

**Searches for Dark Matter Production in Two Fixed Target
Experiments: HPS and LDMX**

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Thomas Brice Eichlersmith

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Jeremiah Mans

Jan, 2025

© Thomas Brice Eichlersmith 2025
ALL RIGHTS RESERVED

Acknowledgements

There are many people that have earned my gratitude for their contribution to my time in graduate school. While I will struggle to name them all, I do wish to specifically thank those people who have helped teach me their knowledge of physics along the way: Tim Nelson, Cam Bravo, Matt Graham, PF Butti, Omar Moreno, Philip Schuster, Natalia Toro, Stefania Gori, Lene Kristian Bryngemark, Einar Elén, Erik Wallin, Christian Herwig, Nhan Tran, Cristina Mantilla Suarez, Lauren Tompkins, Valentina Dutta, Ruth Pöttgen, and Alic Spellman.

Of particular note is my advisor Jeremy Mans. Professor Mans has been wonderfully flexible as I have changed course while navigating the uncertain waters of graduate school. Thank you.

Finally, there are many friends who I have met while in graduate school or with whom I have gotten closer. Your support outside of physics has been necessary for me over the years and I am in your debt.

Dedication

To my friends and family trying to comprehend anything I have done over the past six and a half years, I hope this helps.

To a future graduate student desparately trying to understand these experiments and their code, I also hope this helps and I wish you luck.

Abstract

Astrophysical evidence strongly indicates the presence of particulate Dark Matter (DM) within our universe; however, the specific particle nature of DM is still unknown. The wide variety of possible DM particles produces a similar range of experiments focused on probing these different categories of possible DM. This work describes two experiments taking different approaches to search for Light DM residing in the 1MeV-1GeV mass range being produced by electron interactions. The Light Dark Matter eXperiment (LDMX) is a proposed fixed target experiment designed for a missing momentum search with an additional, orthogonal missing energy search channel described here. The Heavy Photon Search experiment (HPS) is another fixed target experiment designed for a displaced vertex search with distances of O(10cm) which are not probed by longer baseline experiments. Specifically, a search in HPS data for a specific Light DM model with a strongly-coupled dark sector enabling a higher expected production rate while keeping the characteristic decay length within HPS's acceptance is also presented.

Contents

Acknowledgements	i
Dedication	ii
Abstract	iii
Contents	iv
List of Tables	vii
List of Figures	ix
I Laying the Groundwork	1
1 Introduction	2
1.1 Standard Model of Particle Physics	4
1.2 Standard Model Analogies	8
1.2.1 Particle Mixing	9
1.2.2 Displaced Particle Decays	9
2 Dark Matter	12
2.1 Dark Matter?	13
2.2 Evidence for Dark Matter	13
2.3 Particle Nature of Dark Matter	14
2.4 Invisible Signature	20

2.5	Visible Signature	21
II	LDMX	22
3	Light Dark Matter eXperiment	23
3.1	Missing Momentum Signature	23
3.2	The Beam Line	25
3.3	Detector Design	26
4	Mid-Shower Simulation	31
4.1	General Data Process	31
4.1.1	Data Processing Stages	33
4.1.2	Biasing and Filtering Technique	34
4.2	Standard Processes	36
4.2.1	Dominant Contributors in this Analysis	37
4.2.2	Expanding Production of these Processes	39
4.3	Dark Matter Signal	41
4.3.1	G4DarkBreM	42
4.3.2	MadGraph/MadEvent	43
4.3.3	Characterization	43
4.4	Summary	44
5	Missing Energy Search	46
5.1	Selection	47
5.2	Background Prediction	51
5.3	Systematics and Background Uncertainty	55
5.4	Reach	59
III	HPS	61
6	The Heavy Photon Search Experiment	62
6.1	Strongly Interacting Massive Particles	62
6.2	Detector Apparatus	65

6.2.1	Silicon Vertex Tracker	65
6.2.2	Electromagnetic Calorimeter	67
7	Alignment	70
7.1	Alignment Procedure	73
7.2	Detector Visualization	75
7.3	Results	75
8	Data Set	80
8.1	Collected Data	80
8.1.1	Pair 1 Trigger	81
8.2	Simulation	82
8.3	Reconstruction	83
8.4	Analysis Pre-Selection	83
9	Displaced Vertex Search	88
9.1	Signal Yield	89
9.2	Reconstruction Categories	93
9.3	Physics Variables	94
9.4	Selection	98
9.5	Results	100
9.5.1	Search	100
9.5.2	Sensitivity and Exclusion	103
IV	Conclusion	106
10	Conclusion	107
10.1	Simulation Validity	108
10.2	Tracking Reconstruction	109
10.3	Reflection	109
References		110

List of Tables

4.1	Parameters of mid-shower biased samples. In all of the samples used in this study, we use the same value for the Sorting and Biasing Thresholds.	36
4.2	Configuration of the simulation samples used in this analysis. E_{beam} is the beam energy being studied (4 or 8 GeV in this work). m_A is the mass of the A' in MeV, ϵ is the dark brem mixing strength, $E_{\text{primary}}^{\text{ECal Front}}$ is the energy of the primary electron at the front of the ECal, $E_{A'}$ is the energy of the generated A', $E_{\text{tot nuc}}$ is the total energy transferred to nuclear interactions during the event, and $E_{\text{tot } \mu}$ is the total energy of produced muons.	45
5.1	Cut-flow analysis comparing background and various signal hypotheses for the simple cuts used in this analysis. The event yield for the background sample is calculated using the event weights and represent the number of events out of 10^{13} EoT equivalent. The signal efficiency is relative to the full simulation sample. The efficiency and event yield values on a given row are for <i>after</i> the analysis stage of that row. The first table is the cutflow for the 4 GeV beam, and the second table is for the 8 GeV beam.	51
8.1	Values for cuts used within the Pair 1 Trigger during data collection of the HPS 2016 dataset.	82
9.1	Polynomial coefficients for radiative fraction and acceptance functions.	91
9.2	Parameters for Equation (9.14).	100

List of Figures

1.1	Comic #2351 from XKCD[1]. If able to be designed from scratch, the names of things in the standard model could be more helpful.	2
1.2	The Standard Model of Particle Physics showing the twelve fermions and five bosons, their various properties (mass, charge, spin), labels (box and circle colors), and interactions (brown loops). Credit to Cush on wikipedia for providing this diagram freely accessible and usable for any purpose.	5
1.3	Feynman diagram for the bremsstrahlung process where a charged lepton (ℓ^-) interacts with a nucleus (Z) via a photon (γ) and then emits another photon before recoiling.	7
1.4	Energy of the outgoing electron (E_e) for standard processes (dominated by bremsstrahlung) in gray and some dark bremsstrahlung (see Section 2.4) in colors. Figure 3 from [2].	8
1.5	Dipiction of neutral kaon and anti-kaon mixing including the underlying Feynman diagram that results from applying the cro:smStandard Model (SM) to this system. This figure was created by user NikNaks on Wikipedia and is licensed under CC Attribution-Share Alike 3.0 Unported.	10
1.6	Image of CERN's first liquid hydrogen bubble chamber from 1960 [3] with pink annotations added. The pink circle highlights the location described in the text where a Λ baryon decays into two particles (most likely a proton and a negative pion). The pink dashed line traces out a possible path the Λ traveled before its decay.	11

2.1	Comic #1758 from XKCD[1]. Generally, claiming that either Einstein or Newton was wrong will be a bad time.	12
2.2	Depiction of the velocity of stars within a galaxy as a function of their distance from the galactic center. The dotted line is a prediction of this relationship using cro:grGeneral Relativity (GR) along with the mass tabulated from the visible starts while the data points (and the solid line fitted to them) are what are actually seen in galaxies today.	15
2.3	From [4], the co-moving cosmological number density of cro:dmDark Matter (DM) as a function of the universe temperature. As the universe cools, the number density decreases until the DM becomes too sparse to interact with other DM particles, “freezing” to a specific number density until today.	17
2.4	Mass scale of Thermal Relic DM. The regions in red are excluded by applying the thermal relic assumption to our observations of the universe’s early evolution. m_e is the electron mass and m_p the proton mass.	18
2.5	Feynman diagram of how a massive field Φ could allow for a standard photon (γ) to mix with a dark photon (A').	18
2.6	Feynman diagram for the dark brem process.	19
3.1	Diagram showing relative rates of background processes within LDMX along with how they motivate various aspects of the design. / E stands for “missing” energy or energy that is “lost” to neutrinos that are extremely unlikely to be detectable within LDMX.	24
3.2	Diagram of LDMX detector apparatus with a representation of a signal event where a dark brem occurs within the target. Credit to Christian Herwig for original development of diagram.	28
3.3	Rendering of LDMX detector apparatus focusing on tracker, target, and ECal. The magnet would fully encompass the tracker, target and trigger scintillator.	29
3.4	Diagram of LDMX cro:ecalElectromagnetic Calorimeter (ECal) design showing the longitudinal segmentation (top and bottom right) and the transverse segmentation (bottom left). Credit to Joe Muse.	30

4.1	Flow chart of how data is processed in the cro:ldmxLight Dark Matter eXperiment (LDMX) framework. Each processor has the ability to “add” data to the event as well as “get” data from the event. The processors are run in a user-defined sequence. Data can also be loaded from one or more input data files where the event data from those files is loaded into memory before the first processor is run. After all processors are done with an event, it is saved to the output data file.	32
4.2	Diagram of processing stages within LDMX showing the external sources of data or software that are used in those stages as well as which stages are commonly done within the centralized processing framework.	34
4.3	The total reconstructed energy within the ECal as a fraction of the beam energy separated by the total nuclear energy within the event (the energy going into the photon-nuclear and electron-nuclear processes).	39
4.4	Fraction of hits within the cro:hcalHadronic Calorimeter (HCal) that were caused by muons or anti-muons. The events lying within the zero bin still have HCal hits caused by other types of particles.	40
4.5	Longitudinal location (z) where the μ^+ was produced within the simulation for a variety of biasing factors B for the 4 GeV beam (left) and 8 GeV beam (right). The rate R is the equivalent number of unbiased cro:eotElectrons on Target (EoT) divided by the CPU time necessary to produce the sample. We can observe that while increasing B also improves the rate R , we eventually over-bias and lose access to the late tail of the distribution (red, purple, and brown in the left plot and gray in the right plot).	41
4.6	ECal reconstructed energy E_{ECAL} as a fraction of the beam energy E_{Beam} comparing the unbiased and biased samples scaled to the same EoT. The 4 GeV (8 GeV) beam case is shown on the left (right).	42
4.7	Distributions of simulated signal events with a 4 GeV beam. The energy of the produced dark photon is shown on the left while the longitudinal location where the production occurred is on the right.	44

5.1	The total reconstructed energy in all layers of the ECal (E_{ECal}). The signal and background distributions are normalized such that their integral is one. The 4 GeV beam is shown on the left and the 8 GeV beam is shown on the right. The events falling into bins with $E_{\text{ECal}} > 1.5$ GeV (3.16 GeV) for 4 GeV (8 GeV) are omitted from this plot but included in efficiency calculations.	48
5.2	Variables used in signal selection for the cro:eatECal as Target (EaT) analysis channel. In each figure, all other cuts are applied except for the variable in question. Some of the bins are empty in which case the upper Poisson limit given the total sample size is drawn with error bars. The grey line shows the selection cut. The background sample is the enriched nuclear and dimuon samples. The top (bottom) row shows the 4 GeV (8 GeV) beam.	50
5.3	Studies of how the tracker requirement affect the simulation samples. . .	52
5.4	The total reconstructed energy in all layers of the ECal (E_{ECal}) as a fraction of the beam energy (E_{Beam}) for all samples that pass the selection criteria except the selection on ECal energy. The 4 GeV beam is shown on the left and the 8 GeV beam is shown on the right. The gray lines mark the edges of the analysis bins used to estimate the expected exclusion limit and the black line is the upper limit on the ECal energy which also serves as the upper limit of an analysis bin.	53
5.5	Exponential fit of the total simulated background distribution. The shaded region is a 95% confidence band on the fit which is interpreted as the uncertainty on the value of the fit. Some bins in the simulated distribution are empty in which case the upper Poisson limit given the sample size is drawn as an error bar in that bin.	54
5.6	Background prediction within the three final analysis bins (blue) compared to the unconstrained simulation prediction in gray. The uncertainty on the background prediction is taken from the 95% confidence band shown on the fit. Some bins in the simulation prediction are empty in which case the upper Poisson limit given the sample size is drawn as an error bar in that bin.	55

5.7	The total reconstructed energy in the ECal E_{ECAL} comparing ten different smeared calibrations (colors) to the original unsmeared calibrations (black).	56
5.8	Effect of smeared ECal calibrations on the ECal hit RMS for events that pass the original trigger. Outside of statistical fluctuations in the lowest-populated bins, the smearing has a less than 5% effect throughout the range of potential cuts.	57
5.9	Reach (left) and maximum allowed yield (right) comparisons for the 4GeV beam case as we vary the cut on maximum HCal cro:pephotoelectrons (PE). We observe minimal movement in the final reach as a function of this cut; nevertheless, the biggest change occurs around a threshold of 30.	58
5.10	The sensitivity of the EaT analysis channel compared to other experiments led by NA64[5, 6], <i>BABAR</i> [7], and COHERENT[8] (gray), Scalar, Majorana, and Pseudo-Dirac theory expectations (black, top-to-bottom), and LDMX projections (colors). The blue (orange) line corresponds to the 4 GeV (8 GeV) beam studied in this work. The red line is the nominal LDMX analysis sensitivity for the Missing Momentum (MM) analysis channel.	60
6.1	Diagram of dark brem production of a dark photon followed by its decay back into an electron-positron pair through cro:simpStrongly-Interacting Massive Particle (SIMP) composite particle V_D with the emission of a lighter SIMP composite particle π_D . The two-particle vertex $V_D - A'$ is allowed since they share enough quantum properties to mix.	63
6.2	Feynman diagrams for the SM process of trident production.	64
6.3	Full rendering of cro:hpsHeavy Photon Search experiment (HPS). Beam would enter the detector from the left and exit to the right if it does not interact.	66
6.4	Simplified diagram of HPS showing an example displaced decay within the named subsystems. The blue arrow is some DM candidate particle that has a macroscopic lifetime causing the decay vertex to be observably displaced from the production within the target.	66

6.5	Figure 10 of [9]. Rendering of the cro:svtSilicon Vertex Tracker (SVT) with the vacuum enclosure shown in gray, active silicon components in red, and readout electronics in green. Similar to Figure 6.3, the beam would traverse this rendering from left to right.	68
6.6	Figure 27 of [9]. Depiction of an event passing the pair trigger in the 2016 data collection run. This depiction also displays the variables used in order to evaluate events and make the trigger decision.	69
7.1	Momentum error derived from injecting a certain amount of position error (blue and orange) compared to normal distributions (red and purple) with known momentum error. These distributions are normalized to have unit integrals. The trials were done in a simplified experimental model with quantities similar to HPS: a 2 GeV electron curving within a 1 T magnetic field, the sensors measuring the position were separated by 5 cm but their separation were also allowed to deviate from the true value by the position error given.	72
7.2	Plotting all six global coordinates completely specifying position and orientation of each of the sensors in the HPS tracking detector. Two different versions of the in-software detector model are plotted; however, on this scale, they overlap one another.	76
7.3	Plotting all six global coordinates completely specifying position and orientation of each of the sensors in the HPS tracking detector relative to the “Original” detector from Figure 7.2. We can now observe the quantitative differences between the different versions.	77
7.4	The transverse impact parameter of the tracks at the target $d_0 = \sqrt{x_0^2 + y_0^2}$ which should be centered on zero to align with the known beamspot at the target.	78
7.5	The magnitude of the track momenta which should be centered on the known beam value of 2.3 GeV.	79
8.1	Relative efficiency of vertex quality pre-selection requirements.	85
8.2	Number of vertices passing the quality pre-selection.	86

8.3	Relative efficiency of pre-selection of events. N_{vtx} is the number of vertices passing the vertex pre-selection requirements. Since other triggers for different purposes are present in the data, the Pair 1 Trigger is applied to data but not to simulation.	87
9.1	Depiction of dark photon yield calculation using a single run of the data ($\sim 1.6\%$). The gray vertical lines show the bounds of the search using the nominal ratio $m_{A'}/m_{V_D} = 1.66$	91
9.2	Expected signal yield N_{sig} for some rudimentary example efficiencies. Red contour lines are given at 1 and 1.6% (the approximate fraction of this data subsample) to give a sense of scale. Values are “clipped” to the color bar (for example, a signal yield of 2×10^4 is colored yellow which is marked as 1×10^3) so we can focus on the important orders of magnitude.	93
9.3	Diagrams showing L1L1 (left) and L1L2 (right) vertex examples.	94
9.4	Diagrams of examples for how different types of L1L2 vertices effect the observed values of y_0 . Truly displaced vertices (left) have both tracks whose y_0 is far from 0 mm while fake displaced vertices (middle) and not displaced (right) have at least one if not both tracks with $y_0 \approx 0$ mm. The diamonds are the reconstructed vertex, the solid circles are the reconstructed hits, and the empty circles are “missed” hits either because the particle did not pass through those sensors (left) or due to some physical or electronic inefficiencies (center and right).	96
9.5	Distributions of cut variables for a few example mass points and $\approx 10\%$ of the full data sample. The vertices are required to be L1L2, have their momentum sum be within the signal region, and their invariant mass within the specified mass window.	97
9.6	$y_{0,\text{min}}$ distribution after the other selections (L1L2, P_{sum} in SR, reconstructed mass lying in mass window, cro:vpsVertex Projection Significance (VPS) < 4 , and $\sigma_{y_0,\text{max}} < 0.4$ mm).	98
9.7	Binomial significance (left) being maximized after the VPS < 4 and $\sigma_{y_0,\text{max}} < 0.4$ mm cuts leading to cuts (right, blue) which are smoothed into a continuous function (right, orange) as described in the text.	99

9.8	$y_{0,\min}$ distribution for the 10% data sample with the $y_{0,\min}^{\text{cut}}$ function overlaid in red.	101
9.9	Signal efficiency as a function of z and mass scale m_{V_D} for this analysis (left) and L1L1 (right).	101
9.10	Search results for the full 2016 data set.	103
9.11	Signal yield for the L1L2 reconstruction category with the described cuts showing the theoretical expectation (left) and the maximum allowed (right) derived using the cro:oimOptimum Interval Method (OIM). . . .	104
9.12	Sensitivity of this analysis with contour drawn including systematic errors suppressing the expected signal yield.	105

Part I

Laying the Groundwork

Chapter 1

Introduction

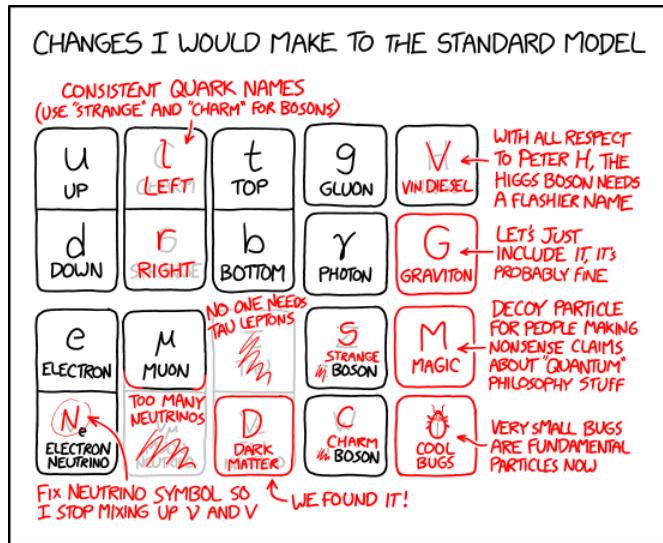


Figure 1.1: Comic #2351 from XKCD[1]. If able to be designed from scratch, the names of things in the standard model could be more helpful.

The long and winding road of a dissertation is not always neatly packaged into a template - this can easily be seen in my own experience. While focusing much of my time on the technical aspects of data processing software for a proposed (yet to be built) experiment, I found myself near the end of my journey without *real* data to analyze from which I can make *real* conclusions about the physical world. This motivated me to participate in another experiment – extremely similar to the original one – sharing

theoretical motivations, technological designs, and even people. This experience of participating in two similar experiments has provided an abundant field of learning opportunities (as well as roadblocks) for me and my journey as a new physicist.

Physics is a lazy science. I would go even further and state that this reductionist perspective is a main attraction for many physicists. We¹ want to avoid memorizing as much as possible; therefore, motivating the idea of condensing sets of observations into “laws” that can be represented in an even more compact mathematical form. We make up these “laws” and the vocabulary surrounding them not to decieve but merely to make communication about our observations and the experiments that make these observations easier. We sometimes debate the origins of these laws and their true philosophical meaning, but often, on a day-to-day basis, the background of them is unimportant. The interesting work comes from *testing* them, breaking them, and remaking them. For our purposes here, this is what a “theory” is: a package of laws and their mathematical forms with which we can make predictions of the observations of our experiments.

While I speak in the context of all physics, in reality, I am residing within a small corner. Primarily concerned with individual particles and how they interact with one another, my field can be described as investigating the foundations of the universe. Our experiments require giving these particles comparatively large amounts of energy, contextualizing the name cro:hep High Energy Physics (HEP). Giving such small particles such large amounts of energy requires extrodinarily large and complex apparatuses. HEP is filled with experiments many stories tall, collaborations consisting of hundreds of institutions and thousands of people, and observations lasting years if not decades. This grand scale is helpful to keep in mind when I speak of the experiments in this thesis – constructed and operated by only a few institutions (less than 10 in each) and only dozens of collaborators. Contrary to many other sciences and experimental methods, these experiments still last longer than a typical doctoral student career. While I report on these experiments, I will not detail their full operation, design, or capabilities. I will focus on the work that I was able to contribute; hopefully providing a building block for these experiments and HEP to grow in the future, and give only the necessary context for parts with which I am less familiar.

¹Here I use the “royal we” as representing my point of view of the culture within the field. It should in no way be construed as scientific or exact statements and does not represent every physicist’s point of view.

Before partitioning this thesis according to the two experiments in which I participated, Section 1.1 summarizes our current laws and its representative theory that best predict our set of observations within HEP, Chapter 2 will explore the theoretical ground on which both of the experiments rest which will also provide necessary vocabulary for discussing these two experiments. Part II presents the work of my initial years as a graduate student developing data processing and realistic simulation software for a proposed experiment. This part includes a description of this proposed experiment in Chapter 3, the simulation infrastructure for it in Chapter 4, and an analysis of this simulated data showcasing this experiment’s performance with $\sim 2.5\%$ of its target data set in Chapter 5. Part III describes the work of my last years showing the difficulties of working with data taken in the real world (with all the complexities that implies). Similar to Part II, this part first describes the experimental setup in Chapter 6, how we can tune our detector model to better match how it was constructed in Chapter 7, the simulation and collected data used are detailed in Chapter 8, and the analysis performed on this data is described in Chapter 9. We then conclude with Chapter 10 which returns back to this high-level-view from which we can discuss what we have learned about physics via these two experiments.

1.1 Standard Model of Particle Physics

What makes up all the stuff around us? This question has been posed, debated, discussed, and studied since the ancient times. Our best current understanding is that stuff is made of particles that interact with one another. While “particle” is difficult to strictly define qualitatively, we have a model for them and their interactions which we creatively named the “Standard Model” (SM). Figure 1.2 displays the known elementary particles, some of their basic properties, and how they interact with one another. Not all of the particles represented within this diagram are of critical importance to this work, but some deserve further description.

The two lowest mass quarks – the up and the down – are the fundamental constituents of protons and neutrons which make up the nuclei of atoms which make up all the day-to-day stuff we interact with. These stay bound together within these nuclei via the strong nuclear force (mediated and represented in Figure 1.2 by the gluon). These

Standard Model of Elementary Particles

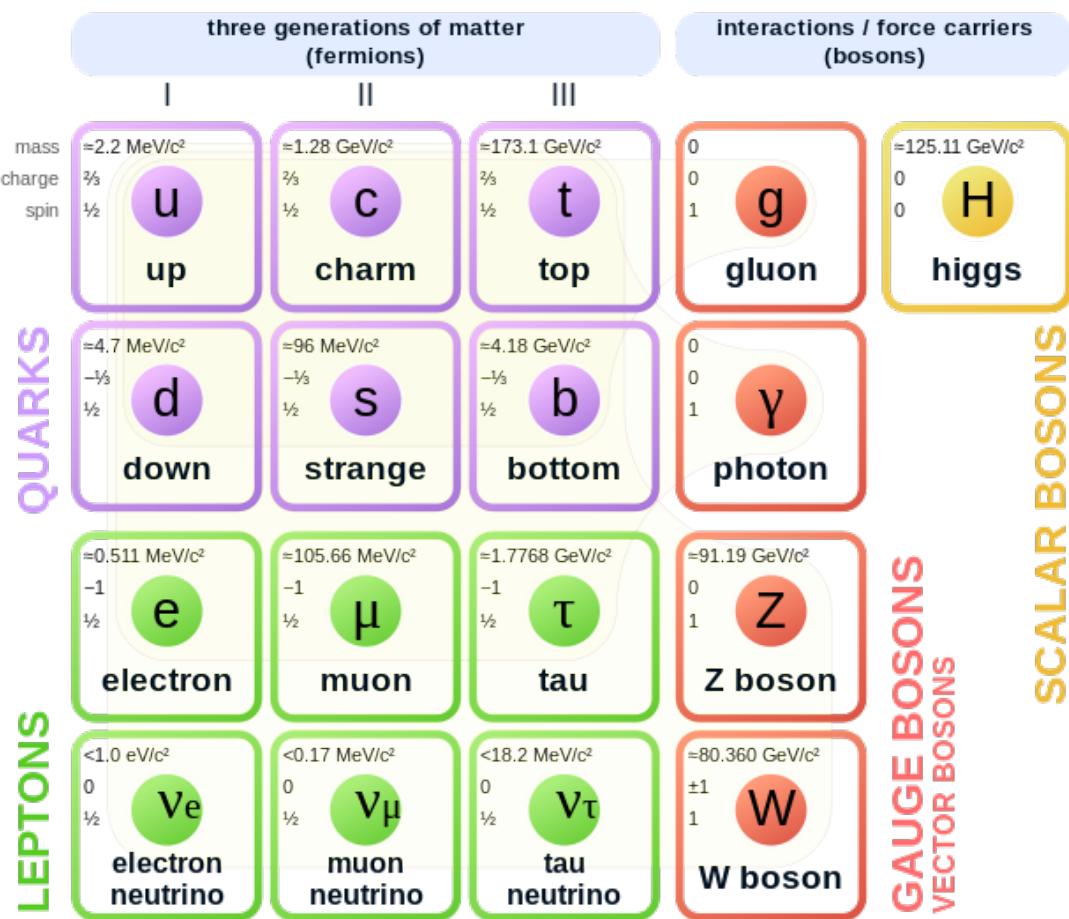


Figure 1.2: The Standard Model of Particle Physics showing the twelve fermions and five bosons, their various properties (mass, charge, spin), labels (box and circle colors), and interactions (brown loops). Credit to Cush on wikipedia for providing this diagram freely accessible and usable for any purpose.

quarks and the top row of leptons also have an electric charge enabling them to interact via with electromagnetic force. The electromagnetic force, represented and mediated by the photon, is the force responsible for light and magnetism. The lowest mass charged lepton (electron) is also very abundant – most commonly residing in orbits around nuclei helping form atoms. Both experiments in this thesis utilize beams of electrons that have been accelerated to high energies and then impinge upon a material of high density to initiate interactions for further study.

The mathematical expressions to calculate how these particles interact are complicated. As mentioned earlier, physicists are lazy and so we have developed a method to represent the specifics of these interactions graphically in a form where many key aspects of the interaction can be understood at an intuitive level without needing to write down any of the long mathematical formulae. These representations are called “Feynman diagrams” after the 20th century physicist Richard Feynman. Feynman diagrams allow us to represent interactions by defining a set of “vertices” that are allowed by our theory and then constructing processes from these vertices that include the in- and out- going particles that we wish to study. Such diagrams are more than visual sketches; they can be algorithmically converted into precise mathematical forms which can be used to directly calculate measurable predictions of the model containing this set of vertices. The conversion of these diagrams into mathematical forms and the calculation of these forms has been written into various computer programs (I say this to emphasize that these diagrams directly represent the mathematics that can be used to calculate them). Figure 1.3 shows an example Feynman diagram representing the bremsstrahlung process where a charged lepton (ℓ^-) interacts with a nucleus (Z) via a photon (γ) and then emits another photon before recoiling. We can see three vertices in this diagram: two “fundamental” vertices where the lepton and photon lines connect and one “effective” vertex where the photon connects with the nucleus. The fundamental vertices are actually strictly defined within the Standard Model, but the effective vertex represents a helpful approximation that is accurate at the energy scales we are studying.²

Due to the complexity of the nucleus itself, the calculation of Figure 1.3 into any

²The distinction between fundamental and effective vertices is not a well defined one. I find it helpful here, but it is not necessarily made elsewhere in physics literature.

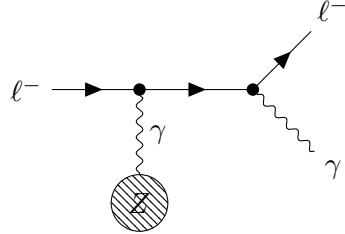


Figure 1.3: Feynman diagram for the bremsstrahlung process where a charged lepton (ℓ^-) interacts with a nucleus (Z) via a photon (γ) and then emits another photon before recoiling.

observables is difficult to do without any approximations. However, one of the initial parametrizations of bremsstrahlung offers a good glimpse of how it behaves [10]. When doing experiments, we often count the number of particles produced with certain criteria. These “rates” can then be translated into estimates of process cross sections (how likely a process is to occur producing particles with the observed properties) with knowledge of how the experiment was conducted.

These cross sections are also calculatable from diagrams like Figure 1.3 and the initial parameterization can be summarized as

$$\frac{d\sigma}{dE_\gamma} \propto \frac{1}{E_\gamma} \quad (1.1)$$

This shows the dominant shape of the cross section – it grows as the energy of the produced photon (E_γ) gets smaller. The full expression prevents the cross section from diverging as the photon energy approaches zero; nevertheless, the general inverse relationship between the rate and the energy of the produced photon holds in both the SM calculations and observations from experiments. Figure 1.4 shows another view of this rate: instead of looking at the energy of the produced photon (E_γ), we can inspect the energy of the outgoing electron after it emits the photon (E_e). This is more readily compared to other processes that do not produce a photon and we can observe a large difference between this SM process and other possible processes that we search for.

This inverse relationship is somewhat special and is mostly due to the fact that photons do not have any mass; however, this special relationship opens a door towards potential discovery. The vast majority of electrons interacting with a material will do

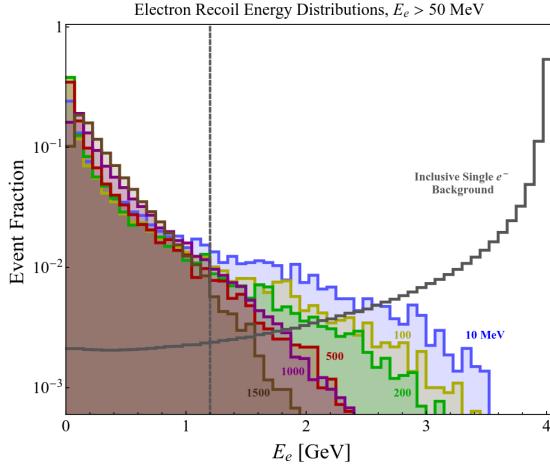


Figure 1.4: Energy of the outgoing electron (E_e) for standard processes (dominated by bremsstrahlung) in gray and some dark bremsstrahlung (see Section 2.4) in colors. Figure 3 from [2].

so softly, only emitting low energy photons, which distinctly separates these electrons from ones that undergo rarer processes (or perhaps so-far undiscovered processes). Both experiments described in this work make heavy use of this fact.

As the work carried out for this dissertation is experimental rather than the definition and calculation of a model, I intend to stay within this diagrammatic realm – representing the theoretical calculations being tested with these diagrams and limiting the scope of mathematical formulae presented. Occasionally, the weight of these vertices which can be interpreted as the strength of the interaction is presented along with the vertices in order to give a sense of scale and connect these diagrams to the formulae that are included here.

1.2 Standard Model Analogies

The SM has been extremely successful in describing the current world of particle physics, including some intricate behaviors of particles that were incredibly confusing when first encountered. When attempting to describe unknown phenomena (Chapter 2), we often draw on this previous experience to form new ideas and models. This section, while a brief detour, will help give context for two key behaviors that the models we are

searching for would give their particles.

1.2.1 Particle Mixing

All of the models tested for in this work use the idea of “particle mixing” in order to connect the realm of theoretical/undiscovered particles with the realm of SM particles. In the quantum arena, we treat individual particles in a probabilistic manner – often times when we observe a particle with certain properties, it can be observed again in a different context with different properties. While this sounds odd, several experiments have supported this random nature of the smallest particles over the course of twentieth century physics. As a well-understood example, neutral kaons have been observed to “swap” with their own anti-particle, which we can describe via this concept of mixing.

Kaons are particles that contain two quarks, one of which is a strange quark (or its anti-quark). A neutral kaon (K_0) and its anti-particle (\bar{K}_0) are nearly indistinguishable – they have the same electric charge (0) and the same mass amongst other shared properties – but we can distinguish them by observing the types of particles they produce when they decay. Describing kaons as consisting of two quarks allows us to understand the mixing of K_0 and \bar{K}_0 , diagrammed in Figure 1.5. From a higher-level view, where we are unaware of the constituents of the kaon, it appears like the K_0 and \bar{K}_0 are swapping places with some probability. Our experiments can observe this probability by counting the frequency with which we see K_0 compared to \bar{K}_0 . We can also use our model of the kaons (as composed of quarks) and calculate this same probability (specifically, using the Feynman diagram shown in Figure 1.5). Comparing the two gives us information about the SM and how the quarks interact with one another.

1.2.2 Displaced Particle Decays

Ever since particle physicists have started designing detectors, we have also observed particles that are “missed” by our detector mechanisms. Figure 1.6 shows an early example of this where a “bubble chamber” shows the paths of charged particles passing through the liquid (the white lines). In the middle of this image, you can observe two lines emerging seemingly out of nowhere (the red annotations were added later). The mechanism that allows a bubble chamber to observe the path of particles requires the

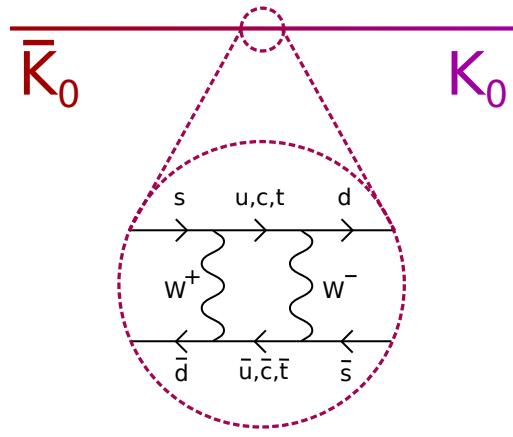


Figure 1.5: Dipiction of neutral kaon and anti-kaon mixing including the underlying Feynman diagram that results from applying the SM to this system. This figure was created by user NikNaks on Wikipedia and is licensed under CC Attribution-Share Alike 3.0 Unported.

particle to have a charge, so a neutral particle (a Λ in this case) does not leave a path itself. The only signal that it was there is the *displaced vertex* from which two other particles (probably a proton and a negative pion) emerge (emphasized with a red circle in Figure 1.6). This is a “displaced decay” of a particle which was produced by a high energy (16 GeV) negative pion interacting with the liquid in the bubble chamber.

While particle physicists have developed other detector apparatuses, some of which can observe neutral particles more directly, this more indirect detection mechanism is still extremely useful especially if a model being tested has particles that have no way of directly interacting with SM matter (the stuff we make our detectors out of). Part III of this thesis focuses entirely on an experiment designed to search for these displaced vertices, the displacement of the vertex coming from the presence of an as-yet undiscovered particle being produced and then decaying back into SM particles.

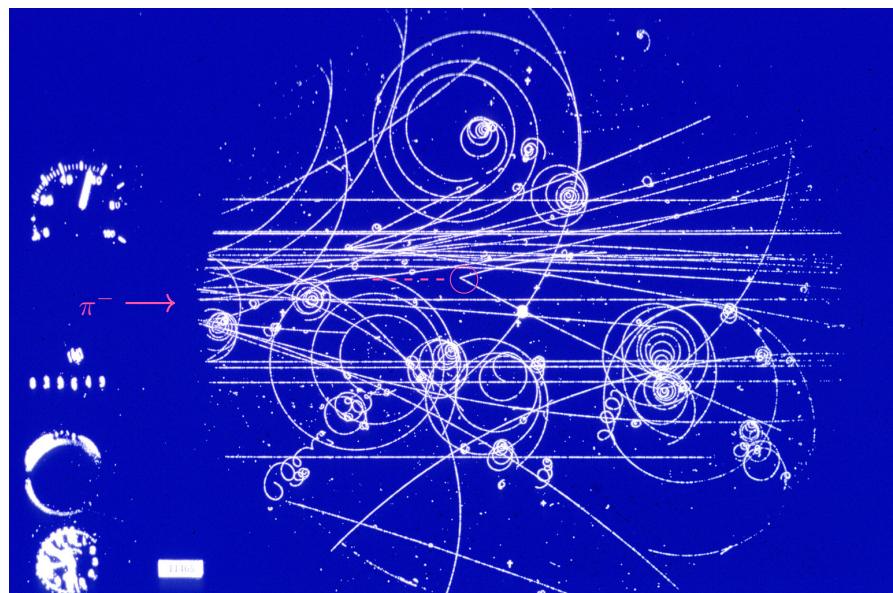


Figure 1.6: Image of CERN’s first liquid hydrogen bubble chamber from 1960 [3] with pink annotations added. The pink circle highlights the location described in the text where a Λ baryon decays into two particles (most likely a proton and a negative pion). The pink dashed line traces out a possible path the Λ traveled before its decay.

Chapter 2

Dark Matter



Figure 2.1: Comic #1758 from XKCD[1]. Generally, claiming that either Einstein or Newton was wrong will be a bad time.

New phenomena have puzzled physicists throughout history, but the confusion surrounding dark matter – estimated to be roughly 85% of the total matter in the universe – is surely one of the biggest puzzles. While many astronomical observations at various scales have confirmed the existence of dark matter, we have not yet seen any observations of its particle interacting with our detectors. This has ruled out many models for dark matter’s particle nature, but there are still many more available which can both explain the current observations of dark matter gravitational and cosmological effects while skirting the limitations defined by the *lack* of observation in particle experiments.

2.1 Dark Matter?

The term DM rose out of a brutally honest description of the present level of understanding. “Matter” originates from comparison to the particles that make ourselves up and with which we are familiar – both this new phenomena and our normal “stuff” interact gravitationally, attracting other groups of “matter” into complex cosmological systems. The adjective “dark” refers to the literal fact that we cannot see it with our telescopes. Other matter out in the universe can be seen via the light that it gives off (or reflects), so the fact that this matter *is not* visible in this way motivates the adjective.

That is it. Dark Matter (DM) is merely a shorthand for this aspect of the universe that is both known to exist within the cosmos and whose particle nature is completely unknown.

2.2 Evidence for Dark Matter

The first (replicable) evidence physicists had for an unseen material floating throughout the universe was the observation of galactic rotation curves [11, 12]. These observations measure the speed of different stars within a galaxy and compare this speed to the distance of that star from the center of the galaxy. We can calculate this relationship using GR [13] and the observations differ drastically from this calculation. The stars within galaxies we’ve observed move much faster than GR would predict (Figure 2.2) leaving us with two explanations: either GR is not the correct theory to use in this situation or there is more un-seen mass floating within the galaxy allowing these stars to move faster without leaving the galactic orbit.

Other indirect measurements give us additional ways to access information about this odd phenomena. Within the framework of GR, since energy and mass actually warp the fabric of spacetime, we expect to see light itself follow a bent path around massive objects - a phenomenon that is called gravitational lensing and is observed and well modeled by GR’s predictions [14]. The accuracy of GR within this context – a mass and distance scale similar to the rotation curve oddities also observed – put more requirements on any modified theory of gravity that could both explain the rotation curves and gravitational lensing. Additionally, measurements on some of the largest scales and from the early universe display signs of a certain mass density attributed to

matter that does not interact in the same ways as our normal matter (called “baryonic” matter).

In the early universe, standard matter was compressed into high densities and existed at high temperature creating a sea of plasma. Gravity attracts while pressure from the squeezing of this plasma repulses which produces oscillations in the density of matter in space and time. The repulsive pressure within the plasma originates from particles interacting electromagnetically with each other, so the oscillations would be disrupted by the presence of extra matter interacting gravitationally but not interacting electromagnetically. These *baryon Acoustic Oscillations* (BAO) are imprinted on our snapshot of this early-universe plasma, the *Cosmic Microwave Background* (CMB), which we can measure with a high degree of accuracy and fit to various models of what existed at this time of the universe. The best fit of these models corresponds to only $\sim 5\%$ of the mass density being normal matter like we see today (“baryonic”) while the rest is composed of material that only significantly interacts gravitationally [15].

Additional astromomy observations from Type1a supernovae [16], fitting models of big bang nucleosynthesis [17], and constraints on non-particulate theories explaining these phenomena (for example mini black holes [18]) all allow us to conclude pretty comfortably that DM exists as a particle in our universe.

2.3 Particle Nature of Dark Matter

The theoretical possibilities explaining DM are broad [19] even when excluding ourselves to the assumption that the DM phenomenon is explained by the existence of a DM particle. The CMB observations are tied to very early on in the existence of the universe, so it is natural to assume that this DM particle has existed since the start of the universe alongside our normal matter particles. With this in mind, we can outline a few criteria that must be met by a proposed DM candidate.

- **Dark** There has been no detection of these particles via the light observed with our telescopes; therefore, the DM has to not interact via the electromagnetic force.
- **Long Living** Measurements of DM’s mass density and presence agree across time (from as early as the CMB era), so DM needs to have a long lifetime.

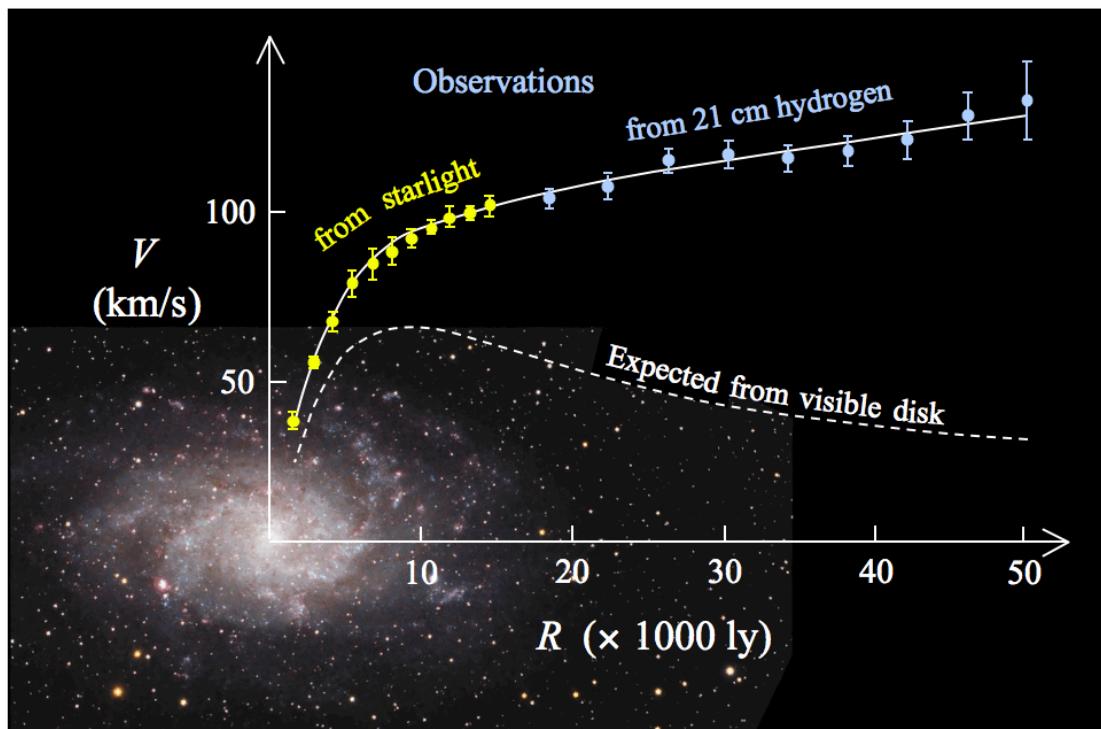


Figure 2.2: Depiction of the velocity of stars within a galaxy as a function of their distance from the galactic center. The dotted line is a prediction of this relationship using GR along with the mass tabulated from the visible starts while the data points (and the solid line fitted to them) are what are actually seen in galaxies today.

- **Universal Density** Since we can indirectly measure a DM mass density on cosmological scales, we impose the requirement that any DM model needs to allow for this density.
- **Thermal Relic** Both DM and standard matter have similar cosmological densities, so we expect some interaction (even a weak one) should connect their origins to the early universe allowing both to exist from the Big Bang. This criterion is not as firm as the others (DM models correctly describing the current observations while avoiding this criterion do exist); nevertheless, this criterion is well motivated and it is satisfied by all of the models for DM studied in this work.

Even with these assumptions and the requirements they imply, there still exist a plethora of theoretical models that can satisfy all of them.

The upside is that a thermal-relic assumption closely connects the mass of individual DM particles to the interaction strength it has with standard matter. In this assumption, the DM evolves along with the universe allowing its number density to follow the density of standard matter until the standard matter does not have enough energy to produce more DM. While the universe continues to expand, the DM continues to decay into standard matter until it becomes too sparse to annihilate with itself and is thus “frozen” at a specific number density. Since this “frozen” density changes depending on how easy it is for the DM to interact with standard matter, the “frozen” number density goes down as the interaction strength increases (Figure 2.3). The additional requirement of the observed astronomical *mass* density connects the mass of individual DM particles to their interaction strength with standard matter.

$$m_\chi \leftrightarrow \text{observed mass density} \leftrightarrow \frac{dN}{dV} \leftrightarrow \text{thermal relic hypothesis} \leftrightarrow \langle \sigma v \rangle \quad (2.1)$$

This connection allows us to define strict “thermal relic targets” which can help be measuring sticks for how well our experiments search for DM (these targets will appear in plots later).

In addition to the connection between interaction strength and particle mass implied by the thermal relic hypothesis, it also puts some loose bounds on mass of individual particles (Figure 2.4). If the mass is too high ($\gtrsim 100$ TeV), the DM will be too strongly coupled to the standard matter and would be over-produced within the early universe.

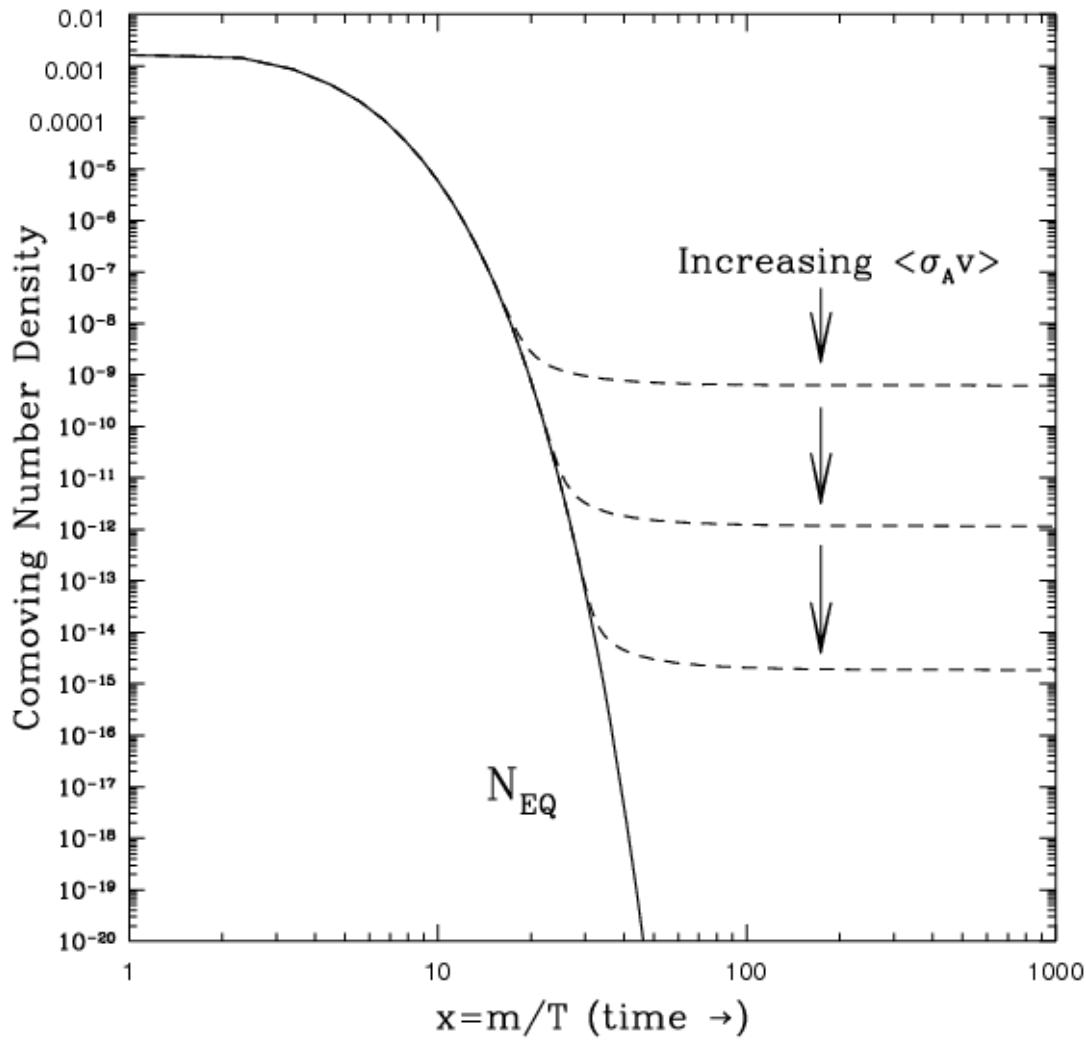


Figure 2.3: From [4], the co-moving cosmological number density of DM as a function of the universe temperature. As the universe cools, the number density decreases until the DM becomes too sparse to interact with other DM particles, “freezing” to a specific number density until today.

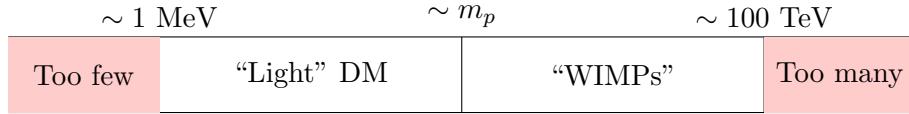


Figure 2.4: Mass scale of Thermal Relic DM. The regions in red are excluded by applying the thermal relic assumption to our observations of the universe’s early evolution. m_e is the electron mass and m_p the proton mass.

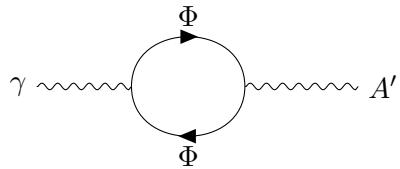


Figure 2.5: Feynman diagram of how a massive field Φ could allow for a standard photon (γ) to mix with a dark photon (A').

Highly precise agreement between nuclear abundance observations and predictions from a model of the early universe with only SM particles provides the lower limit; allowing a thermal relic DM mass $\lesssim 1$ MeV would then break the consistency between this model and several of its predictions [17, 20, 21].

The mass scale of thermal relic DM is further divided by the mass of the proton (m_p). Above this threshold, the DM could be interacting with standard matter through the standard Weak Force (as described in Chapter 1); thus, they are named **Weakly Interacting Massive Particles** (WIMPs). This phase space was first searched due to its theoretical simplicity: no new forces, just an extra particle (or two) creating the clouds of DM we see today. Unfortunately, these searches have not found evidence of a WIMP-like signature[22, 23, 24] and the phase space has become tighter and more excluded after many years of searching. Below m_p , the DM requires a new force which is even weaker than the Weak Force. In comparison to the heavier WIMPs, this category of DM is called “Light” DM.

In general, the definition of a new force of nature is not well constrained; however, we can define a “baseline” model that can represent more concrete theories in the context of our experiments. For this situation, we postulate the existence of a massive gauge boson that represents an additional $U_D(1)$ symmetry of nature. Since this additional symmetry has the same structure as the electromagnetic interaction whose gauge boson

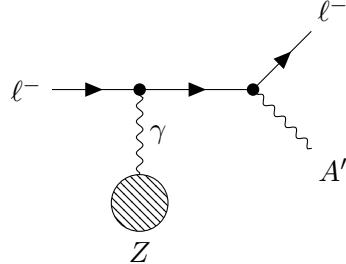
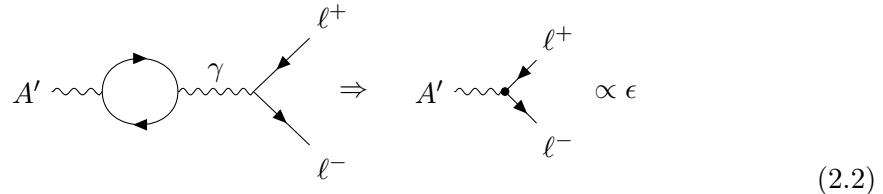


Figure 2.6: Feynman diagram for the dark brem process.

is the standard photon, we generally refer to this postulated massive boson as a dark photon. Without any additional assumptions, this new dark photon does not interact with any standard matter, so we also assume that some previously-unknown massive fields are able to interact with both. These other fields need to be sufficiently massive (or sufficiently weakly coupled) so that they only allow the standard and dark photons to weakly mix. Figure 2.5 diagrams how the standard and dark photons mix which can then be effectively represented by a new vertex in our model.



where the mixing strength ϵ encapsulates the effect of a massive field Φ at the energies of our experiments (presumably low enough to avoid creation of a Φ directly) and is generally small ($\ll 1$).

This new vertex within our model for the universe enables a new process to occur: so-called “Dark Bremsstrahlung” (dark brem) where a charged particle exchanges a standard photon with a nucleus and then emits a dark photon and recoils.

But what happens to the dark photon after it interacts with standard matter via this vertex? It cannot be the long-lived DM we view in the universe today because this vertex allows for it to decay *back* into standard matter eventually (and in some models, rather quickly). This is where we expand on the idea of a “dark sector”. In this description of DM, we already have a dark photon representing some force and a very

heavy field that can interact with it. Suppose this “dark sector” also has other particles (like the standard sector) – one (or more) of which could be long-lived and represent the DM we observe in the universe today (so-called “DM candidates”). We can further partition this category of models depending on what happens to the dark photon after we produce it within an experiment.

2.4 Invisible Signature

One of the simplest options is to hypothesize another particle that can take on the role of long-lived DM and only interact “within” this dark sector (i.e. it only interacts with the dark photon from our point of view). Calling this particle χ (and its anti-particle $\bar{\chi}$), we then have an additional dark sector vertex.



where α_D is a parameter representing the strength of this dark sector interaction, limited to be below 1 but usually much larger than ϵ so that this interaction is more likely than the one with standard matter. If the mass of the dark photon is greater than twice the mass of the χ ($m_{A'} > 2m_\chi$), then this vertex can occur immediately after a dark photon is produced within an experiment. Since the χ does not interact with our normal SM particles, the energy they have is “lost” from our perspective. Part II focuses on a proposed experiment designed to precisely measure all of the *visible* energy so that this “invisible signature” of DM being produced could be observed.

When comparing analyses across DM search literature, it is common to parameterize the DM phase space with an effective interaction strength y and the mass of the candidate DM m_χ that would constitute the astronomical DM observed today. The parameter y is related to the parameters of our “baseline” model by

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

where we make standard, benchmark choices $\alpha_D = 0.5$ and $m_{A'} = 3m_\chi$ when converting from $(\epsilon, m_{A'})$ to (y, m_χ) .

2.5 Visible Signature

There is no a priori reason for $m_{A'} > 2m_\chi$ – the quantum nature of these fundamental particles allows for “virtual” dark photons to have more or less energy than precisely $m_{A'}$ allowing them to mediate the interactions between χ and SM particles, so we must accomodate the possibility that $m_{A'} < 2m_\chi$. In this case, there would be a significant probability that the dark photon, after it is produced, would convert *back* into SM particles. Since these decay products are observable by our detectors, this signature is called visible.

These visible signatures are a bit more complex to model due to the fact that both a production process and a decay process need to occur according to the model of DM we are testing. With this in mind, many models that provide more detail about how particles exist within a dark sector have been created, each of which providing a specific estimate for production and decay rates. Part III focuses on an experiment looking for these visible signatures and my work specifically was investigating one of these more intricate dark sector models in detail. Section 6.1 provides more detail on the dark sector model studied in this work.

Part II

LDMX

Chapter 3

Light Dark Matter eXperiment

LDMX is a proposed fixed-target experiment aiming to definitively explore the thermal relic light dark matter phase space. Even as a proposed experiment, it has a detailed plan for construction, a beam already in construction, and well established connections with current technologies used within HEP. While LDMX is not yet built, it has a well formulated simulation infrastructure that can realistically model how the detector design responds to various types of interactions happening within it.

3.1 Missing Momentum Signature

LDMX aims to search for an invisible signature – with known incident particle kinematics, measuring the outgoing particle kinematics allows for a natural deduction. All of the incoming momentum must exit somehow and so if the detector is unable to observe some momentum (some momentum is “missing”) frequently enough, we can conclude that some other, previously-unknown, process is taking place and carrying momentum away from the experiment.

Precisely understanding both the incident and outgoing momenta requires knowledge of the currently known processes and how to detect the particles they produce. The center of Figure 3.1 displays an ordering of known processes based on how frequently they occur given an incident electron with 4 GeV of energy. This chart can be broken into three regions.

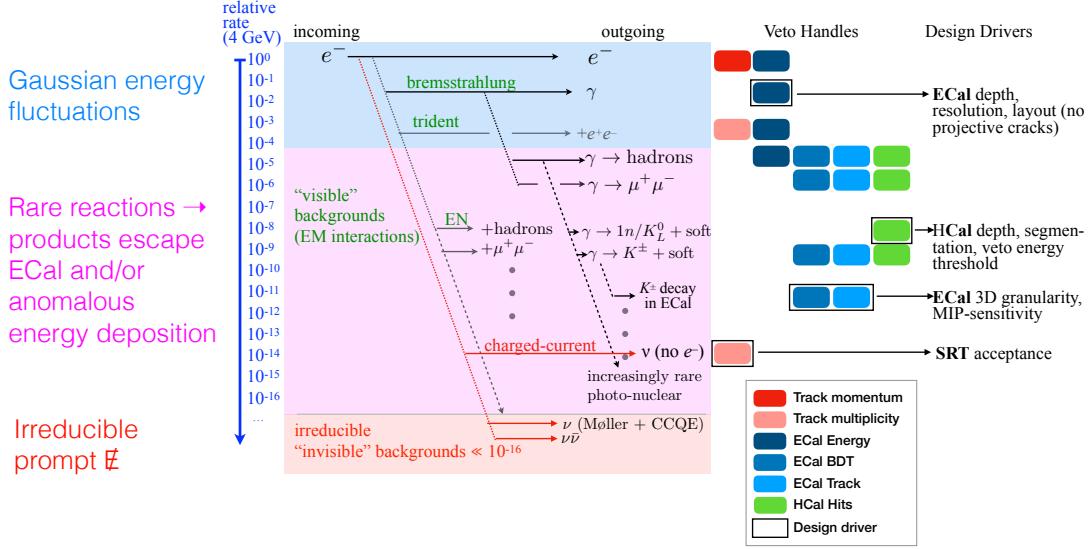


Figure 3.1: Diagram showing relative rates of background processes within LDMX along with how they motivate various aspects of the design. / E stands for “missing” energy or energy that is “lost” to neutrinos that are extremely unlikely to be detectable within LDMX.

The blue region at the top are the most frequent types of processes, but are simultaneously less complicated. The processes in this region (as the right side shows) put requirements on the design of the ECal, making sure it can quickly and faithfully reconstruct the outgoing energy in photons and electrons. The ECal must be able to veto the first several orders of magnitude in basic energy fluctuations including simple interactions in the target like bremsstrahlung or trident production of an electron-positron pair.

Entering the pink region is where the background processes become rarer and more complicated. One of the first ways to partition these backgrounds is whether charged secondary particles are produced within the target. These “prompt” backgrounds help define the recoil tracker’s design and are thus left to be caught by it. More frequently, a photon is produced within the target (which is not observable by the tracker) and then undergoes complicated photon-nuclear interactions within the ECal producing particles that are difficult for the ECal itself to observe. The HCal is able to detect the presence of these hadrons and muons, acting as a “pure” veto in the sense that the signal process

should never create significant activity within it.

All of these subsystems can collaborate to help LDMX reject known SM backgrounds down to $\sim 10^{-16}$ fraction of all incoming electrons where extremely rare known processes that produce invisible products (neutrinos ν) begin to emerge.

Detailed study of the missing momentum search strategy[25, 2, 26] shows that LDMX can reject all simulated backgrounds to a relative rate of $\gtrsim 10^{-14}$. This performance is accomplished through the basic design drivers outlined above and in Figure 3.1, but also through the additional granularity of the ECal enabling the use of a cro:bdtBoosted Decision Tree (BDT) to distinguish between background and signal events using features of the showers within the ECal.

Cycles of detailed simulation and redesign have led to a detailed construction plan for the LDMX detector apparatus. The subsequent sections of this chapter detail this design as well as the beam it is expected to receive.

3.2 The Beam Line

LDMX is planned to receive electrons from the superconducting Linear Accelerator (Linac) at SLAC National Accelerator Laboratory. The SLAC Linac can provide high energy, high rate, and low intensity electron beams for the various experiments it hosts. Specifically, the Linac Coherent Light Source (LCLS) is used to guide the beam (of a certain energy) towards the experimental hall - the upgraded phase of LCLS (LCLS II [27]) is currently under construction and is what will be used for running with LDMX.

The experiment is hosted in End Station A (ESA) at SLAC which requires an additional upgrade to the accelerator complex in order to receive its beam. LESA (Linac to ESA) beamline [28] is also currently being constructed and will be ready for test beam in early 2025. Part of the infrastructure that transfers the beam to ESA (Sector 30 Transfer Line – S30XL) is already constructed and LDMX components are expected to participate in a test beam runs with it.

3.3 Detector Design

As described above, LDMX is a missing momentum experiment and its design is focused on measuring *both* the incoming and outgoing momenta of charged particles interacting with a thin target such that any momentum given to undetectable (dark) particles can be precisely determined. In addition, the energy of neutral particles must be measured. This design has led to four subsystems each with specialized roles.¹

1. **Tracker** Measure charged particle momenta both before (“Tagger”) and after (“Recoil”) the target, using a dipole magnetic field of 1.4 T.
2. **Electromagnetic Calorimeter** (ECal) Measure the total energy of electrons, positrons, and photons.
3. **Hadronic Calorimeter** (HCal) Veto additional particles difficult for other subsystems to measure (muons, pions, hadrons,...).
4. **Trigger Scintillator** Count the number of electrons incident on the target in time to make a trigger decision (more on what a “trigger” is below).

Figure 3.2 displays these subsystems in a diagram along with a representation of a dark brem interaction occurring within the target. Figure 3.3 shows a rendering of the detector design.

The Tracker (purple in Figure 3.2) is a thin silicon strip detector modeled after the HPS tracker. These silicon strip sensors are arranged in layer pairs where one layer is angled slightly askew relative to the other in the pair to enable reconstruction of three dimensional hit locations. The part of the tracker upstream of the target (to the left in Figure 3.2) is named the “tagger” since its purpose is to measure the incident electrons’ momenta, rejecting electrons with momentum below 30 % of the expected beam momentum. The tagger is situated within the bulk of the magnetic field enabling highly precise measurement of the high incident momentum. The other part of the tracker located downstream of the target (to the right in Figure 3.2) is named the “recoil” tracker since its job is to measure the momenta of all charged particles recoiling

¹These subsystems also take on additional roles when the full breadth of the LDMX physics program is taken into account. This description just focuses on the DM search.

from interactions within the target. While it is not located within the magnet volume, it is still situated within the fringe field enabling it to maintain track reconstruction down to the lower-momentum products of interesting processes within the target.

In many HEP experiments, a “trigger” system is necessary in order to filter data defined to be interesting by the experiment from the wealth of uninteresting (or normal) data that is expected to be produced at a much higher rate. These trigger systems are the first filter that any data goes through and are designed to help the experiment obtain a statistically large data sample without collecting an overwhelming amount of data. The ECal is the primary detector subsystem responsible for making this trigger decision due to its excellent energy resolution and fast measurement capabilities. As a search for *missing* momentum, the ECal requires less than 30 % of the incoming beam energy to be observed within the first twenty sensitive silicon layers.

The Trigger Scintillator (yellow-orange in Figure 3.2) is designed to help inform the trigger decision by counting the number of electrons present within the detector. It is made of layers of vertically segmented bars of plastic scintillator. These layers are arranged in pairs where the layers within each pair are offset from one another to cover any gaps between the bars. These bars are readout in time to be used within a trigger decision combined with information from the ECal and HCal.

The HCal (green in Figure 3.2) is a sampling calorimeter made up of alternating layers of steel absorber and plastic scintillator bars. The HCal is further subdivided into the “side” HCal which is situated around the ECal and the “back” HCal downstream of the ECal. The back HCal has the orientation of the scintillator bars alternate between vertical and horizontal so that clusters and tracks can have three-dimensional coordinates more precisely identified.

The ECal (blue in Figure 3.2), as a primary volume of interest within the analysis discussed here, is given its own diagram Figure 3.4. The ECal is also a sampling calorimeter; however, it uses a different absorber material and a different sensing mechanism to more precisely measure the energy of electrons, positrons, and photons. The calorimeter is constructed out of seventeen layers each consisting of tungsten absorber, service materials, and sensitive silicon sensors. Each of the layers of the detector has two sub-layers of sensitive silicon sensors and each of these sub-layers are built up out of the hexagonal High-Density modules designed for the CMS Phase II High Granularity

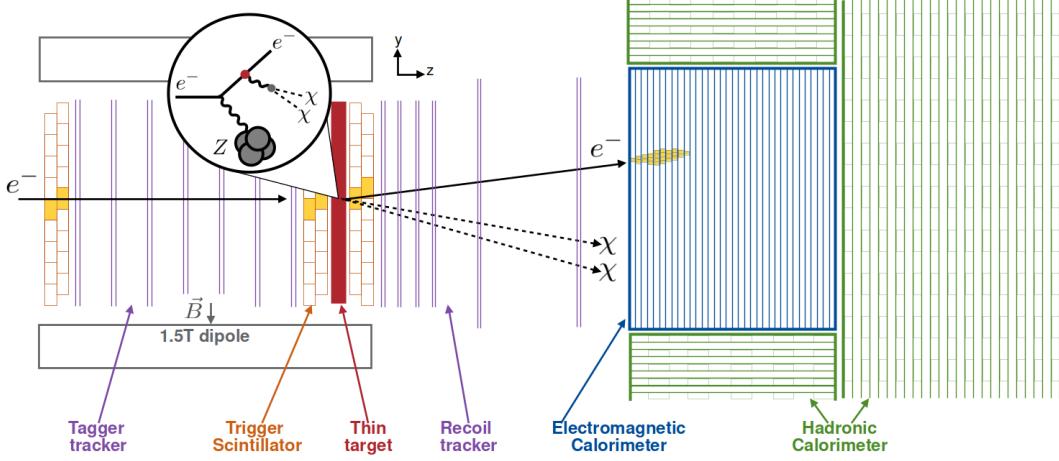


Figure 3.2: Diagram of LDMX detector apparatus with a representation of a signal event where a dark brem occurs within the target. Credit to Christian Herwig for original development of diagram.

Calorimeter upgrade[29]. These hexagons are arranged in a “flower” providing excellent transverse resolution of shower location and shower shape. The layers are built and arranged in order to space the sensitive silicon sub-layers to give good longitudinal resolution of showers as well. In total, the designed ECal has more than one hundred thousand channels that can each individually detect particles depositing energies from ≈ 0.1 MeV up to ≈ 1 GeV. This high granularity calorimeter gives LDMX excellent discrimination power since it can measure the amplitude and location of several incident particles. Due to its high performance, the ECal is a primary tool for both designing a trigger decision (with the aid of the Trigger Scintillator’s electron count) as well as downstream analysis separating SM background processes from potential DM signal.

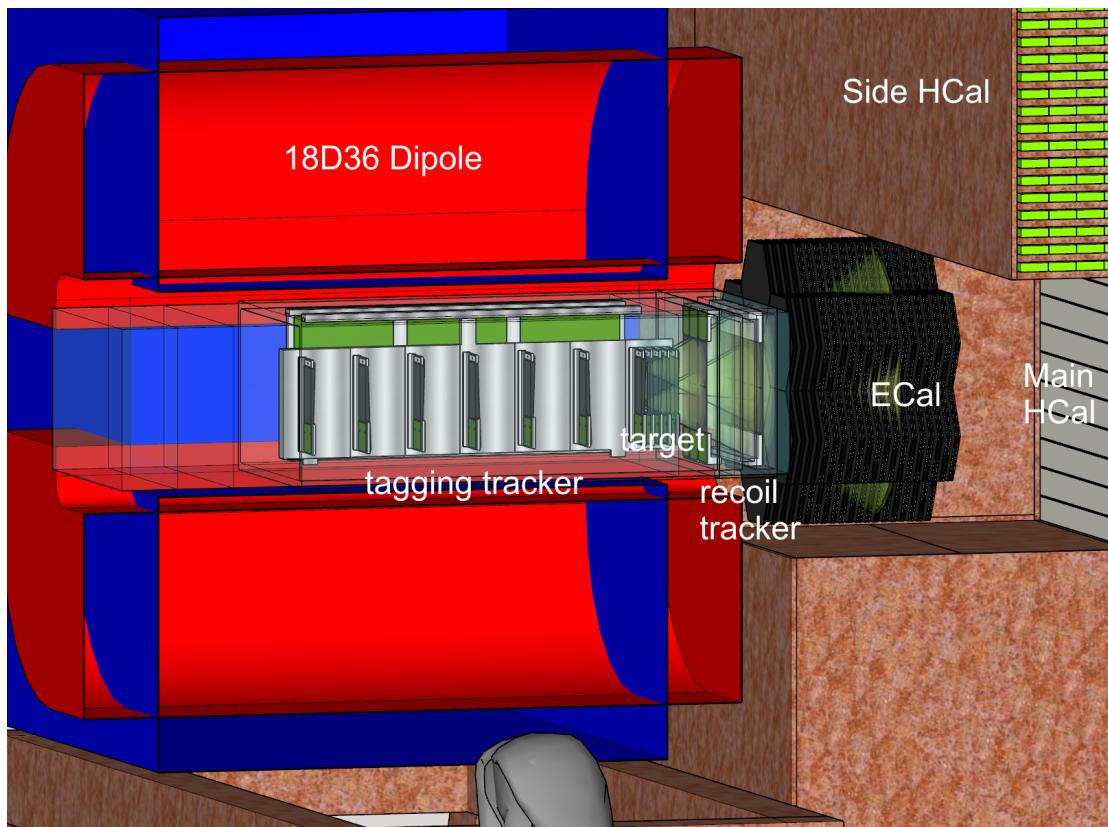


Figure 3.3: Rendering of LDMX detector apparatus focusing on tracker, target, and ECal. The magnet would fully encompass the tracker, target and trigger scintillator.

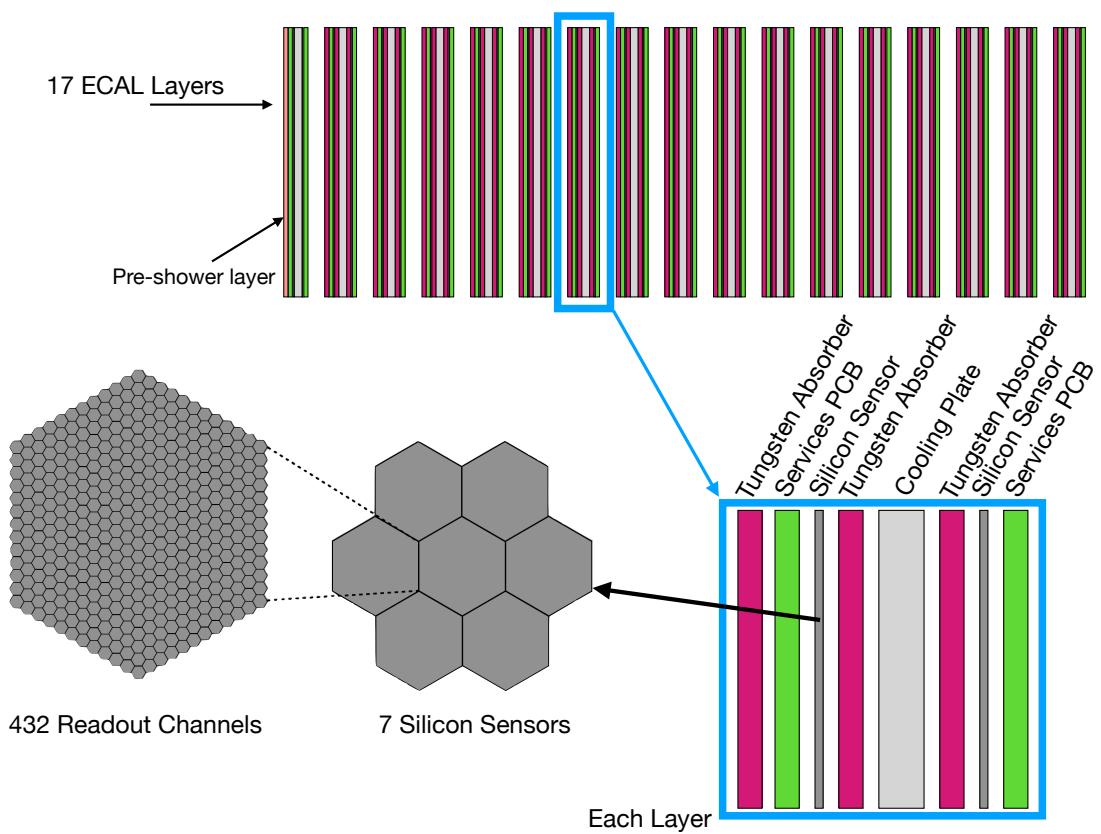


Figure 3.4: Diagram of LDMX ECal design showing the longitudinal segmentation (top and bottom right) and the transverse segmentation (bottom left). Credit to Joe Muse.

Chapter 4

Mid-Shower Simulation

LDMX (like many other HEP experiments) uses an intricate software stack in order to realistically and efficiently simulate particle interactions with the detector, emulate the electronics that would be used to measure these interactions, and reconstruct the output of these electronics into physically-understandable variables. These software tasks are accomplished by a wide swath of different software packages, some of which custom-written for LDMX, most of which written in C++. This chapter is focused on describing this simulation infrastructure – focusing particularly on parts of the infrastructure I was involved in – while also pointing out areas that are expected to remain constant in the presence of data gathered from a real detector. After a discussion of general data processing, I move into discussing the specific samples used to do this simulation study.

4.1 General Data Process

One of the core principles helping organize HEP data is the concept of an “event.” In the context of LDMX, an event is the data collected within a small window of time around the arrival of a beam electron into our detector apparatus. In essence, each event within the software is independent from one another; however, they all share a similar structure to the information they hold. A natural example is a data table: a row has the same variables in each of its columns as all the other rows. Events behave the same way; however, unlike a data table, the structure of an event can be more intricate than simply a series of values corresponding to different column titles. While

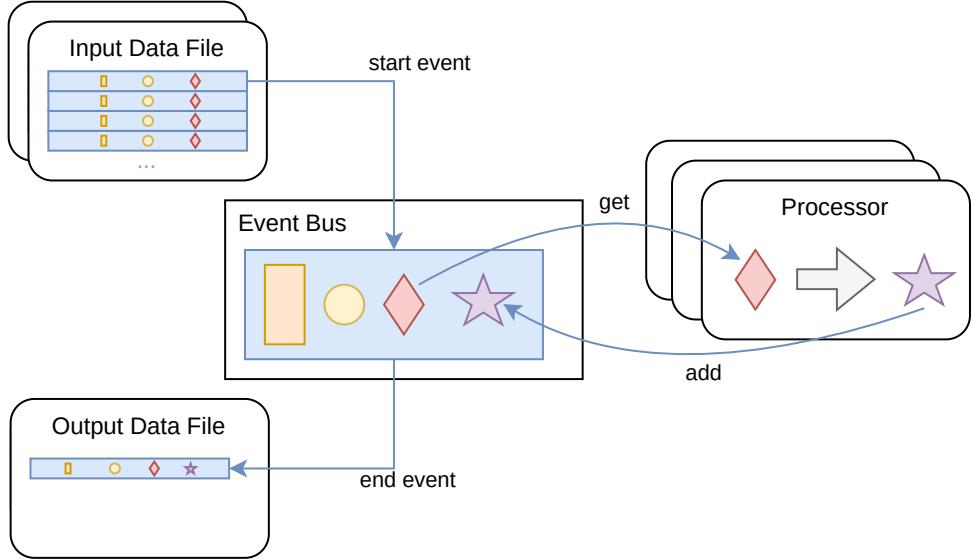


Figure 4.1: Flow chart of how data is processed in the LDMX framework. Each processor has the ability to “add” data to the event as well as “get” data from the event. The processors are run in a user-defined sequence. Data can also be loaded from one or more input data files where the event data from those files is loaded into memory before the first processor is run. After all processors are done with an event, it is saved to the output data file.

our basic “unit” of data is an event, we require many events in order to make statistical conclusions about our data; thus, we have developed an “event processing framework” that allows us to unify the various aspects of processing the data held within an event.

The event processing framework designed, developed, and maintained by LDMX is designed to allow the flexibility necessary to do the wide range of tasks necessary for the experiment. The C++ framework uses CERN’s ROOT [30] for data serialization, Boost for logging, and Python [31] for dynamic run-time configuration. As diagrammed in Figure 4.1, the design is a sequential model: the “event bus” stops at the individual processors in a certain sequence. The individual processors can inspect the data currently on the bus and board more passengers onto the bus for later processors to use and which can be eventually serialized into the output ROOT file. The processors can be built separately from the framework and then dynamically loaded and created at run-time by

an abstract factory. This design choice allows for all of the computationally-intensive software tasks necessary for LDMX (simulation, reconstruction and some analysis tasks) to use this framework and be organized into separate modules which are only loaded into memory when that module is being used.

The serialization portion of the framework is similarly dynamic; focusing on enabling users' code to add data structures from the most simple (e.g. individual booleans) to the most complicated (e.g. containers of custom classes). This wide array of data types is supported by ROOT's dictionary system during the serialization stage and abstract wrapper classes with partially-specialized template derivatives during run-time. Such complexity within the framework is necessary in order to allow a simple interface – one where the user interacts with simple and complex types in the same way.

Combining this highly dynamic serialization library with the sequential-processing model configured at run-time gives a strong foundation for all of the software needs of LDMX. Written in C++, this software framework enables high performance for all of the major data processing tasks necessary for the experiment. Moreover, its design focuses on flexibility and modularity so that seemingly-disparate data processing tasks can be unified under one framework. Everything from simulation to detector emulation to event reconstruction to analysis calculations can be done within this framework, reading and writing files from this framework and enabling our software to be well organized while also centralized in one location.

4.1.1 Data Processing Stages

The centralized nature of the LDMX processing framework makes it much simpler to stay unified as a collaboration. Experts are able to work on their specialized area of the software and share those improvements with the entire collaboration with ease. Since the flexibility of the framework allows for arbitrary groupings of these different data processing stages, we can choose to separate them into natural groups that correspond to the different areas on which experts focus (diagrammed in Figure 4.2).

In this work (as the chapter title implies), we focus mainly on the detector simulation stage where events are produced using random sampling of relevant physical phase spaces in a way geared towards realistically modeling the detector and particles interacting with it. LDMX like most HEP experiments use GEANT4[32] to help perform this

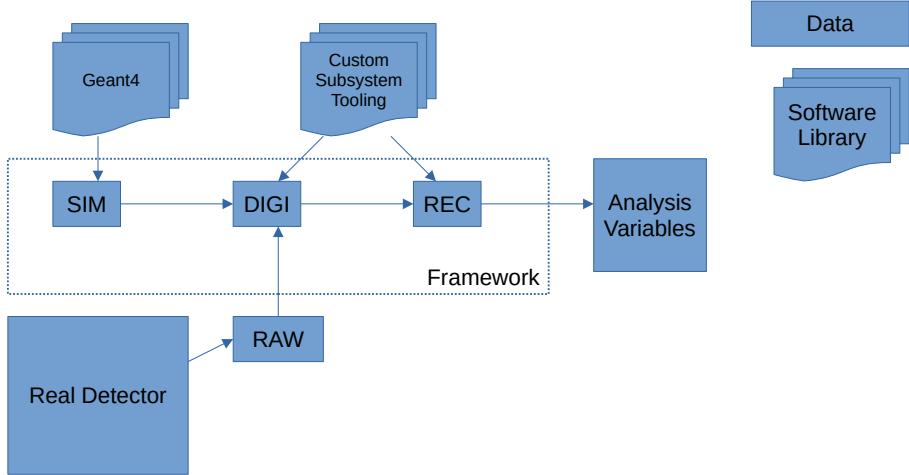


Figure 4.2: Diagram of processing stages within LDMX showing the external sources of data or software that are used in those stages as well as which stages are commonly done within the centralized processing framework.

complicated detector simulation. The downstream stages, namely electronic emulation and reconstruction, are not described further here; however, it should be emphasized that all subsystems of LDMX have their own custom implementations of these stages in order to realistically emulate and reconstruct the data within their subsystem.

4.1.2 Biasing and Filtering Technique

Many of the processes most interesting for a DM search within LDMX are rare relative to processes that are “easier” to reject by normal data processing. In many cases, the rarity of these more interesting processes is a computational roadblock since waiting for the entire detector simulation (hundreds to thousands of particles each with dozens if not hundreds of interactions) to complete before looking for these processes of interest wastes a lot of computer time. Typically, only one out of every ten thousand events actually contains a process we are interested in. To make simulation of these processes more computationally feasible, we turn to the common simulation techniques of biasing (artificially increasing the probability of a certain process occurring) and filtering (proactively ending the simulation of events if certain criteria are not met).

In the analysis channel studied here, all of our interesting processes occur within the developing shower of particles in the ECal. While more complicated than checking for processes happening to the single beam electron, we can select specific processes occurring within this shower in a computationally efficient way by tuning the order with which the simulation processes particles. The detector simulation is focused only on particle-material interactions and the materials are assumed to not change significantly during these interactions, so the order in which particles are simulated does not change how the physics is modeled. We choose to process particles in two groups – “high” and “low” energy – where the high-energy particles are all processed before the low-energy particles. The border between these two (called the “Sorting Threshold”) is configurable by the user and enables a specific GEANT4 function to be called mid-shower after all of the high-energy particles are processed (i.e. simulated until their total energy falls below the Sorting Threshold).

Suppose we know from studying other (unbiased or low-bias) samples that showers need to have a minimum amount of energy entering a specific process of interest before the shower becomes “important.” We can set this minimum energy (here called the “Filtering Threshold”) as a requirement on the ECal shower for the simulated event to be kept within the sample and, in order to improve the speed of the filtering, we can apply this requirement *during* the simulation when the event transitions from the high-energy group to the low-energy group of particles.

Now, this infrastructure of determining a minimum simulated energy and having an energy-based sorting of the processing order already helps improve the computational efficiency of these simulation samples (from $\sim 6k$ CPU-hours to ~ 30 CPU-hours for the “important” events from $\sim 1B$ total events). Nevertheless, it is still extremely inefficient (most simulated events do not pass our filter requirements) and so we add biasing in order to artificially increase the rate of the process we wish the sample to focus on. Since the processes we are interested in biasing are connected to particles that are abundant within the electromagnetic ECal shower, we need to require only particles above a certain energy threshold (the “Biasing Threshold”) to avoid over-sampling of the process. In addition, we also need to define the factor that we will multiply the cross section by when a particle is above the Biasing Threshold and in the ECal volumes (the “Biasing Factor”).

Artificially increasing the rate of a process relative to its true rate measured in nature is obviously unphysical, but we can account for this by weighting the events depending on this Biasing Factor when counting them. The simulation engine we use does this weight calculation for us and it accounts for both the biasing factor (beginning the weight with a factor of $1/B$) and unbiased processes happening to a particle with biasing enabled (increasing the event weight to reflect that physical processes at the natural rate have occurred).

Parameter	Dimension	Symbol	Short Description
Sorting Threshold	Energy	T_S	Minimum total energy to be processed in first group
Biasing Threshold	Energy	T_B	Minimum total energy for a particle to be biased
Biasing Factor	Dimension-less	B	Factor to multiply cross section by
Filtering Threshold	Energy	T_F	Minimum energy transferred into a process for the event to be kept

Table 4.1: Parameters of mid-shower biased samples. In all of the samples used in this study, we use the same value for the Sorting and Biasing Thresholds.

Table 4.1 summarizes the parameters related to this sample generation technique. Practically, we choose to have the Biasing and Sorting Thresholds share the same value for all of the samples in this work in order to reduce the number of parameters needed to tune and to process all of the biased particles (as well as particles of similar energies but different flavors) before making any filtering decision.

Generally, validation of this biasing and filtering technique can be done by checking that the samples generated with this technique match samples generated without this technique and with the filtering applied after the fact. This validation needs to be done on a per-sample basis since the specific aspects of the simulation that need to align may change depending on the specific physics for which we are biasing and filtering.

4.2 Standard Processes

As displayed in Figure 3.1, there are a few categories of physics processes that are described by the standard model and we expect to happen within our experiment. These “background” processes – while interesting for other analyses of LDMX data –

should be rejected by this search for DM since they are understood to *not* be DM by the standard model and other experiments. In order to thoroughly and faithfully estimate our ability to reject these backgrounds, we need to realistically simulate them within our detector volume.

GEANT4 [32] has been developed for precisely this purpose. This study uses a slightly modified copy of GEANT4 v10.2.3 which focuses on improving the accuracy of processes relevant to a DM search within LDMX with the following needs.

1. Updating cross section and sampling of $\gamma \rightarrow \mu^+ \mu^-$ process to align it better with collected data and calculations of the standard model.
2. Update nuclear cascade to align with more recent CLAS data specifically focused on the rate of low-multiplicity forward neutrons (Appendix A.A of [25]).
3. Introduction of *back scattering* π^0 in a $\gamma p \rightarrow \pi^0 p$ within a nuclear cascade (Appendix A.B of [25]).

These updates enabled us to more reliably study the known and expected background processes interactions with the LDMX detector design and compare this data to simulations of a model of a DM production process.

4.2.1 Dominant Contributors in this Analysis

First and foremost, the analysis detailed in Chapter 5 is focused on utilizing data collected with LDMX that is not used in the nominal cro:mmMissing Momentum (MM) analysis. This design goal produces two requirements on the samples we will analyze.

- Same Trigger Requirement – while designing a separate trigger for different analyses is feasible, this specific analysis is focused on working with the same data that is collected for the primary MM analysis. The trigger requirement for the primary MM analysis is one of missing energy within the ECal making it an appropriate trigger for this analysis as well.
- Opposite Tracking Requirement – the primary MM analysis enforces the assumption that the processes of interest occur within the thin target; thus, the electron track leaving the target (colloquially named the “recoil”) is required to have less

than 30 % of the beam energy. In this case, we flip this requirement and instead require the electron to arrive at the ECal with more than 87.5 % of the beam energy. While the application of this requirement would be done after the trigger requirement in real data, for simulation we apply this requirement first so that the simulation samples we are studying can be more efficiently produced.

All of the simulation samples studied here require the primary beam electron to arrive at the ECal with more than 87.5 % of its original energy. Once this requirement is imposed, we then turn our attention towards the processes that occur within the ECal.

Generally, we focus on processes that would cause the energy in the ECal to be estimated at a significantly smaller value than the known beam energy. This “missing” energy is used as a key piece of evidence that the event in question should be stored for further study and thus is the main ingredient in the infrastructure deciding whether a specific event should be stored (the “trigger”). From prior experience with LDMX (e.g. [25, 2]), we expect certain types of standard processes to dominate the backgrounds that pass this trigger requirement on missing energy: interactions with the nuclei of the detector material producing hard-to-detect hadrons (so-called “nuclear” processes) and conversion of photons into a pair of muons (which are also difficult for the ECal to contain).

Nevertheless, we still can use a large, unbiased simulation sample in order to verify these previous conclusions and inform our decisions on how the subsequent biased samples should be produced. Figure 4.3 is a key piece of evidence along these lines where we separate the total reconstructed energy (the estimate from our detector) as a fraction of the known beam energy by the amount of energy going into these nuclear processes (so-called “nuclear” energy). We see that as more energy is given to these nuclear interactions, the reconstructed fraction decreases. Eventually, we reach a point where enough energy has gone to these nuclear interactions that we expect the event to be below the trigger threshold and be kept by the trigger system at a significant rate.

Figure 4.3 shows that a significant fraction of events whose total nuclear energy is greater than 62.5 % of the beam energy fall below the threshold for the trigger. This figure also shows that there are other types of events that do not have much (if any) nuclear interactions (since their nuclear energy is less than 10 % of the beam energy) but still are reconstructed below the trigger threshold. Further investigation reveals

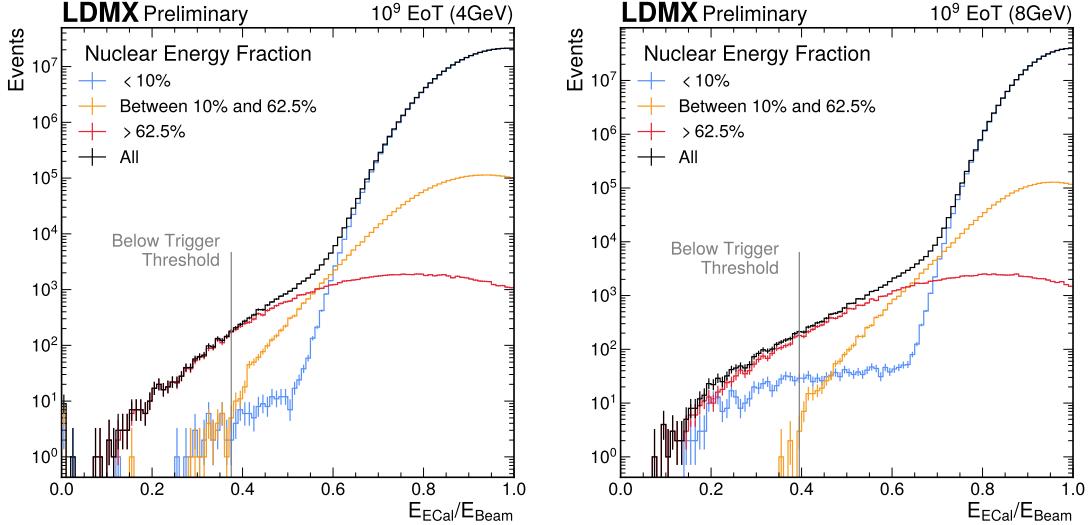


Figure 4.3: The total reconstructed energy within the ECal as a fraction of the beam energy separated by the total nuclear energy within the event (the energy going into the photon-nuclear and electron-nuclear processes).

that these events are indeed “di-muon” events (Figure 4.4). These two categories of processes – nuclear and di-muon – are the primary background processes since they mimic our missing energy trigger selection at a significant rate.

4.2.2 Expanding Production of these Processes

As mentioned in Section 4.1.2, the unfiltered and unbiased simulation is not very efficient. In some sense, this shows that our basic trigger selection is doing a good job of rejecting the background processes that our analysis finds uninteresting; nevertheless, we need to efficiently sample events that contain interesting background processes at a large scale so we can analyze them in detail. This is where the biasing and filtering technique described above is employed – our main goal is to improve the speed with which we can generate events originating from these background processes while maintaining faithfulness to the unbiased and unfiltered distributions.

The process of tuning these simulation variables is largely a guess-and-check one. We test a variety of different settings and see how they perform both in terms of their computational speed as well as their faithfulness to the underlying distributions. For

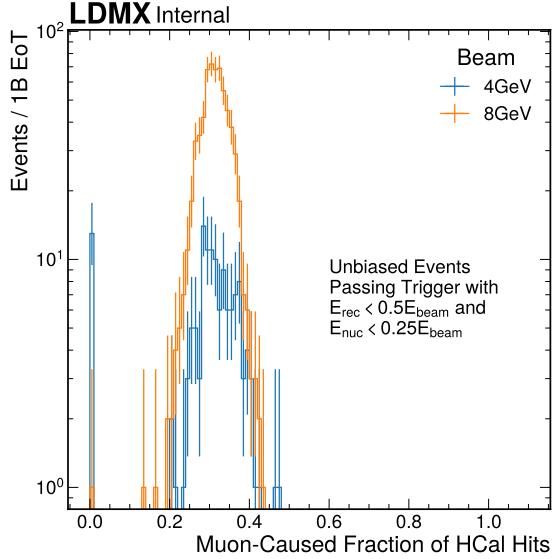


Figure 4.4: Fraction of hits within the HCal that were caused by muons or anti-muons. The events lying within the zero bin still have HCal hits caused by other types of particles.

our purposes, the filtering threshold T_F can be estimated from the large unbiased and unfiltered sample since that sample contains events with these processes of interest. Specifically, Figure 4.3 displays a value of $T_F = 0.625E_{\text{Beam}}$ for the nuclear processes. More detailed study of the events residing in the low-side tail of the blue distribution showed that $T_F = 0.5E_{\text{Beam}}$ was a good choice for the di-muon process; however, since there were fewer events containing the di-muon process within this sample, other choices of T_F were surveyed in order to confirm this value.

The other threshold T_B (which we also set to be T_S in this work) is tightly coupled with the biasing factor B and so we test them both together. A rather simple proxy for checking that the events sampled during biasing and filtering matches the underlying distribution is to inspect the longitudinal distribution of where the process occurred. While the biasing and filtering should improve the speed of the simulation, it should not observably distort this distribution which roughly follows the development of the shower (an example of this is shown in Figure 4.5).

We can confirm that our biased and filtered simulations are properly extending the tail of the distributions in which we are interested. Figure 4.6 shows a comparison

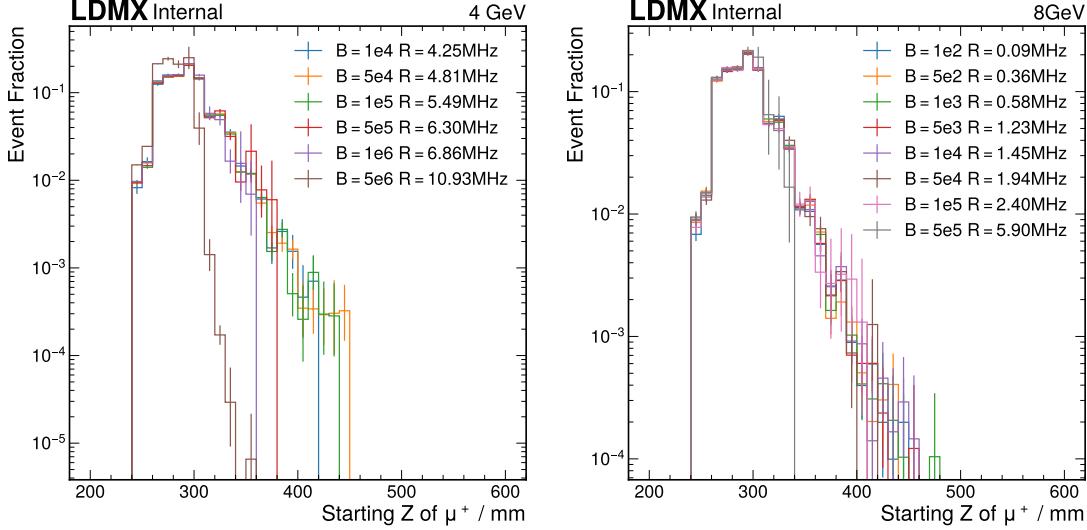


Figure 4.5: Longitudinal location (z) where the μ^+ was produced within the simulation for a variety of biasing factors B for the 4 GeV beam (left) and 8 GeV beam (right). The rate R is the equivalent number of unbiased EoT divided by the CPU time necessary to produce the sample. We can observe that while increasing B also improves the rate R , we eventually over-bias and lose access to the late tail of the distribution (red, purple, and brown in the left plot and gray in the right plot).

between a large unbiased sample that has been scaled to the same equivalent EoT as the biased and filtered samples – there we see that the biased samples properly and smoothly extend the low reconstructed energy fraction tail which corresponds to the signal region in this missing energy search.

4.3 Dark Matter Signal

The particular signal process this analysis channel is looking for is the production of a dark photon followed by an *invisible* decay. In this regime, what happens to the dark photon after it is produced is irrelevant to the analysis since both it and its products are not observable by our detector.

With this focus in mind, we developed a dark bremsstrahlung simulation method that allows for the visible particle (the recoiling lepton) to be distributed according to a full matrix element calculation (via MG/ME) while the incident particle can have

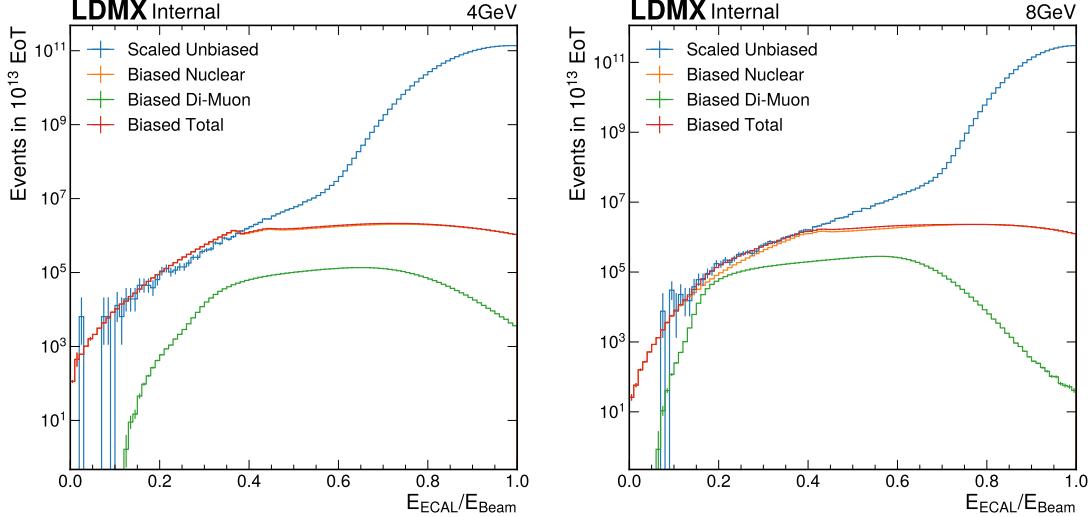


Figure 4.6: ECal reconstructed energy E_{ECAL} as a fraction of the beam energy E_{Beam} comparing the unbiased and biased samples scaled to the same EoT. The 4 GeV (8 GeV) beam case is shown on the left (right).

variable energy and be handled by GEANT4 directly. This novel simulation technique allows for the dark bremsstrahlung process to be treated (from GEANT4’s perspective) on the same footing as the background processes while maintaining the precision of a matrix-calculator method.

4.3.1 G4DarkBreM

To accurately simulate the kinematics of the dark bremsstrahlung process for electrons in thick targets, the process must be included at the level of experimental simulation instead of using initial state event generators to account for the possibility of energy loss through bremsstrahlung or multiple scattering prior to the dark matter interaction. Accurate kinematic simulation of the outgoing electron are required for optimal experimental sensitivity measurements and appropriate design of search strategies.

For this reason, we utilize G4DarkBreM [33] which performs this embedding of the dark bremsstrahlung process into GEANT4. G4DarkBreM calculates the cross section using numerical integrals of the Weizsäcker-Williams approximation, and the kinematics are simulated using a scaling technique of MADGRAPH/MADEVENT event libraries. The accuracy of the total cross section and kinematics is validated using

MADGRAPH/MADEVENT samples at a range of incident lepton energies.

4.3.2 MadGraph/MadEvent

For this study in LDMX, the code package used to develop and validate G4DarkBreM was also used to generate the input reference libraries for sample generation. This code is a custom version of MADGRAPH/MADEVENT4 with the following updates.

- Introduction of basic dark sector particles (massive boson and spin-1/2 fermion) which act as the representatives of the dark sector interacting with the standard model particles.
- Definition of a nuclear particle (electrically neutral, spin-1/2 fermion) with new couplings to the dark photon including the nuclear form factors.
- Updating the definition of the electron to include its small (but non-zero) mass to prevent divergence of the cross section at lower energies.

These updates, along with some wrapping code, enable the generation of dark bremsstrahlung events for a range of target nuclei, incident energies, and either incident electrons or muons.

As suggested by the validation of G4DarkBreM [33], we generate libraries where the incident electron's energy drops in steps of 10% down from the beam energy. Since the dark photon is required to be simulated with at least 50% of the beam energy, we stop the library generation also at 50% of the beam energy. The primary nuclei that could dark brem within the ECal are tungsten, silicon, oxygen, and copper so those are the nuclei for which we generate libraries. The rest of the nuclei within the simulated ECal (sodium, calcium, carbon for example) have atomic numbers close to those nuclei already sampled and so can be faithfully simulated with these libraries. In order to avoid duplicating the same event, a unique reference library is generated for each simulation run even when keeping the dark photon mass and beam energy the same.

4.3.3 Characterization

These samples, as expected, show that the dark bremsstrahlung process can occur following normal shower development peaking at $\sim 1X_0$ into the ECal (with the exception

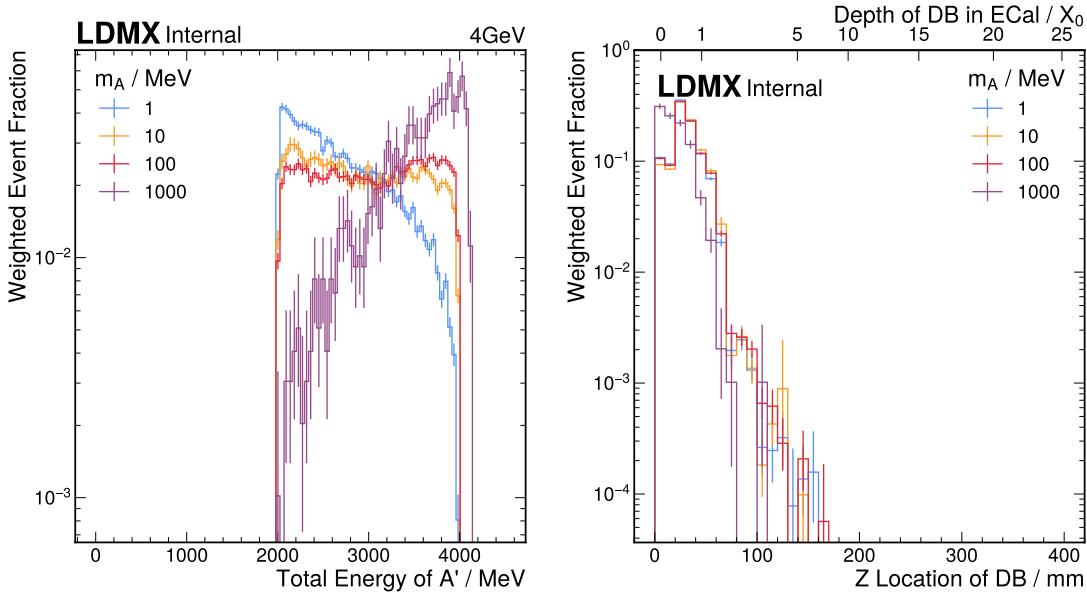


Figure 4.7: Distributions of simulated signal events with a 4 GeV beam. The energy of the produced dark photon is shown on the left while the longitudinal location where the production occurred is on the right.

of higher mass dark photon samples which are kinematically prevented from being generated further into the shower). Figure 4.7 shows some example distributions from these samples. We can observe the normal shower development behavior in the location of the dark bremsstrahlung and we can see the simulation-level cut on the dark photon energy at half the beam energy (in this case 2 GeV) which is imposed to improve computational efficiency without losing events that would otherwise pass the downstream trigger requirement.

4.4 Summary

While this chapter has been a long and winding road through the simulation infrastructure of LDMX, the key feature that I would like to emphasize is that it is able to efficiently and reliably simulate both background and signal processes within this search for DM. These samples, after being thoroughly studied and validated, were enlarged to provide a sufficiently sized sample for the downstream analysis of Chapter 5.

Sample	Biasing Factor	Biasing Threshold
Signal	$m_A^{\max(\log_{10}(m_A), 2)} / \epsilon^2$	$0.5E_{\text{Beam}}$
Enriched Nuclear	200	$0.375E_{\text{Beam}}$
Di-Muon	10^5	$0.5E_{\text{Beam}}$
Sample	Filtering Cuts	
Unbiased	$E_{\text{primary}}^{\text{ECal Front}} \geq 0.875E_{\text{Beam}}$	
Signal	$E_{\text{primary}}^{\text{ECal Front}} \geq 0.875E_{\text{Beam}} \& E_{A'} \geq 0.5E_{\text{Beam}}$	
Enriched Nuclear	$E_{\text{primary}}^{\text{ECal Front}} \geq 0.875E_{\text{Beam}} \& E_{\text{tot nuc}} \geq 0.625E_{\text{Beam}}$	
Di-Muon	$E_{\text{primary}}^{\text{ECal Front}} \geq 0.875E_{\text{Beam}} \& E_{\text{tot } \mu} \geq 0.5E_{\text{Beam}}$	

Table 4.2: Configuration of the simulation samples used in this analysis. E_{beam} is the beam energy being studied (4 or 8 GeV in this work). m_A is the mass of the A' in MeV, ϵ is the dark brem mixing strength, $E_{\text{primary}}^{\text{ECal Front}}$ is the energy of the primary electron at the front of the ECal, $E_{A'}$ is the energy of the generated A' , $E_{\text{tot nuc}}$ is the total energy transferred to nuclear interactions during the event, and $E_{\text{tot } \mu}$ is the total energy of produced muons.

Specifically, the background samples were generated to be statistically equivalent to 1×10^{13} EoT and the signal samples were generated to provide ~ 1 M events for each dark photon mass point to study selection efficiencies. Table 4.2 gives a summary of the configuration used to generate these samples.

Chapter 5

Missing Energy Search

One of the primary strengths of the LDMX detector design is its ability to use the tagging and recoil tracker system to reject a large number of background events by separating a nominal beam with non-standard energy loss from a nominal beam with standard energy loss or even low-energy beam. Moreover, the tagging and recoil tracker system gives LDMX the potential to further suppress backgrounds or potentially study DM properties by studying the transverse momenta of electrons recoiling from dark-bremsstrahlung candidate events.

While this design is optimal for a large number of Electrons on Target (EoT), the strategy has some limitations for a low EoT data run. To limit multiple scattering which ruins momentum measurements, the baseline detector configuration requires a thin ($\approx 0.1X_0$) target. In the early stages of LDMX, when the total EoT will be lower, a different analysis strategy and detector configuration may be optimal to probe the largest amount of the $y-m_\chi$ phase space. The alternative strategy ignores the dedicated target inside the tracker volume and instead uses the ECal as an active target. The EaT analysis channel for LDMX thus has two primary purposes which can be separated by the timeline over which they are relevant.

1. **Short Term:** In early running, when the number of EoT is relatively small, the nominal MM analysis will not have obtained significant reach into new DM phase space (yet). The EaT channel serves here as a way to obtain world-leading sensitivity early in the lifetime of LDMX and give the collaboration a first look at

the data the apparatus has collected.

2. **Long Term:** As LDMX collects data, the MM analysis enters into unexplored phase space and serves as a better discovery mechanism due to its access to the Tagger and Recoil trackers. The EaT channel, while struggling to suppress complicated backgrounds with relatively limited analysis handles, can operate “orthogonally” in the collected data since its primary selection (an approximately beam-energy electron passing through the Recoil tracker) is inverted relative to the MM analysis (the electron passing through the Recoil tracker has significantly less energy than the beam).

An initial study of the EaT analysis channel is the primary focus of Part II of this thesis, focusing primarily on the first (short term) purpose. In this regard, we target an EoT that is reasonable to accomplish early in the running of LDMX and avoids particularly intricate backgrounds. A total of 1×10^{13} EoT fits these requirements by avoiding the charged current production of neutrinos and represents $\sim 2.5\%$ of the first full LDMX dataset which is obtainable within approximately a two weeks of nominal beam time¹. Since EaT is expected to be the *first* physics analysis on LDMX data, we want a *simple* and *robust* analysis that can withstand the test of time and the complexities of real data.

With these design goals in mind, a simple “cut-and-count” analysis has been developed. The simplicity of this analysis is one of its strengths, enabling it to be applicable despite potential surprises arising from first encounters with real data. The bulk of time and effort on this first investigation was focused on making this investigation *possible* via the introduction of midshower process filtering and a dark bremsstrahlung simulation process described in Chapter 4.

5.1 Selection

The core goal of most search analyses is to develop a selection that avoids events known to be standard processes (i.e. backgrounds) while keeping events containing the process

¹ Assuming the LDMX detector apparatus and beam delivery is operating according to specifications, we expect the beam to be delivered on a frequency of 37.5 MHz with a duty cycle of ≈ 0.5 and the number of electrons within each bunch to be Poisson distributed with $\mu = 1$. $37.5 \text{ MHz} \times 0.5 \times P(\mu = 1, 1) \times 2 \text{ week} \approx 1 \times 10^{13}$ EoT.

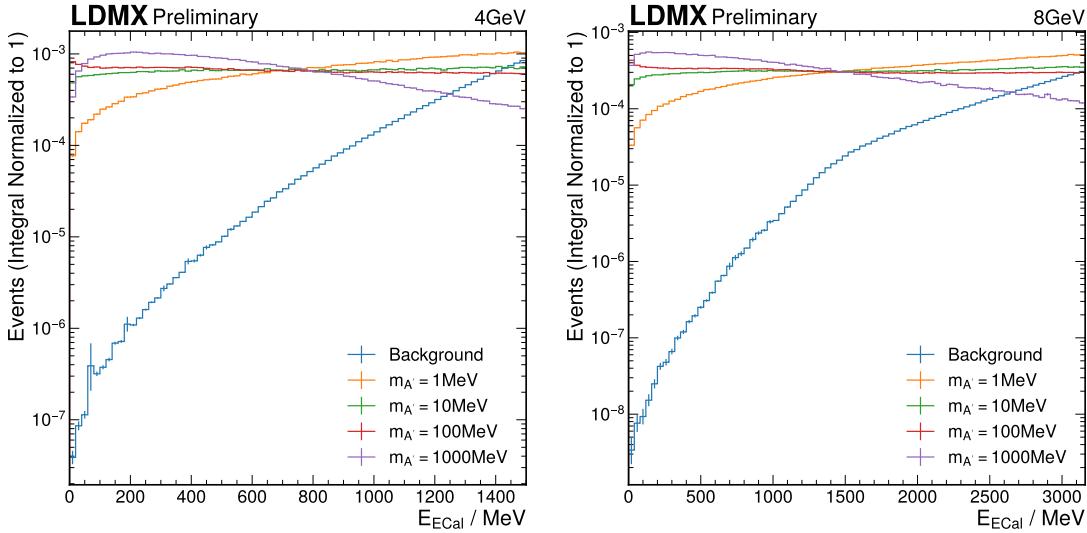


Figure 5.1: The total reconstructed energy in all layers of the ECal (E_{ECal}). The signal and background distributions are normalized such that their integral is one. The 4 GeV beam is shown on the left and the 8 GeV beam is shown on the right. The events falling into bins with $E_{\text{ECal}} > 1.5$ GeV (3.16 GeV) for 4 GeV (8 GeV) are omitted from this plot but included in efficiency calculations.

being searched for (in this case, the production of DM via dark-bremsstrahlung). Both the EaT analysis channel and the primary MM channel share the dark-bremsstrahlung signature of missing energy within the ECal relative to the known incident beam energy. This allows for the first selection made to be shared between these two channels – a requirement that the sum of the observed energy in the first twenty layers of the ECal is less than 1.5 GeV (3.16 GeV) for a 4 GeV (8 GeV) beam. This preliminary selection acts as the primary trigger for both of these analyses – selecting events that resemble dark-bremsstrahlung via their lower-than-average observed energy during data collection – however, the EaT analysis channel in the early-running scenario may not be collecting data with a fully calibrated energy scale at the trigger level. The potential for mis-calibration motivates tightening the analysis missing energy requirement by 400 MeV as well as requiring the sum to be performed over all thirty-four layers of the ECal instead of only the first twenty. Figure 5.1 shows the unit-normalized distributions of this total reconstructed energy within the ECal after the trigger has been applied.

After this total energy requirement, there are still backgrounds left and those that

remain are largely events where a significant fraction of the energy is carried by particles difficult to contain (or even observe) within the ECal (neutrons, K_L , or muons for example). To suppress these types of persistent backgrounds, the event is required to have less than 10 PE deposited in any single bar of the HCal – $\text{max}(\text{PE}_{\text{HCal}}) < 10$ – well below the typical signal of a through-going muon of 80 PE in any single bar.

While this selection does a good job of removing events containing long-living particles, there are some events containing several low-energy particles that range out within the ECal itself. These particles (often due to their number) distribute energy over a range of cells, so we can suppress these events by requiring the spatial distribution of energy within the ECal to be small. We choose to measure the spatial distribution using the energy-weighted root-mean-squared spread of the hits within the ECal² – the ECal Hit RMS – which is typically larger for these background processes than for a single, truncated electromagnetic shower present within a signal event. We therefore require the ECal Hit RMS to be less than 20 mm.

These two selection variables are shown in Figure 5.2 where the distributions are shown after all other selections are applied (including the trigger selection and the tighter selection on the total ECal energy including all layers). Table 5.1 shows the tabulation of this selection for the studied simulation samples. Figure 5.4 shows the signal and background total ECal energy distributions after these selections are applied where we can still observe difference in the shapes of these distributions which we will exploit in the final statistical analysis of these samples.

An additional, sneaky selection has also been applied due to how the simulation samples were constructed. As mentioned in Section 4.2, the simulation samples have a filter requiring the primary electron to arrive at the ECal with at least 87.5% of the beam energy. This simulation filter is attempting to replicate a requirement by the tracking system that would be applied after the trigger requirement when analyzing real data. Studies of other simulation samples without this tracker requirement showed that while the tracker is required to help confirm a near-full-energy electron arrives at the ECal, it keeps over 90% of the events that are interesting for this analysis (i.e. nothing interesting happens in the thin target). Figure 5.3 shows example figures from

²A “hit” is defined as a single cell within the ECal registering a signal corresponding to 50% (or more) of the energy deposited by a typical muon during the event.

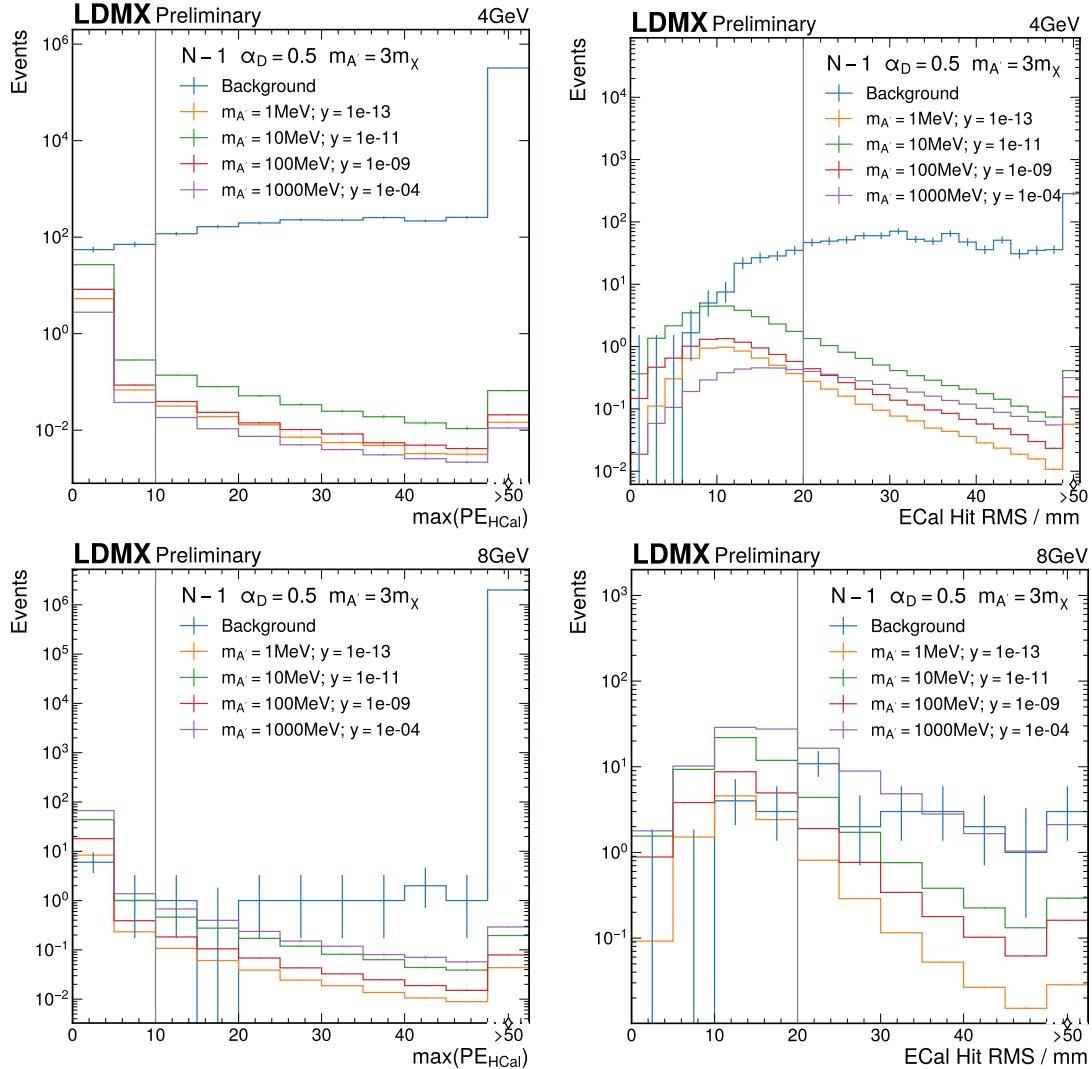


Figure 5.2: Variables used in signal selection for the EAT analysis channel. In each figure, all other cuts are applied except for the variable in question. Some of the bins are empty in which case the upper Poisson limit given the total sample size is drawn with error bars. The grey line shows the selection cut. The background sample is the enriched nuclear and dimuon samples. The top (bottom) row shows the 4 GeV (8 GeV) beam.

Analysis Stage for 4 GeV Beam	Background Event Yield	Signal Efficiency (%)			
		1 MeV	10 MeV	100 MeV	1 GeV
ECal Trigger ($E_{20} < 1.5$ GeV)	4.60×10^7	58	67	71	83
ECal Energy ($E_{\text{ECal}} < 1.1$ GeV)	1.95×10^6	35	48	53	72
max(PE _{HCal}) < 10	1.15×10^3	34	47	52	69
ECal Hit RMS < 20 mm	126	28	37	41	33
Analysis Stage for 8 GeV Beam	Background Event Yield	Signal Efficiency (%)			
		1 MeV	10 MeV	100 MeV	1 GeV
ECal Trigger ($E_{20} < 3.16$ GeV)	6.10×10^7	66	74	79	89
ECal Energy ($E_{\text{ECal}} < 2.76$ GeV)	6.88×10^6	52	63	69	84
max(PE _{HCal}) < 10	31.8	50	61	67	81
ECal Hit RMS < 20 mm	7	43	52	56	52

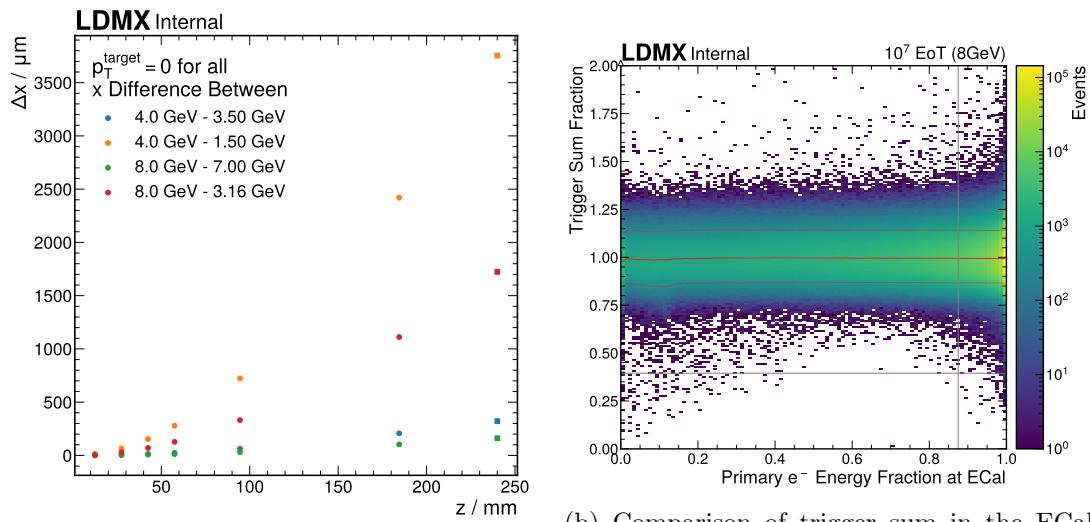
Table 5.1: Cut-flow analysis comparing background and various signal hypotheses for the simple cuts used in this analysis. The event yield for the background sample is calculated using the event weights and represent the number of events out of 10^{13} EoT equivalent. The signal efficiency is relative to the full simulation sample. The efficiency and event yield values on a given row are for *after* the analysis stage of that row. The first table is the cutflow for the 4 GeV beam, and the second table is for the 8 GeV beam.

these studies that confirm the necessity of the tracker while also displaying that it is highly efficient. Due to the tracker requirement’s high efficiency and the large separation between events that are acceptable into this analysis and events that are not, we neglect any systematic error (more on this in Section 5.3) related to this selection and instead plan to set the tracker requirement’s value to whatever threshold is necessary to keep the 90% efficiency when studying real data.

5.2 Background Prediction

When performing a search for new processes, a common technique is to develop a well-motivated prediction for the background processes and then look for excesses above this prediction. Before we go through rigourously defining what an “excess” means and how to quantify what types of new processes we would exclude if no excess is observed, we should more precisely define what our background prediction is.

We begin with fitting the cumulative background distribution with a simple exponential function. Figure 5.5 shows this cumulative distribution, the resulting fit, and the 95% confidence intervals surrounding the fit. This fit is a helpful prediction tool



(a) Difference in x position along the electron's path between different energy electrons. This shows that the change in position is smaller than a single ECal cell width, but it is well within a conservative alignment of the tracking detector.

(b) Comparison of trigger sum in the ECal to the primary electron's energy at the ECal. Below the horizontal gray line are the events that pass the trigger where there is large separation between events where something interesting happens in the target (left) and in the ECal (right).

Figure 5.3: Studies of how the tracker requirement affect the simulation samples.

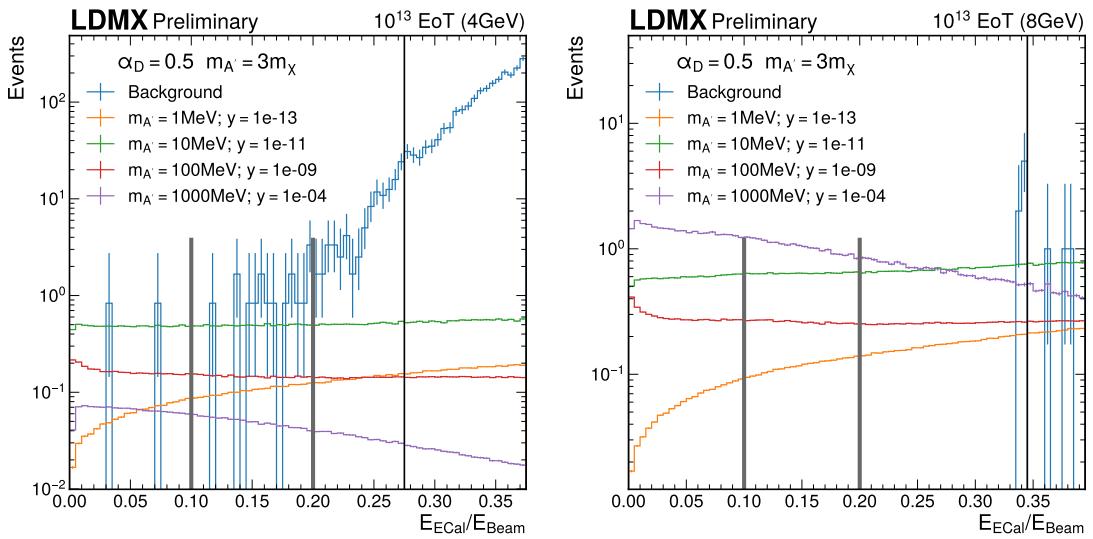


Figure 5.4: The total reconstructed energy in all layers of the ECal (E_{ECal}) as a fraction of the beam energy (E_{Beam}) for all samples that pass the selection criteria except the selection on ECal energy. The 4 GeV beam is shown on the left and the 8 GeV beam is shown on the right. The gray lines mark the edges of the analysis bins used to estimate the expected exclusion limit and the black line is the upper limit on the ECal energy which also serves as the upper limit of an analysis bin.

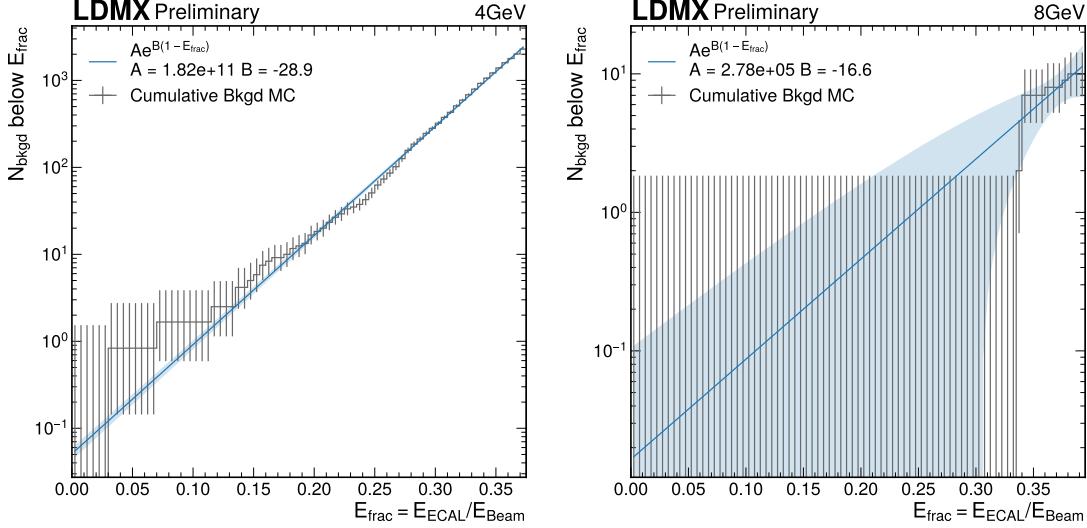


Figure 5.5: Exponential fit of the total simulated background distribution. The shaded region is a 95% confidence band on the fit which is interpreted as the uncertainty on the value of the fit. Some bins in the simulated distribution are empty in which case the upper Poisson limit given the sample size is drawn as an error bar in that bin.

in two keys ways. The fit produces a meaningful prediction across the entire range of observable values which is a helpful way to “smooth” out this sample that is limited in size after the selection criteria already applied. Moreover, the fit uses the 400 MeV-wide control region between the upper limit on the total ECal energy and the trigger threshold applied during readout to help constrain its parameters. This constraint reduces the uncertainty on the background prediction while also serving as an example of what an analysis using real data could do – this same region could be utilized in order to help set the scale of the background in a data-driven way.

The fit is then used to provide a background prediction within each of the three analysis bins, shown in Figure 5.6 along with the unconstrained simulation prediction for comparison.

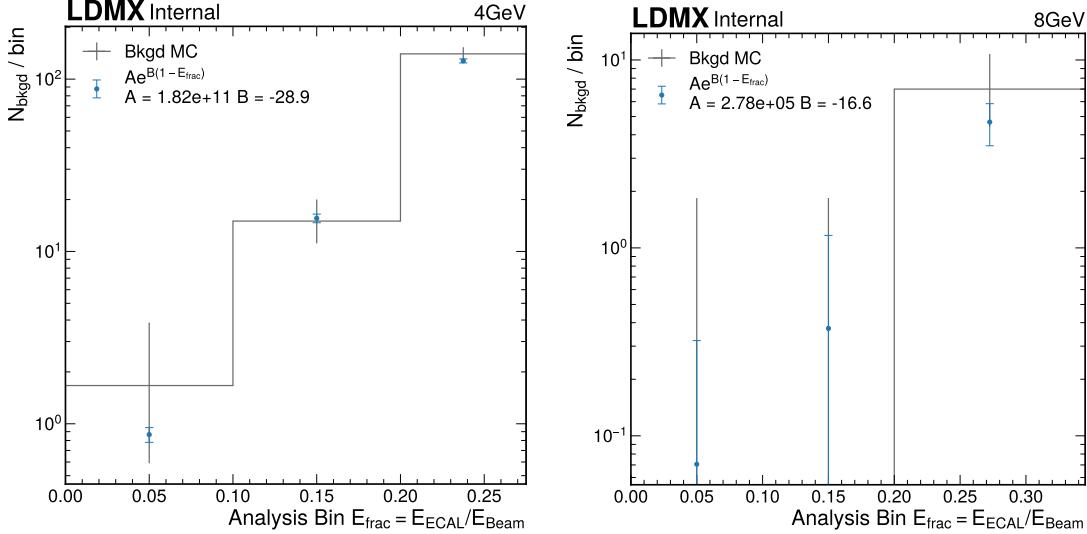


Figure 5.6: Background prediction within the three final analysis bins (blue) compared to the unconstrained simulation prediction in gray. The uncertainty on the background prediction is taken from the 95% confidence band shown on the fit. Some bins in the simulation prediction are empty in which case the upper Poisson limit given the sample size is drawn as an error bar in that bin.

5.3 Systematics and Background Uncertainty

Besides the statistical uncertainty due to the finite size of the simulation sample (whose estimation is represented by the 95 % confidence intervals around the fit), there is additional sources of uncertainty that originate from the experimental design (so-called “systematic” uncertainty). Most prominently in this early-running analysis, we expect mis-calibration of the ECal to be a potentially major source of systematic uncertainty especially since the calibration of the ECal used when applying the trigger selection will probably differ from the calibration used later on during the final analysis.³ Estimating this systematic uncertainty is rather direct; we simply vary the calibration used at the reconstruction and analysis stage within our data processing and see how that changes the results.⁴

³This difference is expected since we can improve the calibration of the detector as we collect more data; however, acknowledging this difference helps us put requirements on the accuracy of the intial calibration used at the trigger level in order for this analysis to function as expected.

⁴The estimate for both the ECal and HCal systematics is done with the 4 GeV beam sample since that estimate is expected to be more conservative because the 4 GeV beam sample allows more background

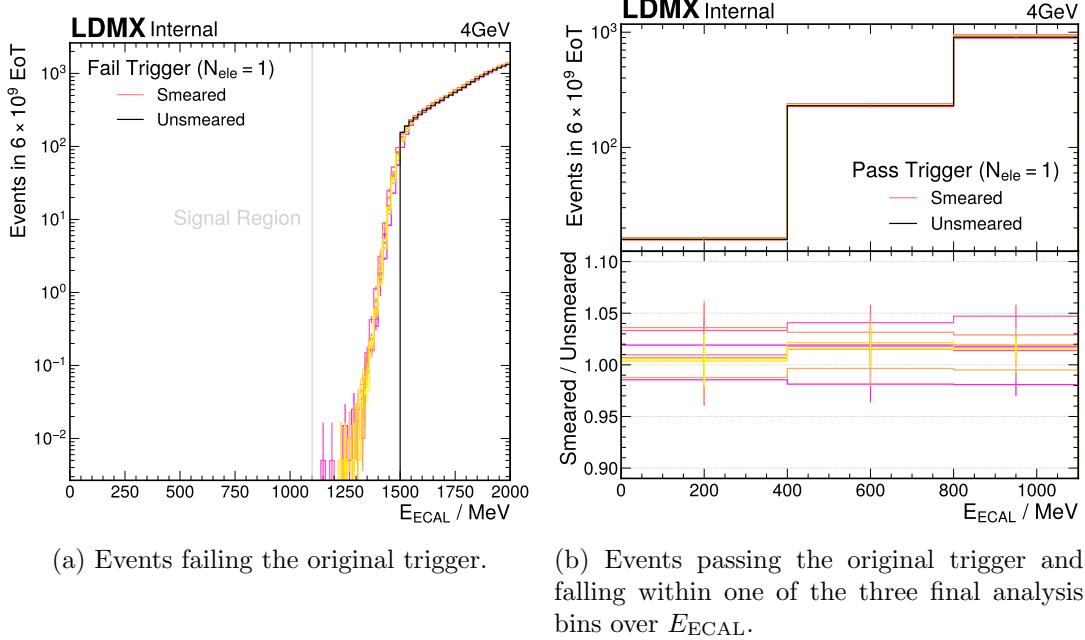


Figure 5.7: The total reconstructed energy in the ECal E_{ECAL} comparing ten different smeared calibrations (colors) to the original unsmeared calibrations (black).

We generated ten different calibrations, smearing the calibration constants by 10 % in an uncorrelated fashion and by an additional, correlated 2 % (5 %) within the central (outer) modules of the ECal. These smearing factors are a conservative estimate of the accuracy of the initial calibrations used by the trigger to make data collection decisions. These ten different calibrations were then used to re-reconstruct the simulated events ten different times yielding Figure 5.7 and Figure 5.8. Figure 5.7a reassures us that our signal region (below 1.1 GeV) is not polluted with events that would have failed the trigger selection even after these rather large calibration changes relative to the trigger calibrations (i.e. we would not have overestimated our signal efficiency). In addition, we can estimate the resulting systematic variation on the background event yield within the three final analysis bins to be 5 % using Figure 5.7b. The final selection on the ECal Hit RMS is only affected negligibly as shown in Figure 5.8 where the distribution only changes by $\lesssim 5 \%$ within the range of potential cuts.

Estimating the systematic uncertainty due to the HCal variable is not as straight
yield past its selections compared to the 8 GeV beam sample.

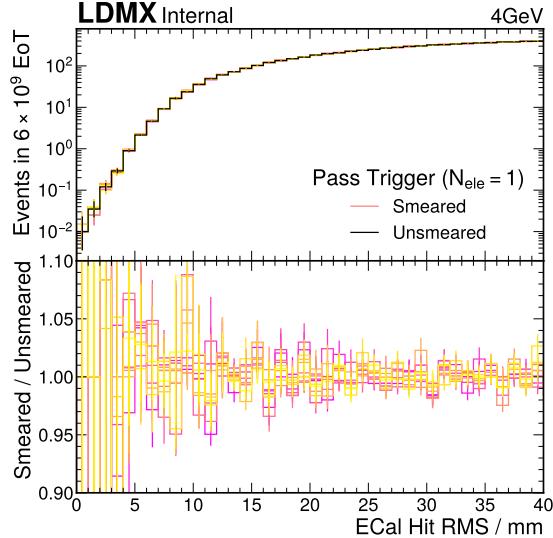
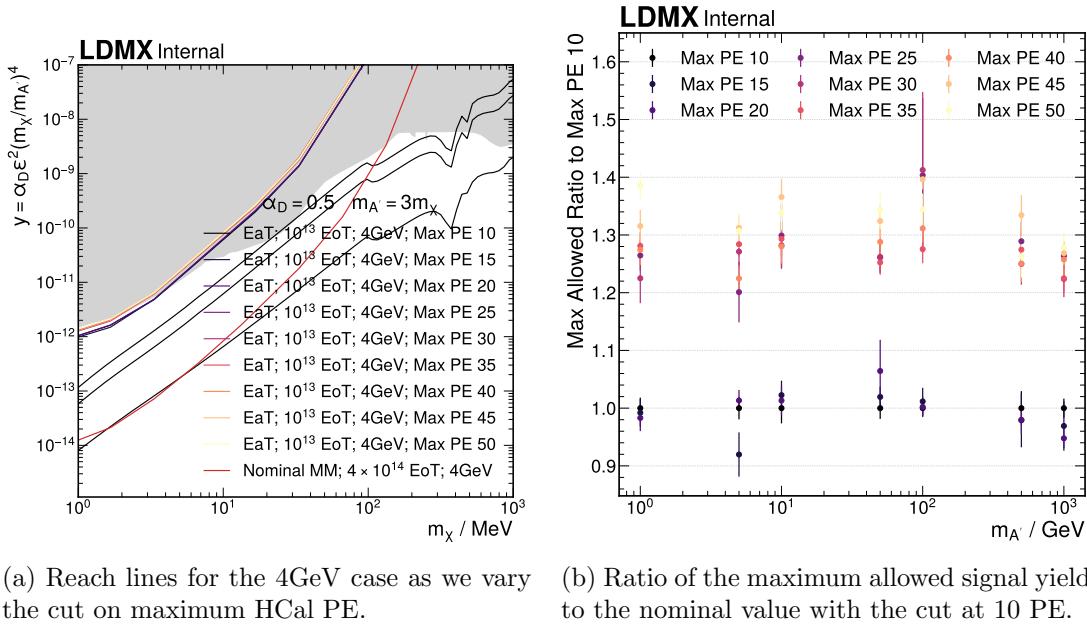


Figure 5.8: Effect of smeared ECal calibrations on the ECal hit RMS for events that pass the original trigger. Outside of statistical fluctuations in the lowest-populated bins, the smearing has a less than 5% effect throughout the range of potential cuts.

forward since we are unsure on how best to directly alter the conditions with which the HCal is reconstructed. We can still estimate the systematic uncertainty due to this cut variable by testing a wide range of cut variables, deducing which ones are “acceptable” and then extracting how the background yield varies between these different cut choices. (What we want to be emulating is the underlying distribution shifting relative to our analysis cut, but we are mimicing this by shifting the cut relative to the distribution.)

Figure 5.9a shows how the final reach (see Section 5.4) changes depending on this cut. While not much movement is observed, we can see separation occurring when the threshold goes above ~ 30 (more clearly seen in Figure 5.9b). With this in mind, we estimate our “tolerance” to a systematic error on the maximum HCal PE to be $1 - B_{<10}/B_{<20}$ where $B_{<M}$ is the background yield with the maximum HCal PE threshold M . This yields an analysis-bin-dependent systematic uncertainty of 0.0, 0.6, and 0.7 for the low, medium, and high E_{frac} analysis bins respectively.

While deduced from inspecting the effect on the final reach, the ultimate systematic variation of 10 PE could also be reasonable determined by thinking about how the HCal would be calibrated. The HCal would use observations of through-going muons to set



(a) Reach lines for the 4GeV case as we vary the cut on maximum HCal PE.

(b) Ratio of the maximum allowed signal yield to the nominal value with the cut at 10 PE.

Figure 5.9: Reach (left) and maximum allowed yield (right) comparisons for the 4GeV beam case as we vary the cut on maximum HCal PE. We observe minimal movement in the final reach as a function of this cut; nevertheless, the biggest change occurs around a threshold of 30.

the overall scale of PE as observed by the HCal bars and electronics. In the simulation a single through-going muon typically deposits ≈ 80 PE, so a variation of $\sim 10\%$ would be equivalent to the 10 PE deduced earlier.

5.4 Reach

We can quantitatively define an excess above a background prediction using statistical techniques like a one-sided Poisson-distributed p-value test; however, in this context it is often more helpful to characterize the signal models that would be excluded if the data was collected and it aligned with the background prediction. These “exclusion estimates” (colloquially referred to as “reach”) rely on statistical tools that are common within HEP and implemented within COMBINE[34]. Figure 5.9 has already shown an example reach that varies depending on the value of the cut we choose to use during the selection; however, the final reach including the statistical and systematic uncertainties discussed above is shown in Figure 5.10. The LDMX EaT analysis channel is able to search through previously-unexplored phase space utilizing simple, physically-motivated cuts and $\sim 2.5\%$ of the planned EoT for the first phase of LDMX.

Figure 5.10 colors the area of phase space that has already been excluded by other experiments in gray. This area is constructed from exclusions by multiple experiments; however, NA64 [5, 6] is of particular importance due to its similarity to this specific analysis within LDMX. NA64 accepts an electron beam (and perhaps a muon beam in the future) on a fixed target and uses an ECal and HCal in order to reconstruct the observable energy and check for any missing energy. NA64 differs from LDMX, mainly in the configuration of the beam the experiments receive and the presence of a tracking detector for estimating momenta of charged particles; however, the similarities – especially within this LDMX analysis channel – motivates collaboration in the future. Previously, I have connected with NA64 researchers to learn from their dark matter simulation package DMG4 [35, 36], and I hope to continue this collaboration and knowledge-sharing in the future especially as LDMX starts to collect data and we can share information about particularly intricate background processes.

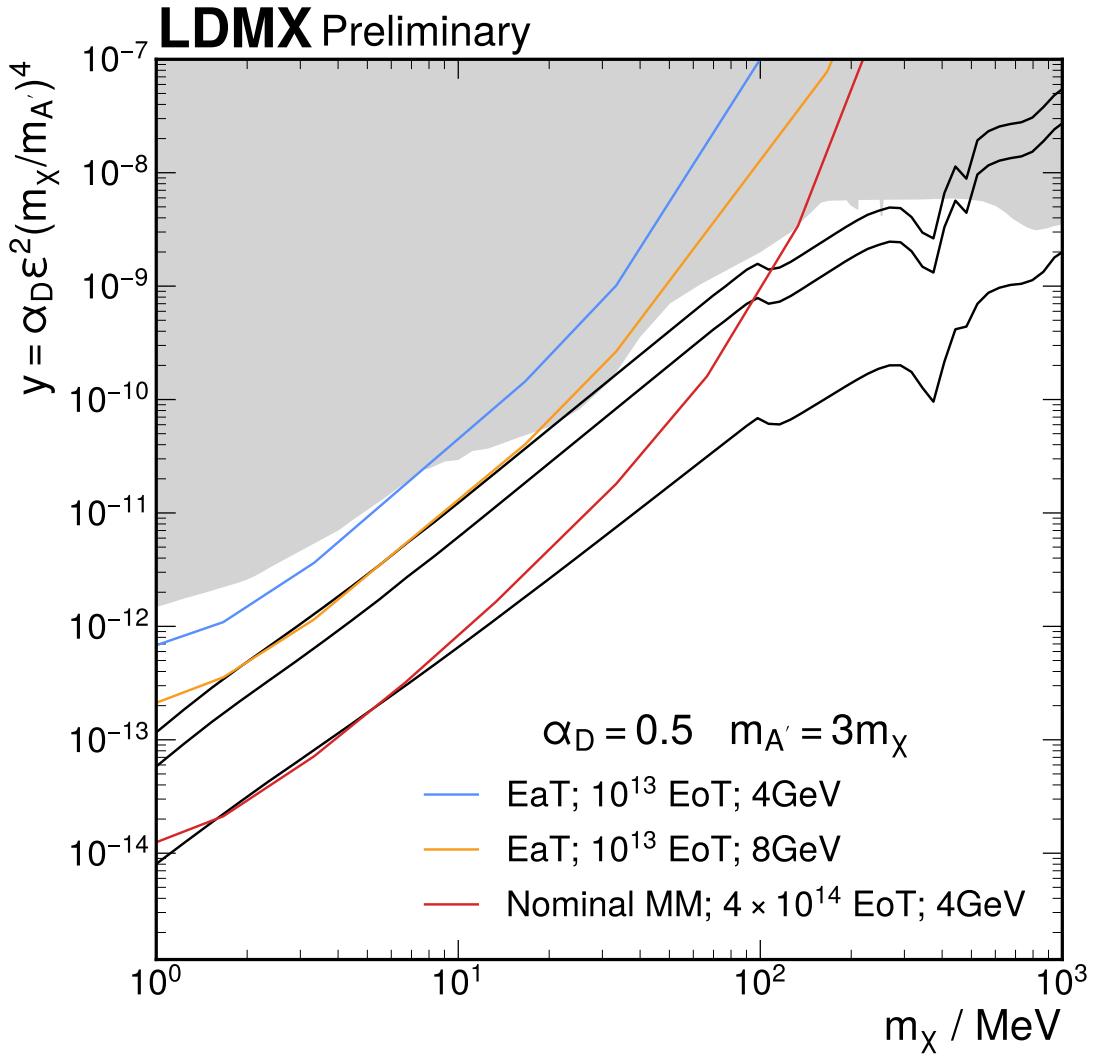


Figure 5.10: The sensitivity of the EaT analysis channel compared to other experiments led by NA64[5, 6], *BABAR* [7], and COHERENT[8] (gray), Scalar, Majorana, and Pseudo-Dirac theory expectations (black, top-to-bottom), and LDMX projections (colors). The blue (orange) line corresponds to the 4 GeV (8 GeV) beam studied in this work. The red line is the nominal LDMX analysis sensitivity for the Missing Momentum (MM) analysis channel.

Part III

HPS

Chapter 6

The Heavy Photon Search Experiment

The Heavy Photon Search experiment (HPS) is a fixed-target experiment currently installed at cro:jlabThomas Jefferson National Accelerator Laboratory (JLab), which is focused on the search for short-length (few centimeter) decays of new particles. In order to search for these visible signatures of new particles, HPS must be able to precisely reconstruct the produced SM particle kinematics and then use this information to estimate their origin point (their “vertex”).

The baseline focus for HPS is to search for rare interactions which produce dark photons that return back to the normal sector after a brief lifetime. This search is the origin for the name of the experiment and the primary motivation for the design of the two key detector components described in more detail below. We will instead consider a model where some of the energy remains within the dark sector.

6.1 Strongly Interacting Massive Particles

As mentioned in Section 2.5, our search for DM must accomodate the possibility that any DM produced within our experiments subsequently decays back into SM particles. In order to model these types of DM, we must account for both how the DM is produced and how it decays back to SM particles. These theories propose new specific vertex rules for constructing the Feynman diagrams and for the vertex factors within the calculations.

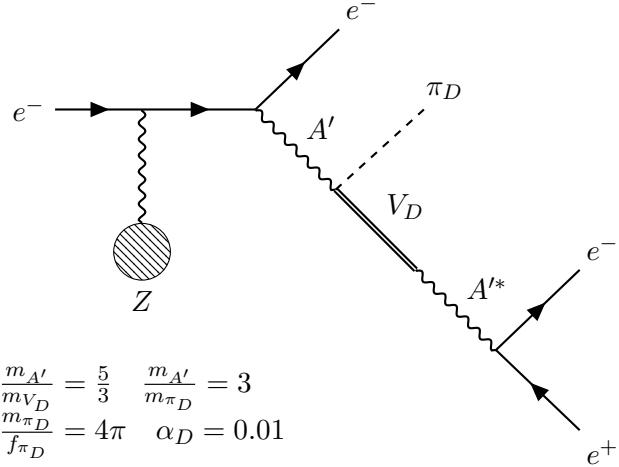


Figure 6.1: Diagram of dark brems production of a dark photon followed by its decay back into an electron-positron pair through SIMP composite particle V_D with the emission of a lighter SIMP composite particle π_D . The two-particle vertex $V_D - A'$ is allowed since they share enough quantum properties to mix.

To be even more specific, we could hypothesize that the dark sector consists of particles that interact with each other strongly in an analogous sense as the standard quarks. These SIMPs [37, 38] would then form composite particles similar to how standard quarks form composite particles like protons and neutrons. Since we expect our experiments to not reach the energies required to separate the constituents of these composite particles, the composite particles themselves would be the dark matter candidates and the ones we represent within our diagrams. A dark sector consisting of SIMPs would have a plethora of composite particles which we could produce; however of particular importance is that there could be a composite particle that has similar enough properties to the dark photon that it could “mix” with it and decay in a way that emits fewer dark particles: Figure 6.1. This decay, since it has less energy “lost” to the dark particles, has a greater potential of being observed within HPS and is discussed here.

Importantly, Figure 6.1 has two vertices where DM and SM particles interact – both of these two vertices carry with them the weak mixing factor ϵ . The first vertex where the A' is produced suppresses the rate that this process occurs by a factor of ϵ^2 similar to the previous case with LDMX, but the second vertex does not affect the rate. Since, in

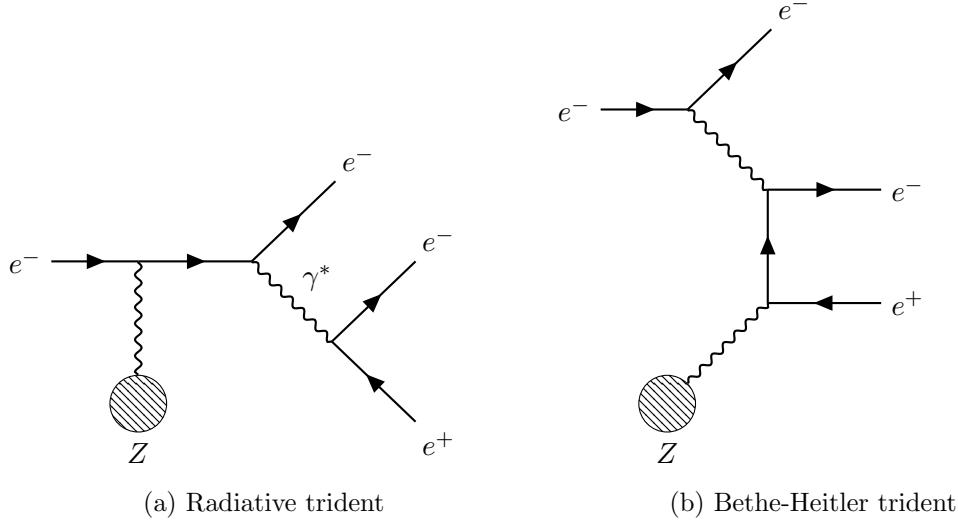


Figure 6.2: Feynman diagrams for the SM process of trident production.

this model and other visible-signature models, the dark matter being produced cannot viably decay to other dark matter (because $m_{A'} < 2m_\chi$ for example), this second vertex forces a longer decay length. HPS has been built to observe a beam whose rate is significantly high in order to help overcome not only the low production rate but also the acceptance and efficiency of observing the electron and positron decay products – high enough to prevent sensitive detector elements from being placed directly within its path.

The observable outgoing particles of Figure 6.1 can be mimicked by existing SM processes. The most prominent background process is trident production. Figure 6.2 shows processes which produce the electron-positron pair through the radiation of the virtual photon (Figure 6.2a) or through direct electromagnetic interaction with the nucleus and a virtual lepton (Figure 6.2b). The difference between a DM signal process and this trident production process is that the trident process is *prompt* – i.e. the produced electron and positron originate from within the target. The position of this vertex is a separation from the signal process Figure 6.1 where the vertex of the produced electron and positron would be observably displaced downstream of the target due to the non-zero lifetime of the dark sector particle V_D .

HPS is designed to focus on visible signatures of DM like SIMPs with a dual strategy

of searching for mass resonances and displaced vertices. Both of these strategies require high data volume and precise reconstruction of a produced electron-positron pair originating from the decay of some DM particle. The precision of this reconstruction is one of the main limiting factors in distinguishing whether a specific pair originated from DM or from some other SM background process. These design drivers have led to the apparatus discussed below.

6.2 Detector Apparatus

HPS is currently installed behind the CLAS-12 experiment in the Hall B alcove at JLab and utilizes the Continuous Electron Beam Accelerator Facility (CEBAF) to provide its beam [39, 9]. Hall B primarily hosts the CLAS-12 experiment in the bulk of its hall; nevertheless, HPS is situated within an alcove behind the CLAS-12 experiment such that it can collect data whenever CLAS-12 is not operational.

CEBAF [40, 41, 42] is able to provide a near-continuous electron beam ranging in energy up to 12 GeV in steps of ≈ 1.15 GeV. CEBAF is built in a oval “race track” where the straight portions consist of linear accelerators and the curve portions steer the beam with dipole magnets either back into the straight portions or into experimental halls. Since CEBAF delivers electrons at such a high rate, HPS must be able to handle these high radiation loads. Designed to be a small experiment to fit within the Hall B alcove, HPS is built in two halves that straddle the primary path of the beam where the highest amount of radiation would be. Both of the subsystems within HPS are split into these two halves as shown in Figure 6.3 and Figure 6.4.

The two subsystems included within HPS are the SVT focused on reconstruction of the momentum and charge of charged particles and the ECal focused on energy reconstruction and online triggering of the data.

6.2.1 Silicon Vertex Tracker

The SVT is made of eighteen modules each of which is a pair of silicon-strip detectors offset from each other by a small angle to enable three-dimensional position reconstruction. The modules are arranged in the two halves separated by the center beam line. In each half, the modules are put into layers with the first three layers consisting of

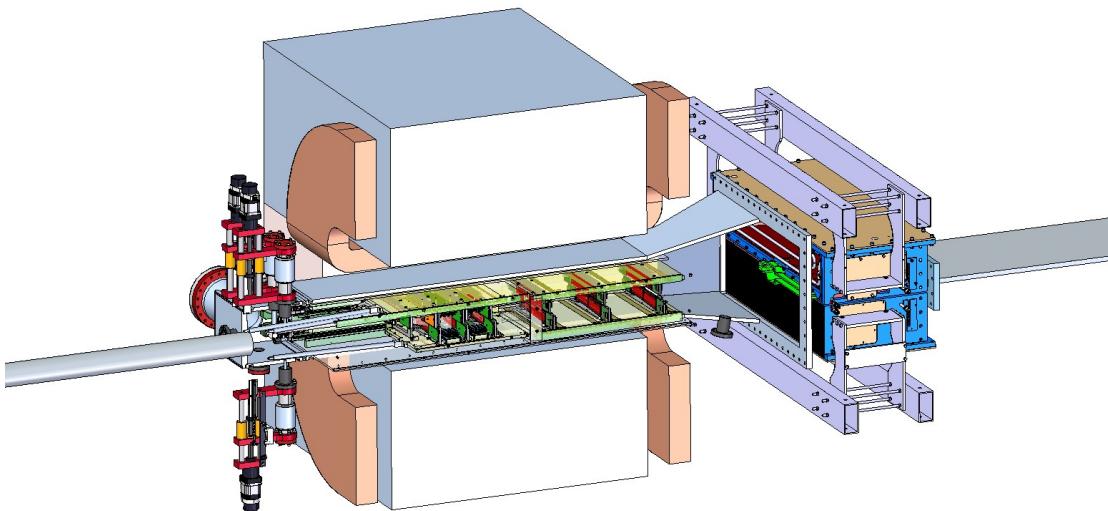


Figure 6.3: Full rendering of HPS. Beam would enter the detector from the left and exit to the right if it does not interact.

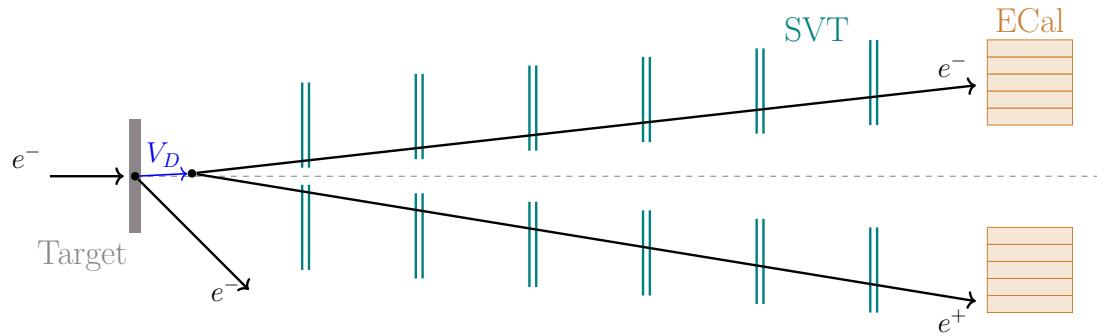


Figure 6.4: Simplified diagram of HPS showing an example displaced decay within the named subsystems. The blue arrow is some DM candidate particle that has a macroscopic lifetime causing the decay vertex to be observably displaced from the production within the target.

one module while the last three layers consisting of two modules to increase angular acceptance relative to the target. Figure 6.5 shows a rendering of the SVT with the sensitive detector elements shown as red rectangles.¹

The two halves of the SVT are positioned to form a 15 mrad angular gap relative to the target in order to avoid the largest amount of radiation from the beam. The first layer (closest to the target, left side of Figure 6.5) is placed 10 cm from the target and 0.5 mm from the beam line in order to maximize acceptance. The entire SVT is enclosed in a vacuum box to limit secondary production of particles and is liquid cooled to help prevent radiation damage.

The SVT modules are readout by APV25 chips which report six samples separated by 24 ns. These samples are only dumped to disk upon receiving a trigger signal from the calorimeter system (Section 6.2.2).

All together, the SVT has been observed to have 10 % momentum resolution, 2.4 ns time resolution, and 6 μm position resolution. These properties enable the necessary momentum and vertex reconstruction for DM searches via visible decays.

6.2.2 Electromagnetic Calorimeter

The ECal consists of 442 lead-tungstate crystals and is a fully-sensitive calorimeter. The crystals are arranged in symmetric grids separated by the beamline similar to the SVT. Figure 6.6 shows this grid in addition to example clusters and ECal-related variables. In addition to a horizontal gap separating the halves, a few more crystals in the highest radiation area are removed to form a “hole” to allow the majority of non-interacting (or minimally interacting) beam electrons to pass through the detector volume. The calorimeter is sized to fit the same angular acceptance as the SVT located 139 cm away from the target.

As mentioned above, one of the primary purposes of the ECal is to be the triggering mechanism for data collection. Many different trigger algorithms with different purposes have been used within HPS and the specific trigger algorithm designed for collection of data relevant to a DM search is studied and motivated in detail in [9]. Figure 6.6 diagrams an event that would pass the pair-wise trigger used for the collection of the

¹This diagram and the description within this thesis is focused on the HPS SVT from its physics run in 2016. Various HPS components have been upgraded for subsequent runs.

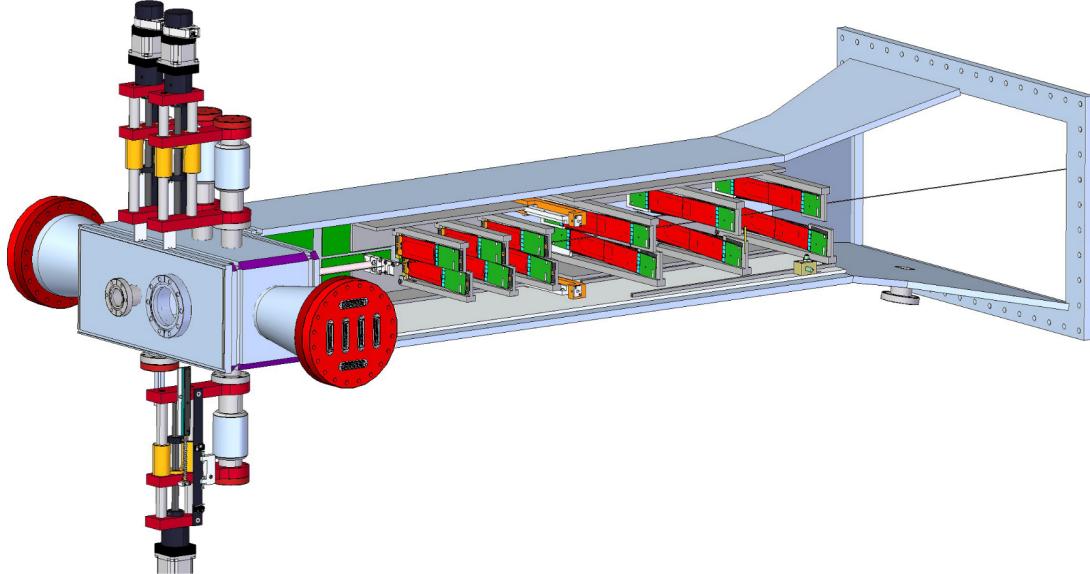


Figure 6.5: Figure 10 of [9]. Rendering of the SVT with the vacuum enclosure shown in gray, active silicon components in red, and readout electronics in green. Similar to Figure 6.3, the beam would traverse this rendering from left to right.

data within this study overlayed on a beam-view of the ECal. Section 8.1.1 provides more detail on the trigger requirements.

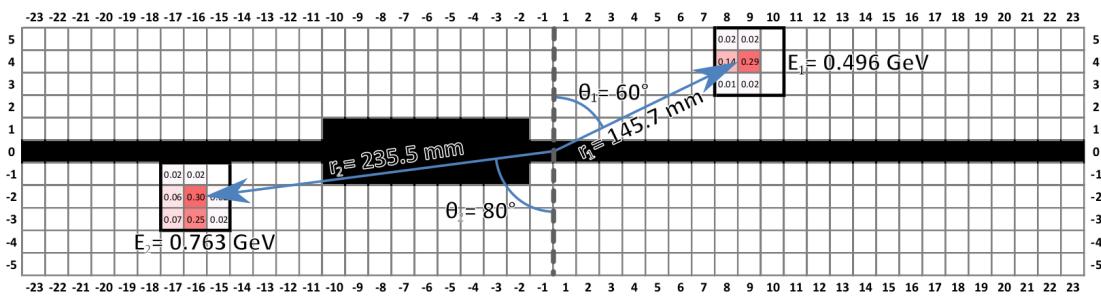


Figure 6.6: Figure 27 of [9]. Depiction of an event passing the pair trigger in the 2016 data collection run. This depiction also displays the variables used in order to evaluate events and make the trigger decision.

Chapter 7

Alignment

Particle detectors used to “track” particles (often called “trackers”) largely operate by precisely measuring the position of the particle at several different points along its trajectory. From these position measurements, we are able to extract other, more physically helpful, measurements such as momentum and charge (as long as the particles move through a magnetic field so their resulting trajectory is curved in a well known shape). This process typically proceeds through two steps. In the first step, called “pattern matching”, a set of hits is collected which could plausibly be from the path of a charged particle. In the second, “track fitting” step, the path and the momentum are determined from the hits taking into account the magnetic field and hit resolution. This serves as motivation: without precise knowledge of position, we are unable to precisely estimate momenta of particles – significantly impacting the capabilities of the experiment. Inaccurate hit positions result in hits being left off tracks, included in tracks where they should not be, or negatively affecting the results of the track fitting.

We need all of these position measurements to be relative to the same reference point (i.e. in the same coordinate system); thus, the process of measuring particle positions is practically done by combining two position measurements.

1. Position relative to individual pieces of the tracker (here called “sensors”)
2. “Global” position of these sensors

More specifically, the software we use to deduce physical measurements from the data output by a tracker consists of a detector model (which stores the positions of all sensors

in the same coordinate system) and a series of algorithms we can apply to this detector model along with data from the sensors.

While there are different strategies to measure position within an individual sensor of the tracker (e.g. pixel vs strip tracking detectors), this measurement is not the limiting factor in the context of this chapter and is often the measurement whose precision is easier to characterize since the individual sensors can be designed, built, and tested individually for a pre-determined precision relative to a well-defined point, such as the center of the sensor.

The second position measurement – the global position and orientation of the sensors themselves – is where a lot of complexity arises. We construct the detector and thus know where we put the sensors, but placing them into position with the precision we desire is extremely difficult (not to mention their position could shift slightly once the magnet is turned on or due to thermal effects as the detector temperature changes).

We can gain some understanding of the position precision required with the help of Figure 7.1. There, we see that the momentum error roughly scales by a factor of five relative to the position error and, as the position error goes up, the momentum’s distribution starts to distort away from a normal shape. In addition to the scaling of the *relative* error, the overall scale of these measurements are an important factor as well.

To achieve a rough estimate of the scale of these position measurements, consider a 2 GeV electron curving within a 1 T magnetic field. This electron would then have a 6.6 m radius of curvature due to this magnetic field, meaning the position measurements within sensors separated by 5 cm would change relative to one another by *less than* 1 mm. This absolute measurement of $\sim 100 \mu\text{m}$, combined with the necessary 1 % relative error, means we need to know the position measurements to the level of 1 μm .

While we can determine the position of these sensors within millimeter precision during construction, we need an “in-situ” method which allows us to determine positions and angles during operation. This motivates another measurement “tool” that allows us to determine (or more accurately, constrain) the position of our sensors.

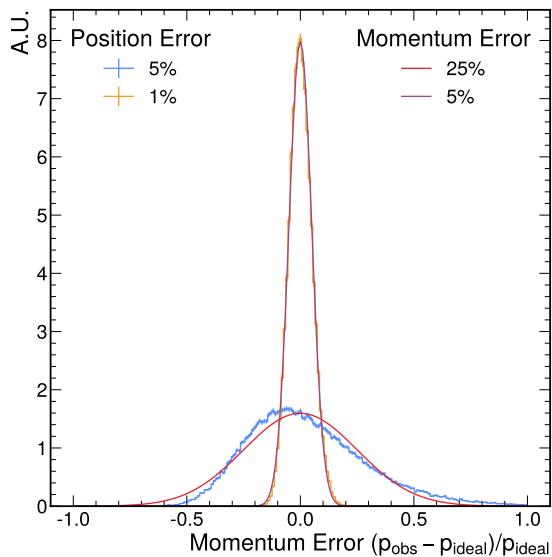


Figure 7.1: Momentum error derived from injecting a certain amount of position error (blue and orange) compared to normal distributions (red and purple) with known momentum error. These distributions are normalized to have unit integrals. The trials were done in a simplified experimental model with quantities similar to HPS: a 2 GeV electron curving within a 1 T magnetic field, the sensors measuring the position were separated by 5 cm but their separation were also allowed to deviate from the true value by the position error given.

7.1 Alignment Procedure

The tool that we use is the same particles that we will use the tracker to study for our experimental goals. The one change is instead of using the tracker to measure the properties of these particles, we will select particles according to our knowledge of their properties so that we can use them to constrain the position of the tracker sensors.

This constraint is born out of concrete connection between the physical kinematics of the particles and the positions they record in the tracker sensors. We can imbue our knowledge of the particle trajectories in the magnetic field into the mathematical form of the equations that are fit to the position measurements. The mathematical fitting process estimates the best values of the starting position and 3D momentum from a set of points given the assumed individual precision of these points. This produces a prediction of where the particle “should have” passed through each sensor. The difference (or “error”) divided by the expected precision of the hits can be summed in quadrature to provide the χ^2 – a measure of how close the fit and the measurements are to one another. If the sensors in the detector model were in slightly different positions than where they are in the real detector, the data would present fits that are not as good (quantitatively, have a higher χ^2). With this in mind, we can slightly “move” the sensors in the detector model in order to improve the fits in the data (lowering the χ^2). In order to account for the fact that individual tracks may still vary due to our imperfect knowledge of their trajectory shapes (e.g. the trajectory shape assumes only soft interactions with the detector material but some tracks could have harder interactions distorting their trajectory), we do this procedure over many tracks at once and focus on improving all of their fits (minimizing the sum of their χ^2).

This technique relies on the movements of the sensors to be small. The process can easily fall into a so-called “Weak Mode” where our quantitative measure of how good the fits are is achieving a local minimum value, but the trajectories do not actually align with data. This motivates a multi-stage approach to aligning a tracking detector.

1. **Survey:** Directly measure positions of the sensors relative to other detector components as precisely as possible.
2. **In-Situ Alignment:** While only allowing for certain types of movements, minimize the total χ^2 of well-known tracks.

Since physical measurements obtaining the precision we require are so difficult, the survey is often split into multiple, cascading measurements as well. For example, the positions of the sensors are precisely determined relative to their mounting brackets which are then measured relative to the holding frame which is then measured relative to the global coordinates within which the experiment operates.

The second stage where we perform small movements of the sensors in order to optimize the fit of our data to our software detector model is also broken into many smaller stages. The most common way is to only allow a certain set of physically motivated movements to occur during certain alignment iterations. This can allow us to “walk” towards a detector model that is more aligned with the physical detector while maintaining physical understanding of the changes being made.

HPS uses a cro:kfKalman Filter (KF) algorithm for track finding, cro:gblGeneral Broken Lines (GBL) [43] V03-01-00 for track fitting accounting for multiple-scattering, and MILLEPEDEII [44] V04-13-01 for determining alignment parameters by minimization of the total χ^2 . However, these tools are not enough to effectively align a tracking detector automatically. The HPS SVT is relatively small but still contains several hundred alignment parameters. As mentioned above, the in-situ alignment procedure generally only allows for certain types of movements (i.e. only a certain set of parameters are allowed to change) during single alignment iterations, but then the natural question is what types of movements should be allowed. While some movements can be well-motivated on physical grounds (for example, moving the sensor along its own sensitive direction), other movements may still be helpful and so many movement trials are required. This is where my specific work occurred – working on testing different sets of movements to see if they could help improve the detector alignment. Along the way, I also helped develop a suite of helpful side-tools primarily focused on visualizing the movements of the model sensors that result from alignment iterations. This visualization is primarily helpful to see if the parameters resulting from the quantitative minimization algorithm produce realistic movements of the detector model.

7.2 Detector Visualization

Figure 7.2 demonstrates a quantitative visualization where the positions and orientations of all the sensors within a specific coordinate system are shown. This visualization is helpful for verifying that the detector model being used within the software is placing the sensors in expected positions and orientations. For example, we observe the y and z positions steadily increasing as the layer number of the sensor increases – representing the opening angle as designed within HPS. Additionally, we apply the designed “flipping” of sensors (the factor of \pm when calculating the angles with the axes) when plotting to make sure that the sensors in the model are within $\pi/2$ of their true orientation.

The differences of these coordinates between multiple versions of the detector model provides deeper insight. Figure 7.3 demonstrates this strategy by subtracting the coordinates of the original detector model (“Original”) from aligned detector model (“Aligned”). Specifically, we can observe that the alignment iterations only allowed certain movements (for example, no translation along or rotation away from the global z axis was allowed). The fact that these movements are very small shows that the physical survey which was done to define the Original detector was close in position and orientation.

7.3 Results

The visualization of the movement of the sensor positions within our detector model is all well-and-good, but we should actually be investigating how these movements affect the reconstructed physical variables of tracks when using the updated detector model. In general, the physical variables that should be investigated can depend on the priorities of the experiment; for this analysis within HPS we wish to inspect two key features.

- **Momentum Magnitude** – the total magnitude of the momentum of the reconstructed track is a key component in all analyses done with HPS data and thus is a priority to be well understood. Often, it is also common to inspect the momentum magnitude for different types of tracks so that it can be understood to in even finer detail. (i.e. p vs $\tan \lambda$, p vs ϕ , p separated by charge, etc.)

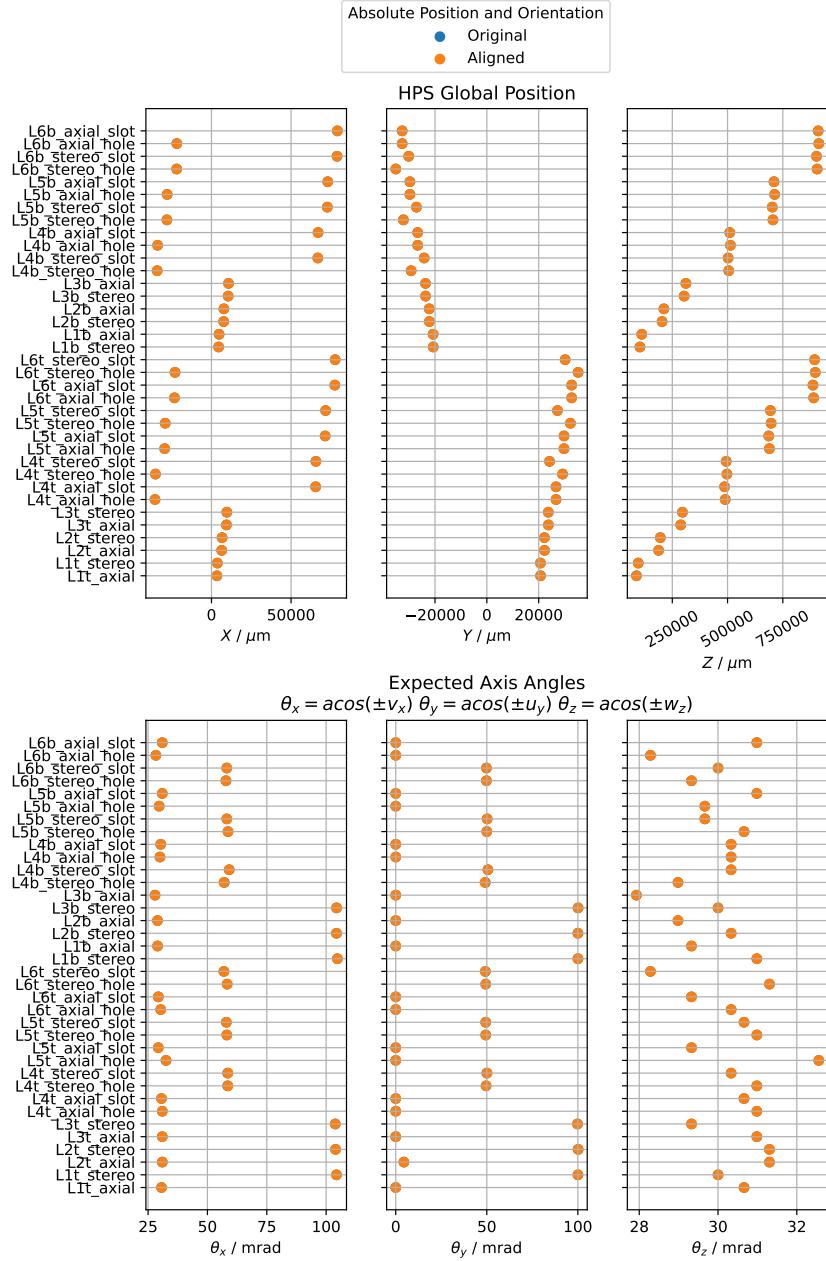


Figure 7.2: Plotting all six global coordinates completely specifying position and orientation of each of the sensors in the HPS tracking detector. Two different versions of the in-software detector model are plotted; however, on this scale, they overlap one another.

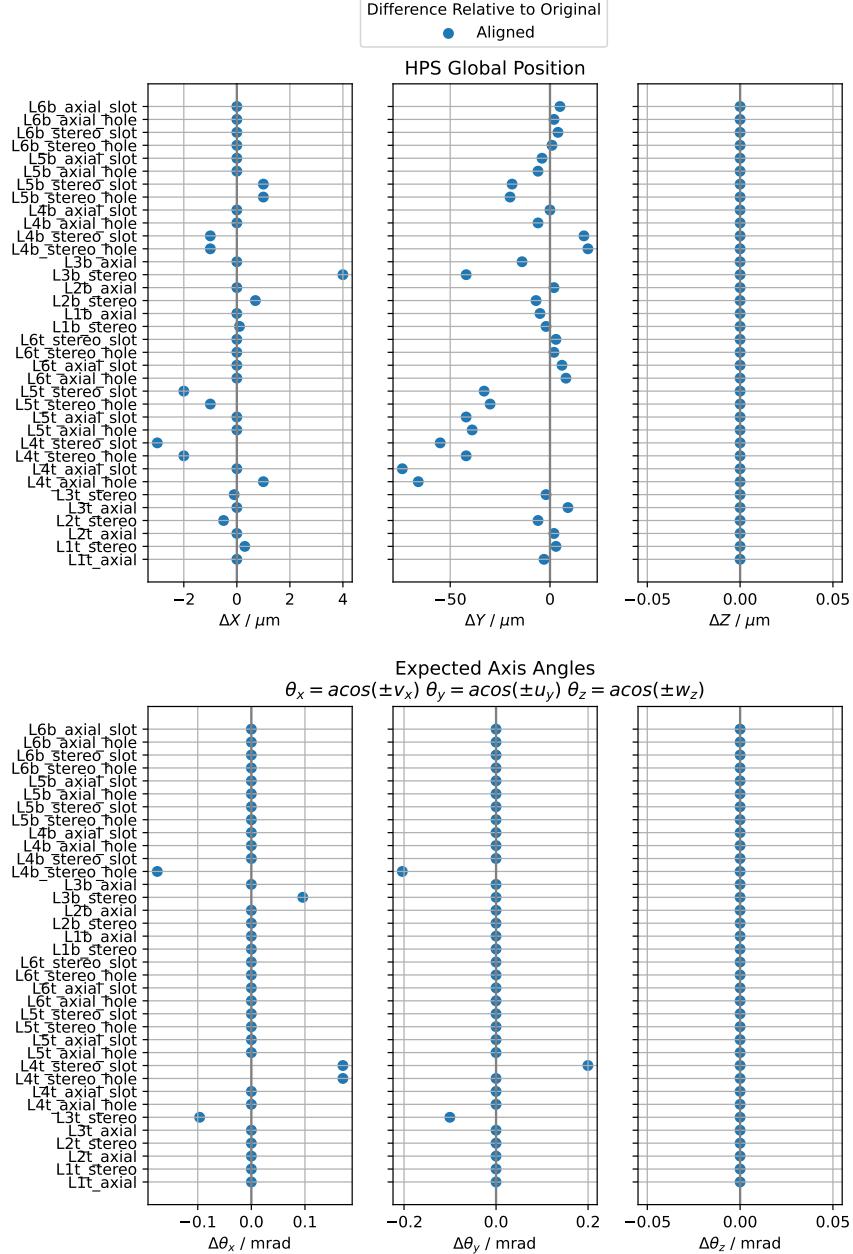


Figure 7.3: Plotting all six global coordinates completely specifying position and orientation of each of the sensors in the HPS tracking detector relative to the “Original” detector from Figure 7.2. We can now observe the quantitative differences between the different versions.

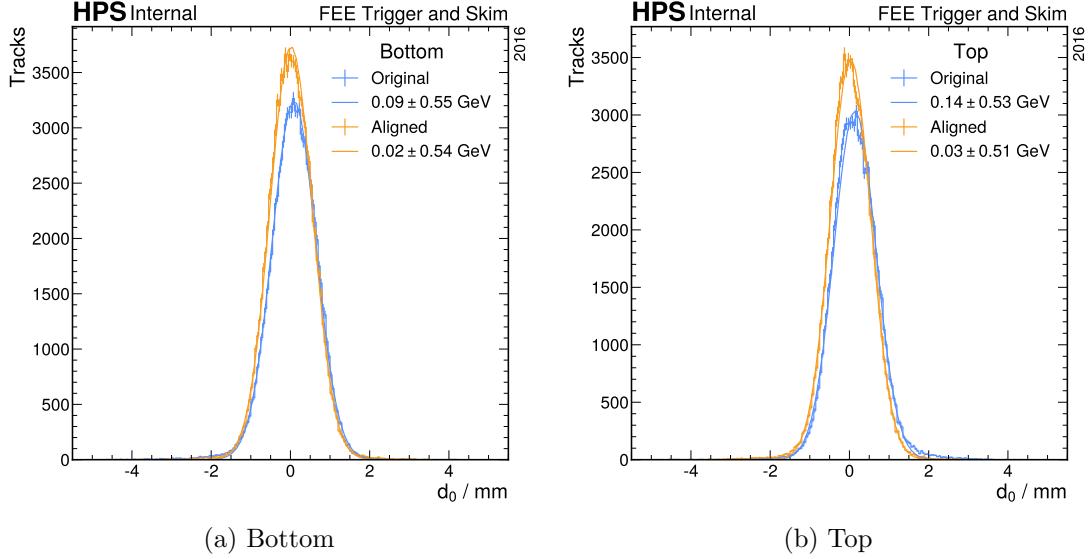


Figure 7.4: The transverse impact parameter of the tracks at the target $d_0 = \sqrt{x_0^2 + y_0^2}$ which should be centered on zero to align with the known beamspot at the target.

- **Beamspot Position** – the extrapolation of the track back to the target is another helpful variable for analyses and can be checked directly since HPS has other means for estimating the beam spot on the target.

In these results, we look at a specific subset of the data collected within HPS during 2016. This subset of the data uses a special trigger and analysis cuts that focus on selecting events where a cro:feeFull Energy Electron (FEE) is captured within the HPS detector volume. An FEE is an electron that passed through the target with minimal interaction and so its total energy is close to the known beam energy. This dataset provides a fertile ground for studying the alignment since we have knowledge about averages of two aspects of the tracks. The momentum magnitude of these tracks should be close to the known beam energy 2.3 GeV and the transverse location of these tracks at the target should be near the origin. Figure 7.4 and Figure 7.5 shows the comparison between the Original and Aligned detectors whose models were shown above. We can see that the Aligned detector achieves a tighter distribution in both of these variables for both the top and bottom half of the SVT meaning that this detector model is able to more precisely represent the true placement of the sensors when the data was collected.

Unfortunately, due to the number of possible parameters that could be tuned, the

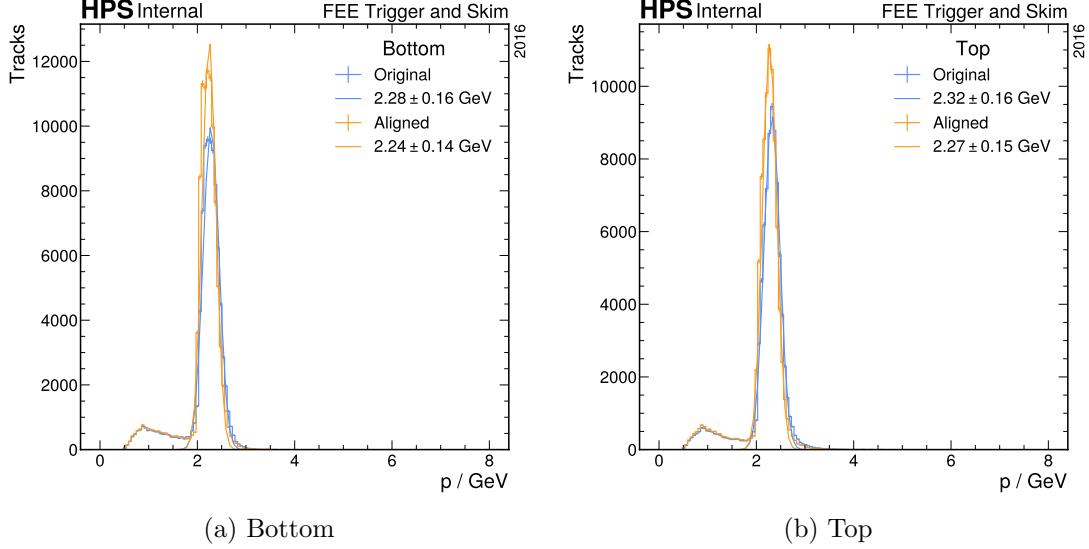


Figure 7.5: The magnitude of the track momenta which should be centered on the known beam value of 2.3 GeV.

possibility of improving the alignment of the tracking detector is something that can always be done. Thus, while there is often a current “best” detector model, future analyzers may need to re-process data with a later (and hopefully improved) model that would be expected to be more accurate to the real detector that data was taken with. The results shown here are just this – a snapshot of the progress in aligning the HPS SVT.

Chapter 8

Data Set

Proper search for DM signals within the HPS detector includes both collection of data with the HPS detector and simulation of how this apparatus responds to the physics of specific processes. This chapter is focused on detailing the origin of these samples. The collected data can be characterized by the known inputs from CEBAF and the run time. The simulation, as one might expect, is a complicated multi-stage procedure in order to appropriately reflect the intricacies of the real data.

8.1 Collected Data

In HEP, the volume of data is often measured in a particular unit that is convenient because it is the inverse unit for a variable that we also often measure (and compare to the calculation from our models): cross section. Simply, rare processes have lower cross sections and therefore we often are looking for processes that have cross sections in units of picobarn (pb). Expressing the data volume in units of pb^{-1} is referred to as the luminosity and allows us to quickly estimate the number of expected events from a given process by multiplying this luminosity by a process's cross section. For example, the radiative trident process (Figure 6.2a) has an estimated cross section of $66.36 \times 10^6 \text{ pb}$, so we expect $\approx 7.1 \times 10^8$ instances of the radiative trident process to occur within the detector given the full luminosity of the data.

The data used in this study was collected over a series of 82 data collection runs within 2016 yielding a total of approximately one week of continuous beam. The full

luminosity of this data is estimated to be 10.7 pb^{-1} . As alluded to in Section 6.2.2, the collected data was triggered in order to focus the sample on specific data of interest.

8.1.1 Pair 1 Trigger

The trigger used for this and other physics analyses of the 2016 dataset is the so-called “Pair 1 Trigger” designed to select events where evidence for both an electron and positron have been found. This trigger algorithm combines requirements on the energies of the clusters formed at the trigger level along with the cluster locations.

After the hits in the ECal are clustered, the clusters are selected such that they are of high enough quality - implemented with energy E and hit multiplicity N cuts.

$$E_{\text{low}} \leq E \leq E_{\text{high}} \quad N \geq N_{\text{min}}$$

With a beam energy of 2.3 GeV , we set $E_{\text{high}} = 1.4 \text{ GeV}$ to help remove near-full-energy beam electrons from being included within this trigger. $E_{\text{low}} = 0.15 \text{ GeV}$ and $N_{\text{min}} = 2$ are set to these values to suppress the effect of noise.

All possible pairs of clusters passing these criteria where one of the pair is in the top half of the ECal and the other is in the bottom are then tested with further pair-wise criteria. Let the top (bottom) cluster have energy E_{top} (E_{bot}) and position $(x_{\text{top}}, y_{\text{top}})$ ($(x_{\text{bot}}, y_{\text{bot}})$) leading to radius $r_{\text{top}} = \sqrt{x_{\text{top}}^2 + y_{\text{top}}^2}$ ($r_{\text{bot}} = \sqrt{x_{\text{bot}}^2 + y_{\text{bot}}^2}$).

The total energy of both clusters $E_{\text{top}} + E_{\text{bot}}$ is bounded below by $E_{\text{sum,low}}$ to remove noise and above by $E_{\text{sum,high}}$ to reduce the common single-bremsstrahlung background process. The clusters are required to have an absolute difference $|E_{\text{top}} - E_{\text{bot}}| < E_{\text{diff}}$ to select cluster pairs originating from the same particle. The energy and position of the lower-energy cluster is required to satisfy $E + rF \geq E_{\text{slope}}$ which accounts for the expected bending of the lower-energy particle’s location in the magnetic field (which is failed by clusters originating from photons which do not bend in the magnetic field). F is determined from the geometry of the detector and strength of magnetic field to be 5.5 MeV mm^{-1} . Finally, the clusters are required to be coplanar within an angle ϕ to again select for clusters originating from the same particle.

$$|\tan^{-1}(x_{\text{top}}/y_{\text{top}}) - \tan^{-1}(x_{\text{bot}}/y_{\text{bot}})| < \phi$$

Description	Variable	Value
Single-Cluster Energy Minimum	E_{low}	0.15 GeV
Single-Cluster Energy Maximum	E_{high}	1.40 GeV
Single-Cluster Hit Count Minimum	N_{\min}	2
Cluster Pair Energy Sum Minimum	$E_{\text{sum,low}}$	0.60 GeV
Cluster Pair Energy Sum Maximum	$E_{\text{sum,high}}$	2.00 GeV
Cluster Pair Energy Difference Maximum	E_{diff}	1.14 GeV
Low-Energy Cluster Slope Minimum	E_{slope}	0.70 GeV
Cluster Pair Azimuthal Difference Maximum	ϕ	35°

Table 8.1: Values for cuts used within the Pair 1 Trigger during data collection of the HPS 2016 dataset.

These requirements and the values selected for their thresholds are summarized in Table 8.1.

8.2 Simulation

As mentioned the simulation goes through many steps in order to account for the different physics processes that are of interest in a realistic fashion. In general, these steps are

1. Generation – using a tool like MADGRAPH/MADEVENT[45, 46] to generate specific events from Feynman diagrams
2. Displacement – if the sample expects to have the decay products be displaced (for example in the SIMP signal process), displace these decay products with a *uniform* distribution of decay lengths to allow for re-weighting.
3. Simulation – simulate the detector response with GEANT4[32]
4. Emulation – emulate the readout electronics and triggering mechanism of collected data

After this emulation stage, we can treat the simulation the same as the collected data, applying the reconstruction and further analysis manipulations and selections.

Both simulated samples of standard processes as well as hypothetical signal processes proceed this way just with a different original generation step initializing the event. For

standard processes, we also use a physical displacement distribution as understood from prior measurements rather than a uniform one that is more helpful for reweighting later in analysis.

8.3 Reconstruction

The two subsystems are first reconstructed separately. The ECal produces clusters of nearby hits whose energy have been corrected by more refined calibration tables. The SVT produces tracks using KF to find tracks and GBL to fit these tracks to the detector model including the possibility of in-detector small scatters. Tracks are then paired with clusters whose position and energy are close to what the track's position and energy would be at the ECal - such pairs are referred to as reconstructed particles. Pairs of reconstructed particles are then used to deduce a vertex by determining a best-fit location for where the two tracks intersect one another. The particle pairs used within this analysis are required to reside within opposite volumes (i.e. one particle has a track (cluster) in the top half of the SVT (ECal) and the other has its constituents in the bottom half); however, vertices from pairs without this requirement are reconstructed and available for other analyses.

8.4 Analysis Pre-Selection

The final stage that all events go through is a rudimentary pre-selection which chooses events that are of higher quality and conform to our expectation of signal topology and simplifies the resulting shape of the data in the event such that final analysis is not as complicated. Specifically, the pre-selection for this analysis is requiring exactly one quality vertex to be reconstructed within the event. This requirement naturally disposes of events which are cluttered with particles from other beam arrivals (thus causing more than one vertex to be reconstructed) and events which do not have a electron-positron pair within acceptance (thus causing no vertices to be reconstructed). The requirements for a “quality” vertex were developed and optimized by prior work[47] and are standardized within HPS analyses. Figure 8.1 shows the relative efficiency of these requirements. Figure 8.2 shows the distribution of number of vertices that pass

these criteria. Figure 8.3 shows the summary of this pre-selection requirement including the relative efficiency of the quality vertex requirements.

The largest drop in signal efficiency is from the requirement that a quality vertex is found within the event. With the relatively high efficiency of the vertex pre-selections, this drop in efficiency can be interpreted as mainly due to the geometric acceptance of the HPS detector. Fortunately, the background processes are removed in higher proportion from this requirement.

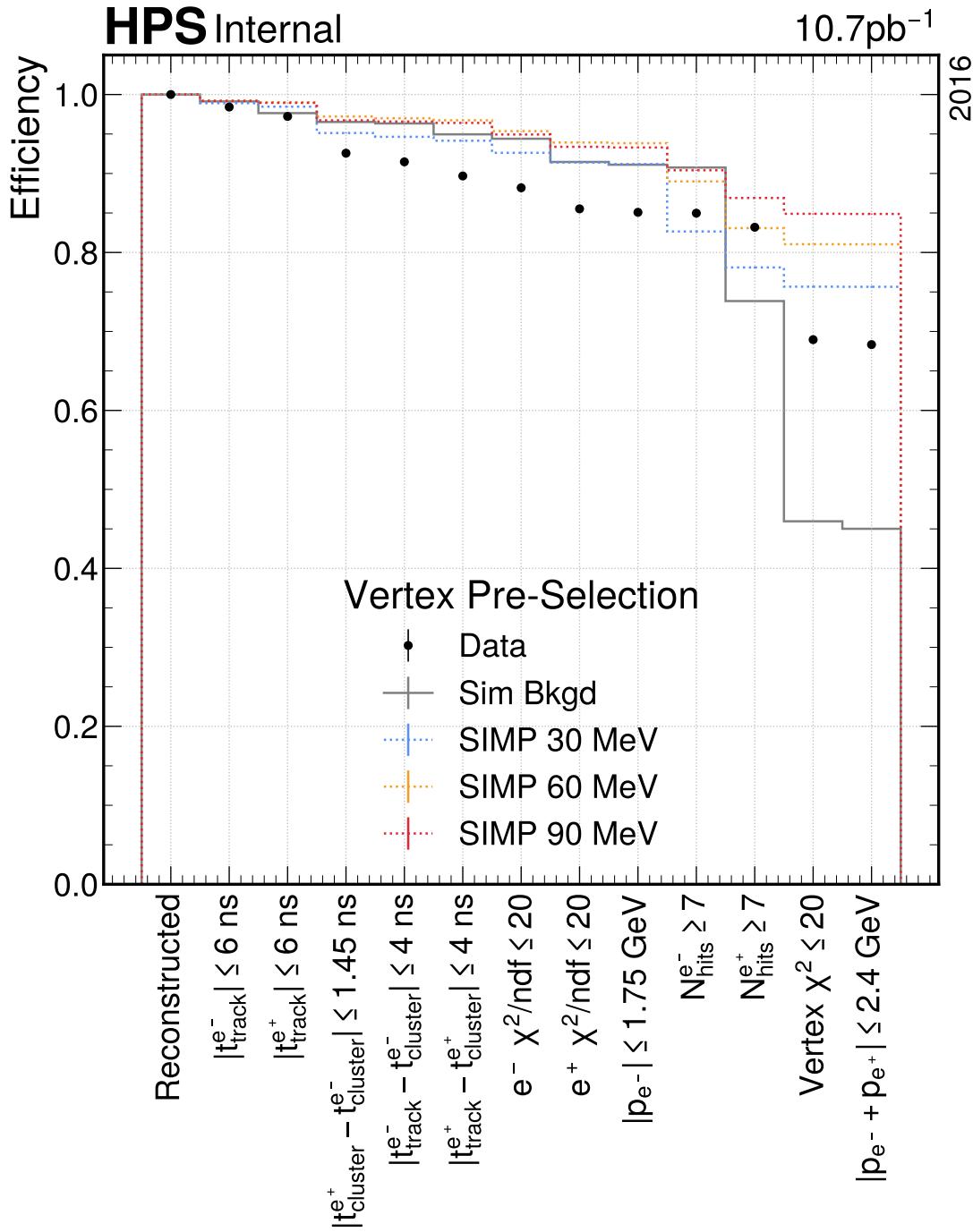


Figure 8.1: Relative efficiency of vertex quality pre-selection requirements.

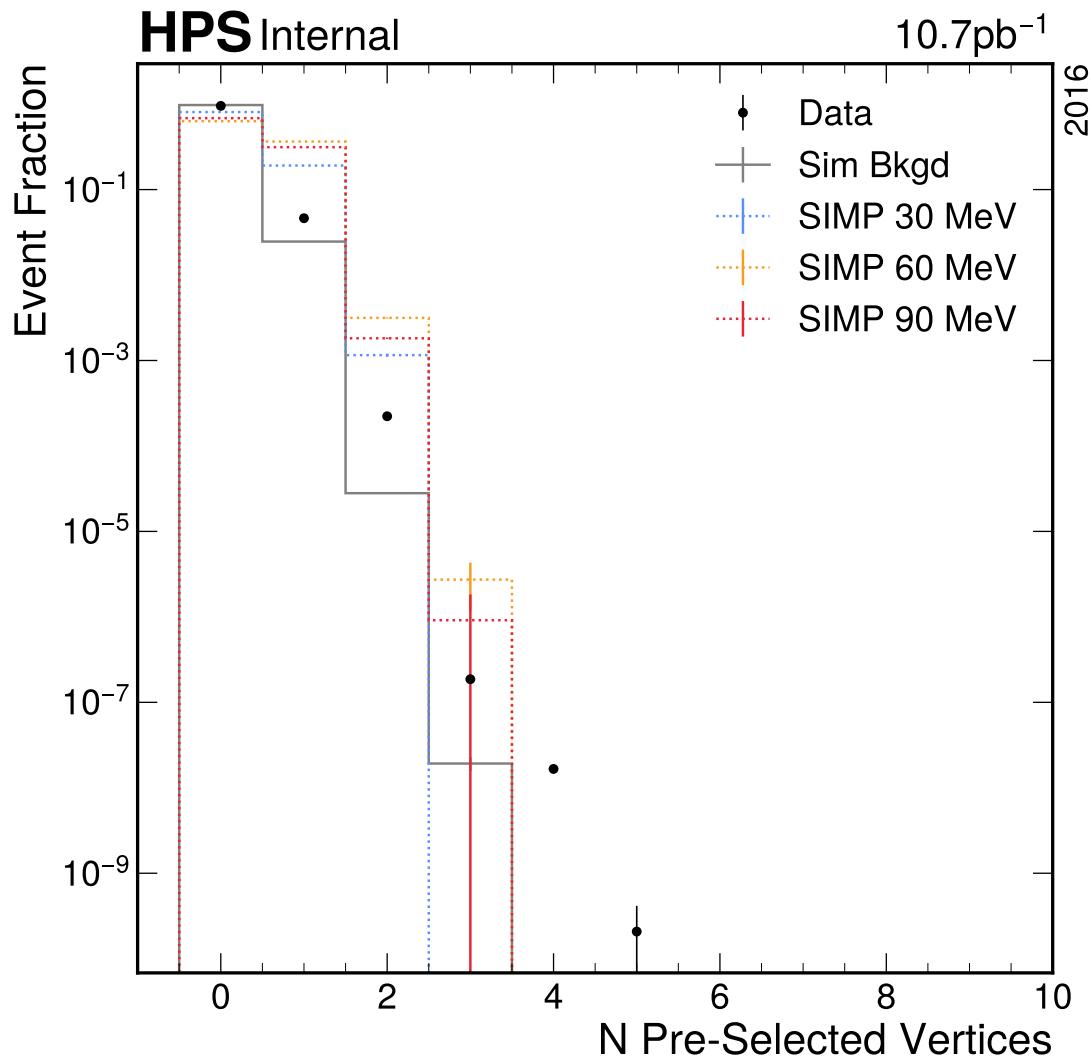


Figure 8.2: Number of vertices passing the quality pre-selection.

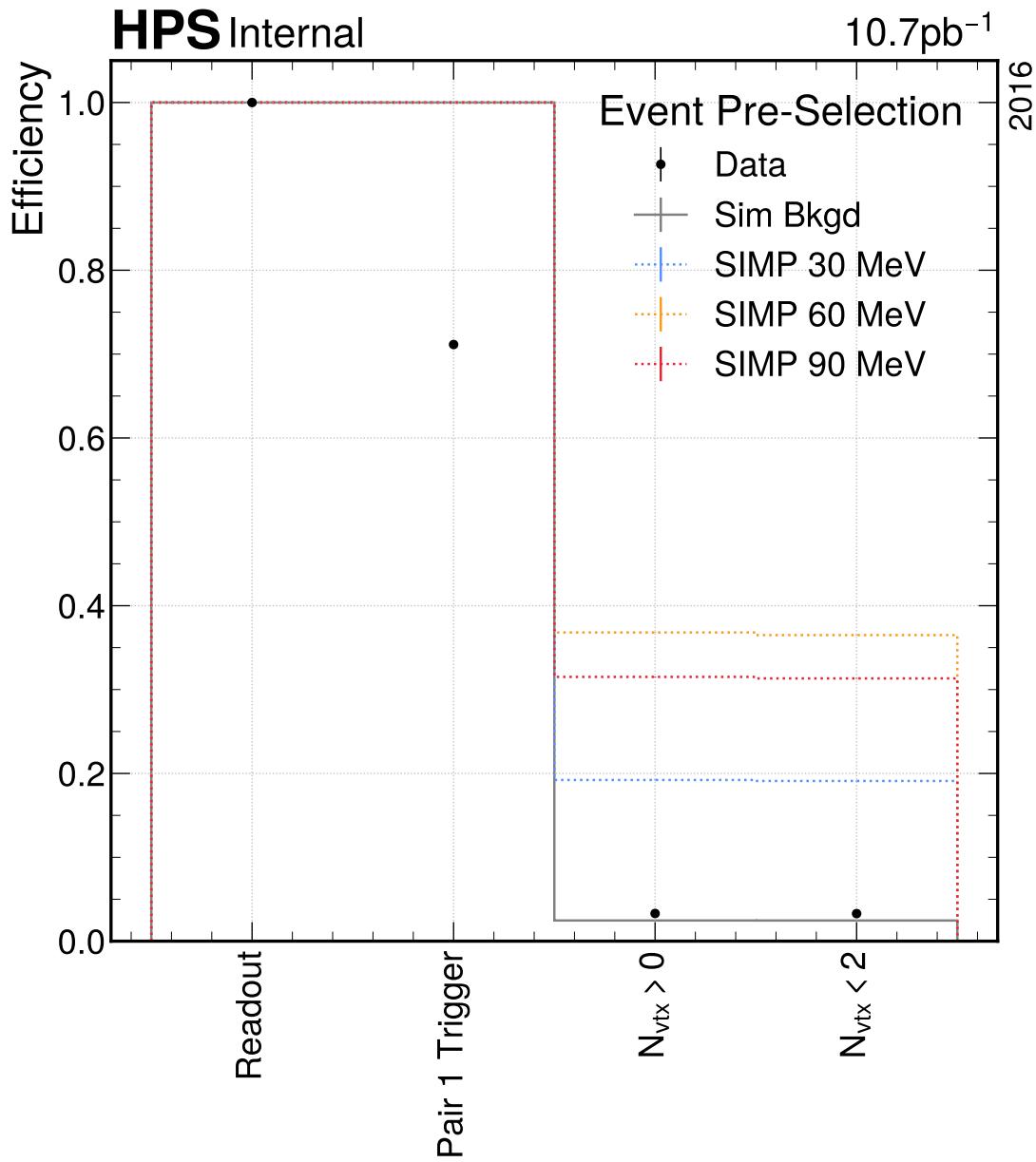


Figure 8.3: Relative efficiency of pre-selection of events. N_{vtx} is the number of vertices passing the vertex pre-selection requirements. Since other triggers for different purposes are present in the data, the Pair 1 Trigger is applied to data but not to simulation.

Chapter 9

Displaced Vertex Search

A search for visibly-decaying DM is not novel, many experiments have been purpose-built for such a search; however, HPS is specially focused on DM that decays within shorter distances. While HPS has previously searched for simple dark sectors more similar to the benchmark model used within the LDMX missing energy search [48], no DM was found and the relatively longer decay lengths of this dark sector model harmed the signal efficiency of the analysis and decreased the resulting sensitivity. Section 6.1 introduced the idea of SIMP DM which has additional interest for HPS. Specifically, the increase in complexity of the DM model slightly decouples the production rate from the decay rate meaning the number of events where DM is produced is not as tightly connected to how displaced the particles are expected to be within the detector. The downside of this decoupling is that some energy is lost to the production of the lighter dark meson π_D when the dark photon decays within the dark sector; however, this also means HPS has not searched this regime since one of the key selections within its previous displaced vertex search was requiring the total momentum of the vertex to be near the beam energy instead of significantly below it.

With the samples presented in Chapter 8, a search for SIMPs within the collected data has been performed. An additional, orthogonal category of data was introduced to be combined with prior work, enabling investigation of a larger portion of the already-collected data.

9.1 Signal Yield

The expected signal yield is a crucial part of any kind of search analysis and this one is no exception. The prior search within LDMX benefited from its missing momentum/energy technique in this regard since the mixing strength ϵ is only connected to the scale of the signal yield and not how the signal events would present themselves in the data. In a visible search, the mixing strength also affects the decay length and thus a more intricate signal yield estimate is required to account for the fact that changing ϵ not only changes the magnitude but also changes which events are more likely to be observed within data. In this analysis, we modify the expected signal yield calculation done previously by HPS [48] to accomodate the two viable decay channels within this SIMP model.

The fundamental premise of this calculation is re-weighting events based on their decay length so that they appear to have been drawn from a specific mixing strength's decay distribution. Specifically, we integrate over a decay weighting factor multiplied by the signal efficiency as a function of the decay length in order to estimate the total signal yield for a specific mixing strength. The details of this calculation are presented in this section since they are relatively novel and not presented fully in another publication; however, this does force me to deviate from my previous commitment to few mathematical equations and less jargony language. Skip to Section 9.2 if this detail does not interest you.

The magnitude of this estimate is set by the relationship between dark photon production cross section and the radiative trident differential cross section[49].

$$\sigma_{A'} = \epsilon^2 \frac{3\pi}{2\alpha} m_{A'} \left. \frac{d\sigma_{\gamma^*}}{dm} \right|_{m=m_{A'}} \quad (9.1)$$

Multiplying both sides by the dataset's luminosity then gives us event yields.

$$N_{A'} = \epsilon^2 \frac{3\pi}{2\alpha} m_{A'} \left. \frac{dN_{\gamma^*}}{dm} \right|_{m=m_{A'}} \quad (9.2)$$

The complexity arises when we remember that specifically radiative tridents are not directly observable – they are intertwined with other standard processes that produce

the same outgoing particles (Bethe-Heitler tridents for example). Thus, we estimate the radiative trident differential yield dN_{γ^*}/dm by modifying the observable trident differential yield $dN_{3e,\text{CR}}/dm_{\text{reco}}$ by two simulation-derived factors. The radiative fraction f_{rad} estimates the fraction of trident events that originate from the radiative process.

$$f_{\text{rad}} = \frac{dN_{\gamma^*}}{dm} \Big/ \frac{dN_{3e}}{dm_{\text{reco}}} \quad (9.3)$$

And the radiative acceptance times efficiency A_{rad} estimates the fraction of trident events that are within the geometric acceptance of the detector and pass the trigger and preselection requirements.

$$A_{\text{rad}} = \frac{dN_{3e,\text{CR}}}{dm_{\text{reco}}} \Big/ \frac{dN_{3e,\text{gen}}}{dm_{\text{reco}}} \quad (9.4)$$

where CR stands for the Control Region in P_{sum} (see Section 9.3 for the definition of this variable). Thus, the expected yield of dark photons created within the detector but not necessarily within its acceptance or passing selection requirements is

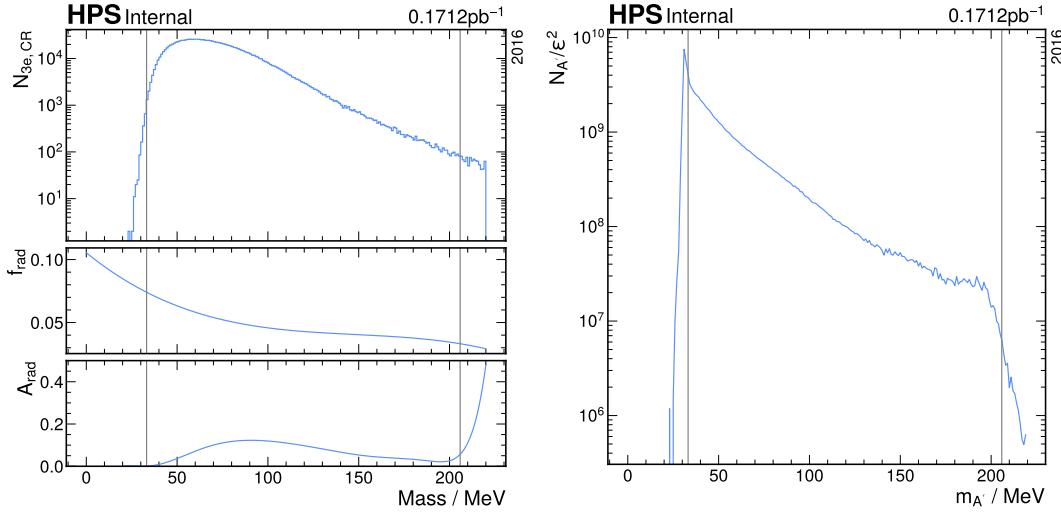
$$N_{A'} = \epsilon^2 \frac{3\pi}{2\alpha} m_{A'} \frac{f_{\text{rad}}}{A_{\text{rad}}} \left. \frac{dN_{3e,\text{CR}}}{dm_{\text{reco}}} \right|_{m=m_{A'}} \quad (9.5)$$

Both f_{rad} and A_{rad} use the event and vertex pre-selection before any downstream analysis selections, so they are common amongst all potential analysis channels that share this pre-selection; however, they both vary as a function of the dark photon mass. Polynomials were fit to the simulation-derived values for these ratios and the parameters for these polynomials are given in Table 9.1. These ingredients along with the resulting $N_{A'}$ is shown in Figure 9.1 using a run corresponding to $\approx 1.6\%$ of the entire dataset.

In this signal hypothesis, we do not observe the dark photon's production or decay, instead, the dark photon decays to an unobservable dark pion and the neutral dark vector meson V_D which in turn is what decays back into a standard model electron-positron pair. The above calculation of $N_{A'}$ accounts for the dark photon's production, but we need to account for the dark photon's decay and the dark vector meson's decay. The first decay has a branching ratio $BR(A' \rightarrow \pi_D V_D)$ which is complicated by the fact that there are actually two different dark neutral vector mesons that fit our requirements.

Coefficient	f_{rad}	A_{rad}
c_0	0.105	-0.489
c_1 / MeV	-1.17×10^{-3}	0.0737
c_2 / MeV^2	7.44×10^{-6}	-4.34×10^{-3}
c_3 / MeV^3	-1.67×10^{-8}	1.34×10^{-4}
c_4 / MeV^4	0.0	-2.36×10^{-6}
c_5 / MeV^5	0.0	2.54×10^{-8}
c_6 / MeV^6	0.0	-1.71×10^{-10}
c_7 / MeV^7	0.0	7.03×10^{-13}
c_8 / MeV^8	0.0	-1.62×10^{-15}
c_9 / MeV^9	0.0	1.60×10^{-18}

Table 9.1: Polynomial coefficients for radiative fraction and acceptance functions.



(a) Ingredients to the total dark photon yield calculation including the observed trident yield (top), radiative fraction (middle), and radiative acceptance (bottom).

(b) Resulting total dark photon yield.

Figure 9.1: Depiction of dark photon yield calculation using a single run of the data ($\sim 1.6\%$). The gray vertical lines show the bounds of the search using the nominal ratio $m_{A'}/m_{V_D} = 1.66$.

The second process is embedded in the decay rate of the dark vector meson to electron-positron pairs $\Gamma(V_D \rightarrow e^+e^-)$.¹ Let $E(z)$ be the signal efficiency of the analysis as a function of the z where the V_D decayed into the electron-positron pair. Then we can sum over the possible V_D and estimate the fraction of $N_{A'}$ that produce a V_D which decays and passes the analysis requirements.

$$N_{\text{sig}} = N_{A'} \int_{z_{\text{target}}}^{\infty} \sum_{V_D} D_{V_D}(z) E(z) dz \quad (9.6)$$

where

$$D_{V_D}(z) = BR(A' \rightarrow \pi_D V_D) \frac{e^{-(z-z_{\text{target}})/(\gamma c \tau_{V_D})}}{\gamma c \tau_{V_D}} \quad (9.7)$$

The branching ratio $BR(A' \rightarrow \pi_D V_D)$ and lifetime τ_{V_D} are taken from [38] (along with the general procedure of this estimate). What is important to remember is that the lifetime is dependent on ϵ^2 , so while increasing ϵ^2 increases $N_{A'}$ it also makes the lifetime shorter and thus the V_D decay “looks” more like standard un-displaced background.

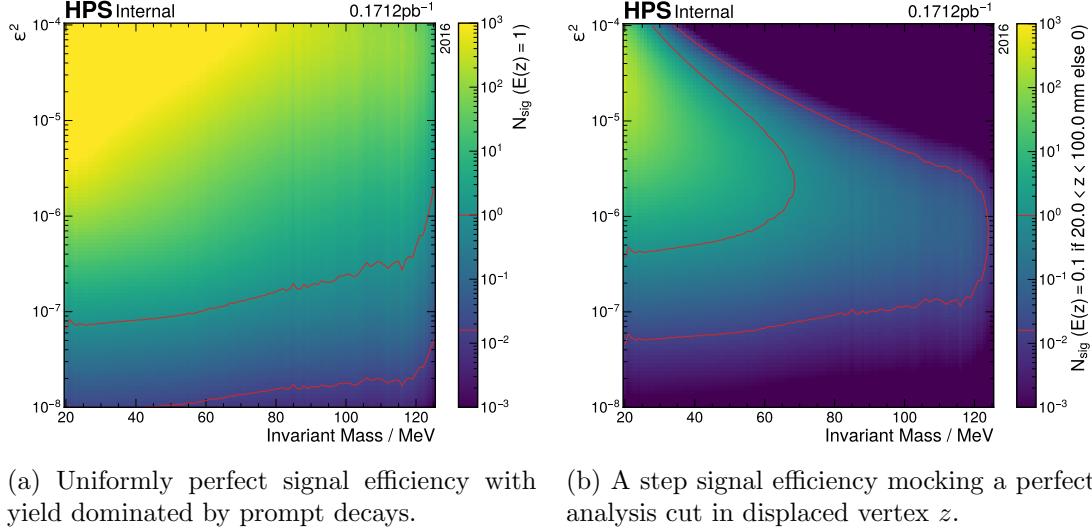
The V_D energy (and thus the relativistic γ) used in $D_{V_D}(z)$ is only distributed over a small range (within $\mathcal{O}(1 \text{ GeV})$) so we replace it with the mean $\langle \gamma \rangle$ in order to make the calculation more practical. Additionally, we impose an upper limit on the decay z to be 25 cm after the target since the efficiency of the reconstruction past $\sim 20 \text{ cm}$ drops to zero.

$$N_{\text{sig}} = N_{A'} \int_0^{25 \text{ cm}} dz' \sum_{V_D} BR(A' \rightarrow \pi_D V_D) \frac{e^{-z'/\langle \gamma \rangle c \tau_{V_D}}}{\langle \gamma \rangle c \tau_{V_D}} E(z' + z_{\text{target}}) \quad (9.8)$$

where $z' = z - z_{\text{target}}$. Figure 9.2a shows N_{sig} with a uniformly perfect signal efficiency $E(z) = 1$ to give a sense of scale, but a slightly more realistic example is using a step-wise signal efficiency $E(z) = 0.1$ if $20 \text{ mm} < z < 100 \text{ mm}$ and 0 otherwise Figure 9.2b.

In summary, the key point to make here is that the signal efficiency as a function of true decay length $E(z)$ is the single analysis-dependent ingredient that effects the resulting signal yield. The rest of the ingredients are either fixed by the dataset being searched (e.g. trident differential production, f_{rad} , A_{rad}) or are parameters that we vary

¹There are other decay processes of the V_D ! Namely, the so-called “three-prong” decay $V_D \rightarrow \pi_D e^+ e^-$ would greatly increase the potential rate at the cost of preventing the $e^+ e^-$ from reconstructing the m_{V_D} mass resonance.



(a) Uniformly perfect signal efficiency with yield dominated by prompt decays. (b) A step signal efficiency mocking a perfect analysis cut in displaced vertex z .

Figure 9.2: Expected signal yield N_{sig} for some rudimentary example efficiencies. Red contour lines are given at 1 and 1.6% (the approximate fraction of this data subsample) to give a sense of scale. Values are “clipped” to the color bar (for example, a signal yield of 2×10^4 is colored yellow which is marked as 1×10^3) so we can focus on the important orders of magnitude.

for the search (ϵ^2 , m_{V_D}). A moderately realistic signal efficiency shows the range of ϵ^2 and m_{V_D} in which this dataset has potential sensitivity (Figure 9.2b) which informs the ranges these values will take below.

9.2 Reconstruction Categories

The design of HPS separates reconstructed tracks into categories that are not necessarily associated with the underlying physics for which we are searching. Specifically, the number of layers (and which layers) that are included within a track has a large effect on the resulting precision of the reconstructed physics variables of that track; thus, we categorize vertices on whether one or both of their tracks contain both sensors in the first layer. Figure 9.3 shows examples of the two categories considered within this work. The L1L1 category whose vertices have both tracks with good precision was first studied for this signal search[47] when optimizing the pre-selection requirements described earlier.

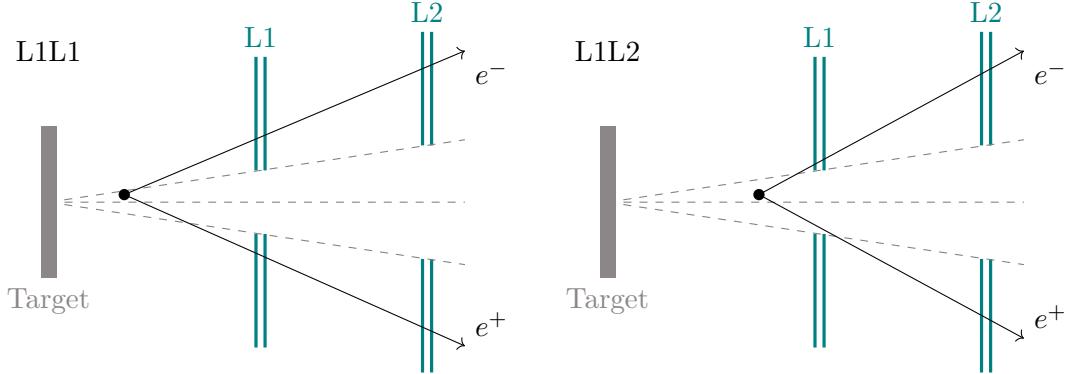


Figure 9.3: Diagrams showing L1L1 (left) and L1L2 (right) vertex examples.

Extending this search to the L1L2 category has a rather simple motivation. While the vertices may degrade in precision, the additional increase of data both in terms of total volume and the potential to observe higher-displaced vertices is expected to improve the sensitivity of the HPS SIMP search.

9.3 Physics Variables

There are many possible variables that could be used to separate candidate signal vertices from the standard process backgrounds. While this section is not an exhaustive list of all possible variables, it does explain the ones used within this analysis as well as motivation for their use.

The SIMP signal vertex will have significantly less energy than the amount delivered by the beam because of the production of a light dark meson when the heavier vector meson V_D is produced. Thus, a selection on the sum of the momentum magnitudes is applied.

$$P_{\text{sum}} = |\vec{p}_{e^-}| + |\vec{p}_{e^+}| \quad (9.9)$$

Specifically, the Signal Region (SR) used for the actual search requires $1.0 \text{ GeV} < P_{\text{sum}} < 1.9 \text{ GeV}$ and the Control Region (CR) used for determining the trident differential production rate is $1.9 \text{ GeV} < P_{\text{sum}} < 2.4 \text{ GeV}$.

Since we are searching for the dark vector boson V_D via its 2-body decay into an electron and a positron, we expect the invariant mass of the vertex m_{reco} to be within

the resolution of the detector σ_m of the mass we are searching for m_{V_D} .

$$p_m = \frac{|m_{\text{reco}} - m_{V_D}|}{\sigma_m} \quad (9.10)$$

Applying an upper limit on p_m is often referred to as a “mass window” since it results in m_{reco} residing within a small range around m_{V_D} . For optimizing the other selections, $p_m < 2.8$ is used. The search for a potential excess of SIMP-like events is using sidebands along m_{reco} in data so no selection on p_m is imposed when searching. The estimate of the resulting sensitivity uses the signal region selection from the search of $p_m < 1.5$.

Each vertex can be projected back to the target using its position and total momentum, and the location of the vertex at the target can be compared to the beam position extracted from 1% of each data-collection run. The separation between the projected position $\vec{x} = (x, y)$ and the beam position $\vec{\mu} = (\mu_{x'}, \mu_{y'})$ is then measured and normalized by the width of the beam in that direction ($\sigma_{x'}, \sigma_{y'}$). The distribution of the beamspot is allowed to be rotated relative to our chosen axes by the angle θ_{beam} , meaning we need to rotate \vec{x} before comparing it to the beam spot. The total “elliptical” separation between the reconstructed vertex projected back to the target and the beam spot can be quantified by

$$N_\sigma = \sqrt{\left(\frac{x \cos \theta_{\text{beam}} - y \sin \theta_{\text{beam}} - \mu_{x'}}{\sigma_{x'}} \right)^2 + \left(\frac{x \sin \theta_{\text{beam}} + y \cos \theta_{\text{beam}} - \mu_{y'}}{\sigma_{y'}} \right)^2} \quad (9.11)$$

I refer to this as VPS since it represents how significantly a vertex deviates from originating at the beam spot.

Since the detector’s tracking modules are oriented to be most sensitive in the vertical direction, the vertical impact parameter y_0 has higher precision compared to the horizontal impact parameter. For truly-displaced signal vertices, both tracks making the vertex would have y_0 far from zero while background vertices would have at least one track with y_0 near zero (undisplaced vertices would have both, but mis-reconstructed fake-displaced vertices could have one far from zero). Figure 9.4 shows some diagrams that illuminate this effect. While there are many ways for a vertex to end up being “fake displaced” (for example, a missing or mis-chosen hit in the first layer like the example shown in Figure 9.4), forcing both y_0 to be far from zero removes these background

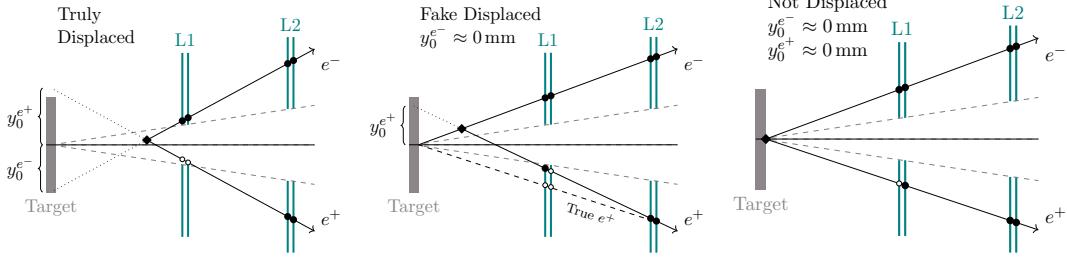


Figure 9.4: Diagrams of examples for how different types of L1L2 vertices effect the observed values of y_0 . Truly displaced vertices (left) have both tracks whose y_0 is far from 0 mm while fake displaced vertices (middle) and not displaced (right) have at least one if not both tracks with $y_0 \approx 0 \text{ mm}$. The diamonds are the reconstructed vertex, the solid circles are the reconstructed hits, and the empty circles are “missed” hits either because the particle did not pass through those sensors (left) or due to some physical or electronic inefficiencies (center and right).

processes. This motivates selecting vertices based on requiring the minimum of the two absolute value y_0 to be above a certain threshold.

$$y_{0,\min} = \min(|y_{0,e^-}|, |y_{0,e^+}|) \quad (9.12)$$

which more sharply distinguishes truly displaced vertices compared to the vertex z often muddled by fake-displaced vertices where one track is mis-reconstructed at high $|y_0|$.

The track fit uncertainty of the vertical impact parameter σ_{y_0} is a helpful quality parameter measuring how confident the track fit is in the y_0 value. Placing an upper limit on this value for both tracks within a vertex effectively requires both tracks to have good vertical resolution, helping remove some highly-displaced vertices presumably arising from mis-reconstructed tracks.

$$\sigma_{y_0,\max} = \max(\sigma_{y_0,e^-}, \sigma_{y_0,e^+}) \quad (9.13)$$

Figure 9.5 displays the distributions of these last three variables for a few example mass points after the other, solidified selections (L1L2 vertex, the momentum sum is in the Signal Region, and the invariant mass falls within the chosen mass window).

Vertex z is left for late-stage statistical analysis of the results and – being highly correlated with $y_{0,\min}$ – is redundant with this variable.

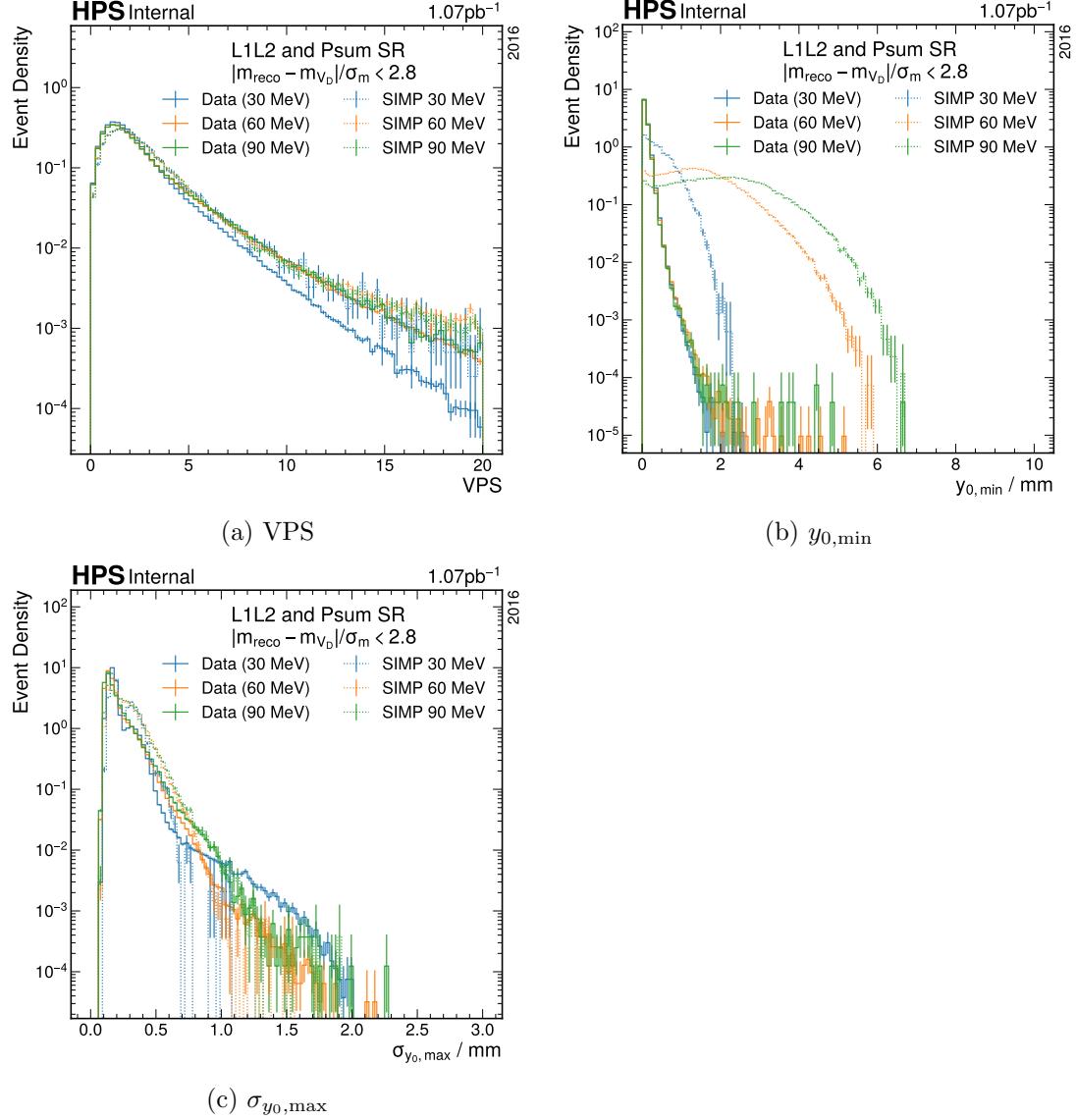


Figure 9.5: Distributions of cut variables for a few example mass points and $\approx 10\%$ of the full data sample. The vertices are required to be L1L2, have their momentum sum be within the signal region, and their invariant mass within the specified mass window.

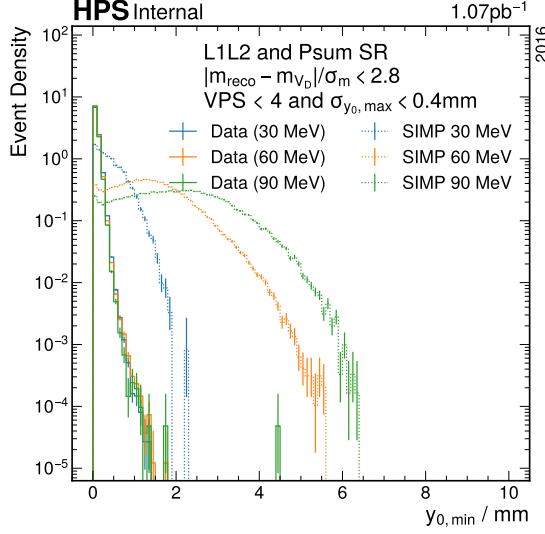


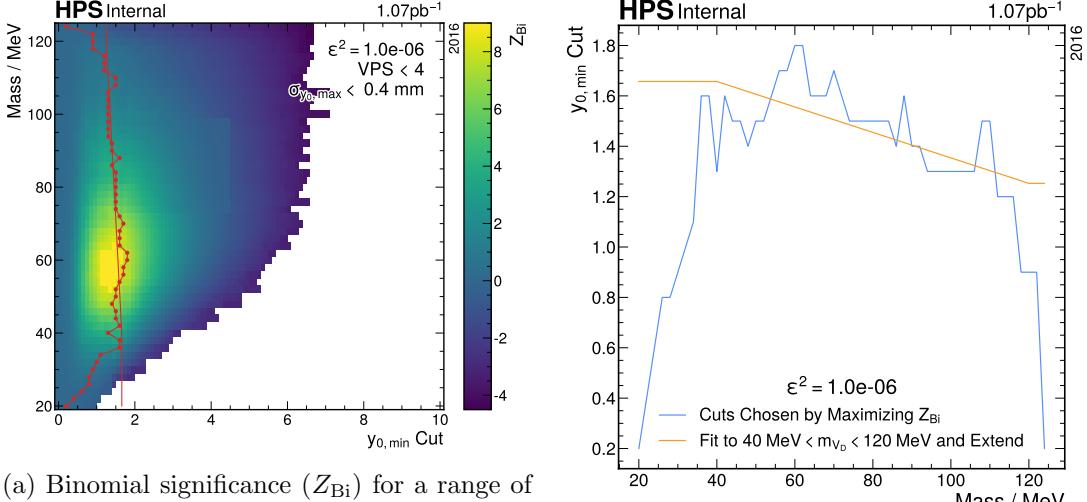
Figure 9.6: $y_{0,\min}$ distribution after the other selections (L1L2, P_{sum} in SR, reconstructed mass lying in mass window, VPS < 4, and $\sigma_{y_{0,\max}} < 0.4$ mm).

9.4 Selection

The selection on these variables was optimized using a subsample of the collected data amounting to $\approx 10\%$ of the full sample. All of the variables except the last, separating variable ($y_{0,\min}$) were varied while keeping the relative signal efficiency high ($> 80\%$), focused on removing the highly-displaced events leading to the long tail in data observed in Figure 9.5b. Figure 9.6 shows the cleaned $y_{0,\min}$ distribution after these selections. There are still a few highly-displaced events left; however, removing them incurs a higher penalty on signal efficiency and leads to a lower resulting sensitivity estimated from this subsample.

The final variable $y_{0,\min}$ was chosen by maximizing the binomial significance of the expected signal at $\epsilon^2 = 1 \times 10^{-6}$ over the observed data for this 10 % subsample. For this optimization, the expected signal was artificially scaled such that it would be on the same order-of-magnitude as the observed data (specifically, a scaling factor of $0.1/\epsilon$ was found to work well although variations in this scaling factor did not vary the chosen selections by much).

The $y_{0,\min}$ values chosen by this maximization (Figure 9.7) are ragged due to the



(a) Binomial significance (Z_{Bi}) for a range of masses and $y_{0,\min}$ cut choices with the points maximizing Z_{Bi} for each mass drawn in red circles and the linear fit with its flat continuation as a red line.

(b) Values maximizing the binomial significance (blue) along with the linear fit and flat extension (orange).

Figure 9.7: Binomial significance (left) being maximized after the VPS < 4 and $\sigma_{y_0,\max} < 0.4 \text{ mm}$ cuts leading to cuts (right, blue) which are smoothed into a continuous function (right, orange) as described in the text.

finite nature of the subsample on which they are being chosen. In order to avoid over-biasing the selection to this arbitrarily-chosen subsample, the actual cuts used in the final analysis are smoothed by fitting a line to the area whose maximum binomial significance is above zero and then continuing this line with flat values outside this region.

In summary, the following cuts developed for this analysis are

- L1L2: In the vertex, one of the particles must have a hit in both sensors in layer 1 and layer 2 while the other must not satisfy this requirement.
- Low Momentum Sum: $1.0 \text{ GeV} < P_{\text{sum}} < 1.9 \text{ mm}$
- Decay after Target: $z > z_{\text{target}}$
- Vertex Projects Back to Beamspot: $\text{VPS} < 4$
- Good Vertical Resolution: $\sigma_{y_0,\max} < 0.4 \text{ mm}$

Parameter	Value
A	1.66 mm
B	1.86 mm
C	-5.1×10^{-3} mm/MeV
D	1.25 mm

Table 9.2: Parameters for Equation (9.14).

A search for a statistically-significant excess of highly-displaced events reconstructed within resolution of a specific mass is then performed at this point. The sensitivity is estimated using two additional cuts.

- Within Mass Window: $p_m < 1.5$
- Truly Displaced: $y_{0,\min} > y_{0,\min}^{\text{cut}}$

where

$$y_{0,\min}^{\text{cut}}(m_{\text{reco}}) = \begin{cases} A & m_{\text{reco}} \leq 40 \text{ MeV} \\ B + Cm_{\text{reco}} & 40 \text{ MeV} < m_{\text{reco}} < 120 \text{ MeV} \\ D & m_{\text{reco}} \geq 120 \text{ MeV} \end{cases} \quad (9.14)$$

and the parameters of this function are given in Table 9.2.

These selections lead to 6 highly-displaced events left in 10% of the data which can be seen in Figure 9.8. The total signal efficiency of this analysis as a function of reconstructed z Figure 9.9a is of the same magnitude but slightly further displaced than the L1L1 analysis Figure 9.9b.

9.5 Results

With the selection fixed, we can analyze the full sample of data without fear of statistical bias.

9.5.1 Search

The first stage in this analysis is to perform the search for an excess of highly-displaced events above what could be expected by the background. The signal model we are searching for would have extra events that are both highly-displaced and centered on

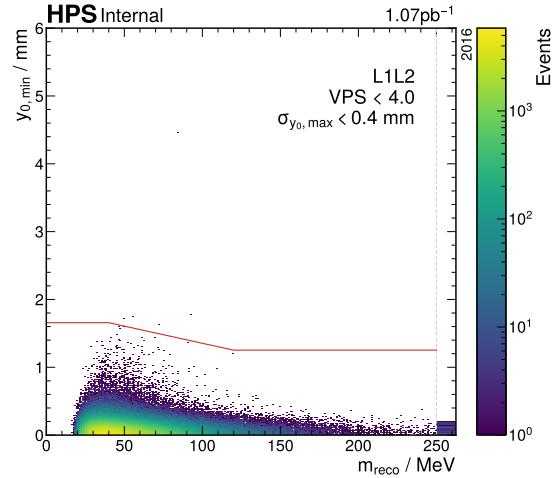


Figure 9.8: $y_{0,\min}$ distribution for the 10% data sample with the $y_{0,\min}^{\text{cut}}$ function overlayed in red.

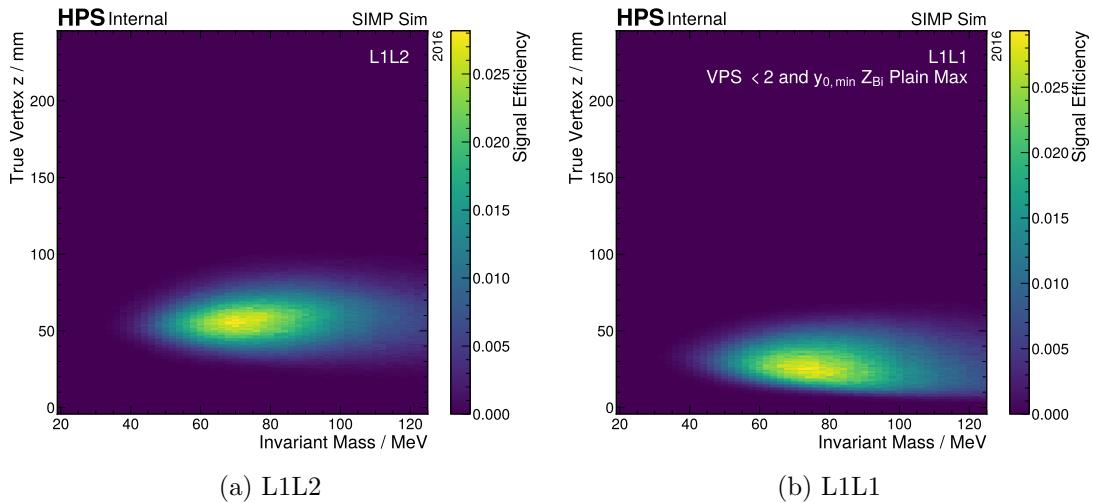


Figure 9.9: Signal efficiency as a function of z and mass scale m_{V_D} for this analysis (left) and L1L1 (right).

Region	m_{reco} Range	$y_{0,\min}$ Range
A	$(m_{V_D} - 4.5\sigma_m, m_{V_D} - 1.5\sigma_m)$	$(y_{0,\min}^{\text{cut}}, \infty)$
B	$(m_{V_D} - 4.5\sigma_m, m_{V_D} - 1.5\sigma_m)$	$(y_{0,\min}^{\text{floor}}, y_{0,\min}^{\text{cut}})$
C	$(m_{V_D} - 1.5\sigma_m, m_{V_D} + 1.5\sigma_m)$	$(y_{0,\min}^{\text{floor}}, y_{0,\min}^{\text{cut}})$
D	$(m_{V_D} + 1.5\sigma_m, m_{V_D} + 4.5\sigma_m)$	$(y_{0,\min}^{\text{floor}}, y_{0,\min}^{\text{cut}})$
E	$(m_{V_D} + 1.5\sigma_m, m_{V_D} + 4.5\sigma_m)$	$(y_{0,\min}^{\text{cut}}, \infty)$
F	$(m_{V_D} - 1.5\sigma_m, m_{V_D} + 1.5\sigma_m)$	$(y_{0,\min}^{\text{cut}}, \infty)$

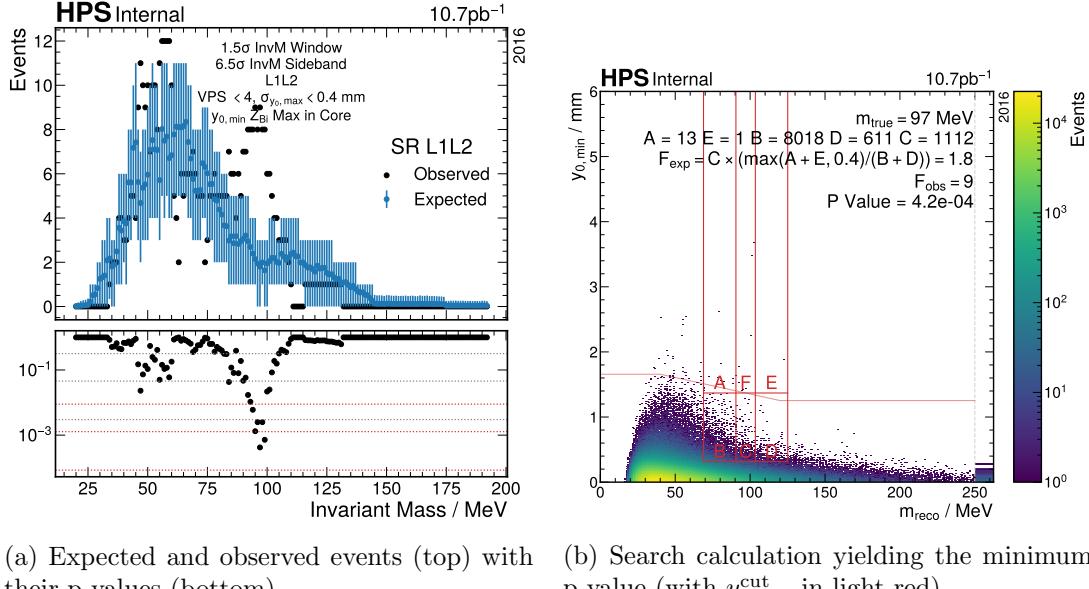
Table 9.3: Region definitions for use in background estimation via sidebands. Region F is the signal region in which we are searching for an excess. m_{V_D} is the mass point we are searching for, σ_m is the detector mass resolution evaluated at m_{V_D} , $y_{0,\min}^{\text{cut}}$ is the optimized cut value evaluated at m_{V_D} , and $y_{0,\min}^{\text{floor}}$ is the maximum value of $y_{0,\min}$ such that region C has at least one thousand events in it.

a specific resonant mass; thus, the search is performed in the $y_{0,\min} - m_{\text{reco}}$ space. The expected number of background events is estimated using a variation of the ABCD technique with two sidebands in mass (one below and one above the mass window in which we are searching) and one sideband in $y_{0,\min}$ (one below the highly-displaced region). The sidebands in mass are defined in terms of the mass resolution of the detector σ_m and the values defining these edges were optimized within the L1L1 analysis. The sideband in $y_{0,\min}$ is defined such that the region within the mass window but at lower displacement (Region C) has at least one thousand events within it to avoid signal contaminating the background estimation.

Table 9.3 gives the definition of these regions. Figure 9.10a shows the search results comparing the expected number of events to the observed along with a corresponding p-value estimated using ten thousand toy experiments.² Figure 9.10b shows an example of the search calculation using the mass point resulting in the smallest p-value observed.

A few mass points had an excess of events observed within the signal region; however, these excesses never exceeded a global significance of 3σ and the masses that do show the largest excess in this channel do not display corresponding excess in the L1L1 channel. Thus these excesses are understood as rare but normal statistical fluctuations.

²Each toy experiment is constructed by sampling a toy values for regions C (normally distributed about actual value), B+D (normally distributed about actual value), and A+E (poisson distributed about actual value) and then re-calculating the expected number in region F. The number of toy experiments whose expected value in F was greater than or equal to the observed number in F divided by the total number of toy experiments is then our estimate p-value.



(a) Expected and observed events (top) with their p-values (bottom). (b) Search calculation yielding the minimum p-value (with $y_{0,\min}^{\text{cut}}$ in light red).

Figure 9.10: Search results for the full 2016 data set.

9.5.2 Sensitivity and Exclusion

With no statistically-significant excess observed, we can then move towards estimating the sensitivity of the analysis and then to exclude certain regions of parameter space. The sensitivity estimate is rather simple; just a ratio between the expected signal yield and the maximum signal yield allowed by the observed data distribution. Section 9.1 already described the expected signal calculation used here as well as previously within the selection optimization. Figure 9.11a shows the final expected signal yield for this selection using the full data sample. The maximum signal yield allowed is calculated using the OIM [50] which leverages the differing shape between signal and background and allows for the presence of unknown background to alter its scale. The OIM result is shown in Figure 9.11b.

Figure 9.12 shows the sensitivity for both this analysis of L1L2 and the L1L1 analysis for comparison along with a contour drawn where the sensitivity equals one after being suppressed by potential systematic errors. Systematic errors arising from the radiative fraction estimate, pre-selection cuts, radiative acceptance estimate, uncertainty in the target position, final selection cuts, mass resolution, beamspot resolution, and difference

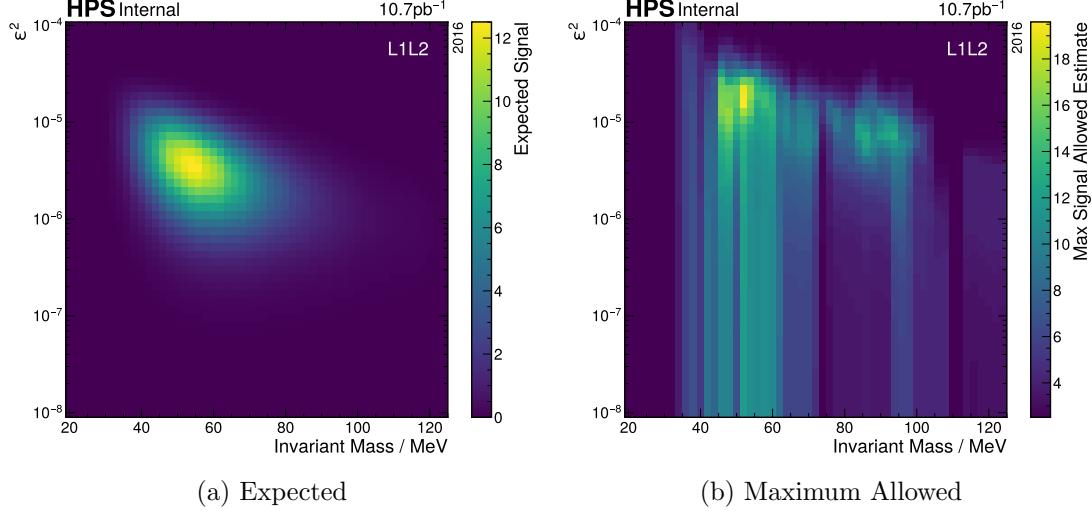


Figure 9.11: Signal yield for the L1L2 reconstruction category with the described cuts showing the theoretical expectation (left) and the maximum allowed (right) derived using the OIM.

in shape of P_{sum} distributions were all evaluated leading to total systematic error of $< 10\%$ for all but the lowest mass points evaluated (rising up to $\sim 18\%$).

This contour encloses the excluded parameter space where the expected yield exceeds what is allowed. In this region of parameter space, *BABAR* [7] has excluded the parameters above $\epsilon^2 = 10^{-6}$, so this analysis does not exclude new parameter space. Nevertheless, this analysis was able to confirm *BABAR*'s results and is the first of its kind (displaced-vertex search in the low- P_{sum} region) for HPS and thus opens the door to further refinement and investigation with later and larger datasets.

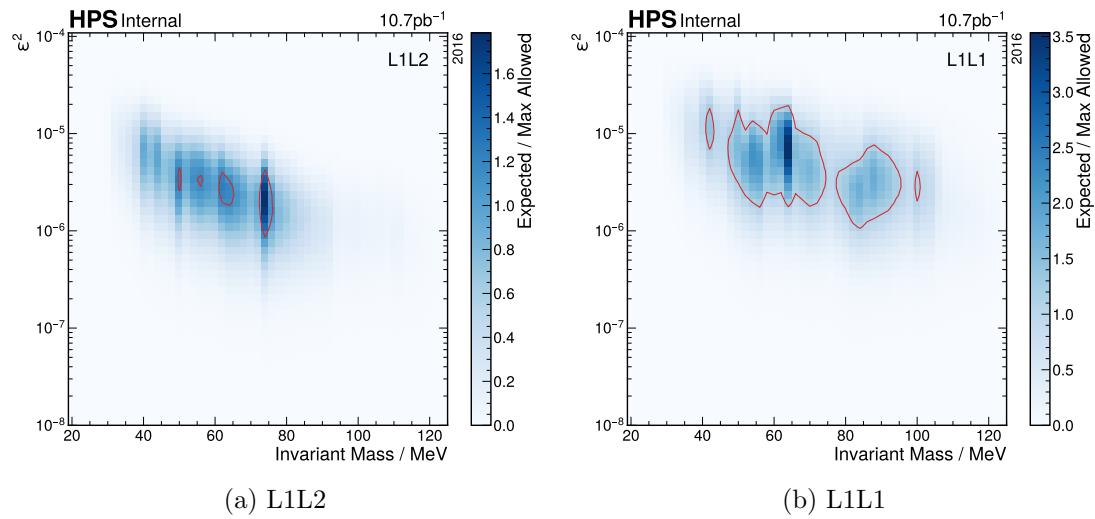


Figure 9.12: Sensitivity of this analysis with contour drawn including systematic errors surpassing the expected signal yield.

Part IV

Conclusion

Chapter 10

Conclusion

Both LDMX and HPS are important and required experiments in our search for the particle nature of DM. While neither has progressed far enough to exclude new parameter space (or potentially discover DM that was previously out of reach), both experiments take novel approaches towards this search. These approaches are required because the traditional tactics have not yielded discovery and are unable to reach into these specific regions of parameter space.

The LDMX EaT analysis presented here is novel for LDMX not only due to the missing-energy search style, but also because we took care to estimate the magnitude of potential systematic errors, form a quantitative non-zero background prediction with its own statistical uncertainty, and performed the statistical analysis over several bins. Previous LDMX analyses focusing on zero-background searches, while important and interesting in their own right, do not have the flexibility necessary to be an analysis used with data collected from a newly built apparatus. The EaT analysis channel is such an analysis and it is well prepared to be the first analysis of LDMX data making new physics conclusions about our universe.

The HPS SIMP L1L2 analysis extended the SIMP search to a previously-unexplored reconstruction category, enabling us to use a larger fraction of the data already collected by HPS. This SIMP search, while not excluding new phase space with the current data set, shows promise. We have been able to show that a displaced-vertex search with the possibility of missing energy is possible. Extending such a search to unexplored phase space, an area where no other experiments have access, is simply awaiting the necessary

reconstruction studies in order to analyze later and larger data sets already collected by the detector.

As newer and smaller experiments compared to other HEP experiments, different excitements and difficulties arise. While participating in the full breadth of a particle physics experiment has taught me a lot and given me opportunity to carve my own path of learning, the lack of larger support structures has revealed some specifically intricate problems that LDMX and HPS share.

10.1 Simulation Validity

The design and initial studies of any experiment requires some simulation of the measurements it could perform. After the experiment has been constructed and collects data, the simulation is still incredibly useful for studying how *known* processes present themselves within the observations. The collected data *could* have a previously-unknown and rare process within it, so we must be careful to avoid biasing our analysis in such a way that would ruin the validity of our results. A reasonable strategy is simply to avoid studying all of the data at once and instead design and optimize the analysis on a particular subset of the data. This process is called “blinding” and was used in Part III; however, this also naturally means the analysis can easily be optimized only for the special events within the subset and the events outside of this subset are unaccounted for.

A valid simulation, one which correctly predicts the scale and shape of the observable distributions, is a valuable tool that could partially or completely replace this blinding procedure. In this case, the simulated observations could be used to help design and optimize the analysis; avoiding bias while also allowing the analysis to explore the full breadth of potential events that could be seen.

While the “valid simulation” I describe here is not necessarily possible, we can move forward towards it in order to make future analyses better. One of the first studies that will be done with new LDMX data will be comparing how our simulated events compare to the events actually being collected by the real detector. Returning to these comparisons for HPS and its subsequent, larger sets of data is being done and will be beneficial for those analyses as well especially given the separation between data and

simulation already being observed within the pre-selection (for example, in Figure 8.1).

10.2 Tracking Reconstruction

Most of my work on tracking and the alignment of a tracking detector was done within the context of HPS. This particular tracking detector is somewhat of a “trial by fire” due to its delicate, two-sided nature required by the specific physics for which we are searching. Nevertheless, this experience has given me a solid introduction to charged particle tracking.

As discussed in Chapter 7, charged particle tracking is very complicated and it almost certainly requires familiarity with the equally-complicated process of alignment. Going forward, both LDMX and HPS require robust and solid tracking software in order to make the interpretations of such intricate algorithms simpler. Without the backing of a larger collaboration to more completely test and validate this complicated software, using a shared tracking framework like ACTS[51] is necessary, a strategy starting to be employed in both experiments.

10.3 Reflection

This dissertation work has shown me the full breadth of a particle physics experiment. From initial design and simulation, to extensions and validations of said simulation, to intricate considerations of potential systematic errors in the experiment, to optimization of selections using various figures of merit, to intentional analysis blinding to avoid statistical bias when developing an analysis, to final unblinding after proposals and evaluations, to post-mortem analysis of the leftover bits of data.

While I have not documented all of these steps within this document, they all brought with them lessons for which I am grateful.

References

- [1] Randall Munroe. xkcd webcomic. Licensed under Creative-Commons Attribution-NonCommercial 2.5.
- [2] Torsten Åkesson et al. (LDMX Collaboration). A High Efficiency Photon Veto for the Light Dark Matter eXperiment. *JHEP*, 04:003, 2020, 1912.05535.
- [3] CERN. The decay of a lambda particle in the 32 cm hydrogen bubble chamber. CERN-EX-11465, 1960.
- [4] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Physics Reports*, 267(5):195–373, 1996.
- [5] D. Banerjee et al. (NA64 Collaboration). Dark Matter Search in Missing Energy Events with NA64. *Physical Review Letters*, 123(12), Sep 2019.
- [6] D. Banerjee et al. (NA64 Collaboration). Improved limits on a hypothetical $X(16.7)$ boson and a dark photon decaying into e^+e^- pairs. *Phys. Rev. D*, 101:071101, Apr 2020.
- [7] J. P. Lees et al. (*BABAR* Collaboration). Search for Invisible Decays of a Dark Photon Produced in e^+e^- Collisions at *BABAR*. *Physical Review Letters*, 119(13), September 2017.
- [8] D. Akimov et al. (COHERENT Collaboration). First Probe of Sub-GeV Dark Matter beyond the Cosmological Expectation with the COHERENT CsI Detector at the SNS. *Phys. Rev. Lett.*, 130:051803, Feb 2023.

- [9] Samantha K McCarty. *Search for an A' Resonance in the Heavy Photon Search 2.3 GeV Data Set.* PhD thesis, University of New Hampshire, 2020.
- [10] H Bethe and W Heitler. On the stopping of fast particles and on the creation of positive electrons. *Proceedings of the Royal Society London A*, 146:83–112, 1934.
- [11] V. C. Rubin, Jr. Ford, W. K., and N. Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). *APJ*, 238:471–487, June 1980.
- [12] Edvige Corbelli and Paolo Salucci. The extended rotation curve and the dark matter halo of M33. *Monthly Notices of the Royal Astronomical Society*, 311(2):441–447, 01 2000, <https://academic.oup.com/mnras/article-pdf/311/2/441/2881340/311-2-441.pdf>.
- [13] Fred I. Cooperstock and S. Tieu. Galactic Dynamics via General Relativity: A Compilation and New Developments. *Int. J. Mod. Phys. A*, 22:2293–2325, 2007, astro-ph/0610370.
- [14] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter. *The Astrophysical Journal*, 604(2):596, apr 2004.
- [15] P. A. R. Ade et al. (Planck Collaboration). Planck 2015 results: Xiii. cosmological parameters. *Astronomy & Astrophysics*, 594:A13, September 2016.
- [16] R. Amanullah et al. (The Supernova Cosmology Project Collaboration). Spectra and hubble space telescope light curves of six type ia supernovae at $0.511 < z < 1.12$ and the union2 compilation. *The Astrophysical Journal*, 716(1):712, may 2010.
- [17] David N. Schramm and Michael S. Turner. Big-bang nucleosynthesis enters the precision era. *Rev. Mod. Phys.*, 70:303–318, Jan 1998.
- [18] Bernard Carr, Kazunori Kohri, Yuuiti Sendouda, and Jun’ichi Yokoyama. Constraints on primordial black holes. *Reports on Progress in Physics*, 84(11):116902, nov 2021.

- [19] Jim Alexander et al. Dark Sectors 2016 Workshop: Community Report, 2016, 1608.08632.
- [20] Kenneth M. Nollett and Gary Steigman. BBN and the CMB constrain neutrino coupled light WIMPs. *Physical Review D*, 91(8), April 2015.
- [21] B. D. Fields, P. Molaro, and Sarkar S. Big Bang Nucleosynthesis. *Phys. Rev. D*, 110(3):030001, 2024. Chapter 24 of PDG Review of Particle Physics.
- [22] R. Agnese et al. (SuperCDMS Collaboration). First dark matter constraints from a SuperCDMS single-charge sensitive detector. *Physical Review Letters*, 121(5), aug 2018.
- [23] A. Aguilar-Arevalo et al. (DAMIC Collaboration). Results on low-mass weakly interacting massive particles from an 11 kg d target exposure of DAMIC at SNOLAB. *Physical Review Letters*, 125(24), dec 2020.
- [24] E. Aprile et al. (XENON1T Collaboration). Dark matter search results from a one ton-year exposure of XENON1T. *Physical Review Letters*, 121(11), sep 2018.
- [25] Torsten Åkesson et al. (LDMX Collaboration). Light Dark Matter eXperiment (LDMX), 2018, 1808.05219.
- [26] Torsten Åkesson et al. (LDMX Collaboration). Photon-rejection power of the Light Dark Matter eXperiment in an 8 GeV beam. *JHEP*, 12:092, 2023, 2308.15173.
- [27] Tor Raubenheimer (LCLS-II Collaboration). LCLS-II: Status of the CW X-ray FEL Upgrade to the LCLS Facility. In *37th International Free Electron Laser Conference*, pages 618–624, 2015.
- [28] T G Buekers, J W Amann, and Y M Nosochkov. A Beam Spreader System for LCLS-II. *NAPAC 2019 Proceedings*, 09 2019.
- [29] D. Contardo et al. The Phase-2 Upgrade of the CMS Endcap Calorimeter. Technical Report CERN-LHCC-2017-023, CMS-TDR-019, CMS, 2017.
- [30] Rene Brun et al. root-project/root: v6.18/02, August 2019.

- [31] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.
- [32] S. Agostinelli et al. Geant4 – a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.
- [33] Tom Eichlersmith, Jeremiah Mans, Omar Moreno, Joseph Muse, Michael Revering, and Natalia Toro. Simulation of dark bremsstrahlung in Geant4. *Computer Physics Communications*, 287:108690, 2023.
- [34] Aram Hayrapetyan et al. (CMS Collaboration). The CMS statistical analysis and combination tool: COMBINE. *Comput. Softw. Big Sci.*, 8:19, 2024, 2404.06614.
- [35] M. Bondi, A. Celentano, R. R. Dusaev, D. V. Kirpichnikov, M. M. Kirsanov, N. V. Krasnikov, L. Marsicano, and D. Shchukin. Fully Geant4 compatible package for the simulation of Dark Matter in fixed target experiments. *Comput. Phys. Commun.*, 269:108129, 2021, 2101.12192.
- [36] B. Banto Oberhauser, P. Bisio, A. Celentano, E. Depero, R. R. Dusaev, D. V. Kirpichnikov, M. M. Kirsanov, N. V. Krasnikov, A. Marini, L. Marsicano, L. Molina-Bueno, M. Mongillo, D. Shchukin, H. Sieber, and I. V. Voronchikhin. Development of the fully Geant4 compatible package for the simulation of Dark Matter in fixed target experiments, 2024, 2401.12573.
- [37] Yonit Hochberg, Eric Kuflik, Tomer Volansky, and Jay G. Wacker. Mechanism for Thermal Relic Dark Matter of Strongly Interacting Massive Particles. *Physical Review Letters*, 113(17), October 2014.
- [38] Asher Berlin, Nikita Blinov, Stefania Gori, Philip Schuster, and Natalia Toro. Cosmology and accelerator tests of strongly interacting dark matter. *Physical Review D*, 97(5), March 2018.
- [39] Matthew Regan Solt. *Searching for long-lived dark photons with the heavy photon search experiment*. PhD thesis, Stanford, 2020.

- [40] Volker D. Burkert. The JLab 12 GeV upgrade and the initial science program - Selected topics. *Proc. Int. Sch. Phys. Fermi*, 180:303–332, 2012.
- [41] Dudek, Jozef, Ent, Rolf, Essig, Rouven, Kumar, K. S., Meyer, Curtis, McKeown, R. D., Meziani, Zein Eddine, Miller, Gerald A., Pennington, Michael, Richards, David, Weinstein, Larry, Young, Glenn, and Brown, Susan. Physics opportunities with the 12 GeV upgrade at Jefferson Lab. *Eur. Phys. J. A*, 48(12):187, 2012.
- [42] Reza Kazimi, John Hansknecht, Michael Spata, and Haipeng Wang. Source and extraction for simultaneous four-hall beam delivery system at cebaf. *Proceedings, 4th International Particle Accelerator Conference (IPAC 2013)*, 1(4), 2013.
- [43] Claus Kleinwort. General broken lines as advanced track fitting method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 673:107–110, 2012.
- [44] Volker Blobel, Claus Kleinwort, and Frank Meier. Fast alignment of a complex tracking detector using advanced track models. *Computer Physics Communications*, 182(9):1760–1763, 2011. Computer Physics Communications Special Edition for Conference on Computational Physics Trondheim, Norway, June 23-26, 2010.
- [45] Johan Alwall, Pavel Demin, Simon de Visscher, Rikkert Frederix, Michel Herquet, Fabio Maltoni, Tilman Plehn, David L Rainwater, and Tim Stelzer. Madgraph/madevent v4: the new web generation. *Journal of High Energy Physics*, 2007(09):028–028, September 2007.
- [46] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), July 2014.
- [47] Alic Spellman. *Searching for Strongly-Interacting Dark Matter With the Heavy Photon Search Experiment*. PhD thesis, University of California Santa Clara, 2024.
- [48] P. H. Adrian et al. (HPS Collaboration). Searching for prompt and long-lived dark photons in electroproduced electron-positron pairs with the heavy photon search experiment at JLab. *Physical Review D*, 108(1), July 2023.

- [49] James D. Bjorken, Rouven Essig, Philip Schuster, and Natalia Toro. New fixed-target experiments to search for dark gauge forces. *Physical Review D*, 80(7), oct 2009.
- [50] S. Yellin. Finding an upper limit in the presence of an unknown background. *Physical Review D*, 66(3), August 2002.
- [51] Xiaocong Ai et al. A Common Tracking Software Project. *Computing and Software for Big Science*, 6, 2022.