, the rate in the Bertini Cascade model is given in parentheses. Again the yield of such events in the Bertini Cascade model is dominated by an unphysically large contribution from dinucleon dissociation. The physical contributions of these reactions are obtained by numerically integrating and interpolating deuteron dissociation cross-section data from [41]; however the forward peak that dominates the single-neutron kinematic region in is likely not present in dineutron dissociation. If the dineutron dissociation cross-section is comparable to diproton dissociation yields measured in [42], a further suppression by $\sim 1/20$ may be expected.

Although GEANT4 overestimates the dinucleon dissociation contributions to this final state, it *underestimates* the contributions of πn final states by neglecting the u-channel pole in these amplitudes, the cross-section for which we have obtained by extrapolating the scaling of back-scattering cross-sections reported in [43] from ≥ 4 GeV down to our reference energy of 3 GeV. The yields of two-step processes (e.g. $\gamma p \to \pi^+ n$, where a forward pion backscatters off a neutron) are not shown — estimates analogous to those in Table **??** indicate that they are subdominant.

These estimates suggest that the total yield of single-neutron-like final states is overestimated by a factor of $\mathcal{O}(100)$ in the Bertini model, but the dominant reaction giving rise to such events is absent from the model. Note that neither our estimates nor the GEANT4 rates listed above account for the survival probability of the final-state neutron (i.e. the probability that it exits the nucleus before scattering into a multi-particle final state), which will mildly suppress the single-neutron rate relative to the above.

The effective number of quasideuterons P_{QD} in heavy nuclei can be inferred from data via photonuclear interaction cross-sections with ~ 100 MeV photon energies (see e.g. [47] Figure 1), and is usually parametrized as $\mathcal{P}_{QD} = L\frac{Z(A-Z)}{A}$, where the "Levinger factor" L agrees within $\sim 20\%$ accuracy with the Local Density Approximation (LDA) expectation $L = 10.83 - 9.76A^{-1/3}$ (for Tungsten, $L \approx 9$) [48]. The differential cross-section for deuteron dissociation has been measured with excellent angular coverage up to 3 GeV, and the differential cross-section for diproton dissociation (in ^3He) has been measured at specific kinematics up to 4.7 GeV [41, 42]. The 2-body dissociation cross-section for photon energies above 2GeV was found to be a factor of 20 smaller for diprotons than for deuterons, which may be attributable to a cancellation between two scattering amplitudes in the hard re-scattering model (HRM) [**?**]. We are not aware of any direct measurements of the dineutron dissociation cross-section; while it seems most reasonable to assume that this is of the same scale as the diproton dissociation cross-section, we will use the deuteron dissociation cross-section for our estimate instead to avoid potentially underestimating yields.

We have also estimated the contributions to dineutron final states of cascades of reactions on

single-nucleon targets, where the first reaction produces a pion-neutron final state, and the pion subsequently elastically scatters off a neutron (inelastic processes such as $\pi^- p \to \pi^0 n$ are significantly lower in rate). Such estimates should not be trusted beyond the order-of-magnitude level, as the actual rate for these cascades depends on the distribution of energy, angles, and intra-nuclear location of the initial reaction — i.e. a more careful calculation must rely on a full intranuclear cascade model. The fact that these processes are not negligible for the final states of interest does nonetheless underscore the importance of using a cascade model for these studies.

In the standard Bertini model, both single- and di-neutron final states are overwhelmingly dominated by dineutron dissociation, for which we have found a dramatically lower rate. This large discrepancy can be traced to three errors in the GEANT4 code:

- The parametrized cross-section for dinucleon dissociation exceeds the measured deuterium 2-body dissociation cross-section starting around 500 MeV photon energies, with the discrepancy growing dramatically at multi-GeV energies, where the dissociation cross-section is expected to scale as $\sim E_\gamma^{-10}$. For example, the cross-section for two-body photodisintegration of deuterium at 2.8 GeV has been measured at Jefferson Lab to be about 5 nb[41]; the Bertini cascade code uses a cross-section of 200 nb for 1.8 GeV $< E_\gamma <$ 5.6 GeV.

- The dineutron density used in the Geant4 model is a factor of ~ 7 larger than the $\mathcal{P}_{nn}$ given above, because all neutrons within a spherical shell zone are allowed to pair with one another to form dinucleons. We note that the quasideuteron number $\mathcal{P}_{pn}$ has been extracted from data in multiple ways, and from a variety of nuclear models, with mutual agreement within $\sim 20\%$ for all of these models.

- The selection of targets (proton, neutron, or dinucleon) for photonuclear interaction in the Bertini Cascade monte carlo is not proportional to the *physical* path lengths for these interactions, but to the path lengths when an interaction is forced — effectively rendering all reactions equally likely irrespective of their specified cross-sections.

In addition, we have found that the processes we believe to be dominant contributions to single-neutron final states, $\gamma N \to \pi n$ where the pion backscatters, are underestimated in the Bertini model. While the total rate for this process is correct, the angular distribution used for $E_\gamma > 1.77$ GeV is simply a parametrization of the t-channel pole. The u-channel pole contribution at few-GeV energies accounts for only 1–2% of the total cross-section, but is highly relevant for the LDMX veto.

b High-Energy Backscatters In large-scale simulations of photonuclear reactions for LDMX, we identified a surprisingly large number of events where much of the incident photon's energy was carried by a hadron produced at backward angles. Conservation of momentum requires that, for such a reaction to occur, the particle must recoil against a multi-nucleon composite which carries multi-GeV forward momentum but negligible energy. The physics of such "cumulative" hadron production is related to intranuclear correlations. Cumulative hadron production has been studied in many experiments — it has been found to be well described by a near-exponential distribution. Published data on cumulative hadrons from photonuclear reactions is limited to Copper and Lead targets and final-state hadron kinetic energies below 0.3 GeV [**?**]. However, there is substantial evidence that the kinematics of cumulative hadrons is only weakly sensitive to the incident projectile and the nucleus [49]. Data on cumulative hadron production in scattering of 400 GeV protons [49] (up to $p = 1.52$GeV [$T = 0.85$GeV] for the emitted hadron), in 14.5 GeV incident electronuclear reactions[36] (up to $T \approx 0.6$GeV), and in 5 GeV incident electronuclear

FIG. 78: Left: Distribution of proton momenta at $\theta_p > 100°$ in 5 GeV electron scattering off ^{208}Pb. The black histogram shows the measured yield in a skim of CLAS eg2 data [35], for events with the electron in the CLAS acceptance. The blue histogram shows the results of a GEANT4 simulation before final-state down-weighting. Right: Distribution of final-state proton momenta at angles from $30°$ to $160°$ for three different reactions (from [36, 49?]), illustrating the approximate universality of these distributions' shapes, and from GEANT4 simulations of photonuclear reactions for LDMX illustrating the unphysical high-energy tail.
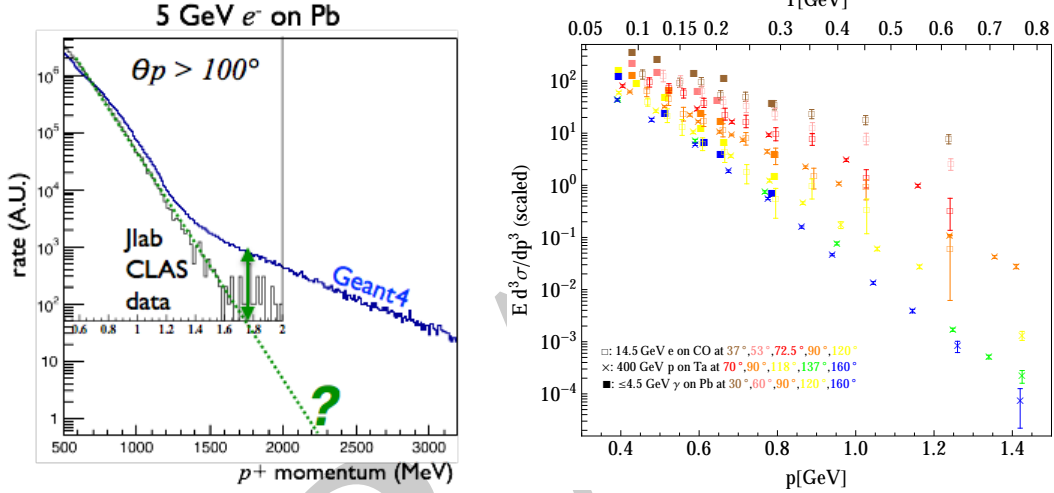
interactions [?] (up to $p = 2\text{GeV}$ [$T = 1.25\text{GeV}$]) all show an exponentially falling hadron energy distribution at angles above $60°$, with a steeper fall-off at more backwards angles throughout the measured kinematic range.

By contrast, GEANT4 simulations of photonuclear and electronuclear reactions (with the physics list described above) show clear excesses over this exponential behavior for $p \gtrsim 1 - 1.3\text{GeV}$. This pattern of over-populating the wide-angle/backwards high energy tails is illustrated in Figure 78. The left panel compares the distribution of back-scattered protons ($> 100°$ polar angles) in electronuclear scattering from the CLAS eg2 experiment at Jefferson Lab [35] to a GEANT4 simulation with the same incident electron energy and target material (but not the same event selection). The right panel shows the back-scattered and wide-angle proton yields in exclusive angle bins, for three different datasets from very different reactions, illustrating in addition both the the angle-dependence of the physical distribution and the universality of these distributions for different reactions. This universality is further discussed in [36, 49].

The artifact seen in Fig. 78 (left) affects only a small fraction of photonuclear events (around 10^{-4}), but events far out on this tail typically have no high-energy forward-going hadrons. Thus, modeling these events more accurately informs the balance in designing the hadron veto between vetoing multi-GeV hadrons and those with much lower energies, as well as a potential floor to the experiment's sensitivity.

By resimulating these events with verbose debugging output turned on in the Bertini cascade, it was found that almost all of the events on the unphysical "cumulative tail" arose from a common (unphysical) mechanism: The GEANT4 Bertini Cascade models the nucleus as a series of nested shells or "zones" with different densities, Fermi momenta, and potentials for nucleons in each region. Accurately propagating low-energy hadrons in the nucleus requires a model for transport from one region to another, with reflections at the zone boundary if they do not have sufficient energy to penetrate the boundary. In a thin-wall model, the condition for penetrating the boundary

depends only on the radial component of the momentum. Therefore, energetic hadrons hitting the boundary at grazing incidence, with large transverse momentum but negligible radial momentum, would reflect repeatedly until they interacted, transferring momentum to the remaining nucleus in the process. This mechanism for reflection and momentum transfer is clearly unphysical — in a realistic (continuously varying) nuclear potential, the angular momentum of a multi-GeV hadron is always more than sufficient to escape the nuclear potential.

We have implemented a modified reflection condition that accounts (conservatively) for the contribution of angular momentum to penetration of the nuclear potential barrier, and are testing this in collaboration with the SLAC GEANT4 group. By re-generate small numbers of events with this modification, we have verified that the tails shown in Fig. 78 (left) is removed. We are still testing this code to see whether the more physical contributions to the cumulative tail, arising from the non-trivial momentum distribution of nucleons in the nucleus, is consistent with data, and plan to eventually incorporate this fix into the main LDMX photonuclear (and electronuclear) simulations.

For the larger-scale simulations, we have implemented a post hoc reweighting scheme in which events with hard nucleons at $\theta > 100°$ and $w \equiv (p + T)/2(\sqrt{1.25} - 0.5\cos\theta) > 1400$ MeV are downweighted to a uniform exponential, removing the "ankle" in the distribution (the form of the weighting function w is motivated by the phenomenological model of [36]). This reweighting is reflected in the event counts throughout the paper, but only affected a fraction of the unphysical events.

[1] J. Alexander et al. (2016), 1608.08632, URL http://inspirehep.net/record/1484628/files/arXiv:1608.08632.pdf.

[2] M. Battaglieri et al. (2017), 1707.04591.

[3] E. Izaguirre, G. Krnjaic, P. Schuster, and N. Toro, Phys. Rev. **D91**, 094026 (2015), 1411.1404.

[4] E. Izaguirre, G. Krnjaic, P. Schuster, and N. Toro, Phys. Rev. Lett. **115**, 251301 (2015), 1505.00011.

[5] D. Banerjee et al. (NA64), Phys. Rev. Lett. **118**, 011802 (2017), 1610.02988.

[6] D. Banerjee et al. (NA64) (2017), 1710.00971.

[7] S. N. Gninenko, N. V. Krasnikov, M. M. Kirsanov, and D. V. Kirpichnikov (2016), 1604.08432.

[8] DASEL Design Team, Tech. Rep. (2016), URL https://confluence.slac.stanford.edu/download/attachments/213891196/DASEL%20LOI.pdf?version=2&modificationDate=1473892060000&api=v2.

[9] C. W. Leemann, D. R. Douglas, and G. A. Krafft, Ann. Rev. Nucl. Part. Sci. **51**, 413 (2001).

[10] T. Akesson et al. (PBC-acc-e-beams working group) (2018), physics.acc-ph/1805.12379.

[11] A. Berlin, N. Blinov, G. Krnjaic, P. Schuster, and N. Toro (2018, to appear).

[12] S. Ritz et al. (HEPAP Subcommittee) (2014).

[13] M. Viel, G. D. Becker, J. S. Bolton, and M. G. Haehnelt, Phys. Rev. **D88**, 043502 (2013), 1306.2314.

[14] K. Griest and M. Kamionkowski, Phys. Rev. Lett. **64**, 615 (1990).

[15] B. W. Lee and S. Weinberg, Phys. Rev. Lett. **39**, 165 (1977).

[16] M. S. Madhavacheril, N. Sehgal, and T. R. Slatyer, Phys. Rev. **D89**, 103508 (2014), 1310.3815.

[17] P. A. R. Ade et al. (Planck), Astron. Astrophys. **594**, A13 (2016), 1502.01589.

[18] A. Berlin, N. Blinov, S. Gori, P. Schuster, and N. Toro, Phys. Rev. **D97**, 055033 (2018), 1801.05805.

[19] E. Izaguirre, G. Krnjaic, P. Schuster, and N. Toro, Phys. Rev. **D88**, 114015 (2013), 1307.6554.

[20] P. Hansson Adrian et al. (HPS Collaboration), *Heavy Photon Search Experiment at Jefferson Labo-*

ratory: proposal for 2014-2015 run (2013), URL `https://confluence.slac.stanford.edu/download/attachments/86676777/hps_2014.pdf`.

[21] CMS Collaboration, Tech. Rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02 (2015), URL `https://cds.cern.ch/record/2020886`.

[22] G. Battistoni, S. Muraro, P. R. Sala, F. Cerutti, A. Ferrari, S. Roesler, A. Fasso, and J. Ranft, AIP Conf. Proc. **896**, 31 (2007), [,31(2007)].

[23] A. Artikov et al. (Mu2e), Nucl. Instrum. Meth. **A890**, 84 (2018), 1709.06587.

[24] J. Mans et al. (CMS), Tech. Rep. CERN-LHCC-2012-015, CMS-TDR-010 (2012), URL `http://cds.cern.ch/record/1481837`.

[25] S. Agostinelli et al. (GEANT4), Nucl. Instrum. Meth. **A506**, 250 (2003).

[26] R. Brun and F. Rademakers, Nucl. Instrum. Meth. **A389**, 81 (1997).

[27] R. Chytracek, J. McCormick, W. Pokorski, and G. Santin, IEEE Transactions on Nuclear Science **53**, 2892 (2006).

[28] J. Alwall et al., Comput. Phys. Commun. **176**, 300 (2007), hep-ph/0609017.

[29] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, JHEP **09**, 028 (2007), 0706.2334.

[30] Y.-S. Tsai, Phys. Rev. **D34**, 1326 (1986).

[31] S. Abrahamyan et al. (APEX), Phys. Rev. Lett. **107**, 191804 (2011), 1108.2750.

[32] S. Andreas, C. Niebuhr, and A. Ringwald, Phys. Rev. **D86**, 095019 (2012), 1209.6083.

[33] D. H. Wright and M. H. Kelsey, Nuclear Inst. and Methods in Physics Research **804**, 175 (2015).

[34] D. Wright, private communication (2015).

[35] S. Stepanyan, private communication (2016).

[36] P. V. Degtyarenko and others, Phys.Rev. **C50**, R541 (1994).

[37] Y.-S. Tsai, Reviews of Modern Physics **46**, 815 (1974).

[38] R. L. Craun and D. L. Smith, Nuclear Instruments and Methods) **80**, 239 (1970).

[39] J. P. Lees et al. (BaBar), Phys. Rev. Lett. **119**, 131804 (2017), 1702.03327.

[40] P. deNiverville, M. Pospelov, and A. Ritz, Phys. Rev. **D84**, 075020 (2011), 1107.4580.

[41] M. Mirazita and T. C. collaboration, arXiv.org p. 014005 (2004), nucl-ex/0405019.

[42] I. Pomerantz, J. Glister, G. G. Petratos, R. Lindgren, Y. Qiang, S. Beck, R. Gilman, C. Dutta, K. Chirapatpimol, K. Allada, et al., Phys.Lett. **B684**, 106 (2009).

[43] R. Anderson, D. Gustavson, J. Johnson, I. Overman, D. Ritson, and B. Wiik, Phys.Rev.Lett. **23**, 721 (1969).

[44] O. Bartholomy and others, Phys.Rev.Lett. **94**, 012003 (2005).

[45] L. Y. Zhu and others, Phys.Rev.Lett. **C71**, 044603 (2005).

[46] S. Michalowski, D. Andrews, J. Eickmeyer, T. Gentile, N. Mistry, R. Talman, and K. Ueno, Phys.Rev.Lett. **39**, 737 (1977).

[47] M. Anghinolfi, V. Lucherini, N. Bianchi, G. P. Capitani, P. Corvisiero, E. D. Sanctis, P. D. Giacomo, C. Guaraldo, P. Levi-Sandri, E. Polli, et al., Nuclear Physics A **457**, 645 (1986).

[48] O. Benhar, A. Fabrocini, S. Fantoni, A. Y. Illarionov, and G. I. Lykasov, Phys.Rev.Lett. **C67**, 014326 (2003).

[49] Y. D. Bayukov, V. I. Efremenko, S. Frankel, W. Frati, M. Gazzaly, G. A. Leksin, N. A. Nikiforov, C. F. Perdrisat, V. I. Tchistilin, and Y. M. Zaitsev, Physical Review C (Nuclear Physics) **20**, 764 (1979).

[50] J. L. Feng and J. Smolinsky, Phys. Rev. **D96**, 095022 (2017), 1707.03835.

[51] Near $m_{A'} \approx 2m_\chi$, the $\chi\chi \to \bar{f}f$ annihilation process is near resonance, so the relic target for this region of parameter space becomes sharply sensitive to the DM/mediator mass ratio, so the relic targets presented in Fig. 4 move down in the y parameter space (see Supplemental Material in [**?**] and [50]

for more details). However, away from the resonant region where $m_{A'} \gtrsim 2m_\chi$, the relic targets in Fig. 4 are robustly insensitive to the mass and coupling ratios that define the y variable

[52] Though keep in mind that the freeze-out targets rapidly decreases on the y vs. m_χ plot in the fine tuned region of parameter space near $2m_\chi = m_{A'}$.