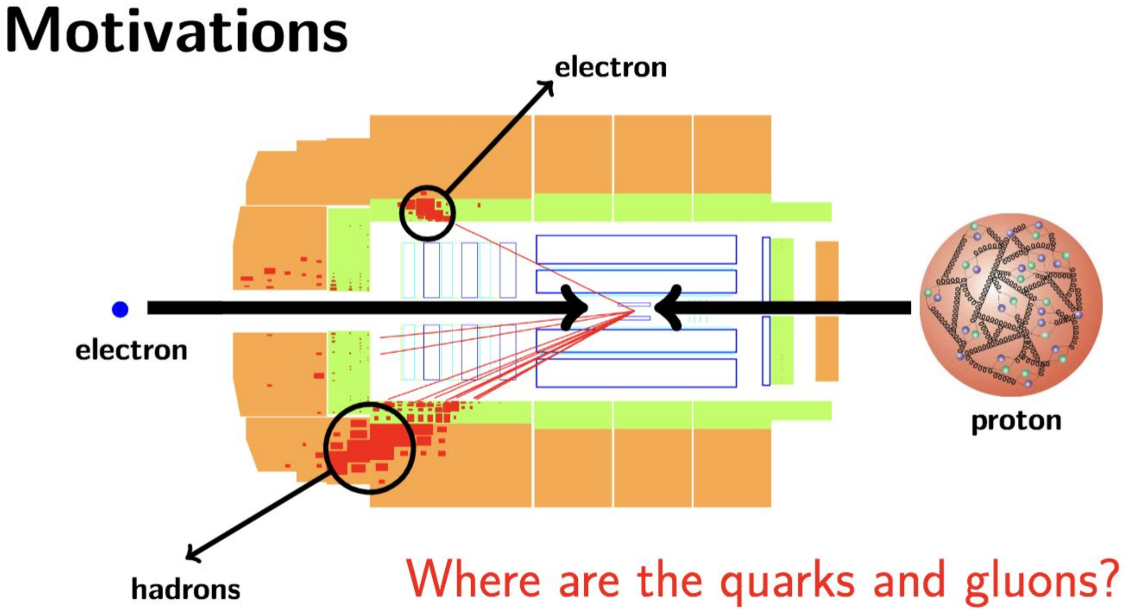
Machine learning techniques to map from experimental cross sections to QCD theory parameters

Eleni Tsitinidi

What is QCD

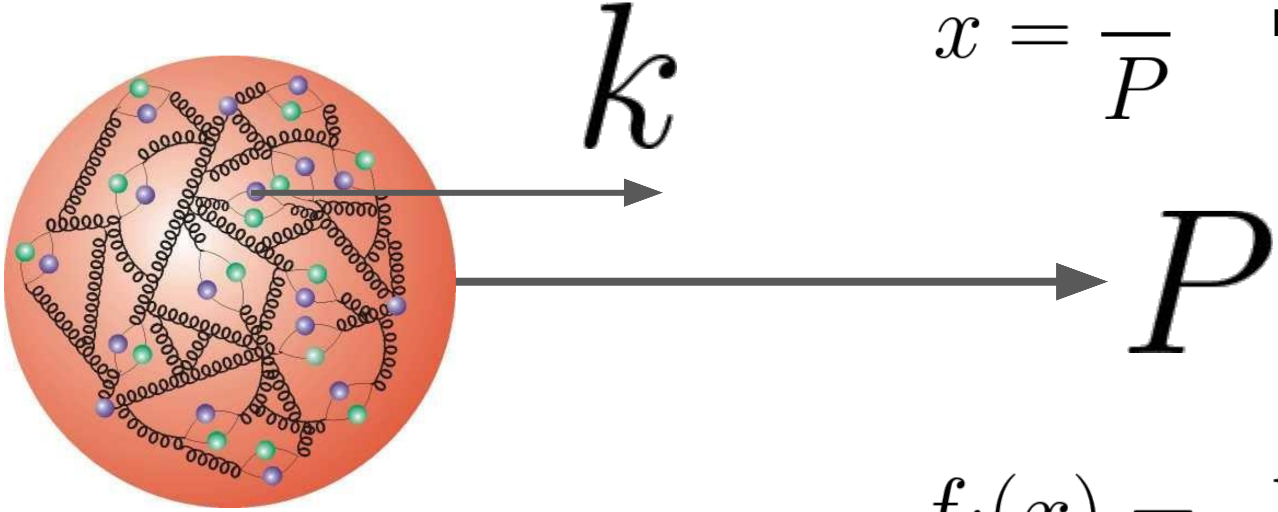
Quantum Chromodynamics (QCD) is the theory that pertains to quark-gluon interactions and the strong force, particularly as components of hadrons.



Quarks and gluons cannot be measured directly

Picture: Nobuo Sato

PDF function : How are quarks and gluons distributed? (number density)



$$x = \frac{k}{P}$$

Momentum fraction

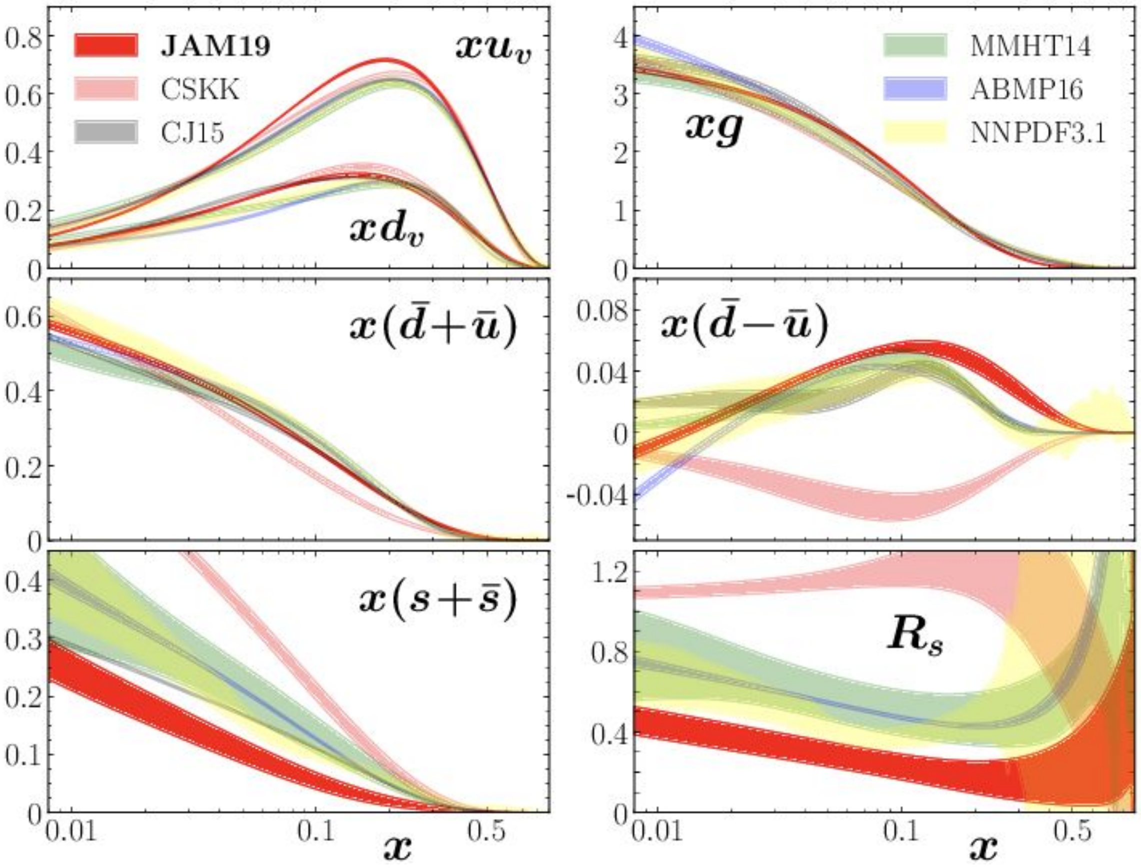
$$f_i(x) = \lim_{\Delta x \to 0} \frac{N_{i,x}}{\Delta x}$$

Parton distribution functions (PDFs)

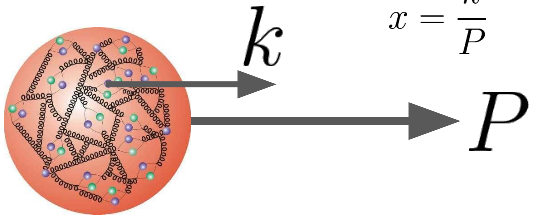
Number of partons of type "i" in an interval x and x+dx

Slide: Nobuo Sato

PDF extracted from experimental data



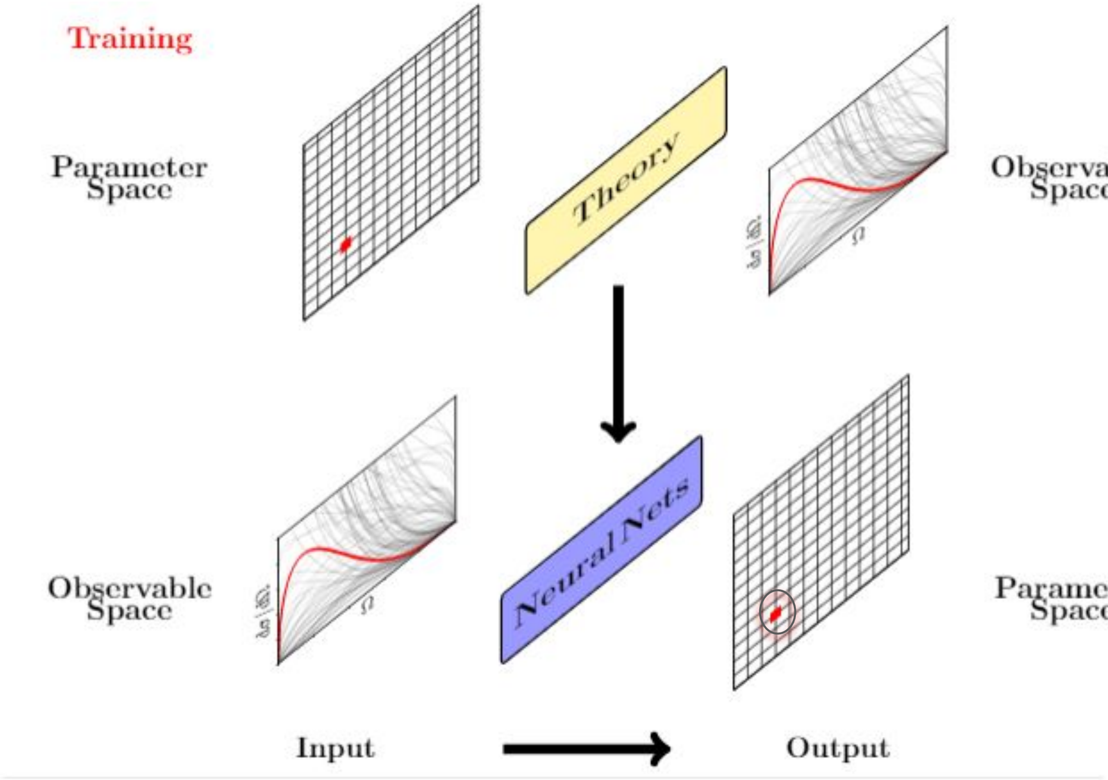
$$x = \frac{k}{P}$$



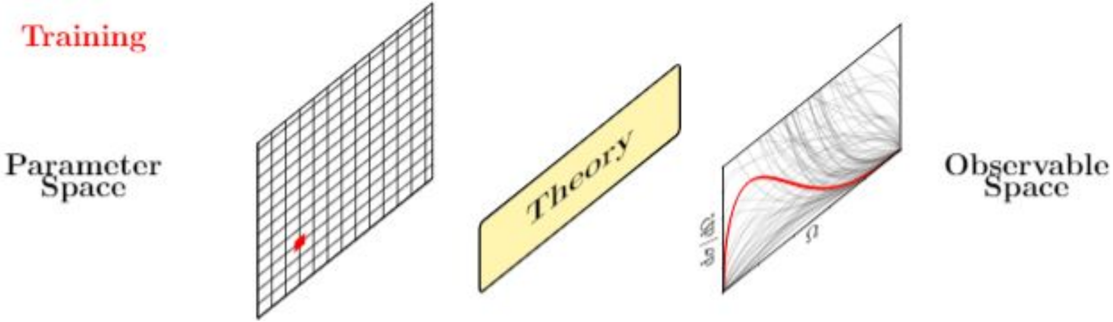
$$u_v = u - \bar{u}$$

$$d_v = d - \bar{d}$$

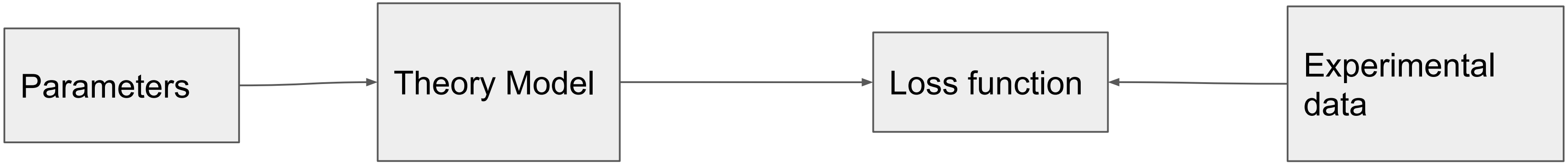
Initial problem - Multiple Solutions



Jefferson Lab Angular Momentum (JAM) workflow



JAM model



How JAM tunes the parameters

Experimental data

Experimental statistical uncertainties

Experimental systematic uncertainties

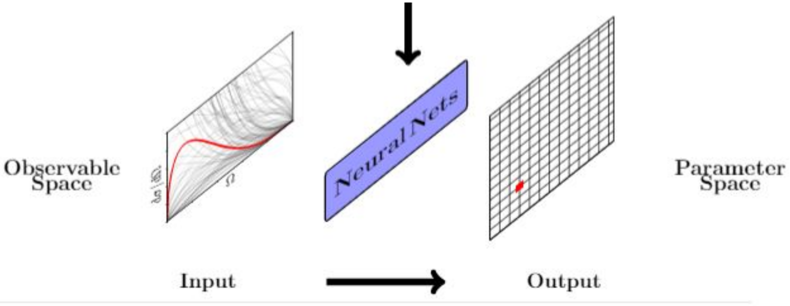
Theory model

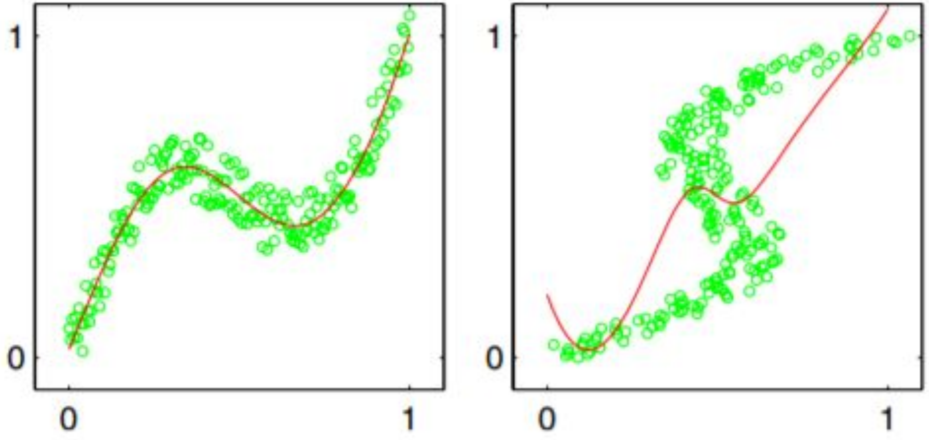
parameters

$$\chi^2(\boldsymbol{a}, \text{data}) = \sum_{e,i} \left(\frac{d_{e,i} - \sum_k r_{e,k} \beta_{e,k,i} - t_{e,i}(\boldsymbol{a})/N_e}{\alpha_i} \right)^2 + \sum_k r_{e,k}^2 + \left(\frac{1 - N_e}{\delta N_e} \right)^2 \qquad (16)$$

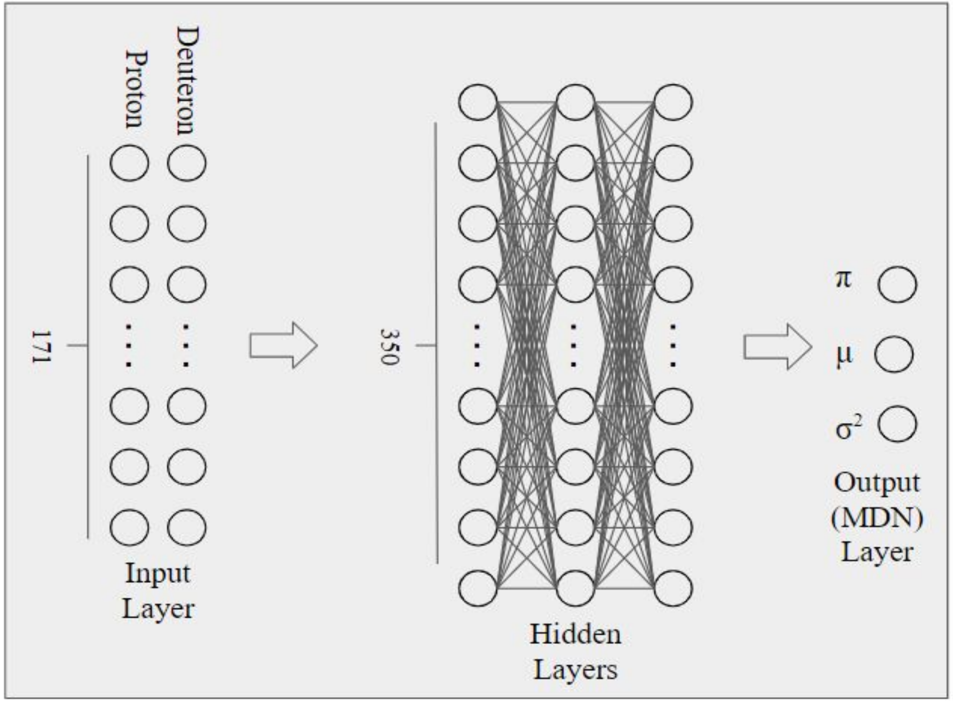
Loss function

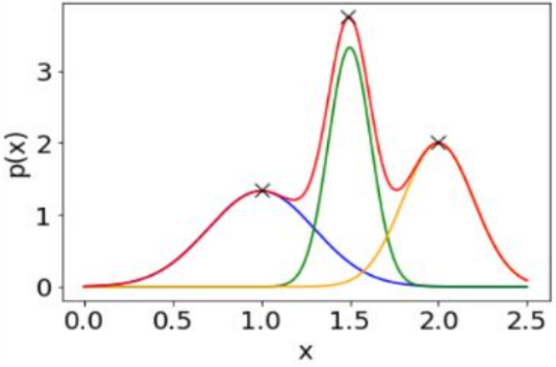
Inverse problem with standard ML





Model Architecture - Mixture Density Network





Output Layer Interpretation:

$$p(\mathbf{t}|\mathbf{x}) = \sum_{k=1}^{K} \pi_k(\mathbf{x}) \mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x})\right)$$

Further architectures

Auto-Encoder Architecture

AE

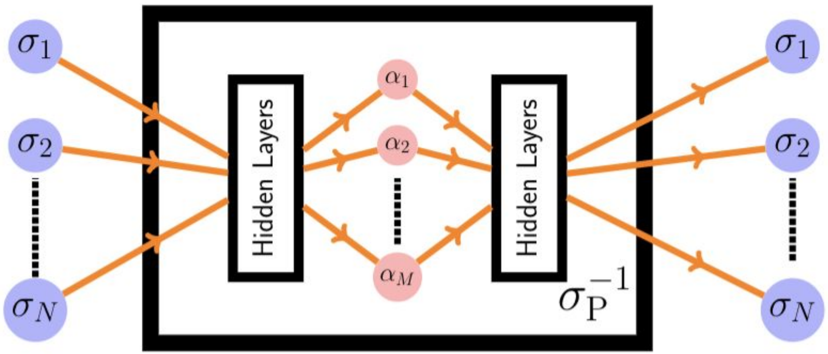


FIG. 3. Inverse mapper based on the parameter supervised autoencoder.

Where a1...aM are the actual parameters

Auto-Encoder with embedded MDN architecture

AEMDN

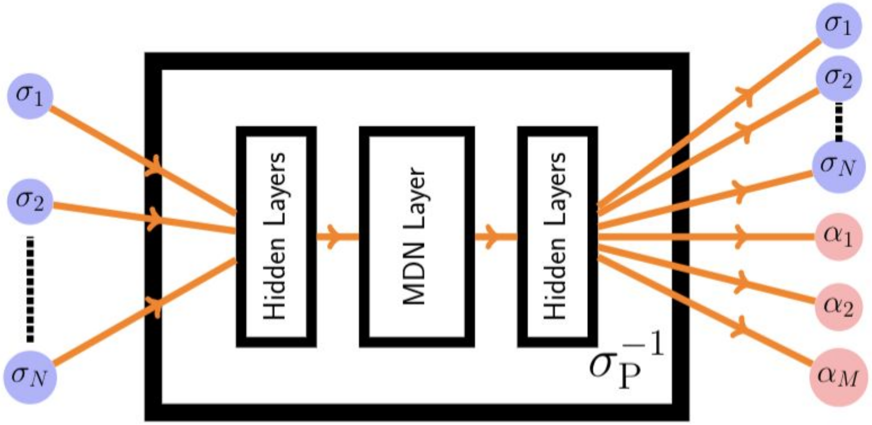
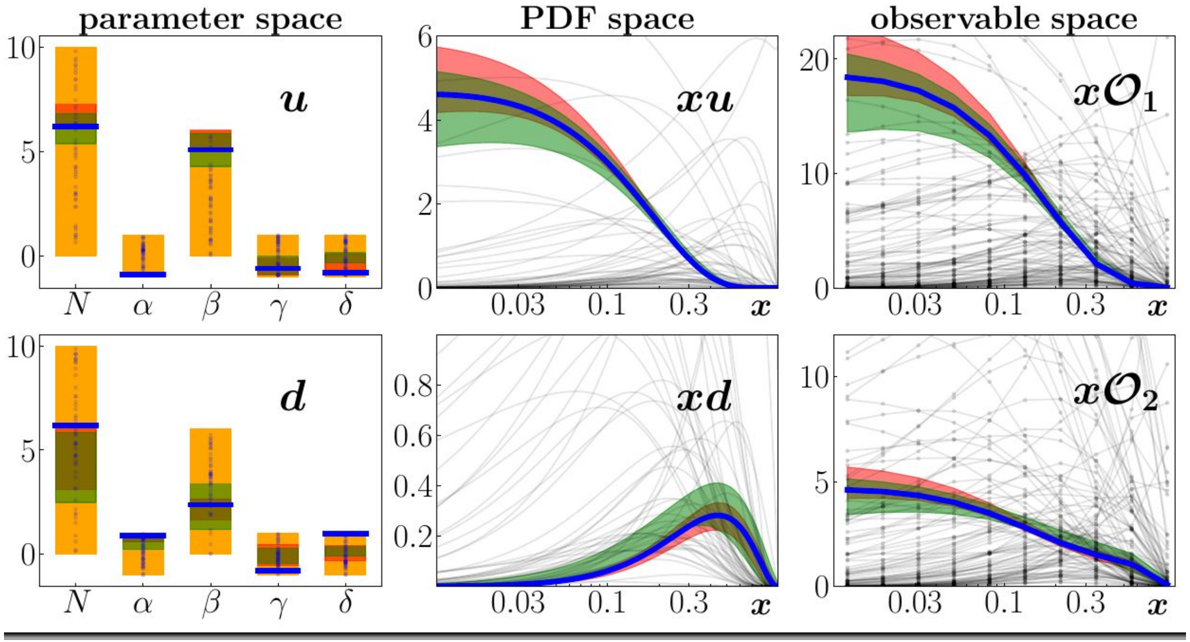


FIG. 4. The parameter supervised autoencoder with the MDN layer.

Data

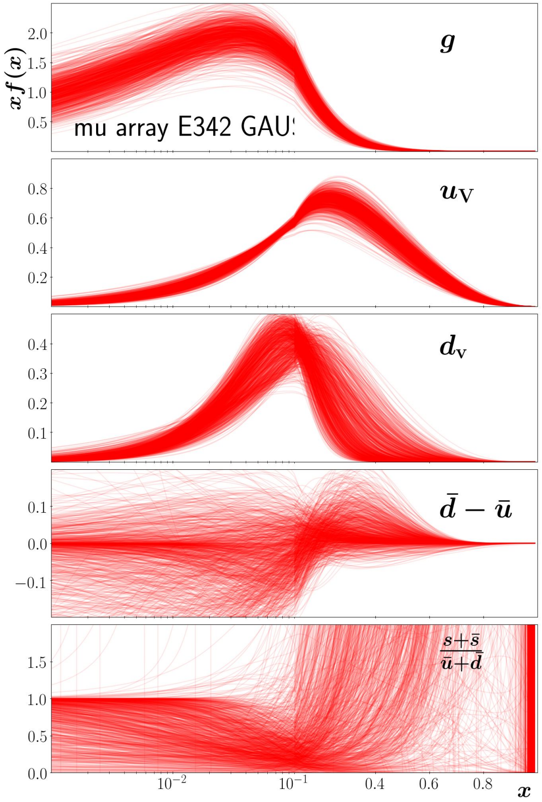
	Simplified model	DIS data
Cross sections (Observable space)	171 σp 171 σn Xsec σp and σn mimic Fp2 And Fn2 structure functions	2680 The cross section space is defined by the kinematics of the world's inclusive DIS datasets from SLAC (p,d), BCDMS (p,d), NMC (p,d/p) and HERA (p)
Parameters	10	25
Parton Distribution Functions (PDFs) "**probability densities"** of parton carrying a momentum fraction x at a squared energy scale Q^2	$i = \bar{u},\ \bar{d},$	$i = \bar{u},\ \bar{d},\ s$ and $\bar{s}.$ $i = g,\ u_v$ and d_v And relationships between them

Simplified Model Results

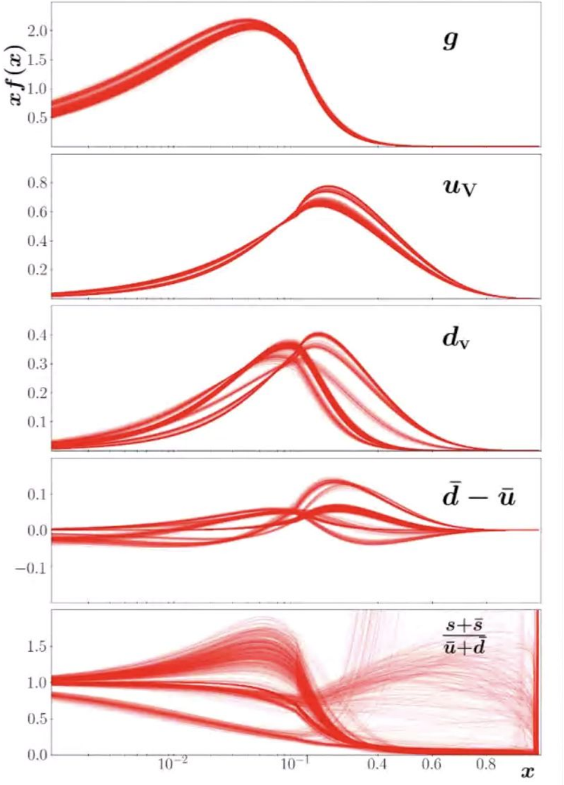


Green=MDN
Red= AE
Blue = JAM
Yellow = full parameter range

Results DIS
Multiple Solutions

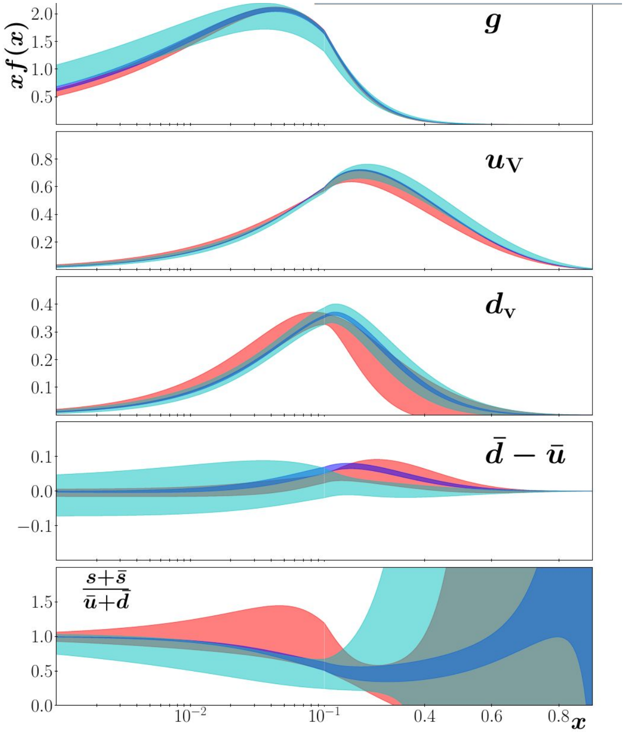


JAM



Average and Error Plot

Chi^2 plot need less yellow circles and more points



Cyan= MDN
Red= Jam
Blue = AE

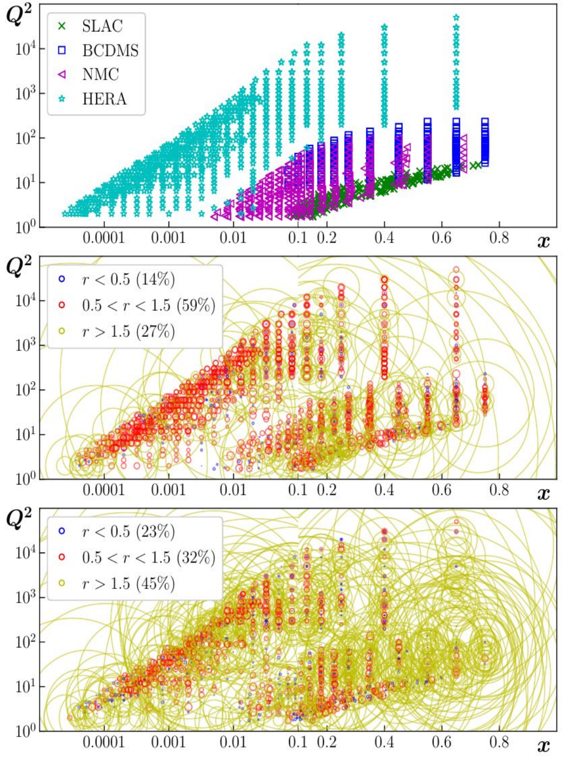


FIG. 8. **Upper pannel**: Kinematics of inclusive DIS global data sets. **Middle pannel**: χ^2 ratio to JAM from the PSA predictions. **Lower pannel**: χ^2 ratio to JAM from the MDN predictions.

$$r = \frac{\chi^2_{\mathrm{ML}}(x, Q^2)}{\chi^2_{\mathrm{JAM}}(x, Q^2)}$$

Taking a step back

- While the results were promising, it was important for us to understand how applicable the model would be to universal fits.

- What happens when we have a different number of free parameters

- Which parameters can be effectively predicted in using our methods.

- What happens when we shrink or expand the trained kinematic region?

Answering those questions

Hyperbox coverage search :

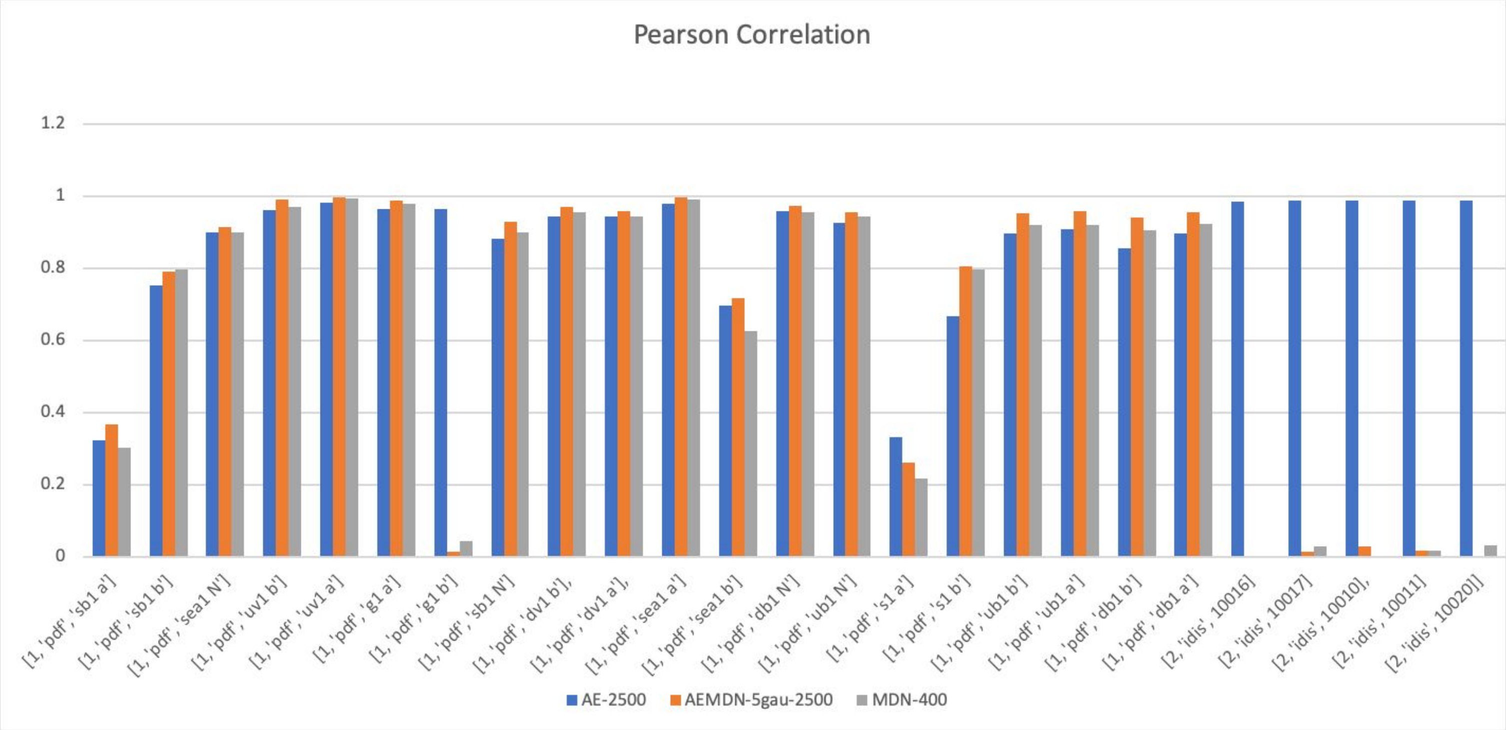
Working up from 10% of the hyperbox of entire experimental kinematic space

Parameter Search:

Working up from 2 to 25 free theoretical parameters

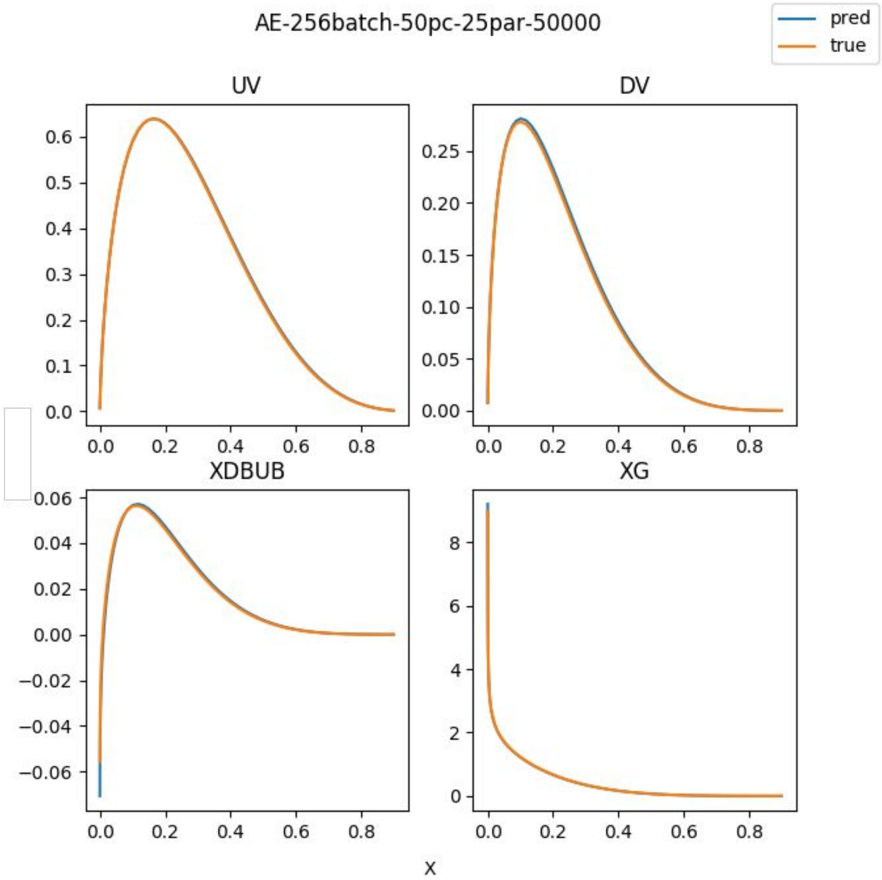
Some initial exploration on 10% of the hyperbox

No params	Model Architecture	Chi2
2	Auto Encoder	1.041
2	Mixture Density Network	1.20
2	AEMDN	1.066
10	Auto Encoder	1.037
10	Mixture Density Network	2.14
10	AEMDN	1.041



Pearson Correlation
1.2
1
0.8
0.6
0.4
0.2
0
[1, 'pdf', 'sb1 a']
[1, 'pdf', 'sb1 b']
[1, 'pdf', 'sea1 N']
[1, 'pdf', 'uv1 b']
[1, 'pdf', 'uv1 a']
[1, 'pdf', 'g1 a']
[1, 'pdf', 'g1 b']
[1, 'pdf', 'sb1 N']
[1, 'pdf', 'dv1 b'],
[1, 'pdf', 'dv1 a'],
[1, 'pdf', 'sea1 a']
[1, 'pdf', 'sea1 b']
[1, 'pdf', 'db1 N']
[1, 'pdf', 'ub1 N']
[1, 'pdf', 's1 a']
[1, 'pdf', 's1 b']
[1, 'pdf', 'ub1 b']
[1, 'pdf', 'ub1 a']
[1, 'pdf', 'db1 b']
[1, 'pdf', 'db1 a']
[2, 'idis', 10016]
[2, 'idis', 10017]
[2, 'idis', 10010],
[2, 'idis', 10011]
[2, 'idis', 10020]]
AE-2500
AEMDN-5gau-2500
MDN-400

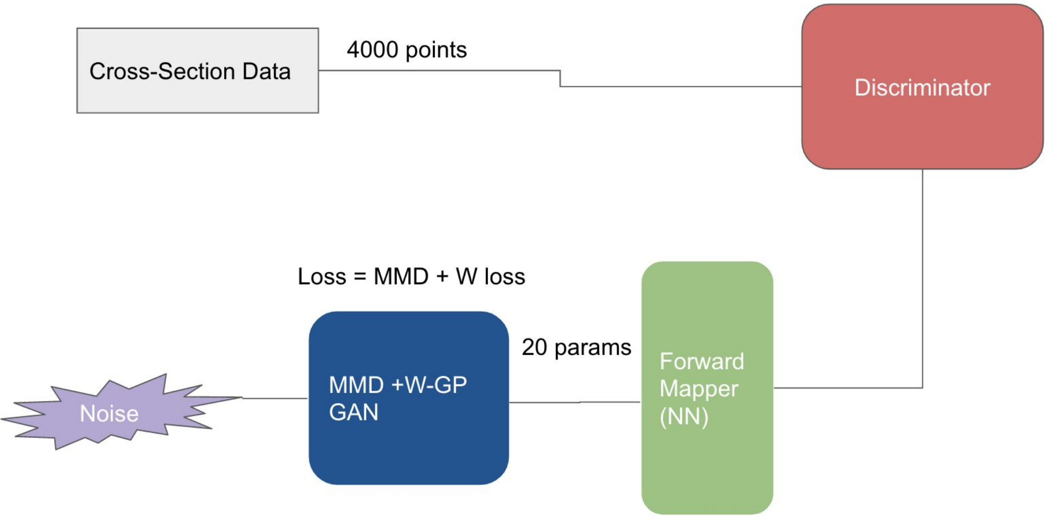
Blind Test Result given one JAM solution -AE



% coverage	No params	Model Architecture	Chi2
75	20	AutoEncoder	1.049
50	25	AutoEncoder	1.173

This semester -- Honors Thesis

GAN for data generation



Preliminary Results

Experiment	Points	$chi2/npts$
SLAC (p)	222	2.69
BCDMS (p)	348	2.09
NMC (p)	274	3.52
HERA II NC e+ (1) (p)	402	2.72
HERA II NC e+ (2) (p)	75	2.29
HERA II NC e+ (3) (p)	259	1.06
HERA II NC e+ (4) (p)	209	1.22
HERA II NC e- (p)	159	2.16
HERA II CC e+ (p)	39	2.27
HERA II CC e- (p)	42	1.66
SLAC (d)	231	5.15
BCDMS (d)	254	2.83
NMC (d/p)	174	1.00

Thank you

DAVIDSON

Dr Raghu Ramanujan Meg Houck

Dr Michelle Kuchera Rida Shahid

Eleni Tsitinidi

ALPhA
DAVIDSON COLLEGE

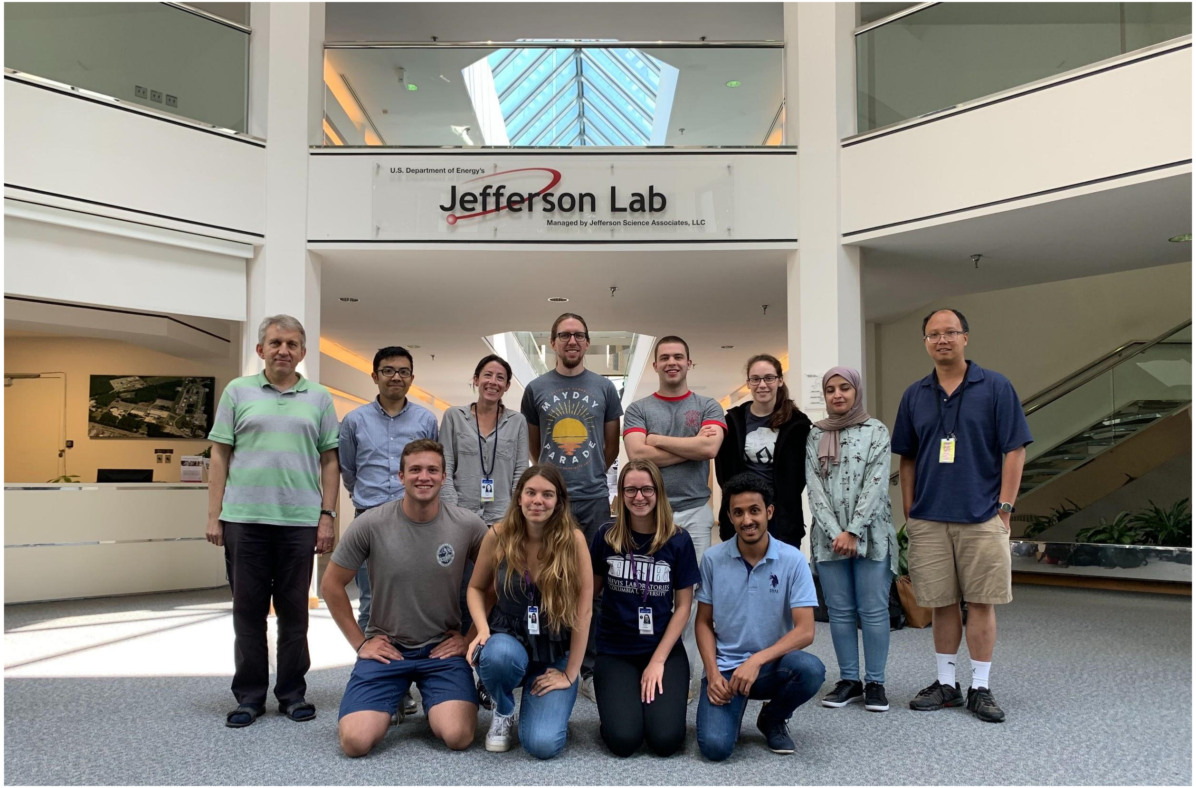
Jefferson Lab
Thomas Jefferson National Accelerator Facility

Dr Nobuo Sato Dr Wally Melnitchouk

ODU

Dr Yaohang Lee

Yasir Alanazi Manal Almaeen



Extra slides

Final model Architectures

Architectures		
AE	**MDN**	**AEMDN**
IN: Original Shape (Cross Section)	Original Shape $\|Tanh\|KR$: 0.01	Original Shape
Dense $\|500\|Relu$	Dense $\|500\|-$	Dense $\|500\|Relu$
Dense $\|300\|Relu$	Dense $\|500\|Tanh\|KR : 0.01$	Dense $\|250\|Tanh$
Dense $\|100\|Relu$	Dropout $\|0.2$	Dropout $\|0.05$
Dense $\|25\|-$	Dense $\|100\|Sigmoid\|KR$:1E-5	Dense $\|100\|Sigmoid\|KR$:1E-3
Dense $\|100\|Relu$	OUT: MDN Layer $\|25$	MDN Layer $\|25$
Dense $\|300\|Relu$		Dense $\|100\|Tanh\|KR$:1E-5
Dense $\|500\|Relu$		Dense $\|500\|Relu$
OUT: Original Shape		Original Shape
Mean Squared Error LR $1E-5$ Batch 256	MDN Mixture Loss 3 Gausians LR $1E-5$ Batch 1024	[MDN Mixture Loss, MSE] 5 Gaussians LR $1E-5$ Batch 256

The data DIS

Dimensions

2680

Cross sections

$$f_i(x,\mu_0) = \frac{N_i\, x^{\alpha_i}(1-x)^{\beta_i}}{B(2+\alpha_i, \beta_i+1)} \qquad i = \bar{u},\ \bar{d},\ s \text{ and } \bar{s}.$$

$$f_i(x,\mu_0) = \frac{N_i\, x^{\alpha_i}(1-x)^{\beta_i}}{B(2+\alpha_i, \beta_i+1)} + \frac{N_S\, x^{\alpha_S}(1-x)^{\beta_S}}{B(2+\alpha_S, \beta_S+1)} \qquad i = g,\ u_v \text{ and } d_v$$

Dimensions

25

Parameters

Normalization coefficients: Ng, Nuv,, Ndv and N̄s

20 free shape parameters

The data -Toy

Cross sections:

Dimensions

101 n	171 n
101 p	171 p

$$\sigma_p(x, Q^2) = 4u(x, Q^2) + d(x, Q^2),$$

$$\sigma_n(x, Q^2) = 4d(x, Q^2) + u(x, Q^2).$$

Here σp and σn mimic what in reality could be the Fp2 And Fn2 structure functions, respectively.

Parameters:

where the Q^2-dependent shape parameters $p = \{N_{u,d}, \alpha_{u,d}, \beta_{u,d}, \gamma_{u,d}, \delta_{u,d}\}$ are given by

Dimensions

10

$$p(Q^2) = p^{(0)} + p^{(1)} s(Q^2), \quad s(Q^2) = \log\left(\frac{\log(Q^2/\Lambda^2_{\rm QCD})}{\log(Q_0^2/\Lambda^2_{\rm QCD})}\right). \quad ($$

Parton Distribution Functions:

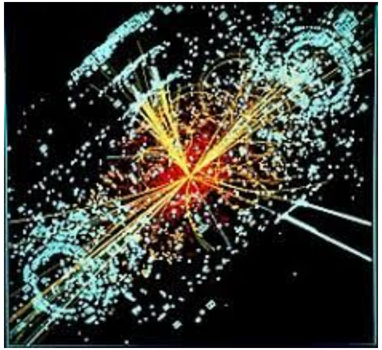
$$u(x, Q^2) = N_u(Q^2)\, x^{\alpha_u(Q^2)} (1-x)^{\beta_u(Q^2)} (1 + \gamma_u(Q^2)\sqrt{x} + \delta_u(Q^2)\, x),$$

$$d(x, Q^2) = N_d(Q^2)\, x^{\alpha_d(Q^2)} (1-x)^{\beta_d(Q^2)} (1 + \gamma_d(Q^2)\sqrt{x} + \delta_d(Q^2)\, x),$$

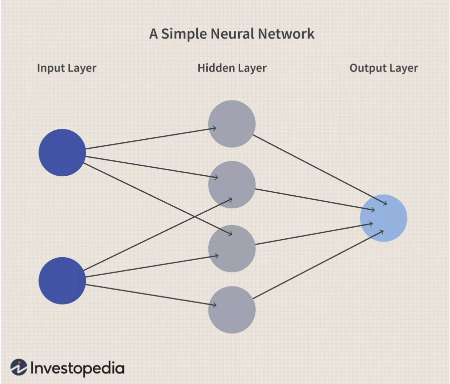
ALPhA
DAVIDSON COLLEGE

What is ALPhA?

Particle Accelerators



Collision Data



Theory

What do these results mean for our theoretical models describing this process?

How can we represent some of these processes in theoretical functions/terms

Experiment

How can we augment our data (ex fill in missing spots on broken particle track)

How can we design better experiments

Simulate experimental process and provide researchers with interface

Data manipulation