Rediscovery of Numerical Luescher's Formula from the Neural Network

Based on arXiv: 2210.02184

Yu Lu (陆宇)

In collaboration with

Yi-Jia Wang (王一佳) Ying Chen (陈莹) Jia-Jun Wu (吴佳俊)

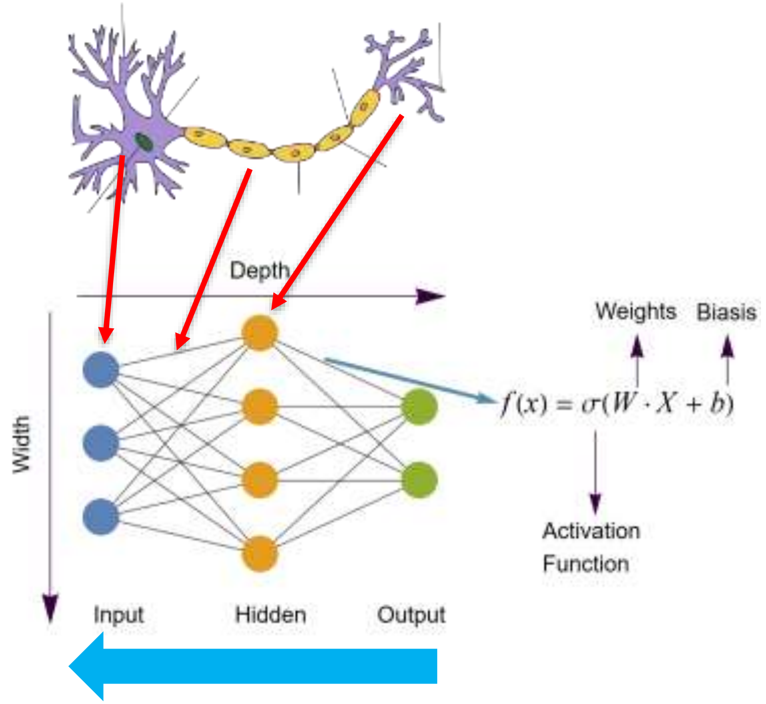
@第二届中国格点QCD年会，上海交通大学李政道研究所

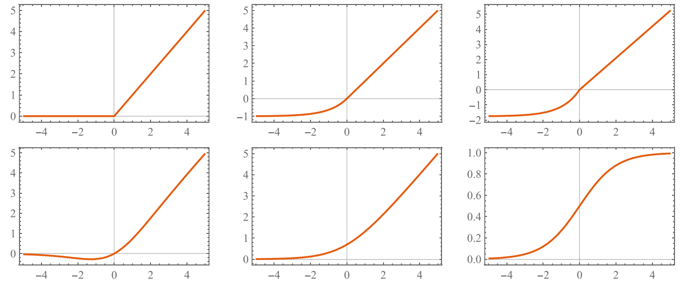
2022.10.10

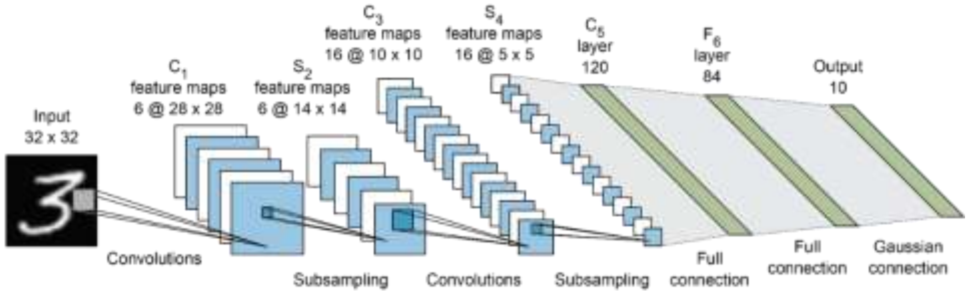
Contents

Introduction to NN

- Artificial Intelligence (AI)=Machine Learning (ML)≈ NN, although ML includes other techniques
- Anatomy/Neural Science Inspired
- Feed Forward Fully Connected NN
  - **Activation function** can be any continuous function
  - Specify suitable **loss functions** for different tasks
  - Optimized by Back propagation

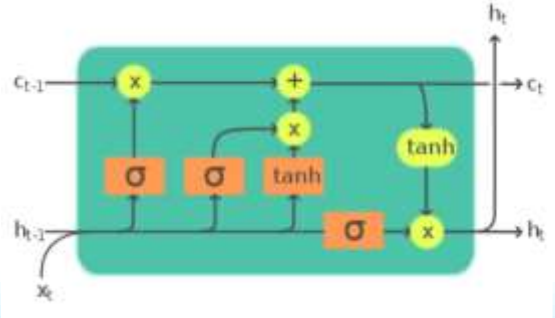


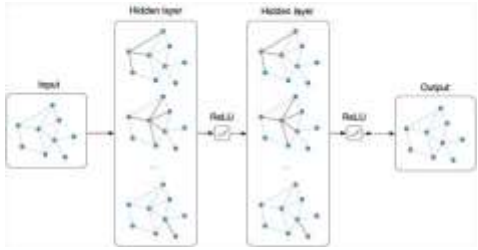




Introduction to NN

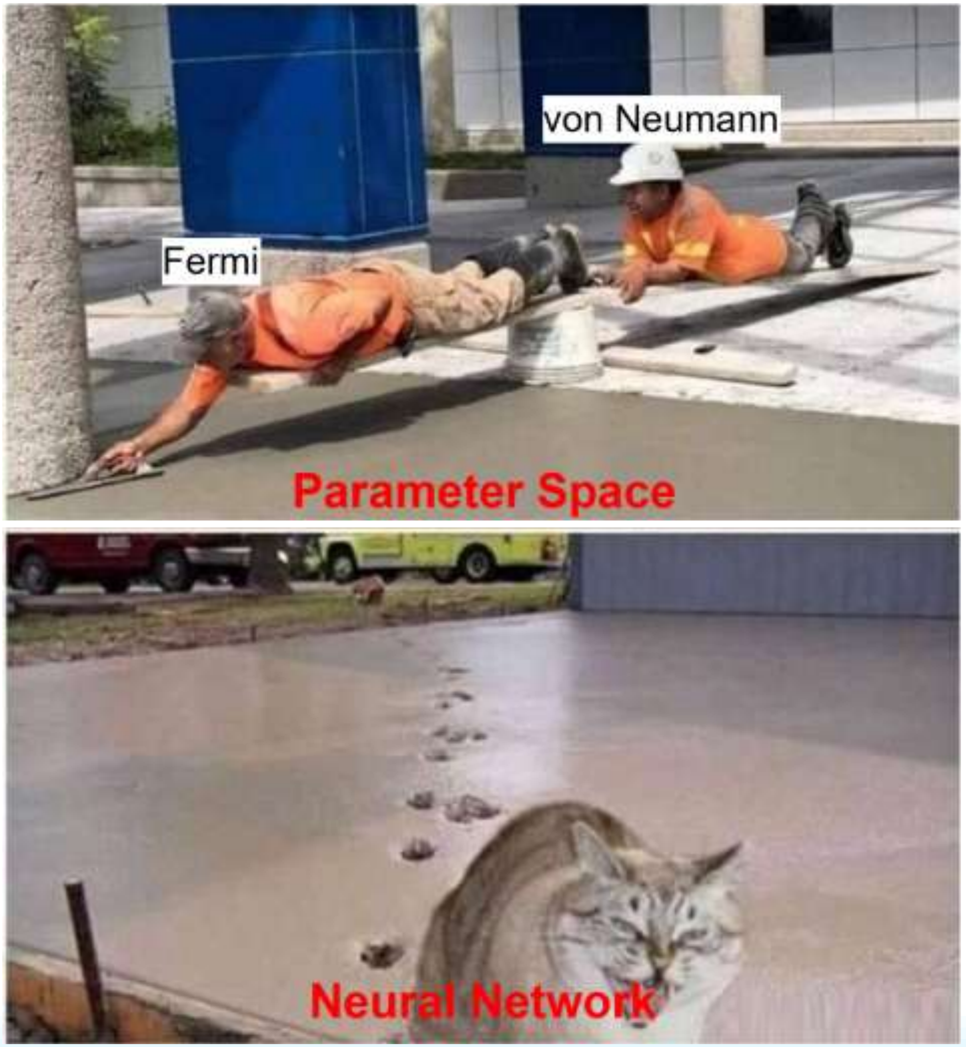
- Modern NN is wide & deep → Deep Learning
  - 1998: LeNet-5, 6×10^4
  - 2022: DALL-E 2, 3.5×10^9
- **Universal Approximation Theorem** (UAT) is the foundation [Hornik *et. al*, Neural Networks 2, 359; 4, 251; 6, 861]
- Beyond FFNN
  - LSTM, RNN, Transformer, GNN …
- NN is the infrastructure of modern digital life
  - Face recognition, recommendation AI, Autopilot, etc.





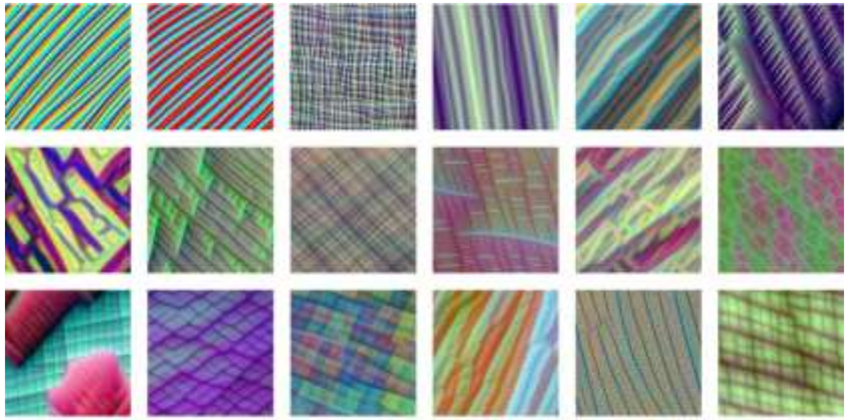
"Teddy bears working on new AI research underwater with 1990s technology"

Possible Questions from Physicists

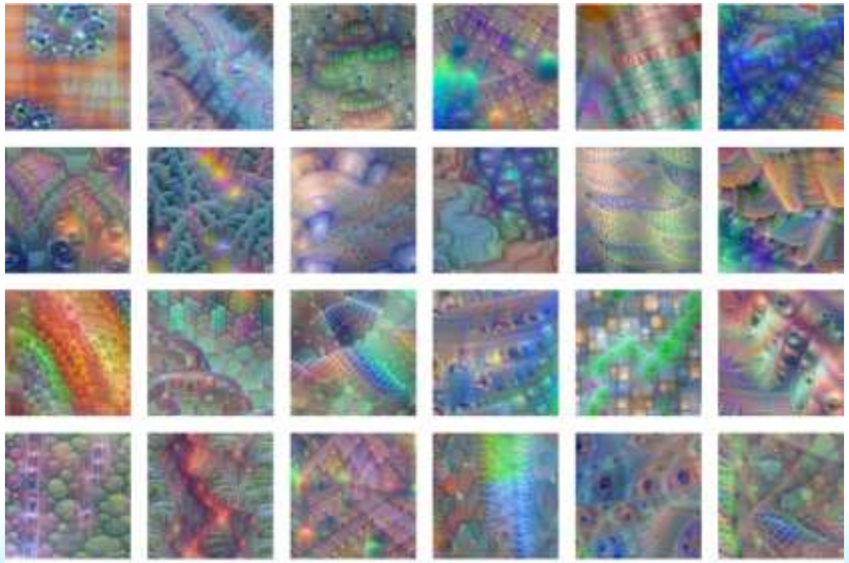


- 10^nparameters?! That's Toooooo Many!
  - Explainability/Interpretability, can be partially explained.
  - Personal comment at the end of this talk.

Possible Questions from Physicists

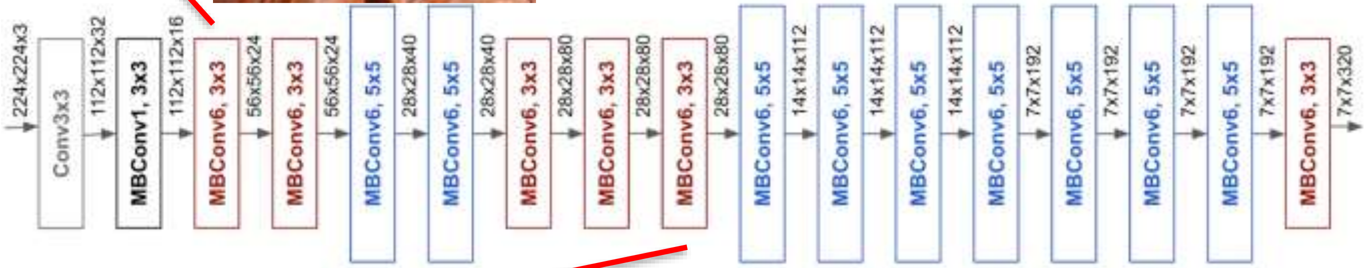


blocks 1a, 2a, 2b, 3a, 3b



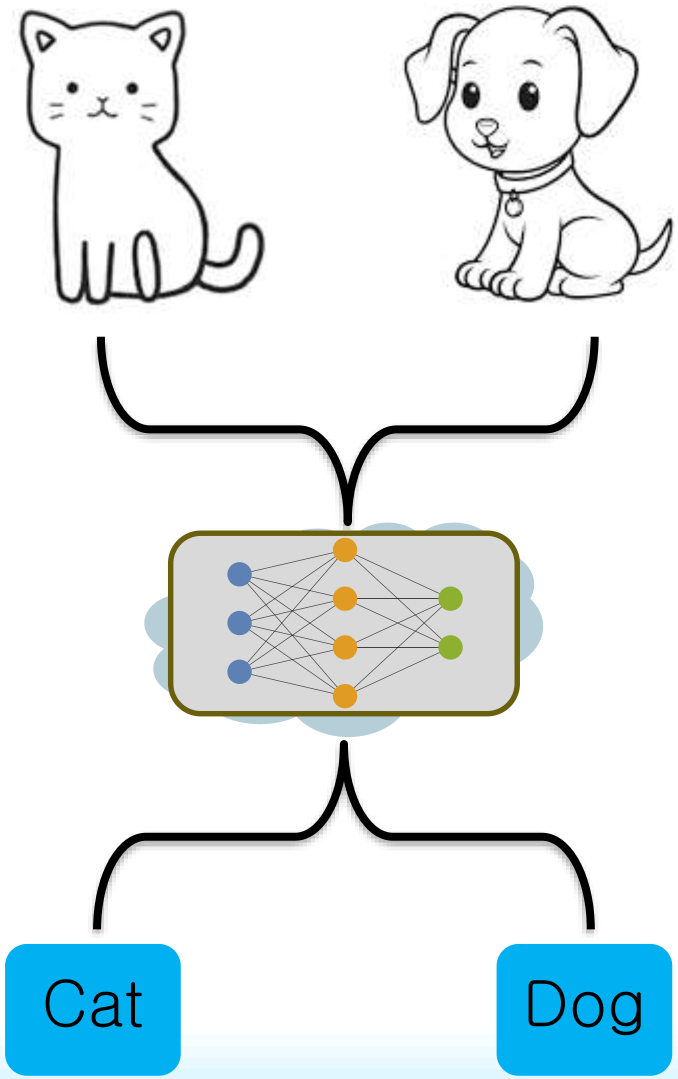
blocks 5a, 5b, 5c

- 10^n parameters?! That's Toooooo Many!
  - Explainability/Interpretability, can be partially explained.
  - Personal comment at the end of this talk.
- Difference between NN and fitting?
  - Fundamentally the same but somehow not that trivial



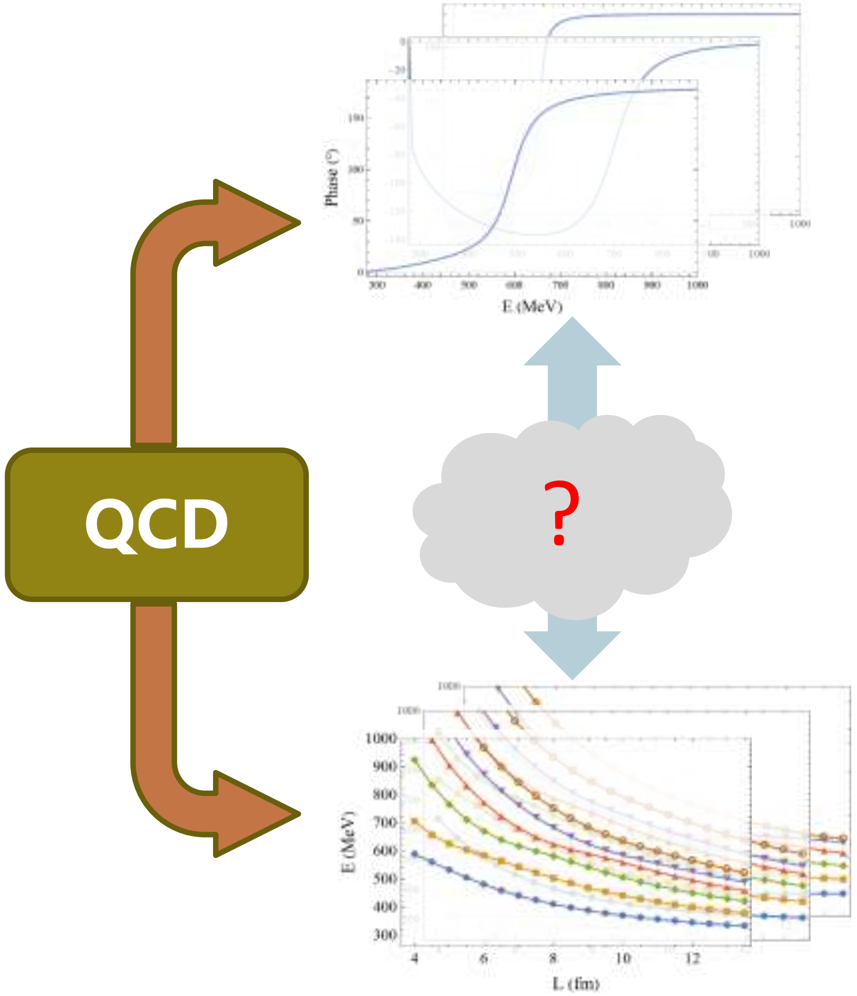
EfficientNet, arXiv:1905.11946

Possible Questions from Physicists



- 10^nparameters?! That's Toooooo Many!
    - Explainability/Interpretability, can be partially explained.
    - Personal comment at the end of this talk.
- Difference between NN and fitting?
    - Fundamentally the same but somehow not that trivial
- Why bother?
    - **Vague** idea becomes **solid**
    - In the spirit of Duck Test -> **NN≈Underline Function**

Motivations



- QCD is hard
  - Phenomenological models/ ChiPT etc.
  - LQCD
- Is there a model-independent link between model-dependent quantities?
  - Exceedingly hard
  - If you know LF, you already know an positive example

Hamiltonian Effective Field Theory(HEFT) & Data Generation

J.J. Wu *et al.* Phys.Rev.C.90.055206

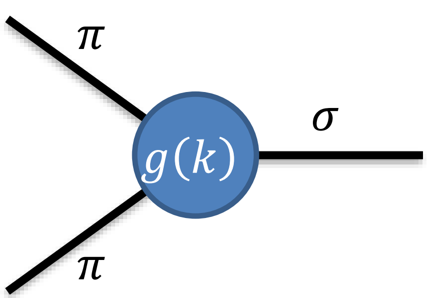
$\pi\pi \to \pi\pi$, s wave elastic scattering

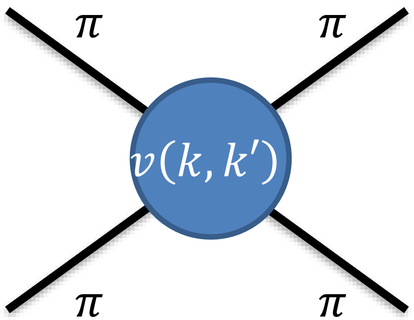
$$H = H_0 + H_I$$

$$H_0 = |\sigma\rangle m_\sigma \langle\sigma| + \int d\boldsymbol{k} \left(|\boldsymbol{k}\rangle \sqrt{m_\pi^2 + k^2} \langle \boldsymbol{k}| \right)$$

$$H_I = \int d\boldsymbol{k} (|\boldsymbol{k}\rangle\, g(k) \langle\sigma| + h.c.) + \int d\boldsymbol{k} d\boldsymbol{k}' |\boldsymbol{k}\rangle\, v(k,k') \langle \boldsymbol{k}'|$$

$g(k) \propto f(k), v(k,k') \propto f^2(a,k) f^2(a,k')$



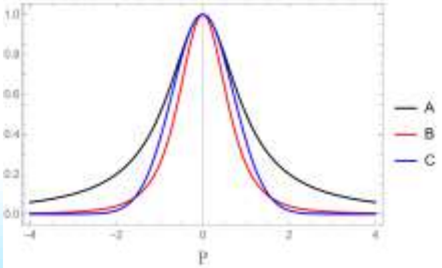


$f(a,k) =$

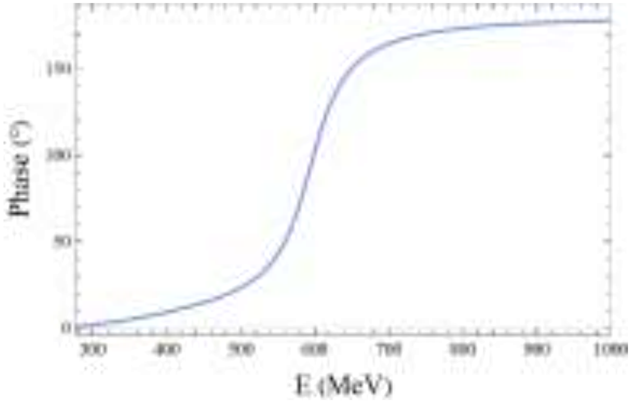
$A: \left(1 + (ak)^2\right)^{-1}$

$B: \left(1 + (ak)^2\right)^{-2}$,

$C: \; e^{-(ak)^2}$

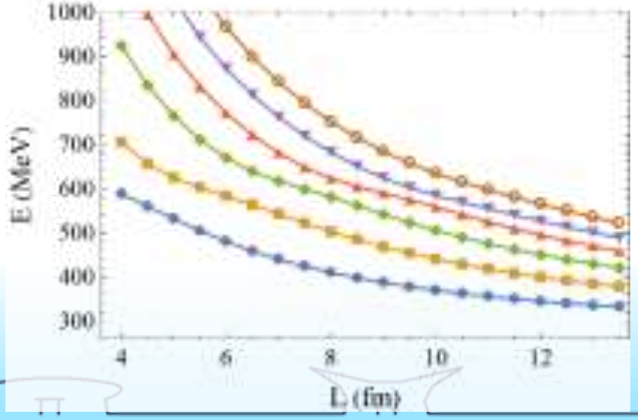


Lippmann-Schwinger equation $\to T \to \delta(E)$



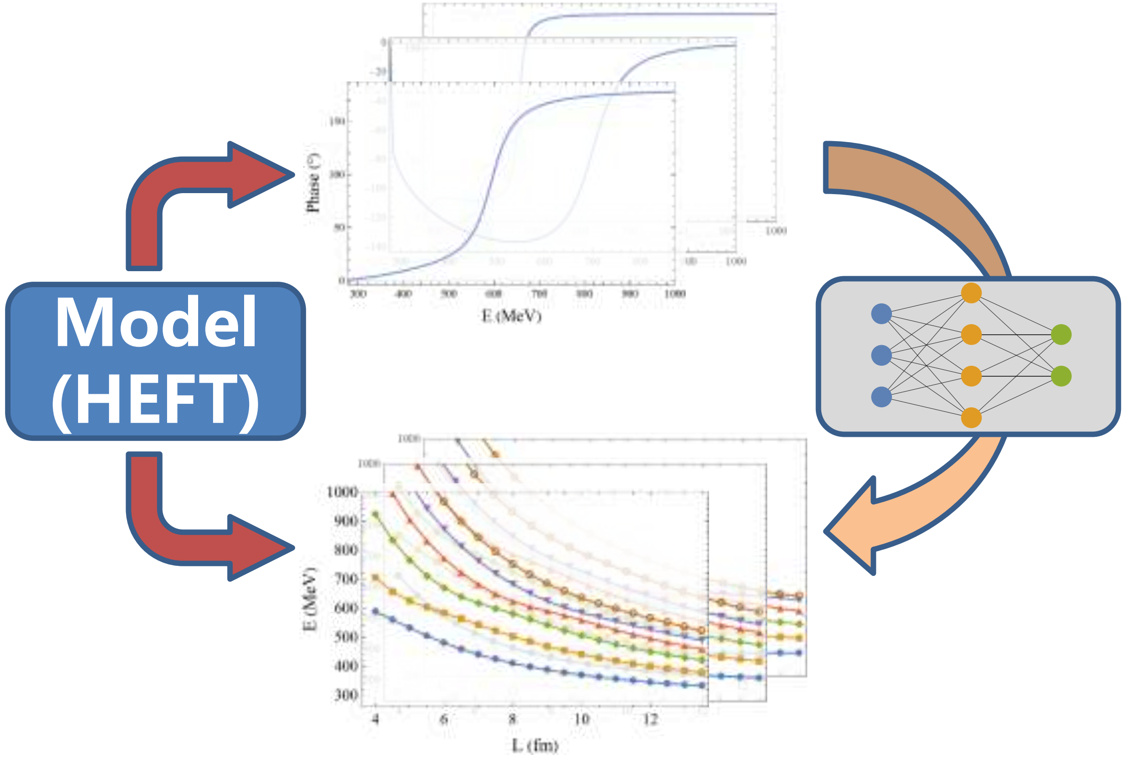
Discretization, Eigenfunction:

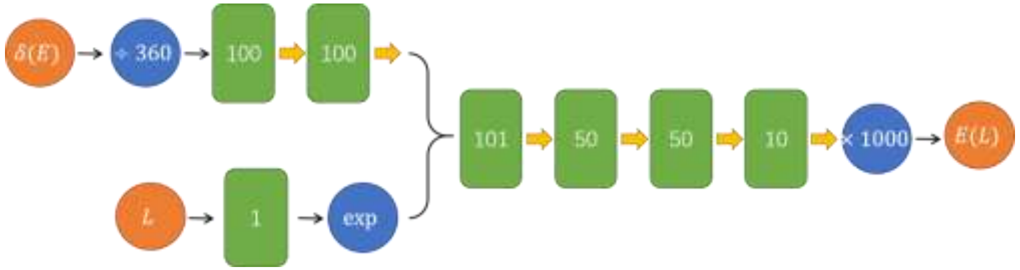
$H(L)|\psi\rangle = E(L)|\psi\rangle$



Hamiltonian Effective Field Theory(HEFT) & Data Generation

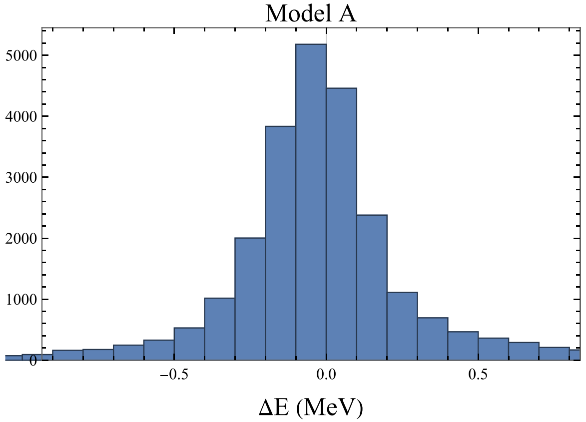
J.J. Wu *et al.* Phys.Rev.C.90.055206

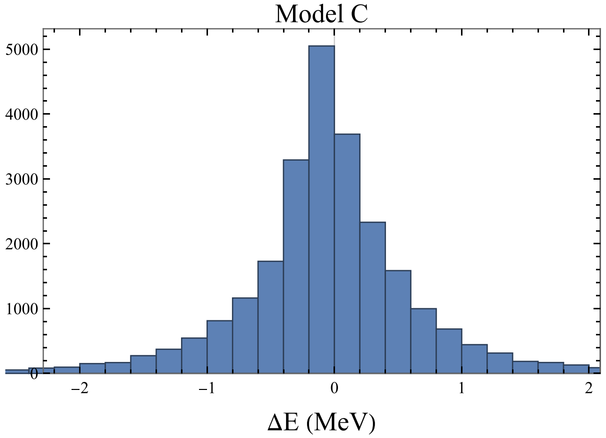


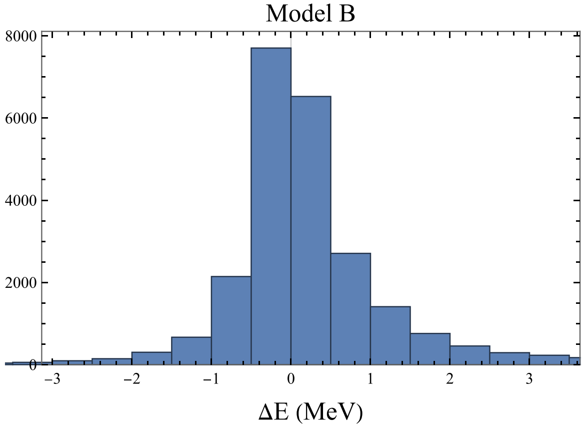


- $\delta(E)$ contains the full information
- SoftPlus Not ReLU
- **Lowest** 10 Energy levels
- LossFunction: mean square error
- 2500 $\delta(E)$ for each model, batch size 10^4, 4*10^4 epoch

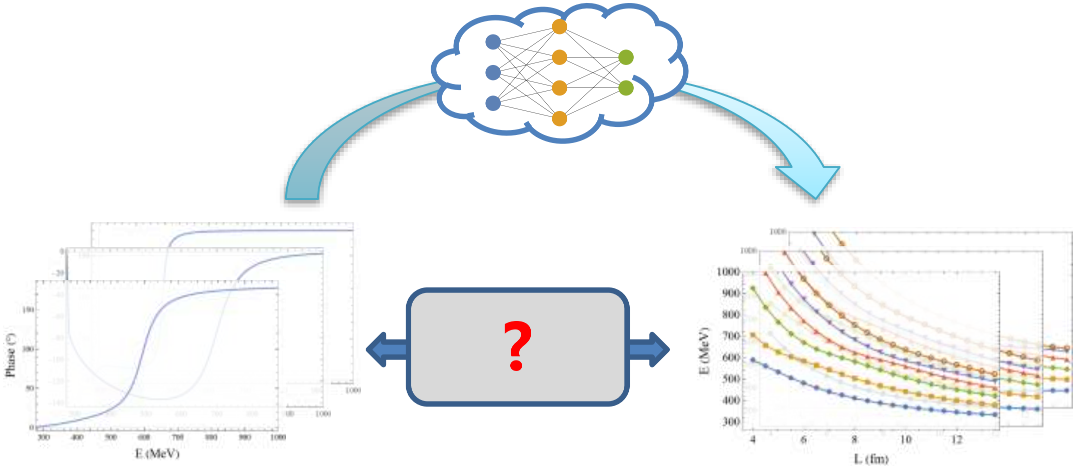
Result Analysis







$$\Delta E \coloneqq E_{model} - E_{NN}$$



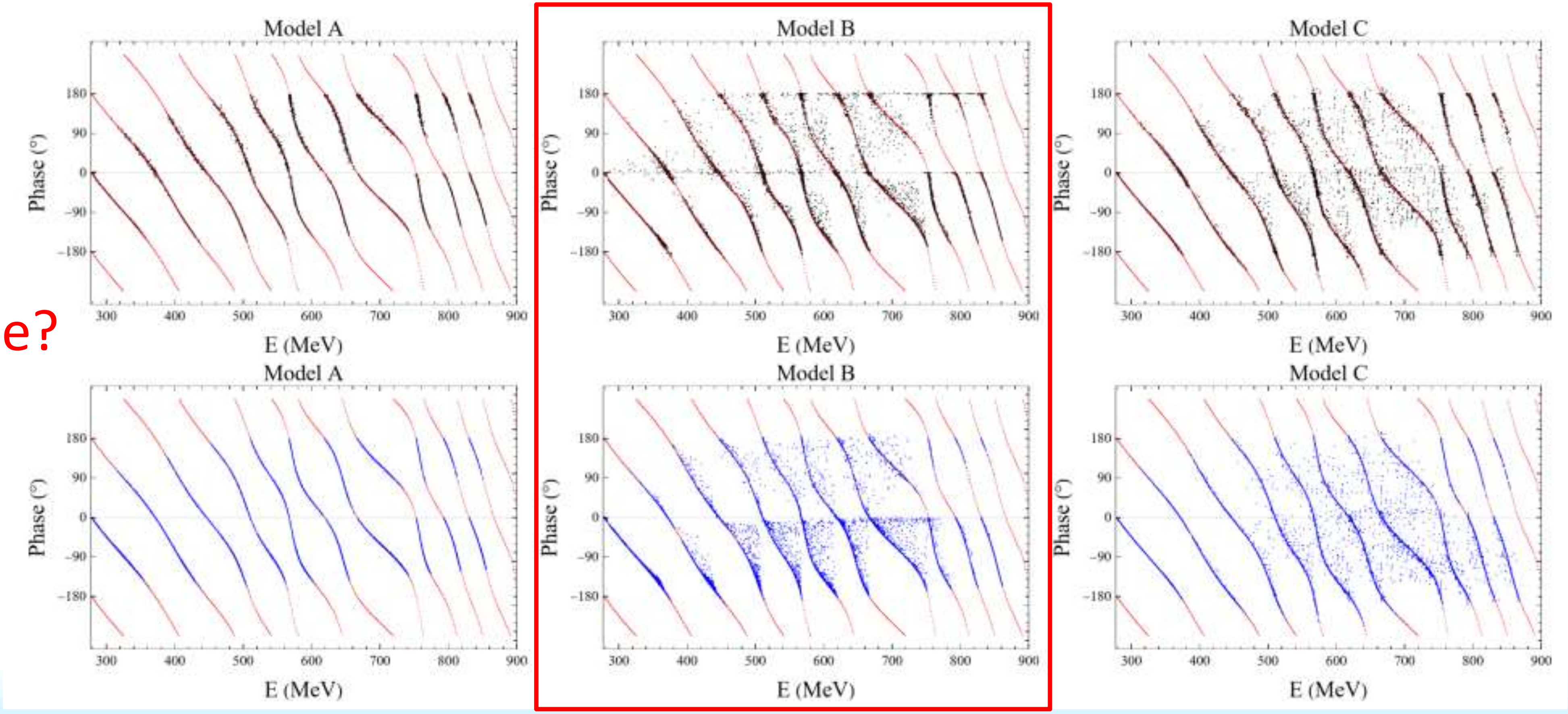
- Decently trained on model A, C
  - $\Delta E(L) < 1$MeV, $E(L) \sim 300 - 900$ MeV
- For model B,
  - as a test set, slightly worse
  - ΔE has heavier-tail on the right
- Signifies the existence of link
- Under the hood, LF is in charge
- -> Check against LF

- LF is model-independent
- Check $\delta(E_L)$ against LF
- NN tries to collect the points towards LF
- UAT-> NN captures model-independent link (to some degree)

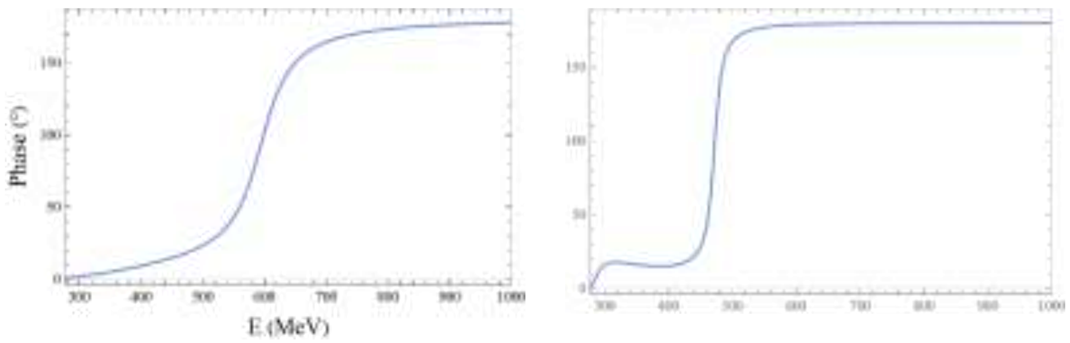
$$\delta(E) = \arctan\left(q\,\frac{\pi^{3/2}}{\mathcal{L}_{00}(1;q^2)}\right) + n\pi \qquad \mathcal{L}_{00}(1;q^2) = \frac{1}{\sqrt{4\pi}}\sum_{\vec{n}}\left(\vec{n}^2 - q^2\right)^{-1}$$

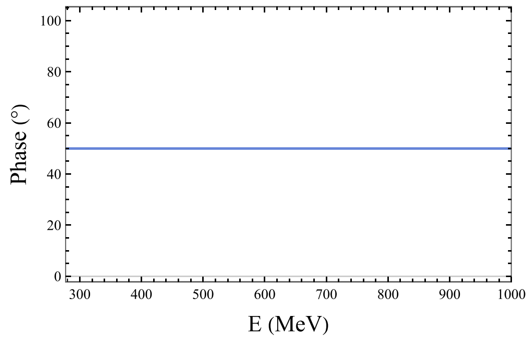
LF NN Model

Stronger Evidence?

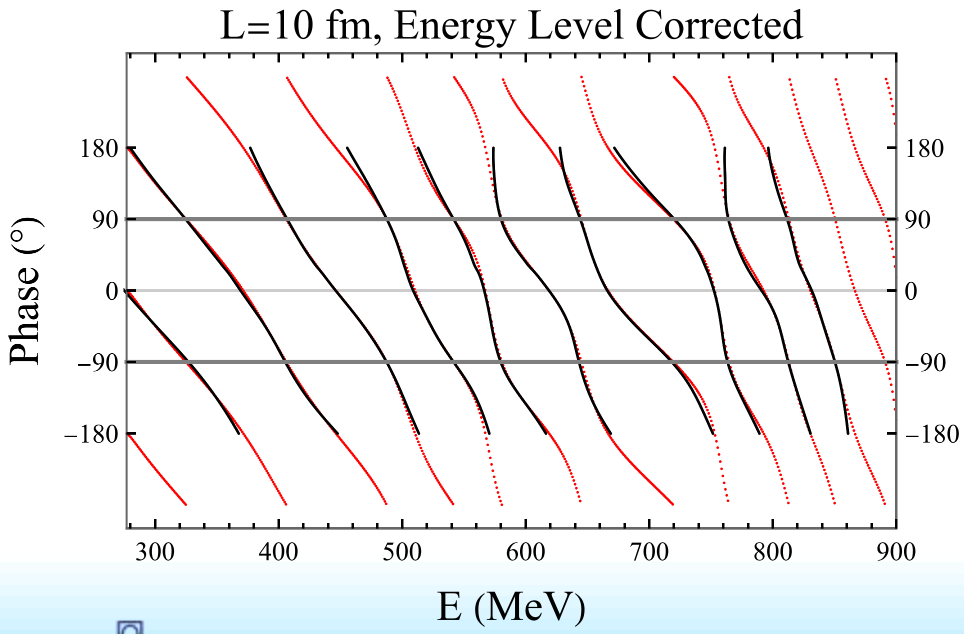


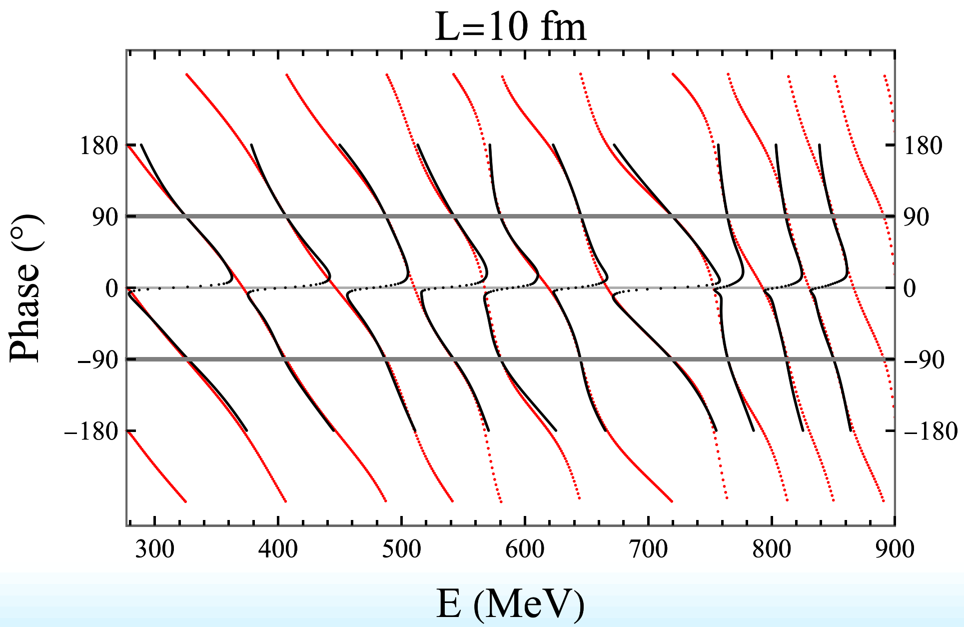
- Go Far beyond training set & challenge the NN with constant $\delta(E)$





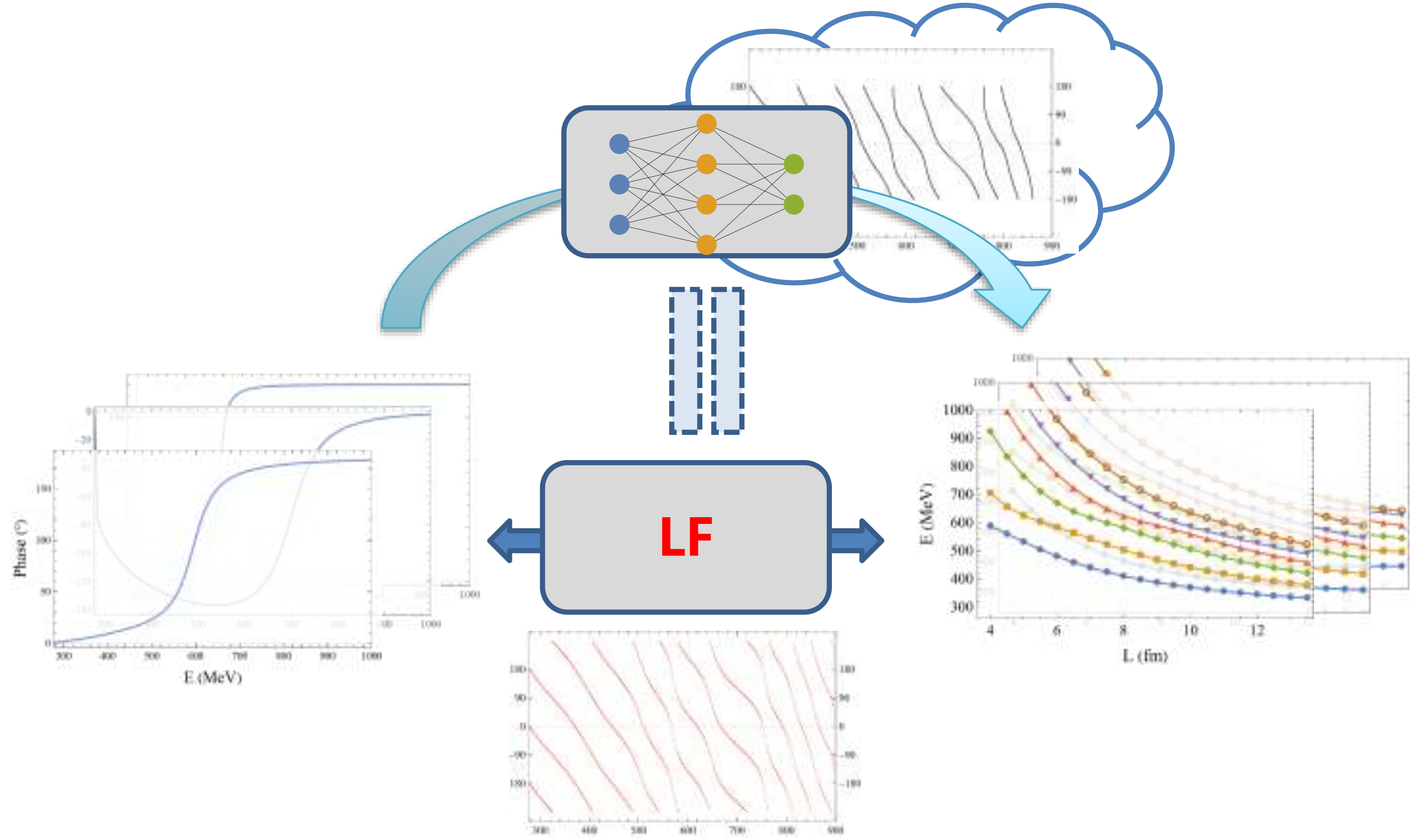
Even closer to LF!





Summary

- Even $\delta(E), E(L)$ are both model dependent,
  NN can extract model-independent link (LF) when $\delta(E) \rightarrow E(L)$



$$\text{LF} + o\left(e^{-mL}\right) \rightarrow E(L)$$

Where there is a link,
there is a neural network :)

Outlook

Sensible mathematics involves neglecting a quantity when it is small

not neglecting it just because it is infinitely great and you do not want it!

→ ★ Renormalization

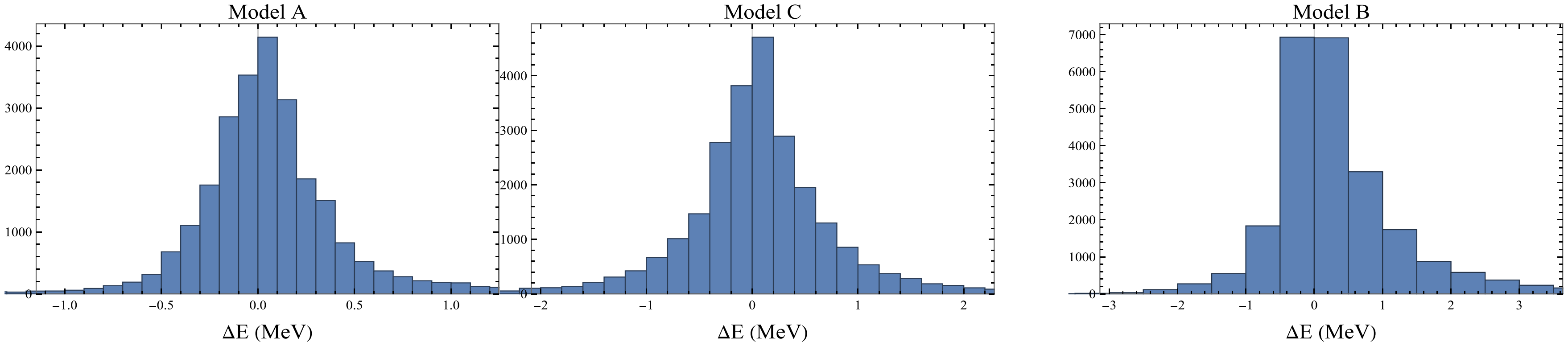
With four parameters I can fit an elephant,

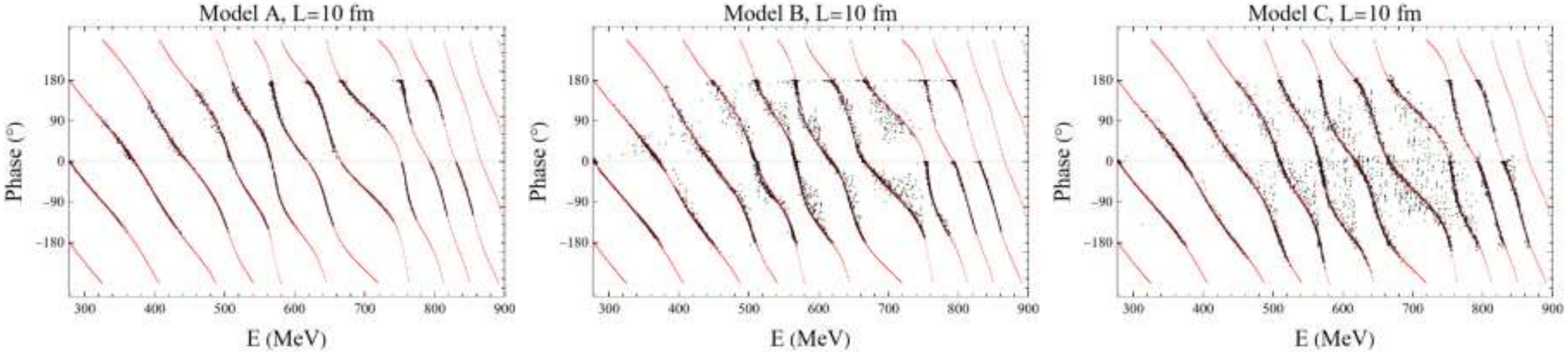
and with five I can make him wiggle his trunk

→

BackUp

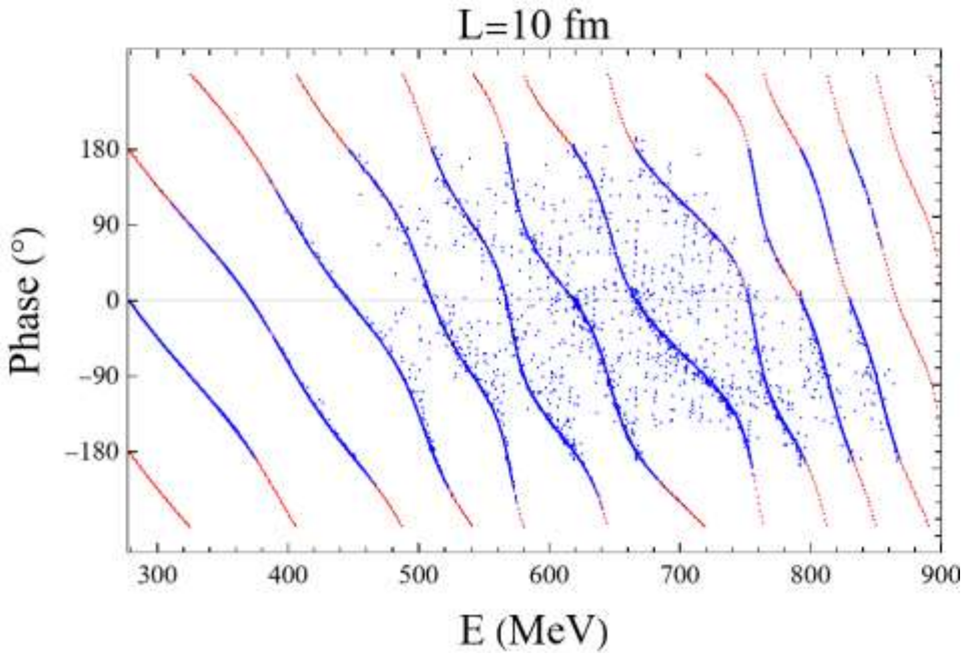
Results after level correction

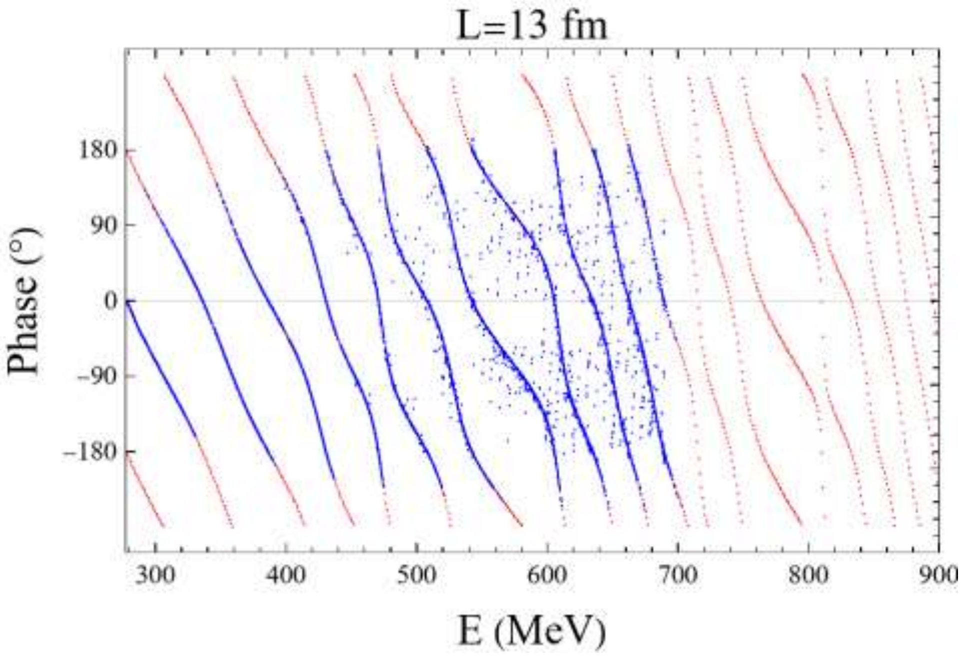




$\delta(E)$ is evenly sampled by 100 points in $[2m_\pi, 1\text{GeV}]$

$m_\sigma \in [350,700], a \rightarrow c, d \in [0.5,2]$





$$\delta(E) = \arctan\left(\mathrm{q}\frac{\pi^{3/2}}{\mathcal{L}_{00}(1;q^2)}\right) + n\pi$$

$$\mathcal{L}_{00}(1;q^2) = \frac{1}{\sqrt{4\pi}}\sum_{\vec{n}}\left(\vec{n}^2 - q^2\right)^{-1}$$