FYSK04 15 hp Bachelors of Science Thesis

Following the Path of Individual Beam Electrons through LDMX

Author: Merit Aerts

Supervisor: Lene Kristian Bryngemark
Co-supervisor: Ruth Pöttgen

Examined May 2025

LUNDS UNIVERSITET

Department of Physics
Division of Particle and Nuclear Physics

Abstract

The Light Dark Matter Experiment (LDMX) is a fixed-target missing-momentum experiment which hopes to detect dark matter in the MeV-GeV mass range. This is done by firing an 8 GeV electron beam at a thin tungsten target, where the electrons produce dark matter through a process called dark bremsstrahlung. Electrons which have gone through dark matter production can then be found by looking for missing energy in the detector. As the electron beam fires incredibly large amounts of particles at the target, many of which are not expected to do anything, a trigger system is set in place, which singles out events with no missing energy and discards the corresponding data. As of now this trigger works by looking at the full energy of all simultaneously incoming electrons. This leads to a decrease in trigger efficiency when several electrons enter the detector at once, as the energy resolution of the detector degrades at higher energies. In order to improve the performance of the trigger, a matching algorithm is developed, which should allow for the study of individual electron energy instead of the grouped energy of all electrons together, thereby removing pile-up electrons which have not lost any energy to dark matter production.

This project focuses on the performance of this matching algorithm in the case when two electrons enter the detector at once. It finds that the algorithm works nominally for the single-electron case, but struggles in the when two electrons enter the detector at once, especially when their energy deposits are spatially close or overlap. In the case when all incoming electrons are matched, studies of the energy distributions point towards the matching algorithm being viable for use for pile-up removal. It is also shown to have a potential use in cases where only part of the incoming electrons are matched.

Acknowledgments

I would like to thank Lene Kristian Bryngemark, my supervisor, for their continued guidance throughout this project. I would also like to thank my co-supervisor Ruth Pöttgen, as well as everyone else in the LunDMX-group. The coding and writing workshops have been infinitely helpful over the course of this project.

I would like to thank my friends for putting up with my complaints about c++ and threats to blow up my computer. I do not think I would have made it this far without you. Special thanks to those of you who took the time to read through my thesis, I greatly appreciate your feedback.

Abbreviations

DM	Dark Matter
ECal	Electromagnetic Calorimeter
EoT	Electrons on Target
Geant4	GEometry ANd Tracking 4
HCal	Hadronic Calorimeter
LDM	Light Dark Matter
LDMX	Light Dark Matter eXperiment
`ldmx-sw`	Light Dark Matter eXperiment SoftWare
MG/ME	MadGraph/MadEvent
PF	ParticleFlow
RMS	Root Mean Square
SM	Standard Model

Contents

1 Introduction

Astronomical measurements point towards inconsistency between the visible mass distribution of the Universe, and the distribution which would be expected to make up for certain gravitational effects. This includes the velocity distributions of rotational galaxies such as the Milky Way, where stars at large distances from the galactic center orbit it at speeds which should be too high when only considering the gravitational effects of visible matter. Additionally, observations of the Bullet Cluster, formed through the collision of two galaxy clusters, point to the existence of some unseen matter causing gravitational lensing. These things, among other such phenomena, signify the existence of some matter beyond the standard model (SM), which interacts gravitationally but not electromagnetically. This is what we call dark matter (DM).

There exist several theories on the creation of DM, all of which ascribe it different properties. In order to determine these properties, one would like to detect DM directly in an experimental setting on Earth. Many experiments are searching for DM within different mass-ranges, yet as of now none have been successful. The Light Dark Matter eXperiment (LDMX) is a future such experiment with hopes of detecting DM in the sub-GeV mass range, which thus far has remained largely unexplored. LDMX is a missing momentum search experiment, where an 8-GeV electron beam is aimed at a fixed tungsten target. As the electrons hit the target they may produce DM in a process called dark bremsstrahlung, thereby losing some of their energy. Behind the target is a tracker and a calorimeter, which measures the total energy deposited by the electron. By searching for events where the full beam energy cannot be recovered, DM can be detected indirectly.

It is paramount that the detector is able to differentiate between signal electrons, which have undergone dark bremsstrahlung, and background electrons, whose lost momentum can be accounted for entirely by SM processes. A trigger system is set in place, which searches for events with missing beam energy. Pile-up refers to multiple electrons entering the detector simultaneously. This causes the trigger efficiency to go down, as accurate measurements of missing energy become less precise. To counteract this, a matching algorithm is developed, which allows one to follow the path of each individual electron through the detector. The focus of this project is the performance of this algorithm during pile-up events. This is a stepping stone towards the final goal, which is making use of the matching algorithm to remove data corresponding to pile-up electrons from missing energy calculations.

2 LDMX

2.1 Dark matter

As early as the 19th century astronomers used the term dark matter to describe matter which could not be observed with any instrument, but whose existence could be inferred from the behavior of surrounding visible matter. At that point in time this was usually in reference to non-luminous stars and planets, matter which interacts electromagnetically but whose light simply does not reach Earth in any meaningful quantity. Yet the concept is not far off from that which we now refer to as DM [1].

In the early 20th century, astronomers began finding discrepancies in the mass-to-light ratio of galaxies, which were too high to be explained by just non-luminous stars. At the same time, measurements of the rotation curve of our and other galaxies were made, which seemed to point to a large amount of mass in the outer reaches of the galaxy that seemingly did not exist. Theories began emerging about invisible, dark, matter which simply does not interact with light. At first, DM was considered a fringe theory, but throughout the late 20th and early 21st century, more and more unexplained phenomena started cropping up, such as gravitational lensing much too intense than that which should be caused by visible matter. And so DM came to be widely accepted [1].

Even so, very little is known for sure about DM. There exists a wide range of theories which ascribe it different properties. Relevant for LDMX is the thermal freeze-out theory [2], which serves as an explanation for the proportion of DM in our universe. The theory assumes that DM is made up of particles which interact, through production and annihilation, with SM matter, only with a very low cross-section. In the early hot and dense Universe, DM and SM matter existed in thermal equilibrium, with DM production and annihilation occurring at the same rate. As the universe expanded and cooled down, the production decreased as it started becoming kinematically forbidden due to lack of energy, yet annihilation remained approximately the same. As expansion continued the probability of two DM particles finding each other decreased, annihilation decreased to the same rate as production, and an equilibrium was reached once more. This is called the freeze out. The resulting DM abundance is called the thermal relic abundance [3].

2.1.1 The LDMX search for dark matter

The thermal freeze out allows for dark-matter particles somewhere in the MeV-TeV mass range. Particles that fall in the upper half of this range, GeV-TeV, are compatible with Weakly Interacting Massive Particles (WIMPs). This mass range has been thoroughly explored by previous experiments, without any results, largely excluding WIMPs as candidates for dark matter. Particles in the lower mass range, MeV-GeV, are compatible with Light Dark Matter (LDM). This range remains largely uncharted, as conventional experiments built to detect WIMPs struggle to probe the region. LDMX is designed to probe this region specifically [2].

The specific process LDMX hopes to detect is called dark bremsstrahlung. Dark bremsstrahlung works much like regular bremsstrahlung; an electron is accelerated in the electric field in the vicinity of a heavy nucleus, causing the emission of a DM-mediator called a dark photon, A', which in turn decays into two LDM particles, $\chi\bar{\chi}$. A Feynman diagram of this process can be seen in Figure 1. As A' is massive, this corresponds to a drop in the energy of the electron, as well as a shift in the direction of the momentum, unlike regular bremsstrahlung [2].

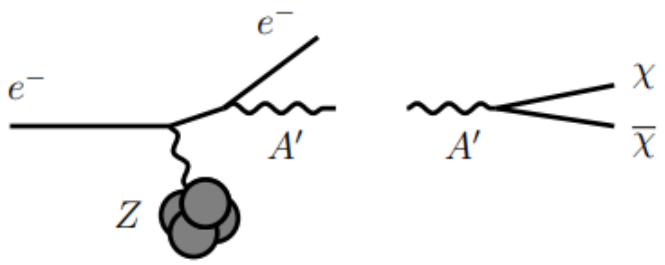


Figure 1: Feynman diagram of dark bremsstrahlung [2].

LDMX is a fixed target missing momentum experiment. An electron beam, provided by the LCLS-II accelerator at SLAC National Accelerator Laboratory in California, USA, is pointed at tungsten target. By measuring the momentum of the electrons before and after they pass the target, events with missing momentum as a result of the electron having undergone dark bremsstrahlung can be detected. These events are called signal events. There are two runs planned for the LDMX: phase 1 and 2. Phase 1 will make use of a 8-GeV beam, giving a sample of $4 \cdot 10^{14}$ electrons on target (EoT) over the course of a few years. Phase 2 will increase the intensity of the beam, and will give a sample of up to 10^{16} EoT. The projected

exclusion zones of both of these runs, as well as a preliminary missing energy analysis, compared to those from previous experiments, are shown in Figure 2 [4].

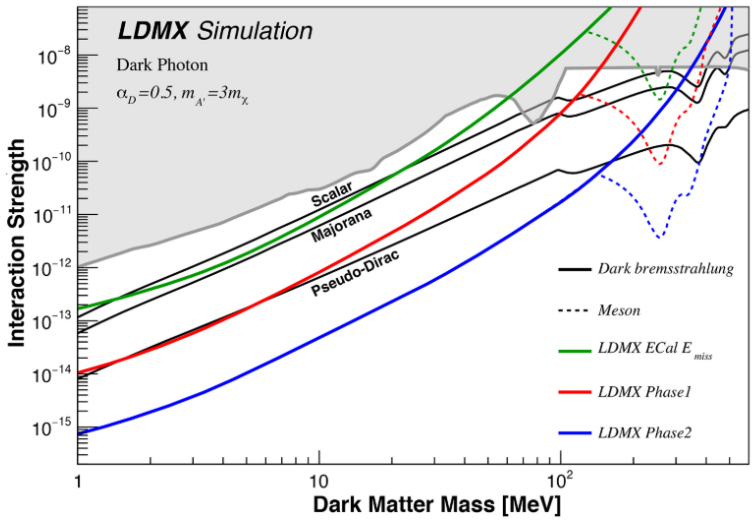


Figure 2: Projected exclusion prospects for the different runs of LDMX. Benchmark thermal relic targets for DM with different particle properties are shown as black lines, and constraints from previous experiments are represented by the gray area [4].

2.2 Setup

The goal of the LDMX detector is to get an as detailed as possible view of the kinematics of entering electrons before and after they pass the target. This is both such that signal events can be identified and studied, but more importantly such that non-signal events may be vetoed and removed. Thus more storage space is saved for events of interest, and less time is spent wading through useless data. Figure 3 contains a diagram of the LDMX detector with all components marked out. The setup contains two trackers, a tagger tracker upstream and a recoil tracker downstream of the target, which track the trajectory of the electron. A 1.5-T dipole magnet induces a strong magnetic field over the trackers, causing the electron trajectories to curl and allowing one to determine their momentum. Further downstream is the electromagnetic calorimeter (ECal), where the electron showers and is absorbed, such that its total energy can be measured. Surrounding the ECal is the hadronic calorimeter (HCal), which catches stray electrons as well as detects various other particles such as muons and hadrons for veto purposes. Surrounding the target, as well as upstream of the tagger tracker, are trigger scintillators, which alow for quick electron counting, used for triggering.

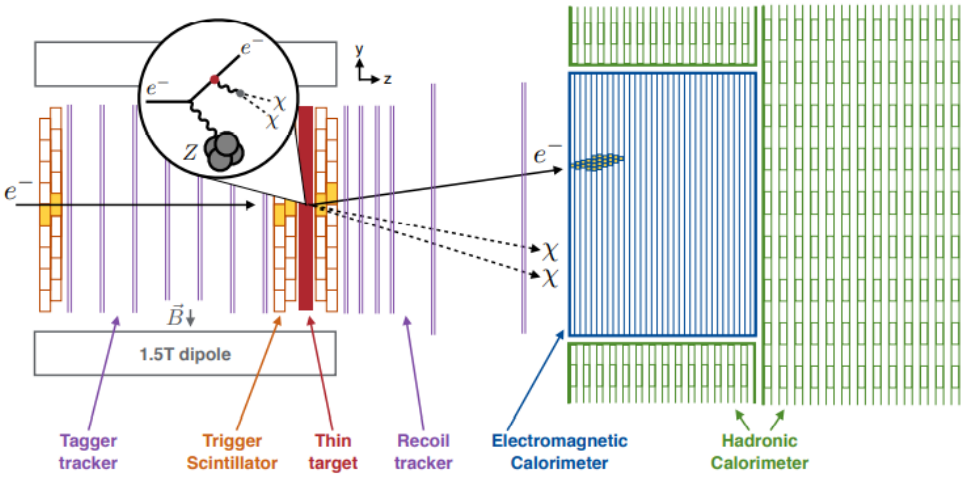


Figure 3: Diagram of the LDMX detector, during DM production [4].

An orthogonal coordinate system is defined for the detector, which is pointed out in Figure 3. When observing the detector from the perspective of an incoming electron, the y axis points upwards, x towards the left, and z inwards towards the detector. The z direction is often referred to as the depth, or downstream.

2.2.1 Electromagnetic Calorimeter

The ECal in LDMX is a silicon-tungsten (Si-W) sampling calorimeter [2]. This means that it is made up of alternating layers of silicon, which is the active material that measures energy deposited, and tungsten, which absorbs the brunt of all energy without measuring. The fact that most of the energy from an electron is not measured but disappears into the tungsten layers must be accounted for. This, however, allows for a more compact detector [5].

As an electron hits the ECal it creates a shower, as it repeatedly goes through bremsstrahlung, and the resulting photon forms an electron-positron pair, which undergo bremsstrahlung, and so on. This leads to a multitude of hits in the ECal which all correspond to the same electron [5].

2.2.2 Tracker

The recoil tracker is made up of six layers, four double sided and two one sided, of silicon microstrip modules, which measure the position of a passing electron with a resolution of $\sim 6\,\mu\text{m}$ in the horizontal plane and $\sim 60\,\mu\text{m}$ in the vertical plane. It is of larger importance

to have a high resolution in the horizontal plane as this is the plane in which the electron will curl under the influence of the magnetic field. It is from this curl that the momentum of the track is determined. The recoil tracker is positioned in the fringe field of the magnet. This is done to grant it higher sensitivity for particles up to two magnitudes softer than the incoming beam electrons [2].

2.2.3 Trigger

Once the experiment is run with an electron beam, there will be a lot of data output all at once, much more than can reasonably be stored. This is where triggering comes in. The goal of the trigger is to remove as much data as possible, while keeping as much signal data as possible. Triggering is very time sensitive, it has to be done quickly as to keep up with the rate of incoming data [5]. Often triggering is performed in several steps, although in LDMX it is done in just one. The trigger scintillators count the amount of incoming electrons, N_{e^-}, and an energy threshold is set based on the amount of expected energy. If the energy measured in the ECal falls above this threshold, no energy is missing and the data is discarded. The threshold is determined by multiplying the beam energy by N_{e^-}, and subtracting it by some amount of missing energy allowed. The allowed missing energy is governed by the amount of storage space available: the lower it is, the more data will be saved [2].

At higher energies, the energy resolution of the ECal degrades. This means that the trigger efficiency becomes worse with increasing amounts simultaneously incoming electrons. Figure 4 illustrates the change in trigger efficiency for different N_{e^-}. In the single-electron case, the slope is very steep, and a fairly low threshold will catch most signal events. But as N_{e^-} increases, the slope flattens out, and more signal events are missed. By using a matching algorithm, the goal is to be able to treat multi-electron events as a collection of single-electron events, thereby circumventing the issue of decreasing trigger efficiency.

2.3 Pile-up

Pile-up occurs when several particles hit the calorimeter at once, cluttering each others signal. This becomes especially problematic when a signal electron hits at the same time as a non-signal electron, called a pile-up electron. The hits corresponding to each individual electron must be separated from the other. This is done with a matching algorithm, which compares tracks from the tracker with showers in the calorimeter in order to determine which energy belongs to which electron. Although it has not been implemented yet, a goal is to use this to remove contributions to the measured energy from pile-up electrons while

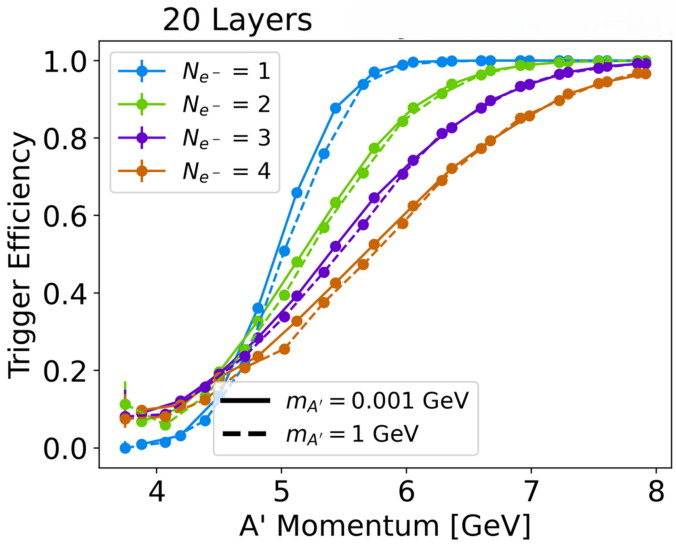


Figure 4: Current trigger efficiency at different values of missing momentum, for different amount of electrons entering the detector simultaneously (N_{e^-}). Figure courtesy of E. Berzin [6].

keeping data corresponding to signal electrons.

As mentioned in section 2.1.1, during phase 2 of LDMX there will be a significant increase in EoT compared to phase 1. In order to achieve this within a reasonable time span, the amount of electrons entering the detector at once must increase considerably, making pile-up removal even more important.

2.4 Software

LDMX uses its own software, `ldmx-sw`, in order to simulate events in the detector apparatus. The simulations can be used to analyze the detector output, and develop accurate reconstruction methods such that the events in the detector can be understood from only detector output once the experiment is run. `ldmx-sw` is also used for analysis of the resulting data, such as veto-processes which make sure only relevant data is saved [2].

`ldmx-sw` is a wrapper around the Geant4 toolkit, with substantially extended functionality. Geant4 is commonly used in high-energy physics for detector simulations. It uses Monte Carlo methods to simulate the passage of particles through matter, based on the four-momentum and particle type [7]. In `ldmx-sw` Geant4 handles non-signal event generation and the detector response to simulated particles [8].

Signal event generation in the target is handled by the MadGraph/MadEvent (MG/ME) [9] event generator. MG/ME allows for the input of Feynman rules for arbitrary physics processes, and calculates the tree-level amplitudes for those processes. For `ldmx-sw` a new-physics model is implemented of a dark photon, the mass of which can be adjusted. The DM production process generated in LDMX are $e^- W \rightarrow e^- W A'$, where W is the tungsten nucleus in the target. The decay of A' into a $\chi\bar{\chi}$ pair is not considered, as the DM does not interact further with matter in the detector. After the particles have been generated, their particle type and four-momentum, matched to incoming beam electron momentum via a procedure described in [10], is passed onto Geant4 [2].

Finally, `ldmx-sw` makes use of ROOT [11] for data storage and visualization. Data from simulations are stored as .root files, in a TTree structure. The TTree is made up of branches, which contain data of the same class. The branches in turn contain leaves, which contain the actual data associated with each class. For instance, a tree may contain the branch *EcalRecHits*, containing data about reconstructed hits in the ECal. This branch contains leaves such as the xyz-coordinates of the hits and their energy.

3 Method

3.1 Event reconstruction

When LDMX is run, the output will be in the form of digitized pulses, corresponding to energy deposited in the different channels of the detector. From this one does event reconstruction, translating the signals into information such as where the particles struck, which path they took, and which hits belong to one cluster. When event simulation is run through Geant4, the output is a bunch of simulation records, containing information such as the exact energy and path taken by each individual particle. We then translate this into data that mimics the output from the actual detector. This way the event reconstruction can be developed, and analysis methods which will work with the detector output.

As a particle hits a calorimeter, it does not deposit all its energy in one go, but creates a shower. Part of event reconstruction includes finding hits which belong to the same shower, and connecting them into clusters. Ideally, one cluster should correspond to one shower, and contain the entire energy of the incident electron. Then relevant information can be found from the clusters, such as their energy, the position of their centroid, and the

root mean square (RMS) of their distribution.

The RMS is a measurement of the average distance between a hit belonging to a cluster, and that cluster's centroid. It is found by taking the square root of the arithmetic mean value of the squared distance.

3.2 ParticleFlow

ParticleFlow (PF) [2] is a pre-existing matching algorithm which is the focus of this project, the purpose of which is to match tracks with clusters in the ECal and HCal. This way the path of an electron can be followed all the way through the detector. As the code was developed before proper tracking-algorithms were developed for `ldmx-sw`, it currently works using the "real", simulated values for track momentum, instead of those found by the software itself. This means that the current algorithm is not entirely accurate as to what will be used once the experiment is run, as it works "too well" as it is implemented right now. In the future it will be adjusted to use reconstructed tracking data.

PF produces objects of the class PFCandidates. Depending on the success of the matching, these objects contain some combination of the reconstructed tracker, ECal, and HCal hits. The type of information contained in the object is given by a variable called pid [12]. This is elaborated upon in Table 1. For instance, a wide angle electron scattering event where a track has been matched to a cluster in the HCal will have pid=1+4=5. For this project, events with pid 3 or 7 are of interest, as they correspond to a match between a track and ECal-cluster.

Table 1: Data contained in a PFCandidate object based on matched detector part, and corresponding pid.

Detector part	Data available	pid
Tracker	Momentum Intersect of extrapolated track and Ecal face	+1
Ecal	Cluster position Energy (calibrated and uncalibrated) Cluster shape	+2
Hcal	Cluster position Energy (calibrated and uncalibrated) Cluster shape	+4

The matching of tracks and ECal clusters works by comparing both distance and energy. The track is extrapolated to the z-coordinate of the cluster centroid, and the distance between the track and cluster is calculated in the xy-plane. If the distance falls within 2 RMS of the cluster centroid, and the energy of the cluster falls between 0.2 to 3 times the energy of the track, the cluster and track are considered matched. Figure 5 contains a schematic representation of this matching process. Matching between the ECal and HCal works similarly, although it is only based on distance and not energy. The spread of the ECal cluster is extrapolated to the z-coordinate of the HCal cluster. If the distance in the xy-plane is within 5 RMS of the HCal cluster centroid, the clusters are matched to each other [12].

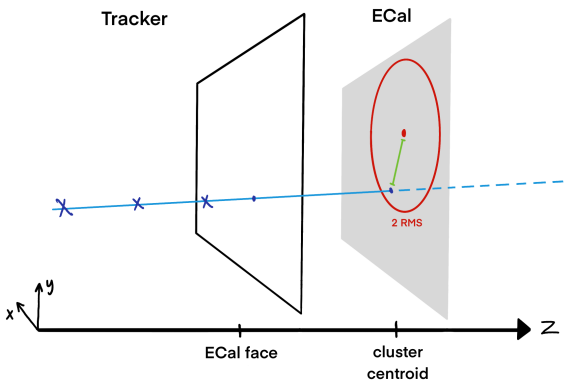


Figure 5: Schematic representation of PF matching between a track (blue) and an ECal cluster (red).

ParticleFlow outputs two different energy measurements for Ecal and Hcal clusters: *Energy* and *RawEnergy*. *RawEnergy* is the energy as it is measured by Geant4, and is closest to the "true" energy of the electron. When making real measurements with the calorimeter, the energy is expected to be reduced by a calorimeter response factor. This expected reduced energy is what is stored in *Energy*. When LDMX is run this response will be measured such that the real energy deposited, corresponding to *RawEnergy*, can be found. Therefore most analysis in this project is done using *RawEnergy*, as it more closely resembles the final energy that will be output by LDMX, and on which most analysis will be based.

3.3 Method

For this project, we are interested in the behavior of the PF algorithm when multiple electrons enter the detector at once, in order to see how well it identifies individual electrons for future pile-up removal. We begin by investigating the single-electron case, in order to establish the baseline performance of the algorithm. This is done by generating 10,000 events where one electron enters the detector, and then running PF on the output. The variables of interest are the energies and distances between track and cluster at which matching happens between the tracker and Ecal.

Once the algorithm's performance in the single-electron case has been established, one can move on to the more relevant case where two electrons, one signal electron which has undergone dark matter production in the target and one pile-up electron, enter the detector simultaneously. 10,000 events were supplied for this, at four different A' mass-points: 1, 10, 100, and 1000 MeV. These events were made by generating signal and pile-up samples individually, which were then combined and run through PF. It is of interest to consider several different masses as the kinematics differ between the different cases. A higher $m_{A'}$ means that the signal electron, which starts off at 8 GeV loses more energy during dark matter production, and enters the recoil tracker with considerably lower energy. This in turn means that it will have more curl compared to the pile-up electron, which at 8 GeV moves in an approximately straight line. As a result one would expect there to be a larger separation between the two electrons when $m_{A'}$ increases, which may improve the performance of the matching algorithm.

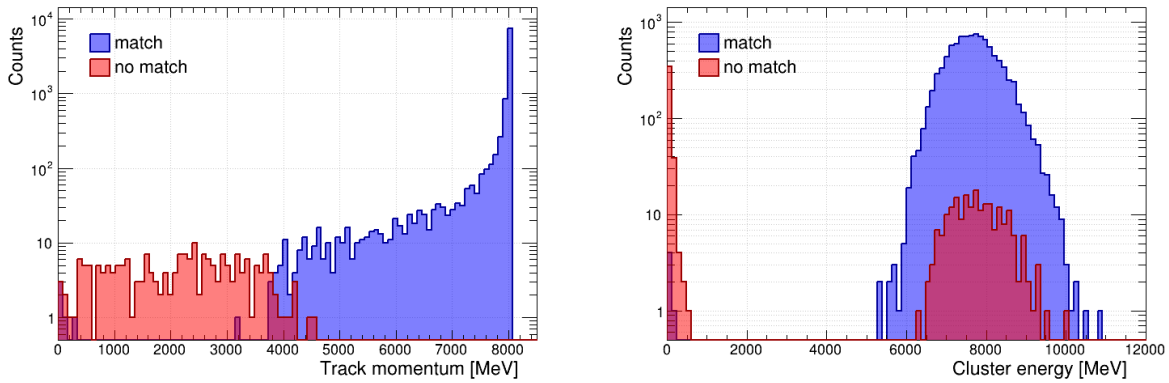
The purpose of this is to eventually utilize PF for pile-up removal. Initially this will be done in the analysis stage, in order to develop a working algorithm where the same events pass the veto algorithms regardless if pile-up is added or not. After that the algorithm will be implemented as part of the trigger.

Code written for this project can be found in [13].

4 Results and Discussion

4.1 Single-electron case

Out of the 10,000 generated single-electron events, the algorithm fails to match a track and ECal cluster in 209 events. Figure 6 compares the energy distribution of both track and ECal cluster energy of events where the matching between ECal and tracker is successful to events where it is not. In events in which no match is made, the energy of the highest energy ECal cluster is drawn in place of the matched cluster, operating on the assumption that the cluster left behind by the electron must have much higher energy than any other cluster. Most notably in Figure 6 (a) there is a sudden drop in matched tracks around 4 GeV. As seen in Figure 6 (b), most matched clusters are collected in a peak around 8 GeV. Quite a few unmatched clusters also have energies around 8 GeV, but compared to the matched case there are considerably more low energy clusters.



(a) Track momentum.
Binning: 80 MeV/channel

(b) ECal cluster energy.
Binning: 120 Mev/channel.

Figure 6: Energy distribution of the tracks and ECal clusters in the single-electron case, for events with successful and unsuccessful matching between the tracker and ECal. The y-axis is shown in log-scale.

Figure 7 shows the distance between the ECal cluster and point at which the extrapolated track intersects the Ecal face, in the xy-plane. Both the matched and unmatched histograms contain peaks around 2.5 mm, although more tracks fall outside the peak in the unmatched case.

From these results it seems the factor most significant in determining the success of the matching is the track momentum, as Figure 6 (a) has the clearest distinction between

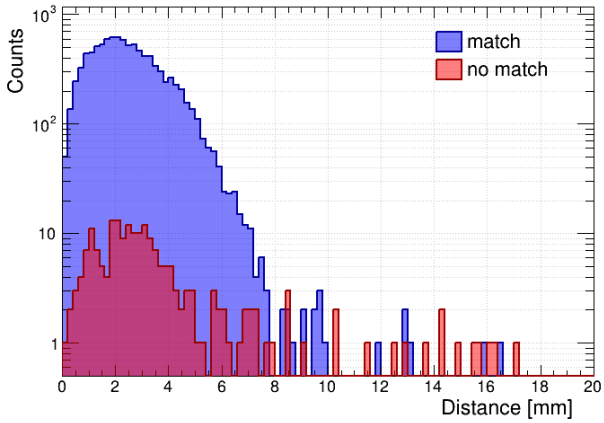


Figure 7: Distance in the xy-plane between matched and unmatched track-cluster pairs. In the case where matching is unsuccessful, the distance used is that between the track (if it exists) and the highest energy cluster. Binning: 0.1 mm/channel

matched and unmatched events. Comparing the track- and cluster energy distributions, and considering the unmatched distances do not differ much from the matched distances, it seems probable that it is in the energy comparasion stage that PF fails. In Figure 6 (b), there are several unmatched events where the highest energy cluster energy is well below 2 GeV. Presumably these are events where the clustering algorithm fails to capture the full shower in one cluster. However, there are also many unmatched clusters with energies around 8 GeV, much like the matched events and the energy of the electrons. Probably these are events where the PF algorithms "misses". All in all one may conclude that the PF algorithm works very well for the single electron case.

4.2 Two-electron case

Table 2: Amount of events out of 10,000 where 0, 1, 2, or 3 matches have been made, at different values of $m_{A'}$.

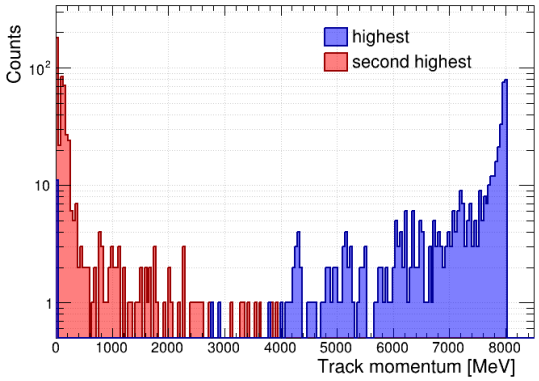
pid 3/7 count / A' mass	0	1	2	3
1 MeV	1685	7817	498	0
10 MeV	1250	7473	1274	0
100 MeV	1300	7398	1301	0
1000 MeV	929	6507	2561	3

Table 2 shows in how many events, out of the 10,000 generated, 1, 2, or 3 matches were

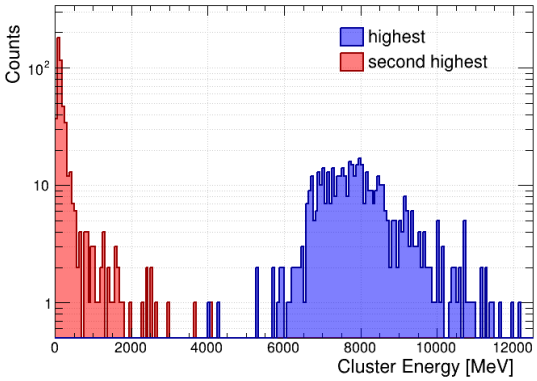
made between a track and an ECal cluster, for different values of the A' mass. As expected, matching is more successful at higher values of $m_{A'}$.

4.2.1 Energy distributions

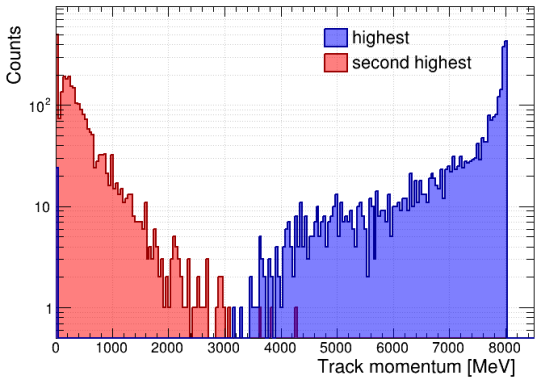
Considering only events in the third column of Table 2, where presumably both the signal electron and pile-up electron have been matched, Figure 8 shows the energy distributions of both the matched tracks and ECal clusters. The energies are differentiated by highest and lowest energy, sorted by track energy. Once again the assumption is made that this is interesting because (in general) the lower energy hit corresponds to the signal electron, and the higher energy hit corresponds to the pile-up electron, which has not lost any energy to DM production.



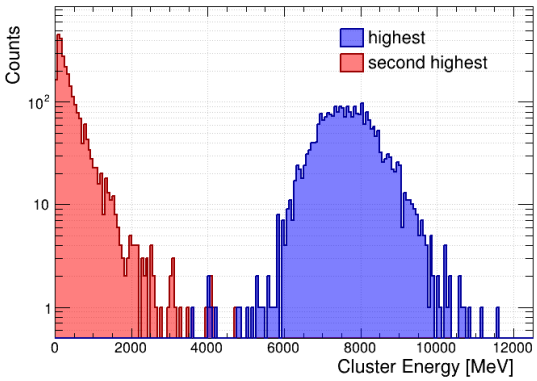
(a) 1 MeV - track momentum
Binning: 42.5 MeV/channel.



(b) 1 MeV - Ecal energy
Binning: 62.5 MeV/channel



(c) 1000 MeV - track momentum
Binning: 42.5 MeV/channel.



(d) 1000 MeV - Ecal energy
Binning: 62.5 MeV/channel.

Figure 8: Energy distribution of the highest and second highest clusters and tracks in the two electron case, for the instances when $m_{A'}$ is 1 MeV and 1000 MeV.

There is a clear distinction between the two cases, with little overlap. In the ECal, the highest energy clusters form a peak around 8 GeV, which agrees with the assumption that they are pile-up electrons, which haven't lost much energy. The difference between signal tracks and pile-up tracks becomes clearer at higher $m_{A'}$, which is to be expected.

As mentioned in section 2.2.3, a goal of using the matching algorithm is being able to treat multiple simultaneously incoming electrons individually, in order to set a low energy threshold without losing out on trigger efficiency. In these plots there is a clear separation between signal and pile-up electrons, supporting the idea that the algorithm could be utilized in this way. An effective threshold would be around 4000 MeV in cluster energy: all signal electrons fall under this value, while most pile-up electrons fall above it.

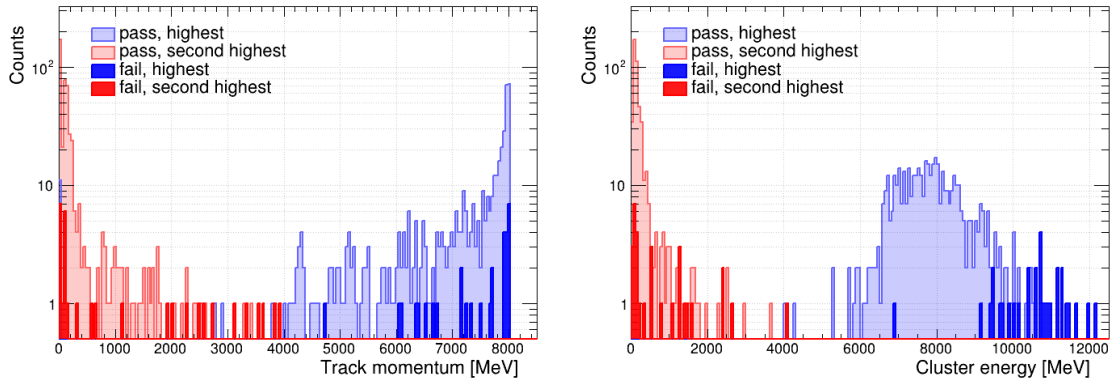
Figure 9 shows energy distributions similar to those in Figure 8, for events which pass and those which do not pass the current triggering process. As all events are signal events, were the triggering to work optimally all events should pass. As it is, 288 events do not pass in the 1000 MeV case, and 1608 do not pass in the 1 MeV case. The difference between the mass cases is expected, with a lower $m_{A'}$ the signal electron more closely resembles the unaltered pile-up electrons, and are less likely to be recognized, as the triggering works based on energy. Most events which do not pass the trigger have pile-up clusters in the higher energy range which is especially obvious in Figure 9 (b) and (d).

So far the triggering works on a global scale for each event. If it is identified as a signal event, all data is saved, including the pile-up, and if the event is mistakenly identified as non-signal because of a high-energy pile-up electron, the signal data is discarded. If the trigger could be adjusted to work on a matched cluster basis, much more pile-up data could be removed, while keeping signal events that may otherwise have been discarded.

4.2.2 When does ParticleFlow (not) work?

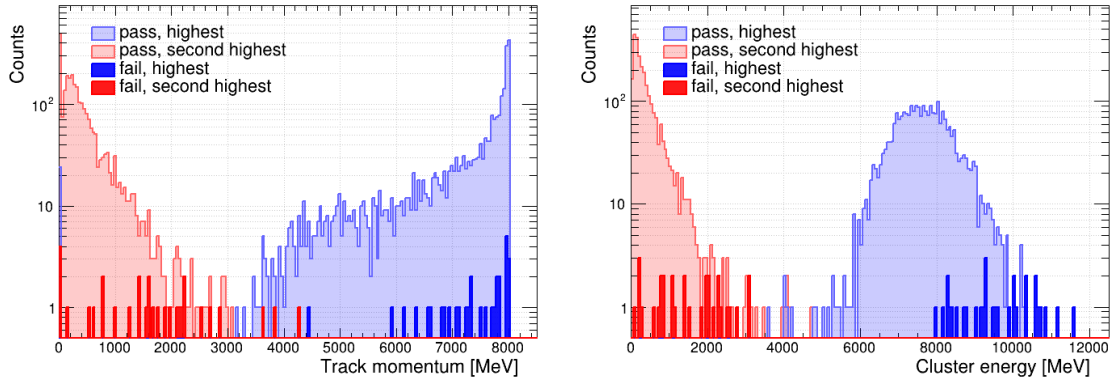
So far it seems like PF, when it works optimally, is fairly effective for identifying pile-up. The largest issue is the rate at which matches are made successfully. Looking back at Table 2, two matches only happen in about one fourth of events when $m_{A'} = 1000$ MeV, and about 5 % of events when $m_{A'} = 1$ MeV. It is of interest to find out the conditions under which PF fails and succeeds, such that necessary improvements can be made.

Figure 10 shows distances between tracks and clusters in the different cases when two, one, or no matches have been made. When one or no match has been made, the



(a) 1 MeV - track energy
Binning: 42.5 MeV/channel.

(b) 1 MeV - Ecal raw energy
Binning: 62.5 MeV/channel.

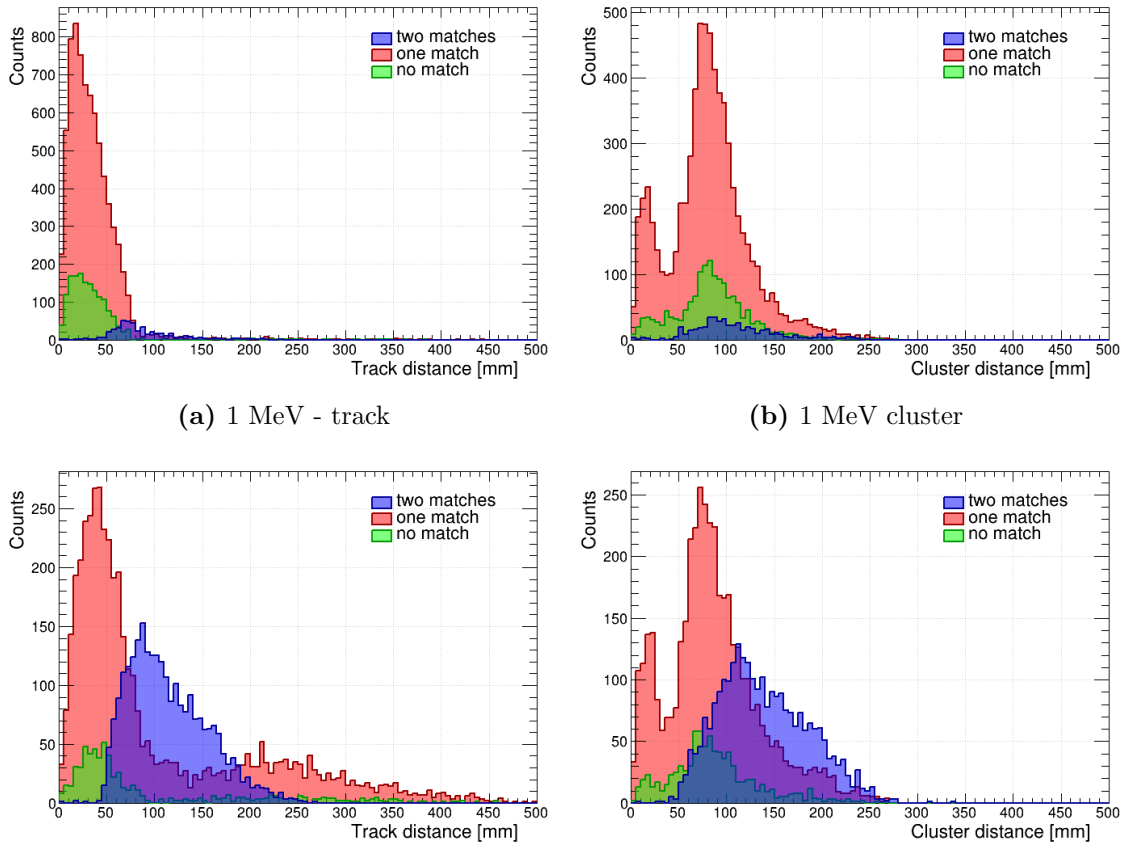


(c) 1000 MeV - track Energy
Binning: 42.5 MeV/channel.

(d) 1000 MeV - Ecal raw energy
Binning: 62.5 MeV/channel.

Figure 9: Energy distribution of the highest and second highest clusters and tracks in the two electron events which pass the trigger, and those who do not, for the instances when the mass of A' is 1 MeV and 1000 MeV.

position of the highest energy cluster is used. The distance between clusters is calculated in the xy-plane, since that is how the matching is performed. While small trends are visible in Figure 10 (a) and (b), they are much clearer in (c) and (d), thanks to the larger number of matched events. When tracks and clusters are too close together only one match is made. At distances of about 50 mm the matching algorithm starts working for two electrons, and it is at its most efficient around 100 mm. It seems that a major issue is the ability to differentiate between tracks and clusters in close proximity. It is likely that there are issues with the clustering algorithm when two showers occur close together. This does not bode well for the case of a low $m_{A'}$. Additionally, in Figure 10 (c) there are many events with a large distance between tracks where only one match has been made. It is likely that one of the electrons has strayed enough and entered the side HCal instead of the ECal.



(a) 1 MeV - track

(b) 1 MeV cluster

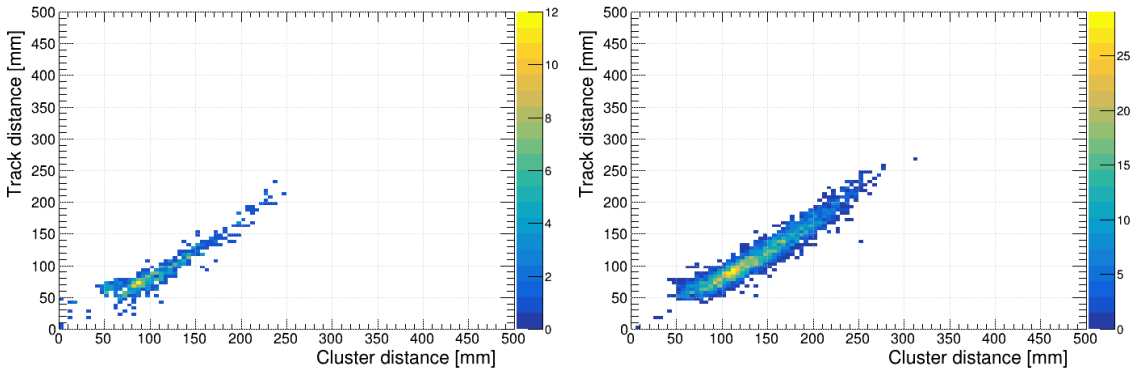
(c) 1000 MeV - track

(d) 1000 MeV - cluster

Figure 10: Distance between the two tracks and clusters when the algorithm makes 2 matches, 1 match, or no matches between tracker and ECal. If one or no match is made, use highest energy cluster. Binning: 0.5 mm/channel.

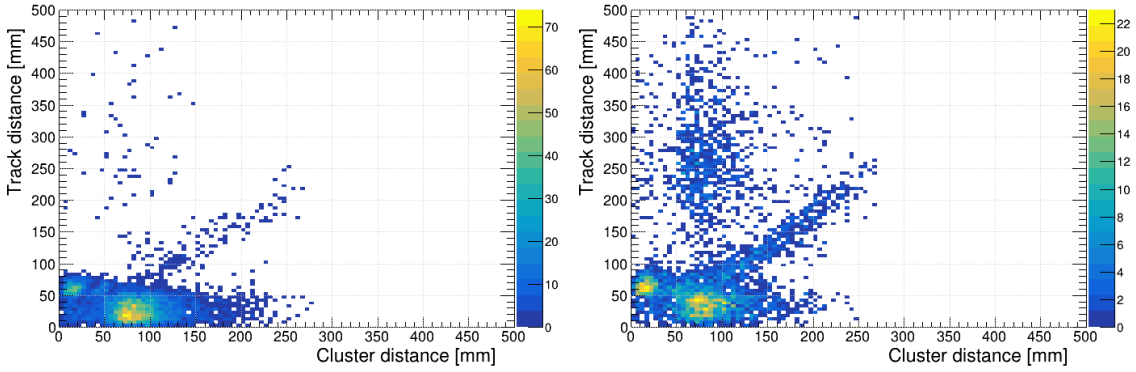
The values from Figure 10 were used to make the histograms in Figure 11, where the correlation between cluster and track distance are shown. In Figure 11 (a) and (b), when the algorithm finds two matches, there is a very clear linear correlation. This correlation can also be seen in Figure 11 (c) and (d), as well as very vaguely in Figure 11 (f). Presumably these are good distance metric matches that have been missed by the algorithm.

In both Figure 10 (b) and (d), the one- and no-match cases contain an additional small peak around the 20 mm mark. This peak is not present in the case where both clusters are matched. This implies that the peak arises from the unmatched cluster in some way. In Figure 11 (c) and (d), these peaks are also visible, and seem to correspond to a small track, which is larger than the cluster distance. While no concrete answer can be given as to why this is, a plausible explanation is an issue with the clustering algorithm. If the



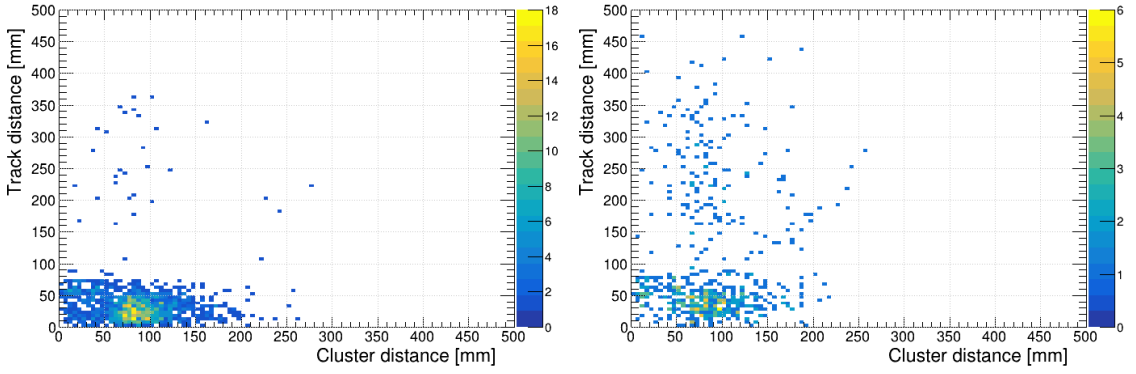
(a) 1 MeV - 2 matches

(b) 1000 MeV - 2 matches



(c) 1 MeV - 1 match

(d) 1000 MeV - 1 match



(e) 1 MeV - no match

(f) 1000 MeV - no match

Figure 11: Distance between tracks with respect to distance between ECal clusters. Binning: 0.5 mm/channel.

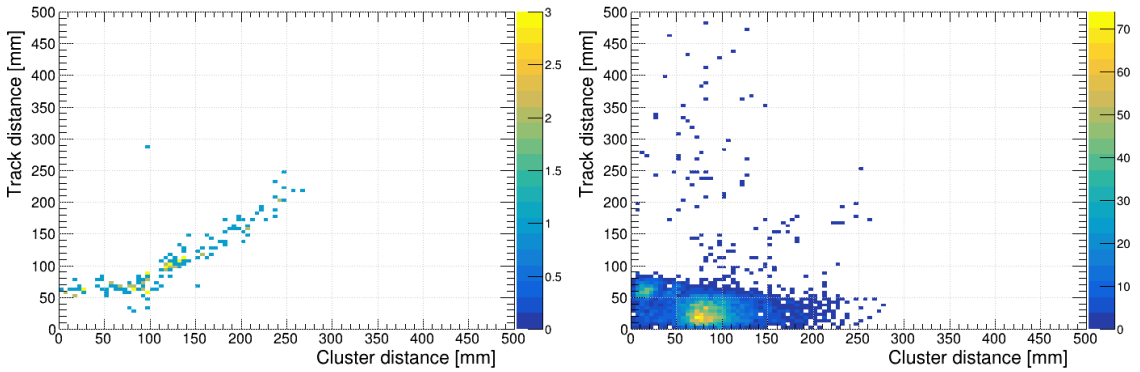
highest-energy cluster is split in two, or two high energy clusters are in close proximity, this may lead to a misinterpretation by the matching algorithm.

In order to use PF to remove pile-up, the matching does not necessarily have to work on the signal electron. Only the pile-up electron has to be matched in order for it to be removed. This means that PF could be utilized for pile-up removal in the cases where not all incoming electrons are matched. A possible solution is implementing the new trigger algorithm for matched electrons, and utilizing the old algorithm for the leftover unmatched electrons.

It is of interest then to find out how often the matched electron is a pile-up or signal electron. In Figure 12, Figure 11 (c) and (d) have been split up into two cases: when the successful matching involves the low energy track, and when it involves the high energy track. In the 1000-MeV case, 3 940 out of 4 294 events the matched track is that with a higher energy, meaning that the pile-up electron is indeed most often matched. In the 1-MeV case, this number increases to 6 831 out of 6 987 events. The fact that, presumably, the pile-up electron is matched in a majority cases is promising, and agrees with the results from the single electron case, which show that higher momentum tracks are more likely to be matched.

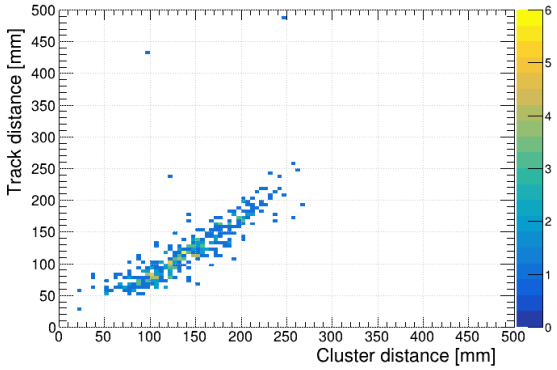
Figure 12 (a) and (c) both resemble the linear correlation in Figure 11 (a) and (b). This seems to suggests that these matches are mostly correct. Figure 12 (b) and (d) on the other hand are less correlated. This could be the assumption that the highest unmatched cluster is the second electron, in this case the pile-up electron, failing. It is reasonable that this happens more often when searching for the signal cluster, which has a considerably lower energy. A likely error is the showers being divided into several clusters, which then do not accurately represent the energy of the full shower. This would mean that searching for the highest energy cluster is likely to produce a cluster in the position of the pile-up electron, but searching for the second highest cluster is about as likely to produce a cluster in the position of the pile-up electron as the signal electron, which would be the desired result. This explanation accounts for the low cluster distances, as clusters corresponding to the same shower will be spatially close together. It is worth noting that many of the cases where the track distance is high, which earlier were theorized to correspond to cases where one electron misses the ECal, are uncorrelated, which agrees with this theory.

The fact that Figure 12 (b) and (d) are largely uncorrelated does not have to mean that the matches are incorrect, only the assumptions made about the unmatched electron, which is not relevant for the prospect of pile-up removal.

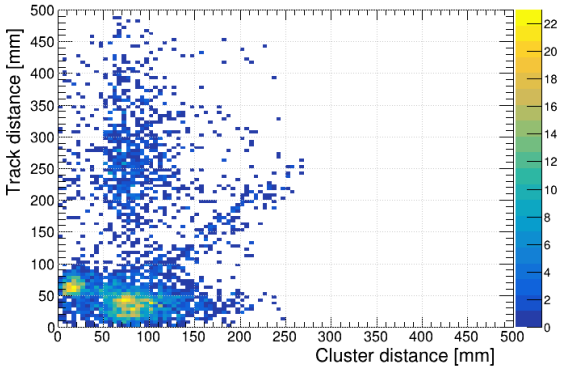


(a) 1 MeV - Low-energy match

(b) 1 MeV - High-energy match



(c) 1000 MeV - Low-energy match.



(d) 1000 MeV - High-energy match

Figure 12: The single match case from Figure 11, separated into whether the matched track/cluster has higher or lower energy than the unmatched. Binning: 0.5 mm/channel.

5 Conclusion

In the most basic case, when a single electron enters the detector at a time, the PF matching algorithm works very well. It only fails to match a track and an ECal cluster in about 2% of cases. In these cases, it seems to be the track energy which is the deciding factor, and once it reaches below a certain point a match cannot be made.

When two electrons enter the detector at once the amount of successful matching go down considerably. Here the mass of the dark photon plays in; more successful matches are made in the cases where the mass is higher, as the spatial separation between signal and pile-up electrons increases. Spatial separation seems to factor heavily into the matching success. When electrons hit the ECal close together it becomes more common for one or no match to be made. This seems to be both an issue with PF itself, as well as the clustering

algorithm, which struggles to separate showers in close proximity. The clustering algorithm also sometimes struggles with dividing showers into several clusters, which causes further issues in the matching stage.

In the cases where both electrons are matched, the algorithm yields promising results. Energies corresponding to signal and pile-up electrons are clearly separated, and with the implementation of a threshold almost all pile-up data could be removed while keeping signal data. This would be an improvement compared to the current trigger algorithm. There is also potential for pile-up removal in cases where not all incoming electrons are matched. By discarding data corresponding to any matched electron falling above a determined threshold, the current trigger algorithm could be applied to the remaining data, thereby improving its efficiency. It is much more common for the pile-up electron to be matched than the signal electron, which benefits this suggested method.

5.1 Outlook

This investigation has not been exhaustive, and there aspects of PF which could be investigated further. Considering cluster RMS when investigating distances between tracks and clusters may yield new insights. When investigating the different cases where two, one, or no matches were made, only track and cluster distance was taken into account. One could also look into track and cluster energy to get a more complete picture.

The performance of PF, especially in cases where electrons hit the detector in close proximity, should be improved in order to reach a higher trigger efficiency. This could be done both by adjusting the code of PF itself, but also the clustering algorithm, in order to prevent scenarios where showers become divided into several clusters, or clusters meld together.

Next PF should be used to perform pile-up removal, in order to evaluate how this improves the trigger efficiency. An energy threshold can be implemented, and any track-cluster pair that falls above this threshold is removed from the data before putting it through the veto algorithm. This should be done even for the cases where only one electron is matched, to see if a combined trigger method is viable.

References

[1] Gianfranco Bertone and Dan Hooper. "History of dark matter". In: *Rev. Mod. Phys.* 90 (4 Oct. 2018), p. 045002. DOI: 10.1103/RevModPhys.90.045002. URL: https://link.aps.org/doi/10.1103/RevModPhys.90.045002.

[2] Torsten Åkesson, Asher Berlin, Nikita Blinov, et al. *Light Dark Matter eXperiment (LDMX)*. 2018. arXiv: 1808.05219 [hep-ex]. URL: https://arxiv.org/abs/1808.05219.

[3] Yong Du, Fei Huang, Hao-Lin Li, et al. "Revisiting dark matter freeze-in and freeze-out through phase-space distribution". In: *Journal of Cosmology and Astroparticle Physics* 2022.04 (Apr. 2022), p. 012. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2022/04/012. URL: http://dx.doi.org/10.1088/1475-7516/2022/04/012.

[4] Torsten Åkesson, Nikita Blinov, Lukas Brand-Baugher, et al. *Current Status and Future Prospects for the Light Dark Matter eXperiment*. 2023. arXiv: 2203.08192 [hep-ex]. URL: https://arxiv.org/abs/2203.08192.

[5] G. Barr, R. Devenish, R. Walczak, et al. *Particle Physics in the LHC Era*. Oxford Master Series in Physics. OUP Oxford, 2016. ISBN: 9780191065453. URL: https://books.google.se/books?id=LrgoCwAAQBAJ.

[6] Lene Kristian Bryngemark. Private communication. 2025.

[7] Tullio Basaglia, Zane W. Bell, Daniele D'Agostino, et al. "Geant4: A game changer in high energy physics and related applicative fields". In: *Future Generation Computer Systems* 159 (2024), pp. 411–422. ISSN: 0167-739X. DOI: https://doi.org/10.1016/j.future.2024.05.042. URL: https://www.sciencedirect.com/science/article/pii/S0167739X24002747.

[8] Torsten Åkesson, Cameron Bravo, Liam Brennan, et al. *Photon-rejection Power of the Light Dark Matter eXperiment in an 8 GeV Beam*. 2023. arXiv: 2308.15173 [hep-ex]. URL: https://arxiv.org/abs/2308.15173.

[9] J. Alwall, R. Frederix, S. Frixione, et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. URL: http://dx.doi.org/10.1007/JHEP07(2014)079.

[10] Tom Eichlersmith, Jeremiah Mans, Omar Moreno, et al. "Simulation of dark bremsstrahlung in Geant4". In: *Computer Physics Communications* 287 (June 2023), p. 108690. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2023.108690. URL: http://dx.doi.org/10.1016/j.cpc.2023.108690.

[11] Rene Brun, Fons Rademakers, Philippe Canal, et al. *root-project/root: v6.18/02*. Version v6-18-02. Aug. 2019. DOI: 10.5281/zenodo.3895860. URL: https://doi.org/10.5281/zenodo.3895860.

[12] LDMX developers. *LDMX software*. Accessed: 2025-01-29. URL: https://github.com/LDMX-Software/ldmx-sw/.

[13] Merit Aerts. *Aerts_thesis_25*. URL: https://github.com/merit-aer/aerts_thesis25.