Identifying Light Dark Matter Production Events Using Machine Learning For LDMX

Aleczander Paul || Professor Valentina Dutta
Carnegie Mellon University

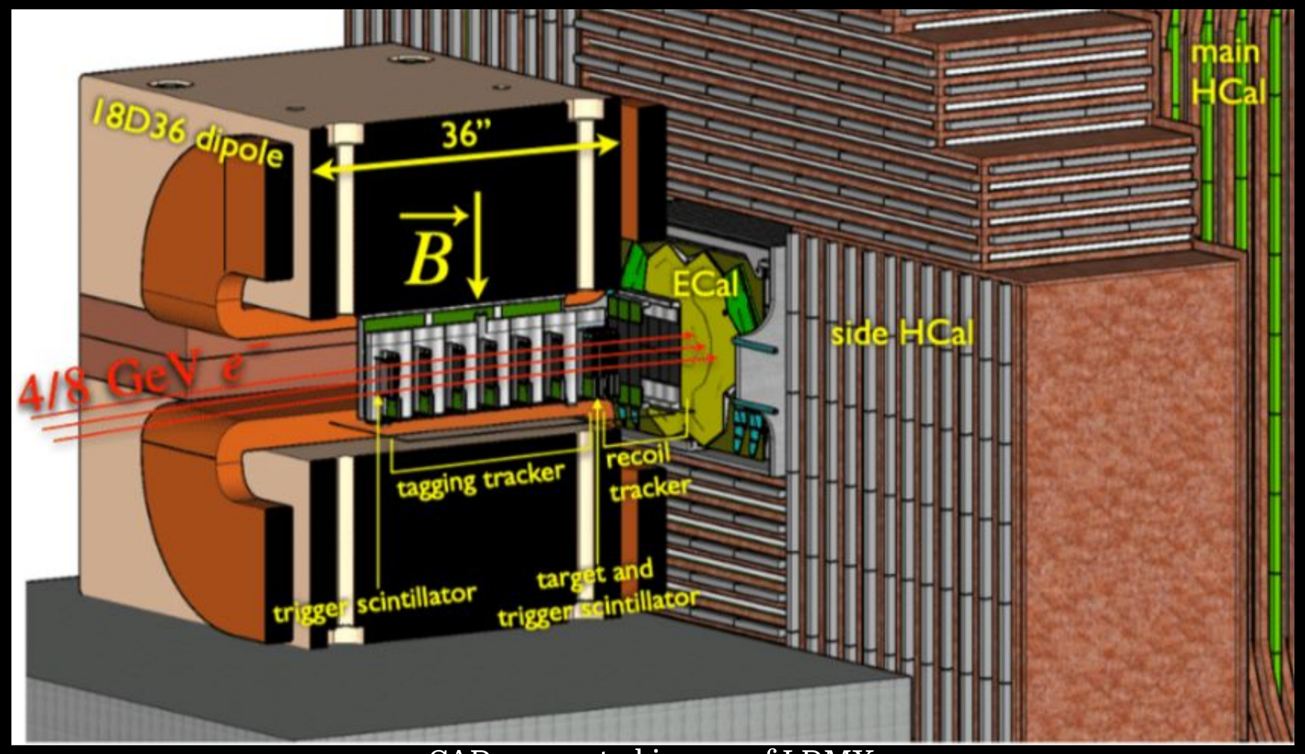
Introduction

What is dark matter (DM)?
A theoretically predicted type of matter that has the following properties:
- Interacts gravitationally
- No electromagnetic interaction
- Very little coupling to particles in The Standard Model (SM)

Evidence for the existence of dark matter:
- Galaxy rotation curves
- Gravitational lensing
- CMB power spectrum

It makes up ~85% of the matter in our universe! However, it is notoriously difficult to detect.

The Light Dark Matter Experiment (LDMX):

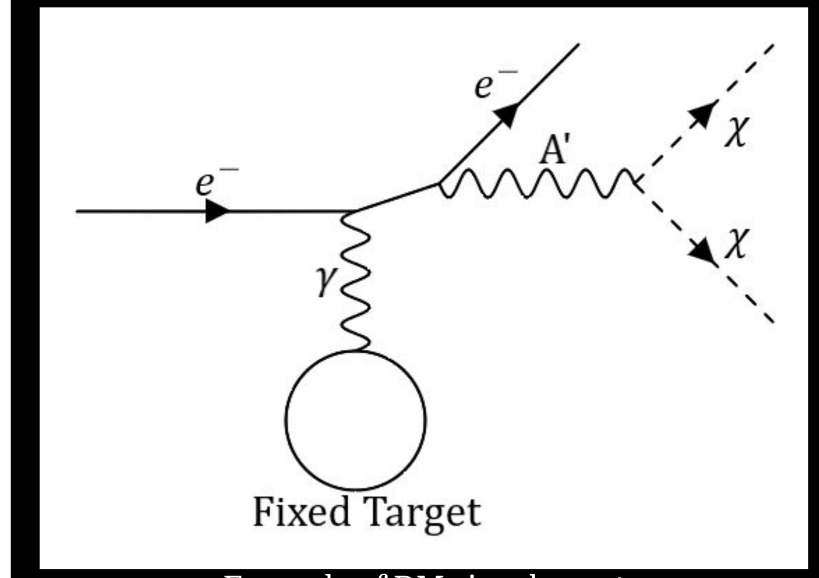


CAD generated image of LDMX

LDMX is a fixed-target experiment that utilizes a high energy electron beam (4 GeV, upgraded to 8 GeV in the future).

Goal is to generate 1e16 events total over a few years.

LMDX Features:
- In front of target:
  - Momentum trackers
  - Strong B-field
- Behind target:
  - Momentum trackers
  - Electromagnetic calorimeter (ECal)
  - Hadron calorimeter (HCal)



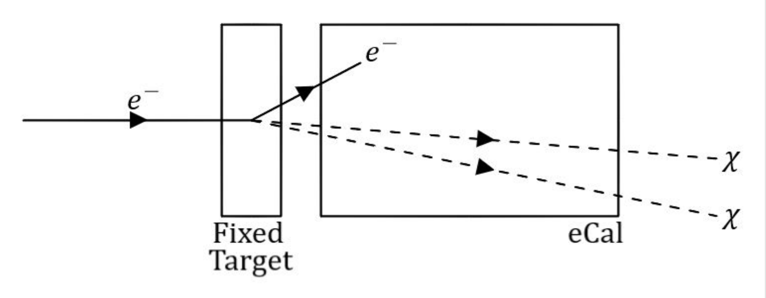
Example of DM signal event

Timeline of a DM signal event:
1. Electron approaches target nuclei
2. Through portal interaction, electron recoils and creates a dark photon (A')
3. Dark photon may produce dark matter particles (χ), A' or χ leave no visible trace

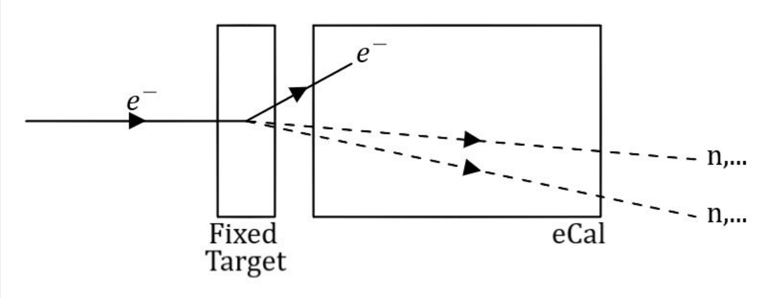
Other possible events (not DM signal):
- No electron-nuclei interaction
- Interaction creates ordinary photon
  - Photon may produce SM hadrons, called the photo-nuclear background

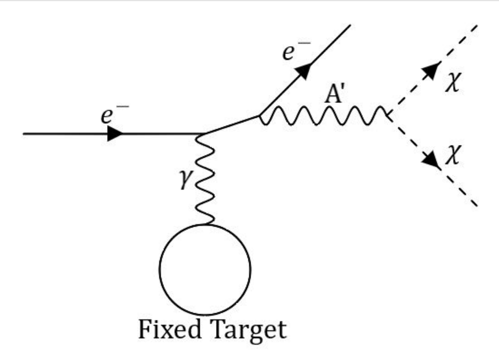
Methodology

Identify these ✓



Reject these ✖





DM production event

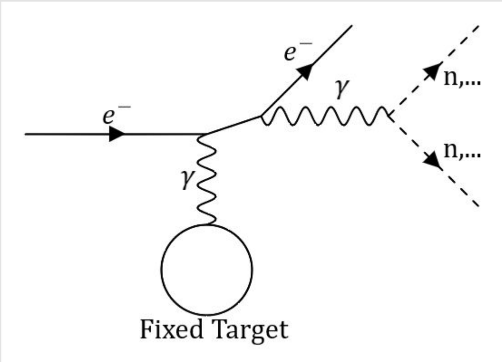
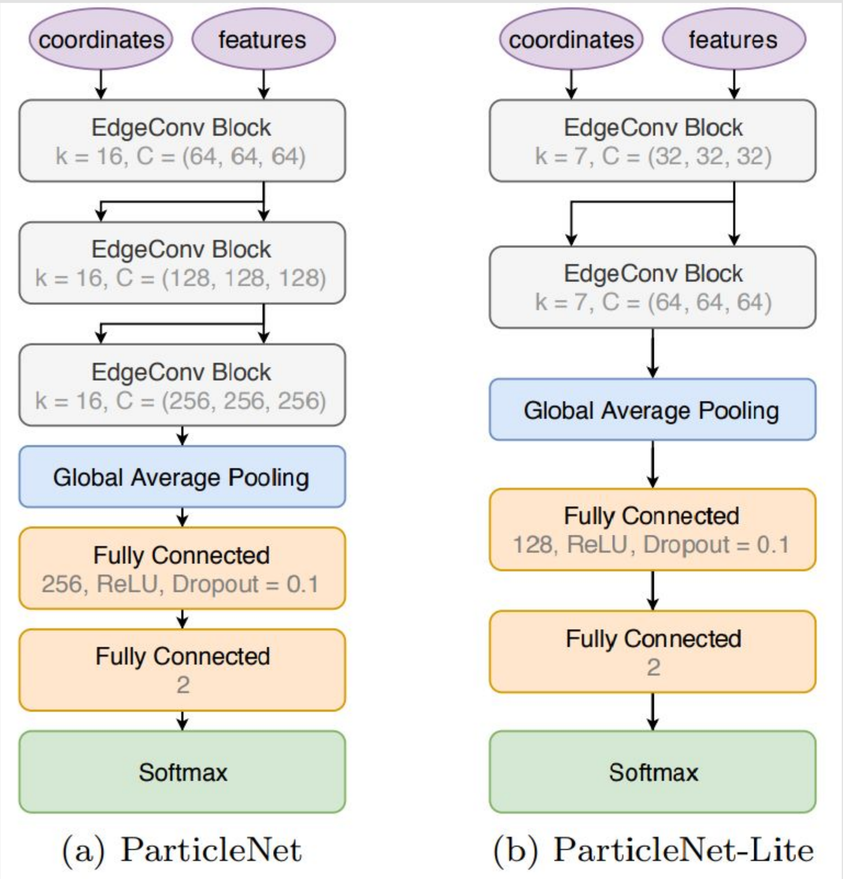


Photo-nuclear background event

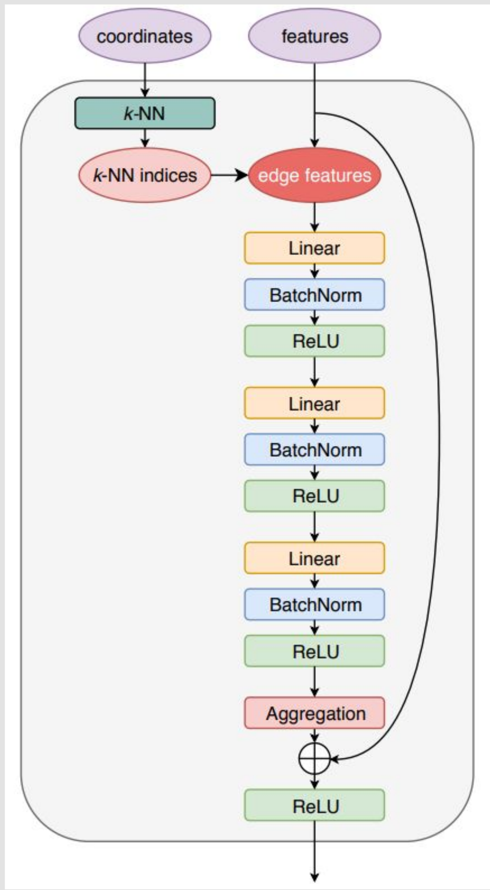
My work on LDMX is centered on training an AI model that can identify DM events in the ECal and veto background events in cases where there are 2 incoming electrons at a time.

This is done with simulated LDMX data for DM events of a specified mass or background photo-nuclear events.

LDMX has two main backgrounds:
- E/M shower (high rate)
  - Easy to identify without AI, do not need to train model to recognize these
- Photo-nuclear (rare)
  - Hard to distinguish from DM event



ParticleNet and ParticleNet-Lite architecture



EdgeConv Block architecture

Neural network model trained on the LDMX data is called ParticleNet, with the architecture shown above.

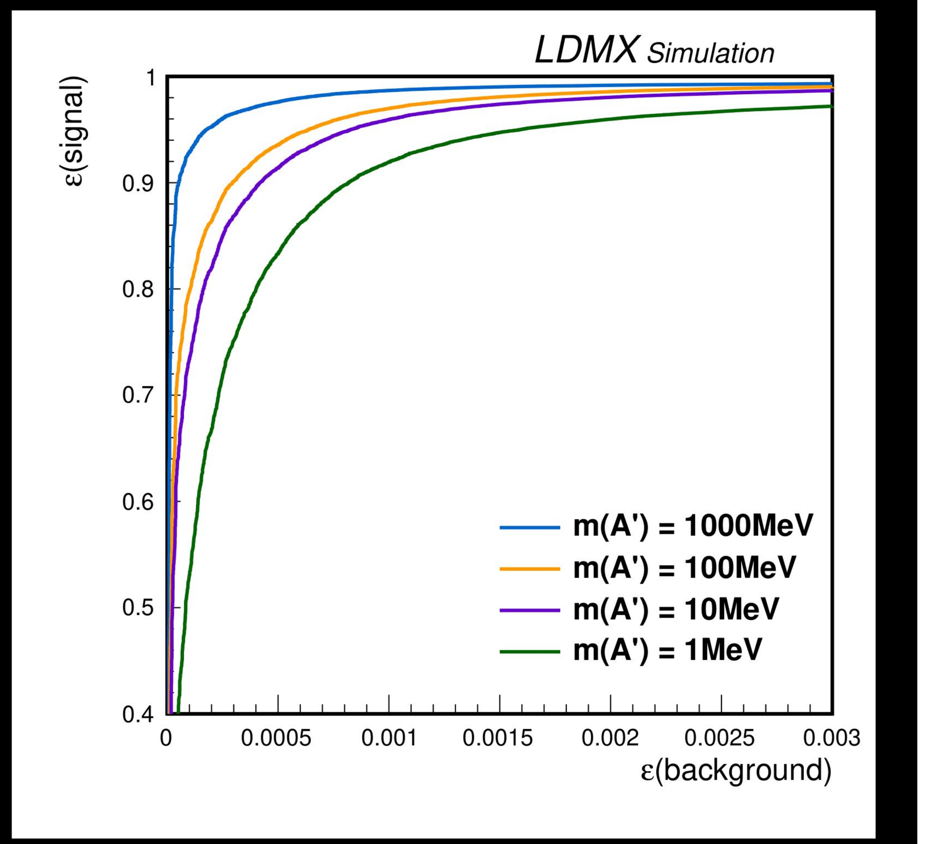
ParticleNet is a model originally made for jet events in particle physics

Trains on "particle cloud" data:
- Data represented as points in space
- Points are unordered, can be distributed any way
- Uses patterns in the cloud to classify events

Results

A Receiver Operating Characteristic (ROC) curve is a plot of signal vs background efficiency for different thresholds on the discriminator output

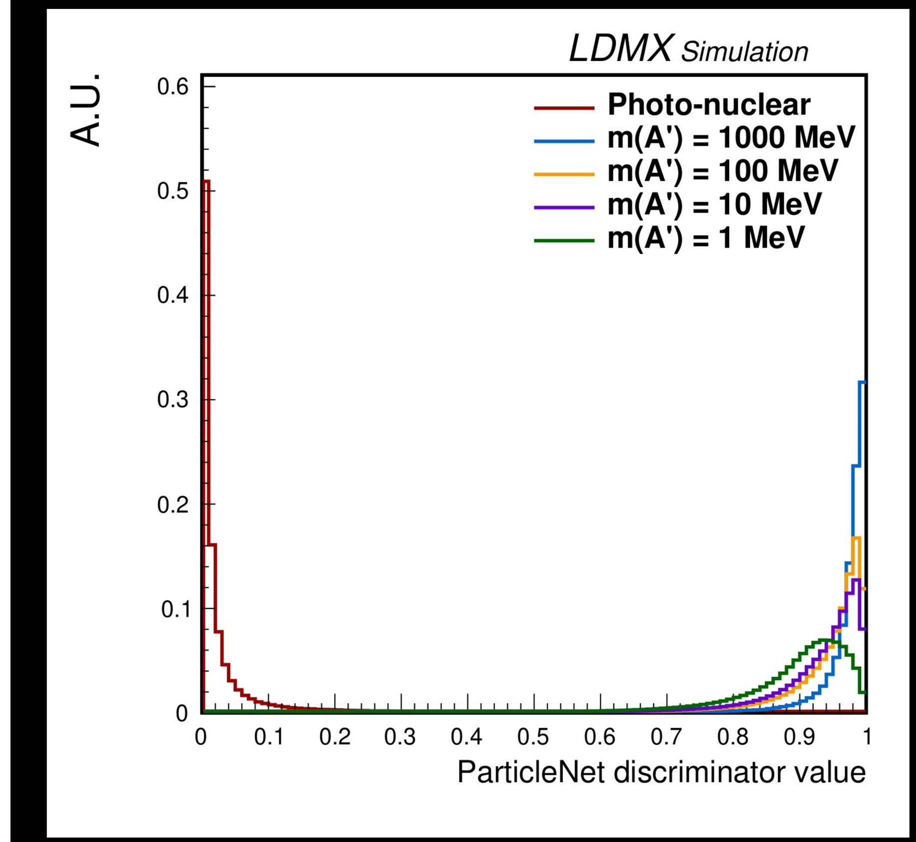
Neural network trained on 300,000 background events and 300,000 DM events. Each DM mass point had 75,000 events.



ROC curve of different A' mass points

		Background Efficiency	
		1E-03	1E-04
Mass Points	1 MeV	92%	53%
	10 MeV	96%	73%
	100 MeV	97%	80%
	1000 MeV	99%	92%

Table of signal efficiencies per background efficiency for each mass point



BDT discriminator value histogram

Histogram showing discriminator values for each event

1 means the network identifies the event as more signal-like (right), and 0 is more background-like (left)

As expected, most signal events are closer to 1, and most background events are closer to 0

As mass of A' increases, network is more confident in identifying the events as more signal-like

References

[1]T. Åkesson et al. Current Status and Future Prospects for the Light Dark Matter eXperiment, arXiv (2022).
[2]H. Qu and L. Gouskos. ParticleNet: Jet Tagging via Particle Clouds, Phys. Rev. D 101, (2020).