The Inverse Swarm Problem with Neural Networks

Vinit Kumar Singh
Indian Institute of Technology, Kharagpur.

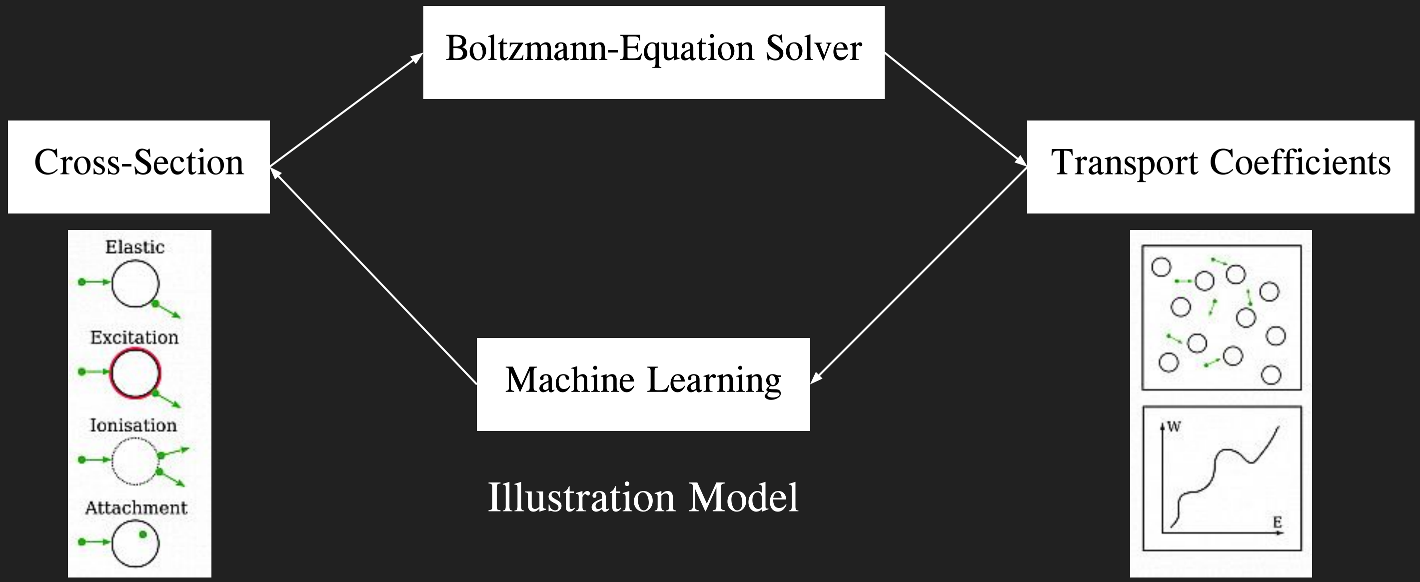
Supervisors: Daniel Cocks, James Sullivan, and Joshua Machacek
Australian National University

Outline

- Inverse Swarm Problem
- Neural Network
- PCA vs VAE
- Surge Function
- Mixture Density Networks
- Recurrent Neural Networks
- Future Applications

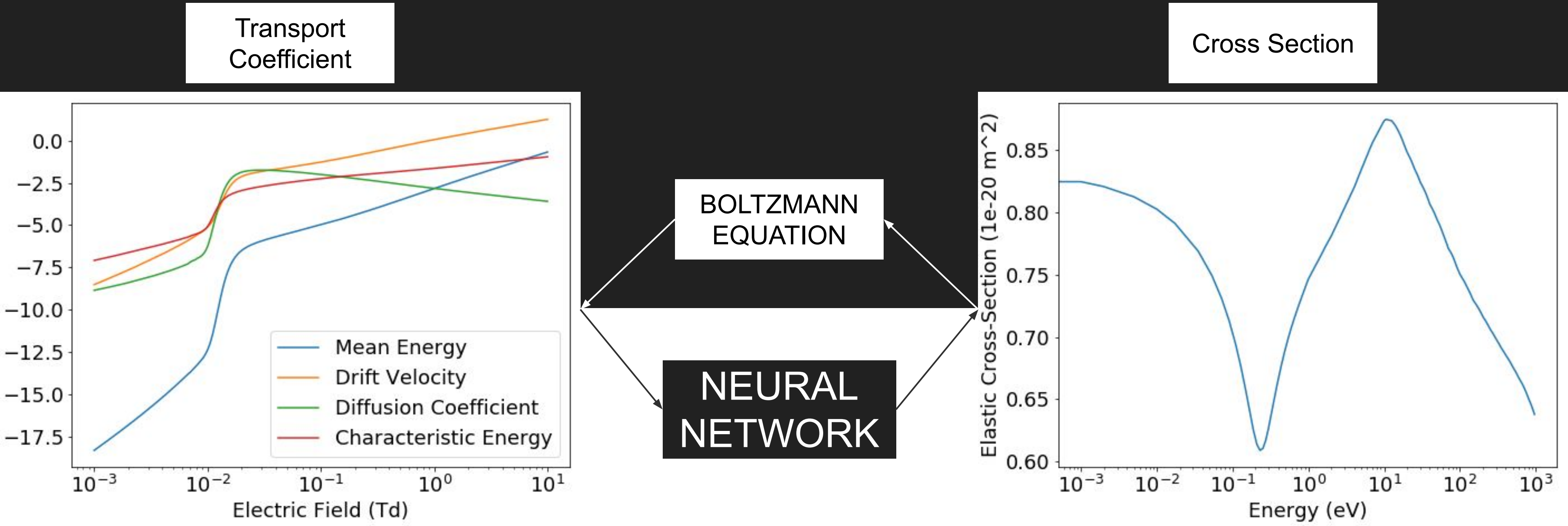
Inverse Swarm Problem

- Unknown if a unique functional exists
- Regardless, not an invertible problem due to sensitivity and uncertainty



Sequence to Sequence Prediction

- Logarithmic spaced grid points.



Boltzmann Equation

- Use cross section in Boltzmann Equation find Electron Distribution.

$$\left[\frac{q^2E^2}{3m^2\boxed{\nu_m(v)}}\right]\frac{d^2f^{(0)}(v)}{dv^2} + \left[\frac{mv\nu_m}{m_0} + \frac{2q^2\boxed{E^2}}{3m^2v\nu_m}\right]\frac{df^{(0)}(v)}{dv} + \frac{m}{m_0}[3\nu_m(v) + v\nu'_m(v)]f^{(0)}(v) = 0$$

- Transport Coefficients: functionals of EDF.

$$\boxed{\epsilon} = \frac{m}{2n}\int_0^\infty dv v^4 f^{(0)}(v)$$

$$W = \frac{1}{3}\int_0^\infty dv v^3 f^{(1)}(v)$$

- Calculating from cross section is easy.
- Calculating cross section from transport coefficient is impossible?

$\nu_m(v)$: collision frequency
f(v): Electron Distribution Function
E: Electric Field
ε: Mean Energy
W: Drift Velocity
V: Velocity

Bolsig (Forward Problem)

Boltzmann Equation Solver
Have also written my own test code to understand the working.

INPUT

ELASTIC
Argon
Mol wt. = 0.136e-4
EFFECTIVE MOMENTUM TRANSFER
CROSS SECTION

U(eV)	σ(Å^2)
0	0.75E-19
0.01	0.75E-19
… …	
1E5	0.49E-22

CONDITIONS
0.1E-2 / Electric field / N (Td)
300. / Gas temperature (K)
200 / # of grid points
200. / Manual maximum energy (eV)
1e-10 / Precision
1000 / Maximum # of iterations

OUTPUT

E/N (Td)	Mean energy (eV)
0.1000E-02	0.3863E-01
0.1012E-02	0.3863E-01
0.1023E-02	0.3863E-01
… …	

E/N (Td)	Mobility *N (1/m/V/s)
0.1000E-02	0.2156E+27
0.1012E-02	0.2156E+27
0.1023E-02	0.2156E+27

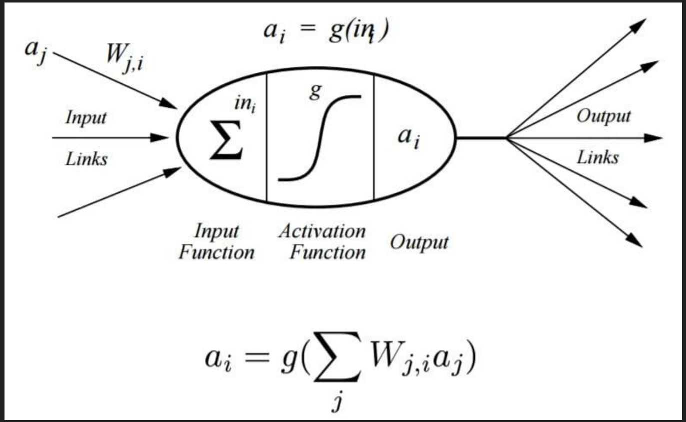
… …

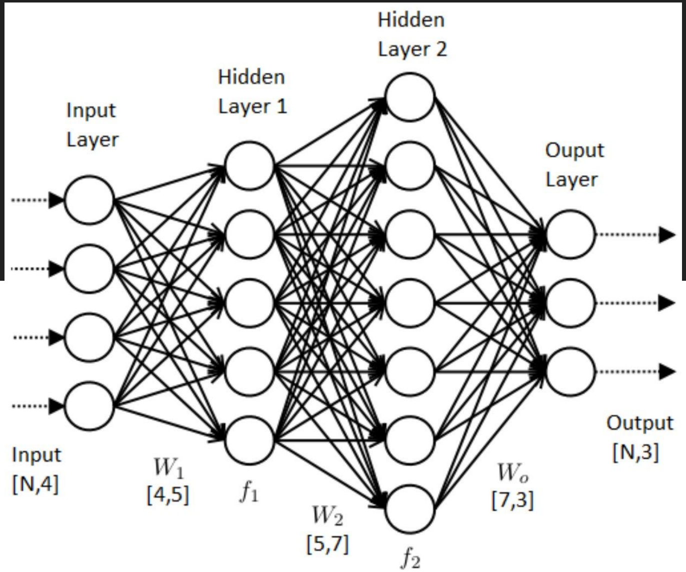
E/N (Td)	Diffusion coefficient *N (1/m/s)
0.1000E-02	0.5586E+25
0.1012E-02	0.5586E+25
0.1023E-02	0.5586E+25
… …	

Neural Network

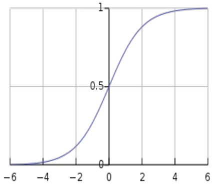
- Widely used
- Function Approximator.
- Perceptron: Artificial Neuron
- Architecture.
- Non-linearity (Activation Function)

Ex. DFT

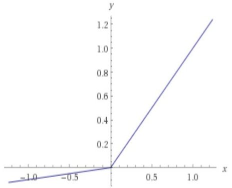




Sigmoid

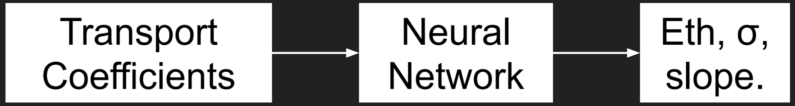


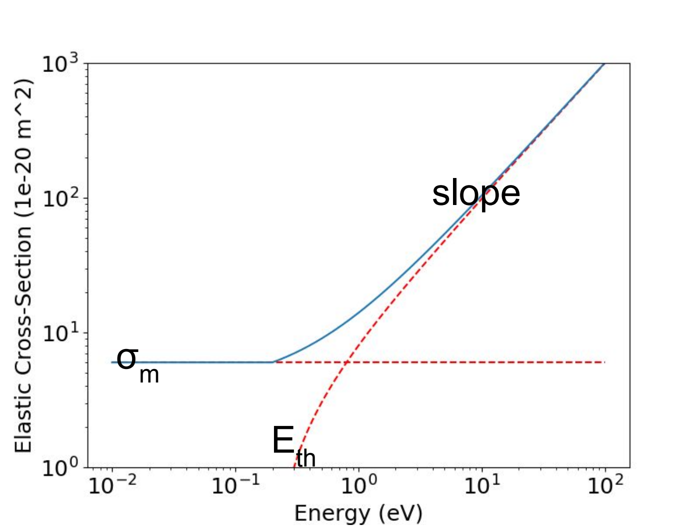
Leaky ReLu



Reid's Ramp Model

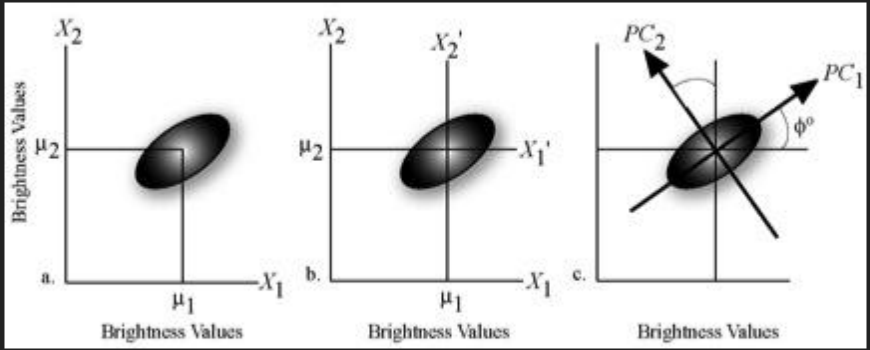
- A simple cross-section for a start.
- Sequence to Parameter prediction

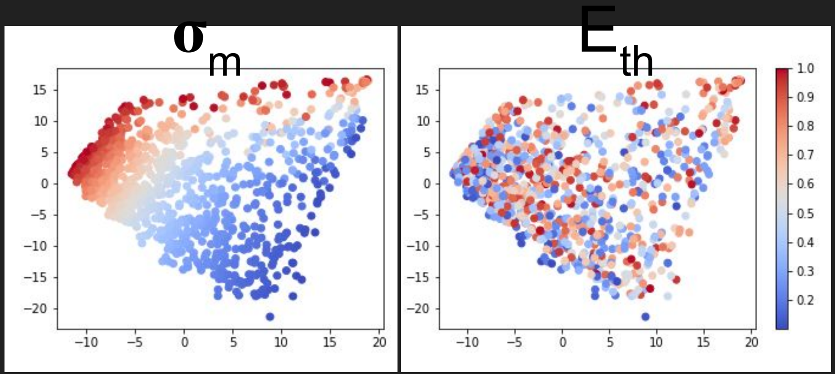




Principal Component Analysis

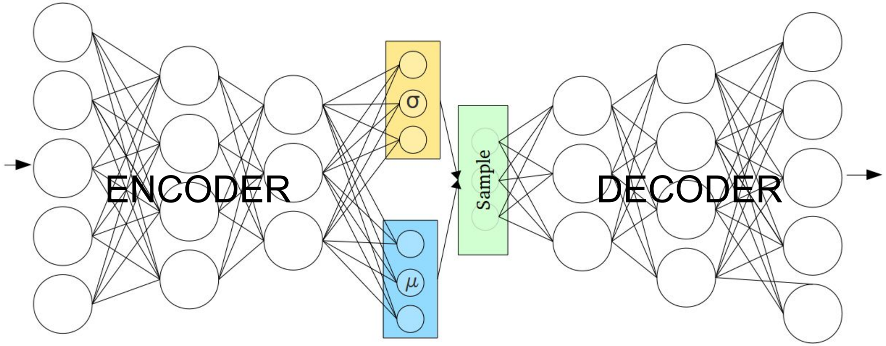
Linear

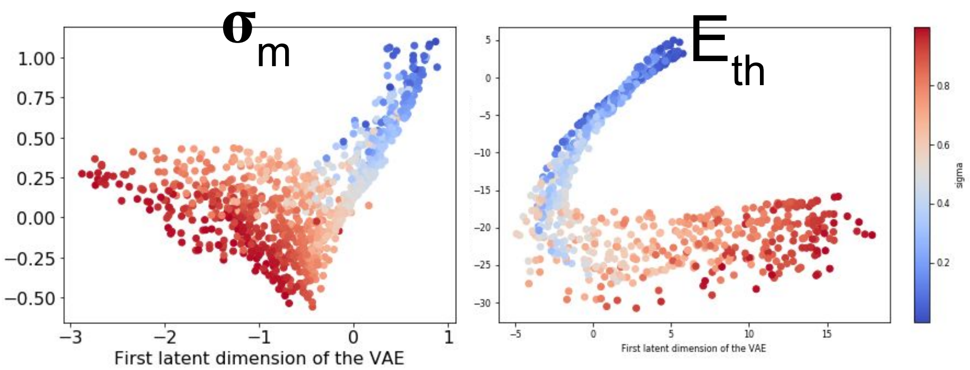




Variational Autoencoder.

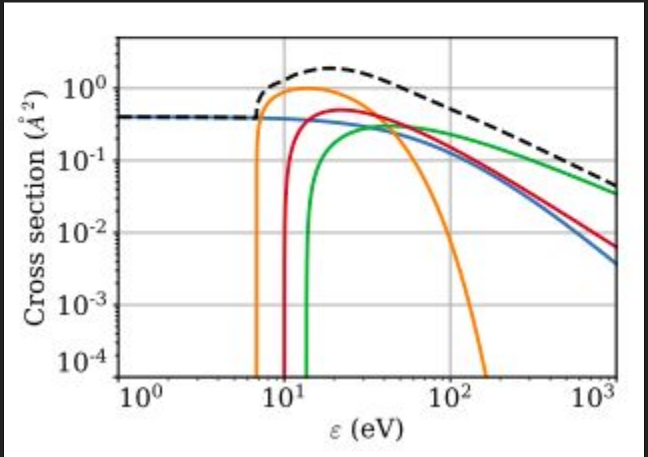
Non-linear





Training Set: Surge Function

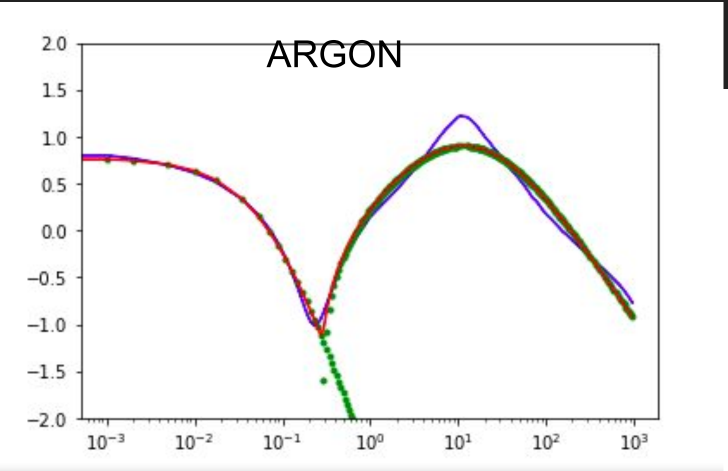
- A realistic model for cross-sections
- Parameters:
  - A : magnitude
  - λ : Width
  - E_{th} : Threshold Energy
  - P : Power-law decay
- Combination of Surge Functions fits realistic cross-sections. (LXCat)



$x = E - E_{th}$

$$S_{\mathrm{pwr}}(x;p,\lambda) = \begin{cases} 0 & x<0, \\ \lambda^p (p+1)^{p+1} \frac{x}{(x+p\lambda)^{p+1}} & x \geq 0, \end{cases}$$

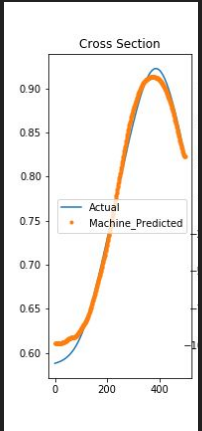
$$S_{\exp}(x;p,\lambda) = \begin{cases} 0 & x<0, \\ \left(\frac{e}{\lambda}\right)^p x^p e^{-px/\lambda} & x \geq 0, \end{cases}$$



Fully Connected Neural Network

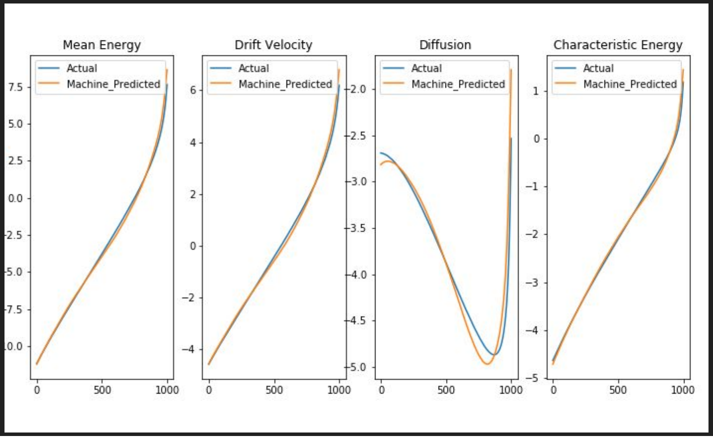
- Fairly good prediction for simple cross sections.
- Predicted cross-sections fed into Bolsig to see how well it machine does on the transport coefficient prediction.
- Smooth curves. Behaves like a low-pass filter.
- Fails to capture deep minimas and sharp peaks.

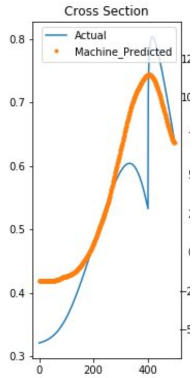
Cross Section

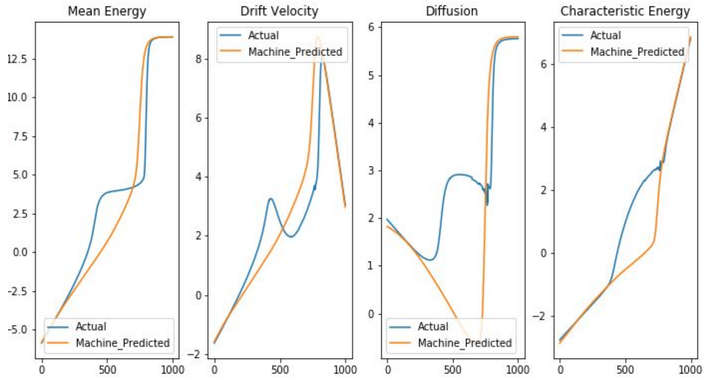


BOLSIG

Transport Coefficient

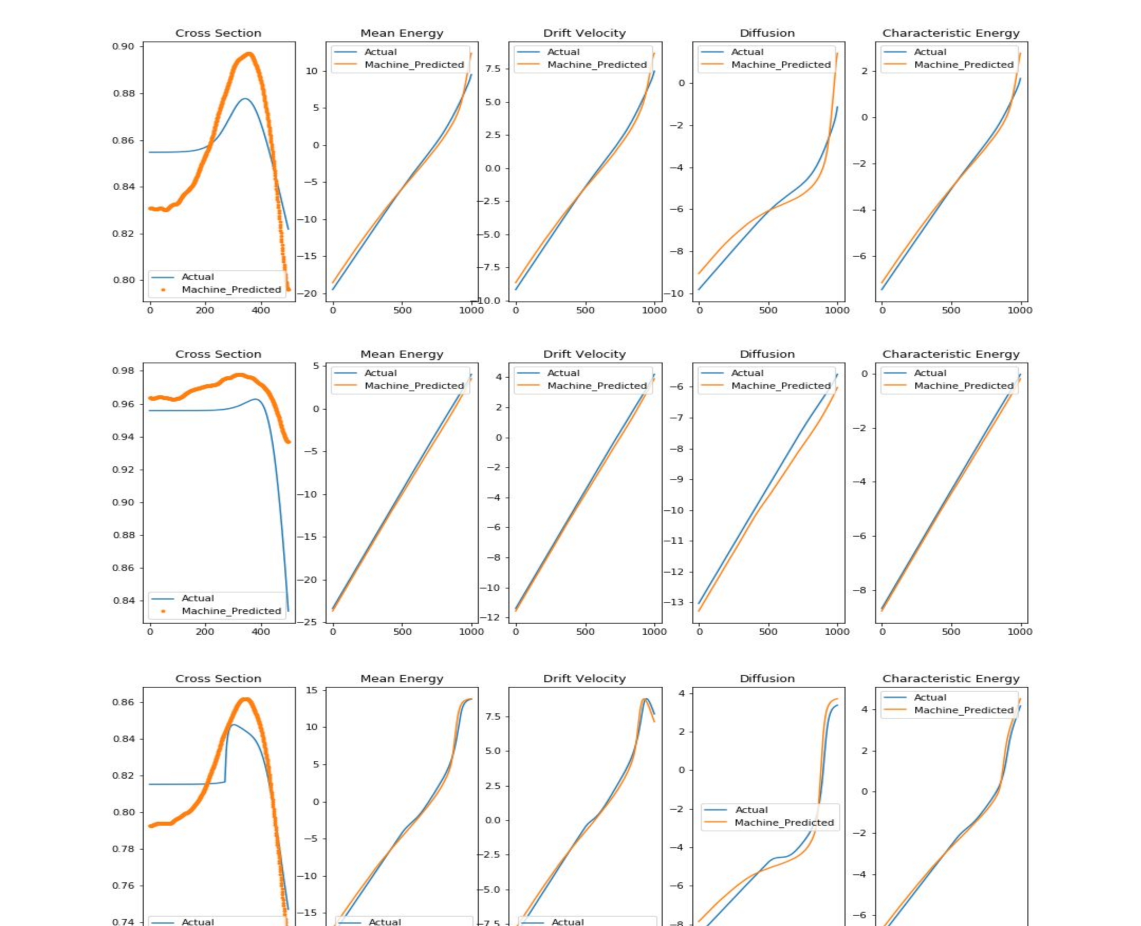






Is the problem Invertible?

- Is the Inverse Swarm mapping many-to-one?
- Similar transport coefficient for very different looking cross-sections.
- Is it a good idea to try to exactly predict the cross-section?

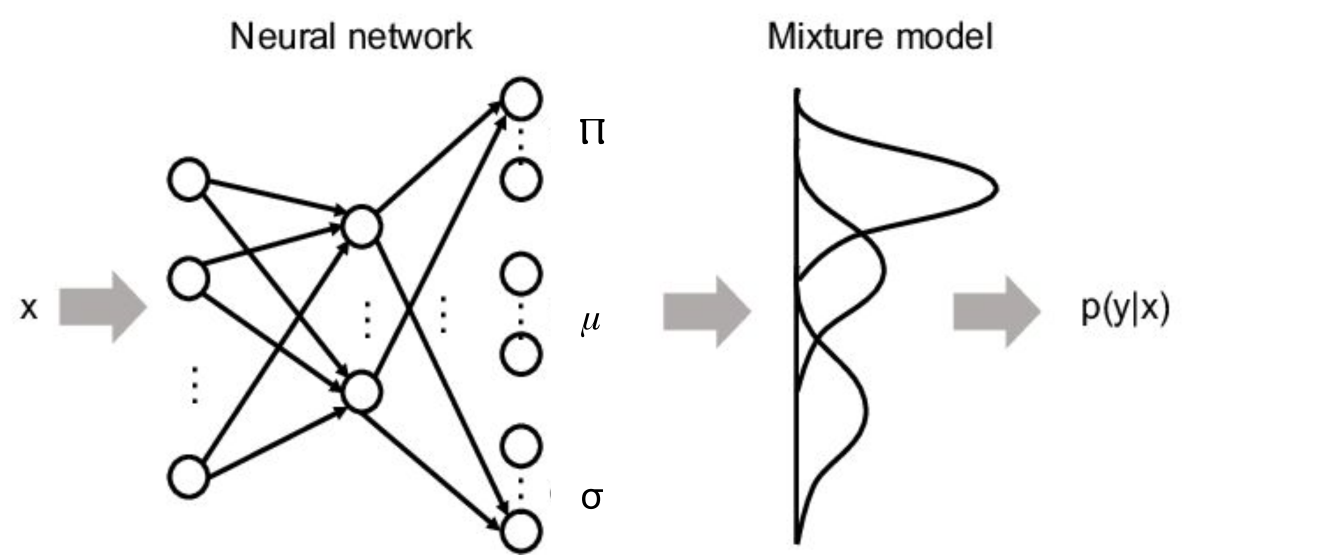


Mixture Density Network

- Successful in predicting many to one functions.
- Predicting Probability Distribution instead of exact values.
- MDN Architecture.

Normal Distribution

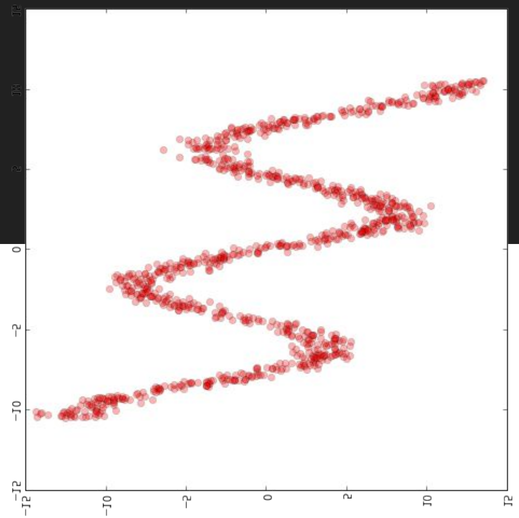
$$P(\hat{y}_i) = \Sigma_j \Pi_j N(\hat{y}_i | \mu_i, \sigma_i)$$



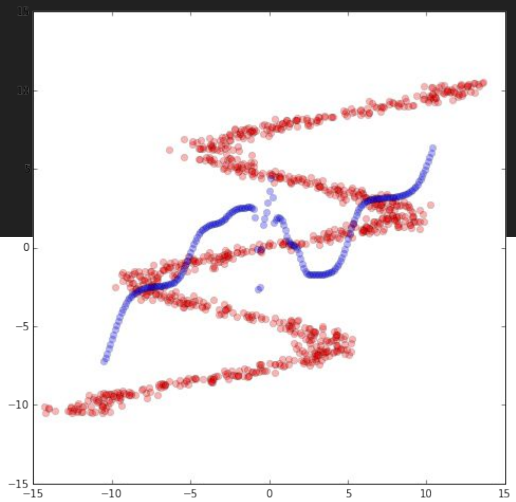
Mixture Density Network

Toy Problem: One-to-Many function.

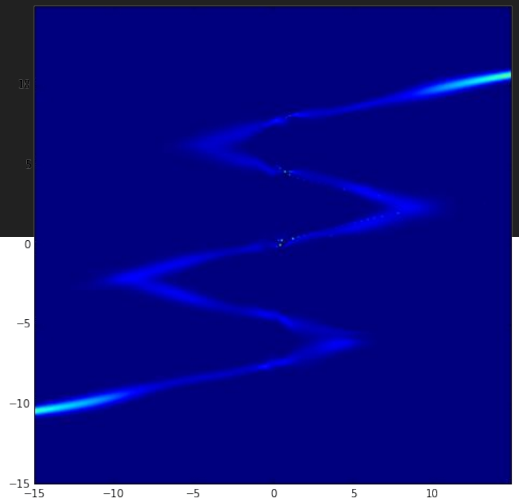
Training input



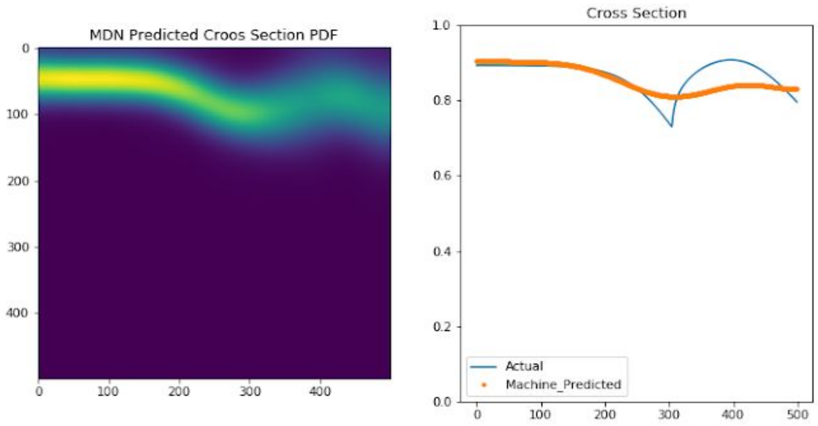
Simple NN prediction

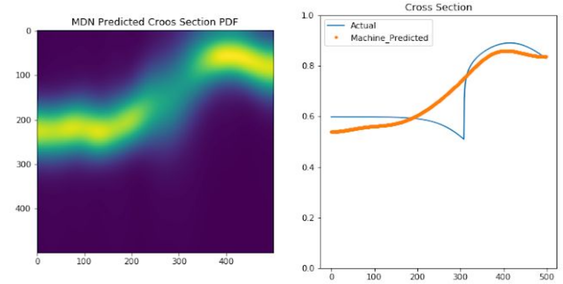


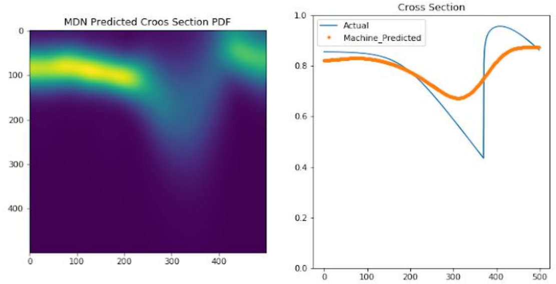
MDN prediction



Mixture Density Network







Post-MDN Optimization

(Work in Progress)
After constraining the value of cross-section with the Probability Distribution Function, we can tweak it slightly away from the expectation value using Bolsig such that they evaluate the transport coefficients accurately.

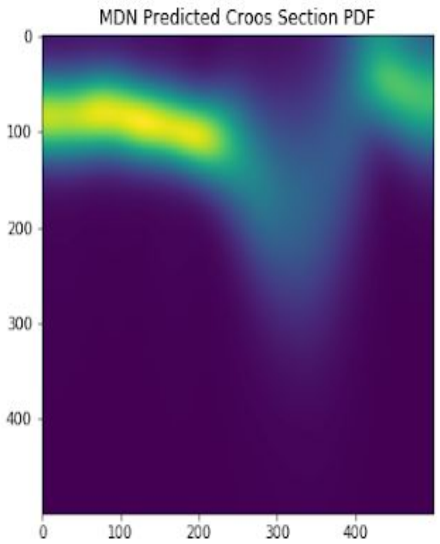
Cost Function = $\Sigma(W-W')^2 - \lambda * \log(PDF_\sigma(\sigma'))$

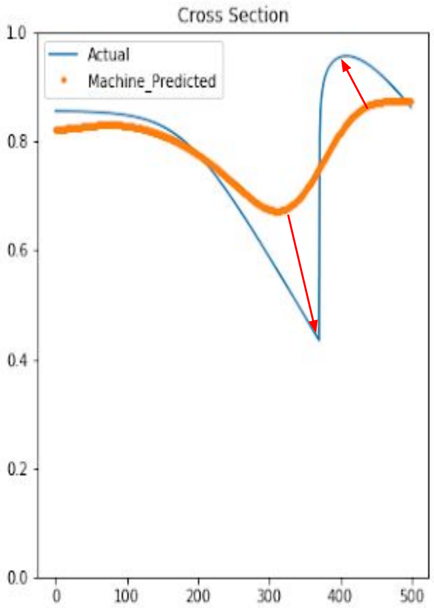
W: Transport Coefficients

σ: Cross section

(‘) denotes Machine Predicted values

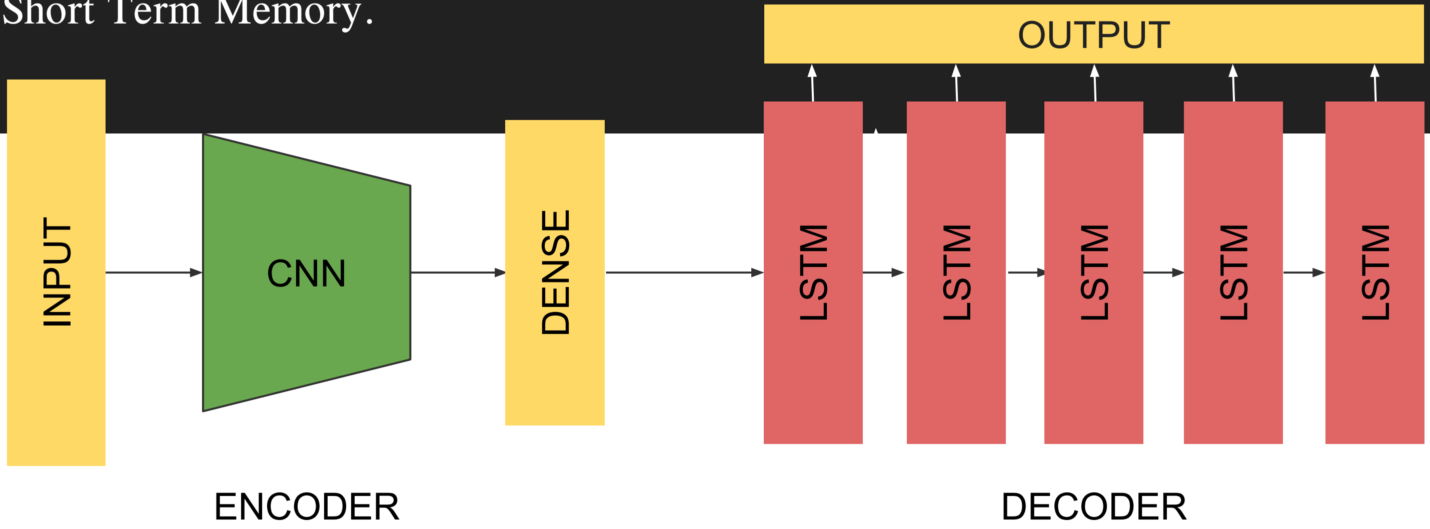
λ: A tunable hyperparameter





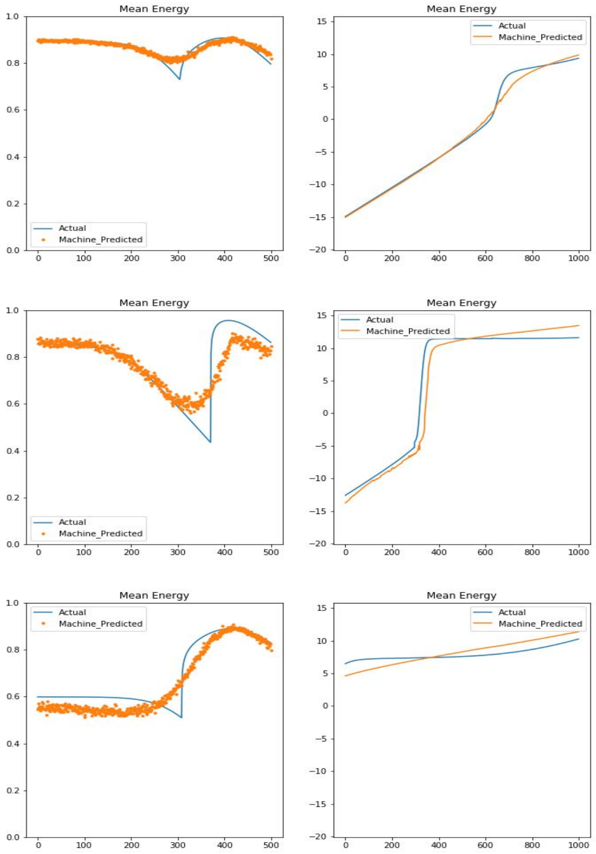
Recurrent Neural Network

- Dealing with sequential data.
- Natural Language Processing Analogy.
- Attempt to capture Ramseur Minima better.
- Convolutional Neural Network.
- Long Short Term Memory.

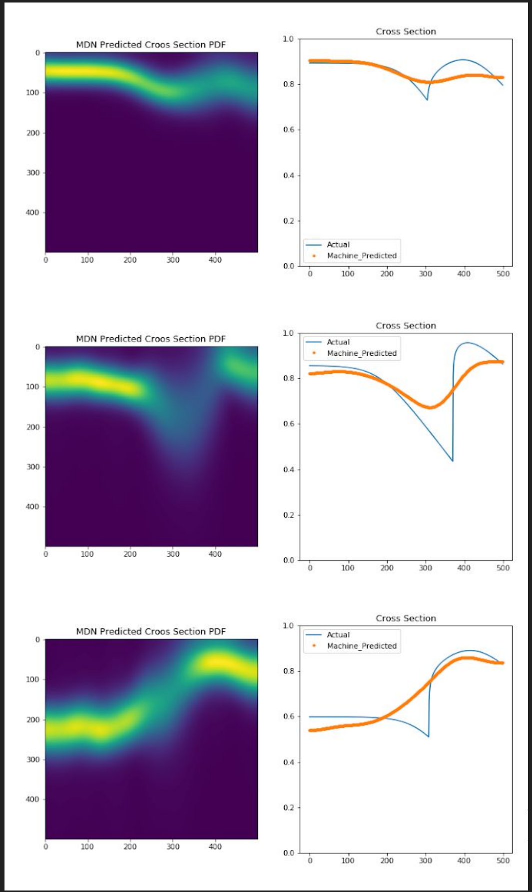


CNN-LSTM Encoder-Decoder Model.

Recurrent Neural Network

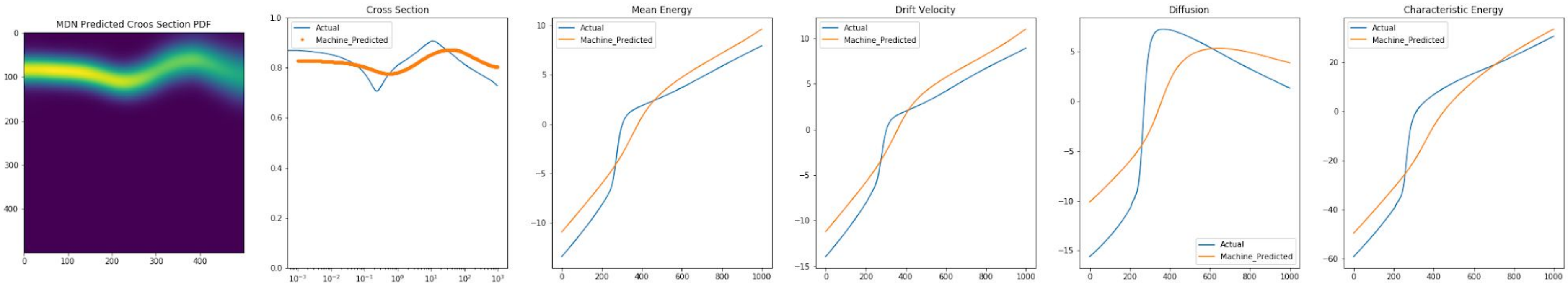


Mixture Density Network



Argon

Machine learned using simulated data but predicting real data.



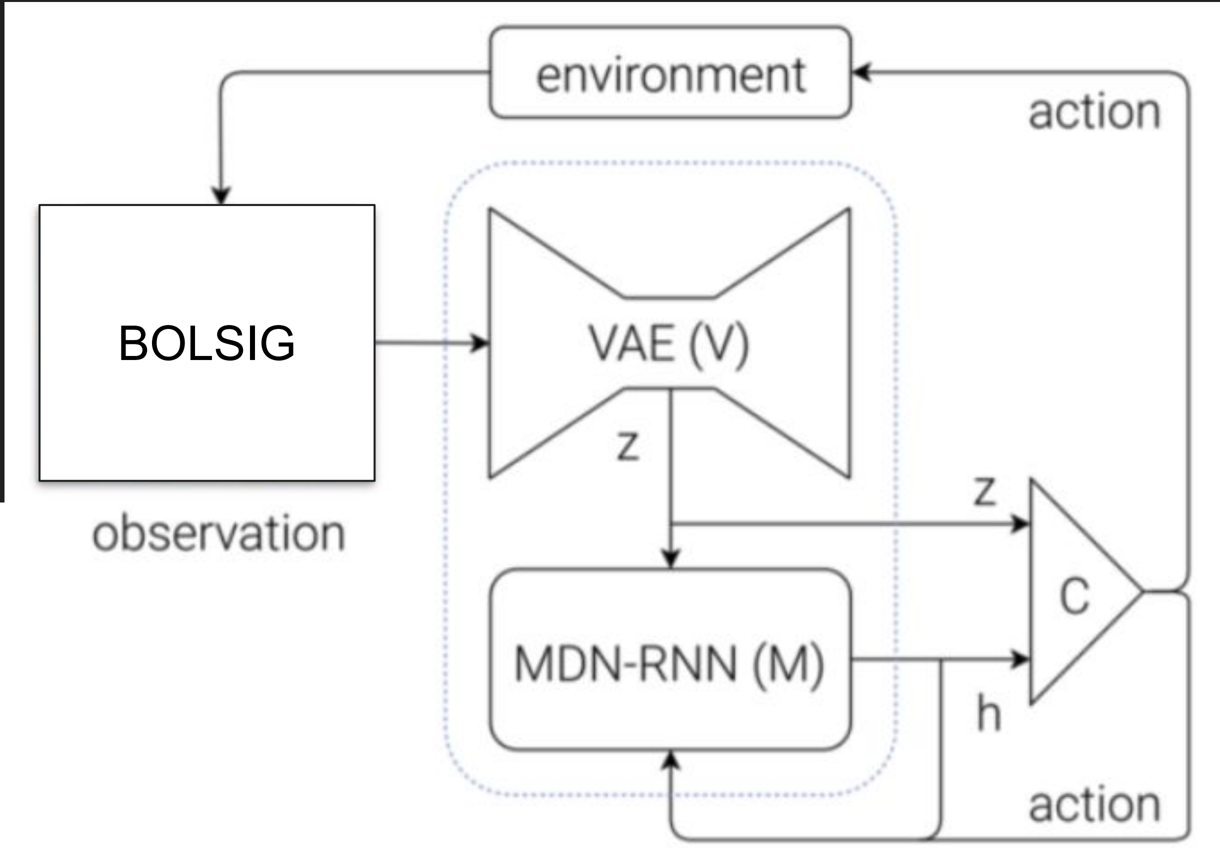
Challenges

- **Is the transformation Invertible?** There is a chance that the transform is many to one. That is a number of different cross-sections can give rise to very similar transport quantities.
- **Function to function mapping** where the basis changes. A trivial mapping with same basis can be simply predicted using one neural network layer. (or if we have some special relations between the bases: Ex. Discrete Fourier Transform.)
- **Finite number** of grid points is insufficient to completely specify a function.
- **Transport quantities are functionals** and require integral over all possible energy values.
- Capturing the **Ramsauer Minima** and the peak of cross-section.
- Transport coefficient run-offs due to steep-decay of cross-section at high energies.
- **Computation Time** required by RNNs is long.
- **Memory** occupied by datasets is huge.

Applications

- HV Circuit Breakers
- Plasma Research
- PET Scans

Further Exploration: Reinforcement Learning



Questions?

Thank You