Energy Reconstruction with VAEs in SuperCDMS

Ashley Li
ashleyli@phas.ubc.ca
University of British Columbia
SuperCDMS Collaboration
August 5, 2020

(image courtesy of Scott Oser)

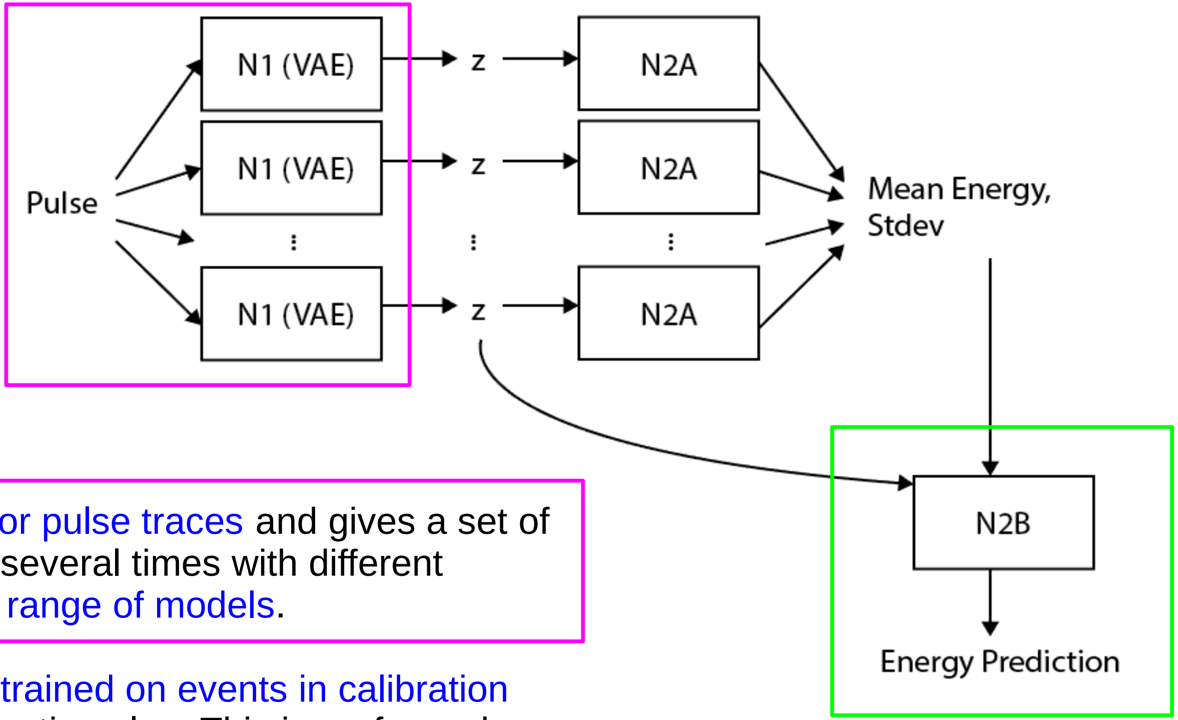
Introduction

VAEs (Variational Autoencoders) are generative neural networks that can perform feature extraction via dimensionality reduction on a dataset.

They are “generative” as they attempt to model how data is generated, and can in turn generate new samples.

Our current energy reconstruction algorithm at SuperCDMS assumes linearity between fit parameters and physical values, which may not apply across the whole spectrum, especially in low energy areas of interest. It is also inflexible with variable pulse shapes.

Structure



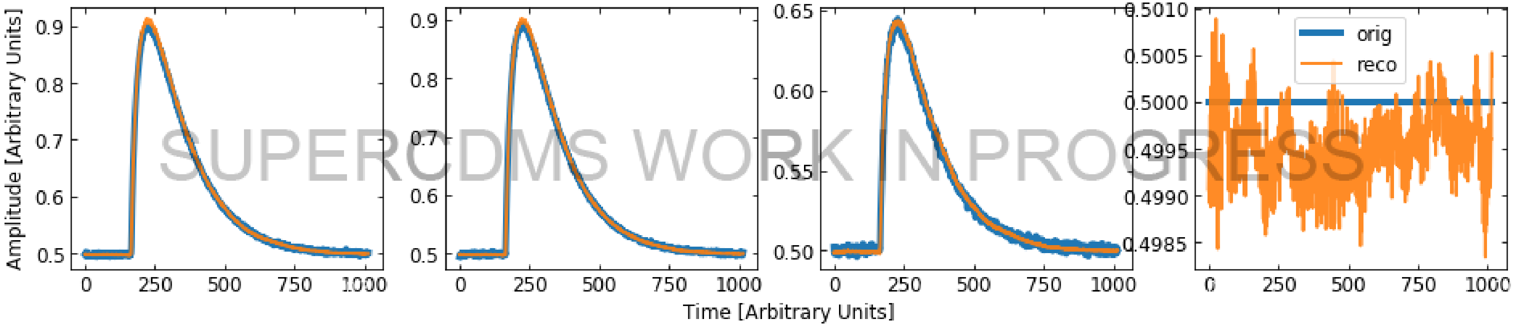
The VAE (Network 1 or **N1**) is trained on detector pulse traces and gives a set of latent variables (**z**) per pulse. Performed several times with different hyperparameters to generate a range of models.

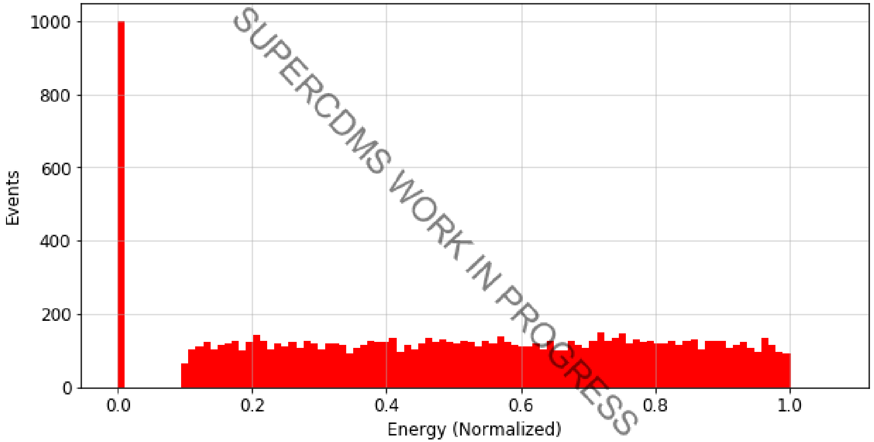
The **z** vectors go to a linear regressor (**N2A**) trained on events in calibration peaks alone, labeled using old energy reconstruction algo. This is performed on each N1 model. Mean energy and std deviation between models are recorded.

A final regression network (**N2B**) is trained on all events. Calibration peak events are weighted highest, and all other events are weighted by their N2A stdev.

Progress - VAE

Sample comes from a toy detector simulation with a flat energy spectrum (bottom left).
1000 empty events are added at zero energy to create a fourth peak.





Five models were generated, each using different numbers of latent variables (z_length = 4, 8, 12, 16, 20). Batch size starts at 32, but is doubled when loss plateaus, up to 1024.

A selection of original and reconstructed pulses are shown above, including a zero energy pulse.

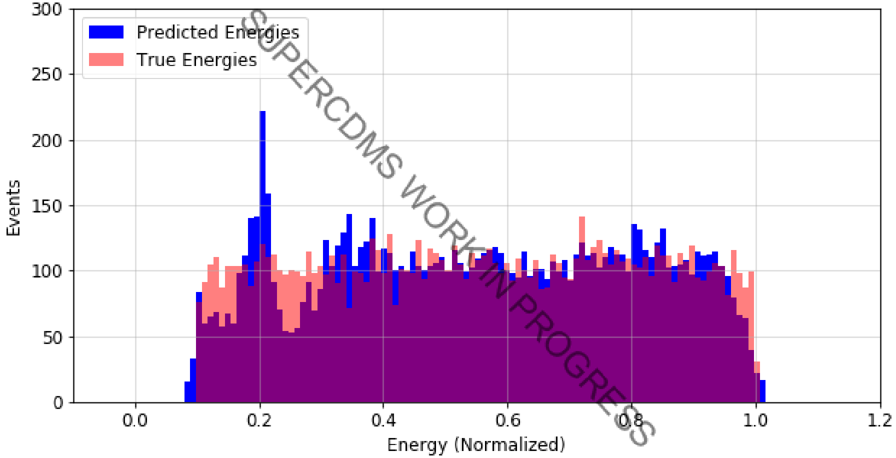
Progress – Energy Predictions

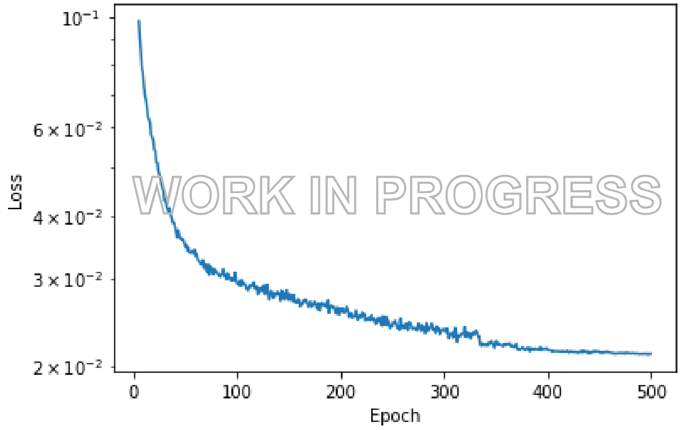
There are no calibration peaks in this data. We choose somewhat arbitrarily to train using events within 0.01 of 0.2, 0.4, and 0.8 in normalized energy, as fake "calibration regions"

Final energy predictions are made w/ a regression network (N2B) with 2 hidden layers of 12 nodes, trained with mini-batches of size 64 with learning rate halved when loss plateaus.

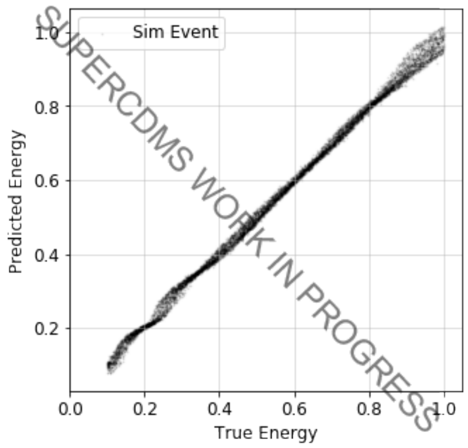
Calibration region events are weighted at 1; all other events are weighted at 0.01*(max_stdev – stdev) with energy labels provided by linear regression step (see backup slide)

Output spectrum shown on bottom left. A peak is can be seen at 0.2, one of the "calibration regions" we selected.





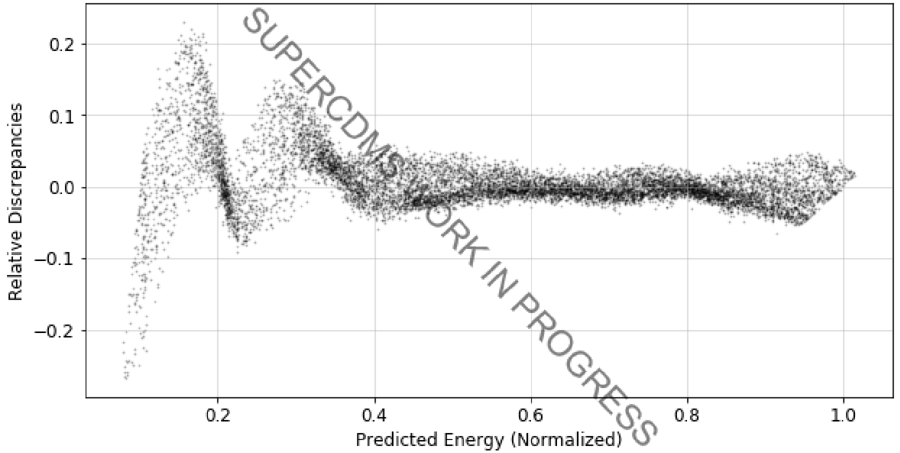
Progress – Energy Predictions

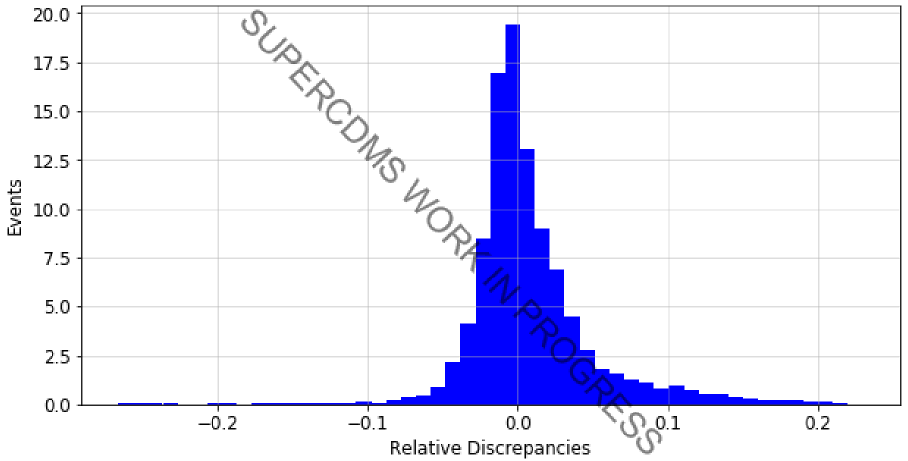


Predicted energy deviates most at edges and around the 0.2 peak.

Mean discrepancy is 0.007, showing no strong bias positive or negative. Standard deviation is 4% overall, and as low as 2% for events above 0.4.

Promising, as it is comparable with existing uncertainties in the OF. However, the energy dependence of discrepancy is concerning.

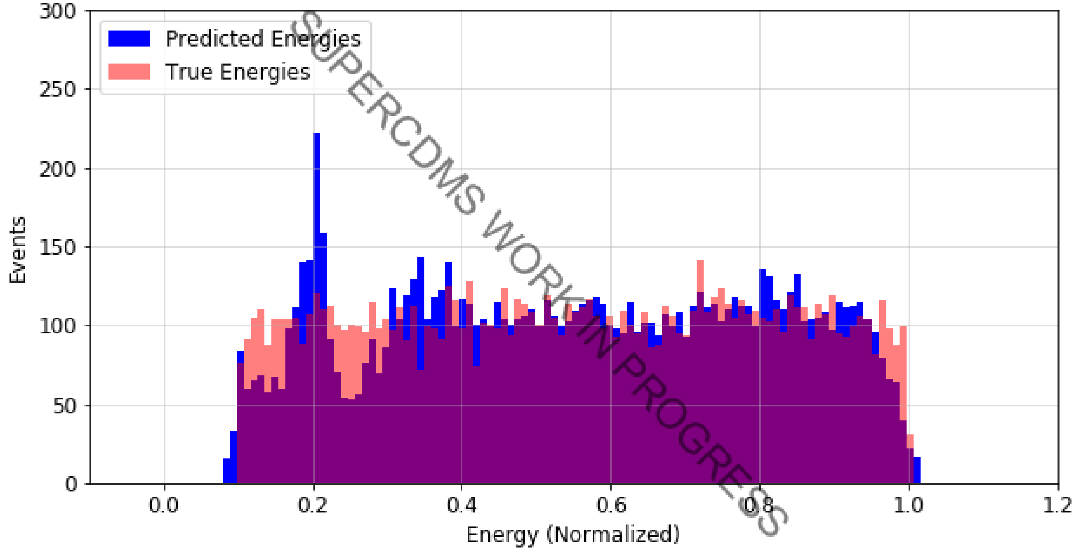




Challenges

It's difficult to get rid of the "peakiness" of predicted spectra, even when true spectrum is flat. It's not always the same peak (sometimes it's at 0.8 instead), but at least one usually appears. Improving semi-supervised part might help.

Extrapolating beyond calibration peaks is difficult. Sharp drops at 0.1 and 1.0 are sometimes reproduced, but not consistently. Momentum/mini-batch training has been implemented to help avoid local minima, but results can still vary widely.



Progress - “N2A”

N2A is a simple multivariate linear regression model. Each latent variable is treated as an independent variable w/ a corresponding parameter.

Individually, they show erratic peaks, but they smooth out satisfyingly when applied over a range of N1 models and averaged. The averaged N2A output shown below.

Predicted spectrum (blue) is relatively flat, conforms to true energy (red) well except at the edges.

Bottom right plot shows that models agree most when close to calibration peaks, and don’t agree well at the edges.