Patching Solutions with Machine Learning

Jiheon Lee
KIAS

In collaboration with
Doyoung Kim, Donghee Lee, Hye-Sung Lee, Jaeok Yi
[arXiv:2601.05889]

$$X'' + \left[k^2 + a^2 m_X^2 - \frac{k^2 \mathcal{H}^2}{k^2 + a^2 m_X^2} \left((n+2)(n+1) - 3(n+1)^2 \frac{a^2 m_X^2}{k^2 + a^2 m_X^2} \right) \right] X = 0$$

$$m_X \propto a^n$$

$$X'' + \left[k^2 + a^2 m_X^2 - \frac{k^2 \mathcal{H}^2}{k^2 + a^2 m_X^2} \left((n+2)(n+1) - 3(n+1)^2 \frac{a^2 m_X^2}{k^2 + a^2 m_X^2} \right) \right] X = 0$$

$$\text{For} \quad am_X \ll k \ : \ \left[k^2 - (n+2)(n+1)\mathcal{H}^2 \right] X = 0$$

$$\text{For} \quad k \ll am_X \ll \sqrt{\sqrt{(n+1)(2n+1)}k\mathcal{H}} \ : \ \left[\frac{(n+2)(2n+1)k^2\mathcal{H}^2}{a^2 m_X^2} \right] X = 0$$

$$X'' + \left[k^2 + a^2 m_X^2 - \frac{k^2 \mathcal{H}^2}{k^2 + a^2 m_X^2}\left((n+2)(n+1) - 3(n+1)^2 \frac{a^2 m_X^2}{k^2 + a^2 m_X^2}\right)\right] X = 0$$

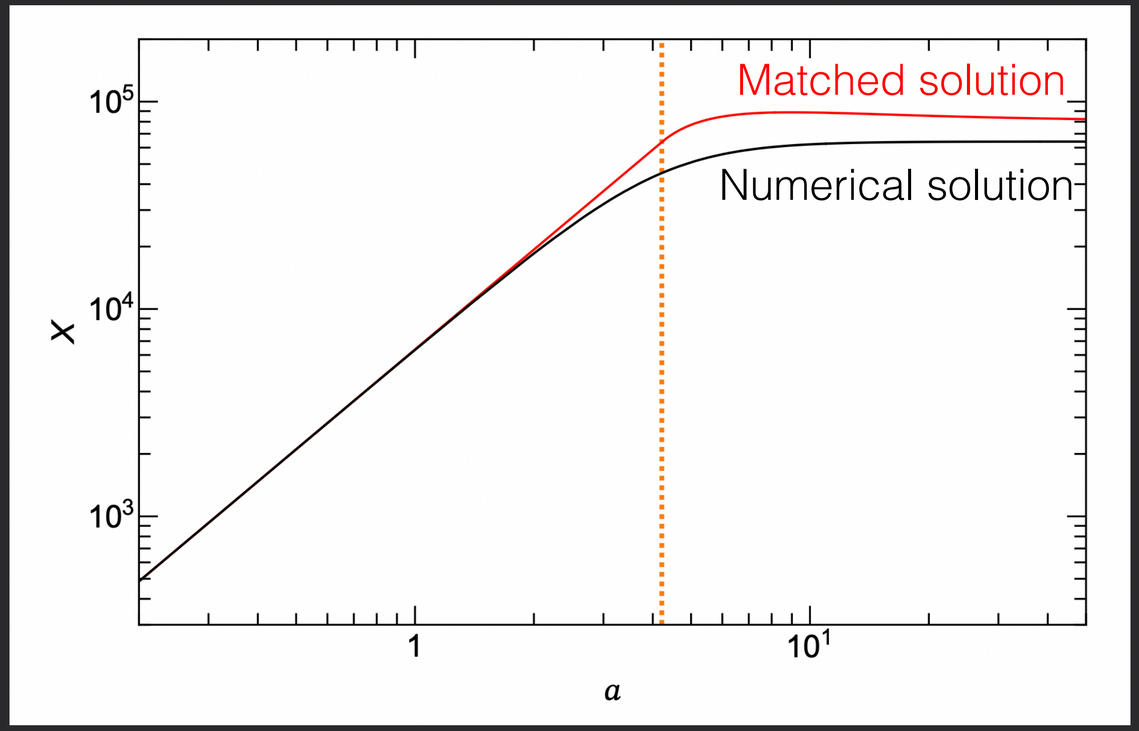
For $\quad am_X \ll k \ :$

$$X = \sqrt{\frac{\pi}{4k}} e^{i\pi(2n+1)/4} \sqrt{\frac{k}{\mathcal{H}}} H^{(1)}_{n+3/2}\left(\frac{k}{\mathcal{H}}\right)$$

For $\quad k \ll am_X \ll \sqrt{\sqrt{(n+1)(2n+1)}k\mathcal{H}} \ :$

$$X = a^{-1/2}\left[c_1 J_{1/2(n+1)}\left(\sqrt{\frac{2n+1}{n+1}}\frac{k}{am_X}\right) + c_2 J_{-1/2(n+1)}\left(\sqrt{\frac{2n+1}{n+1}}\frac{k}{am_X}\right)\right]$$

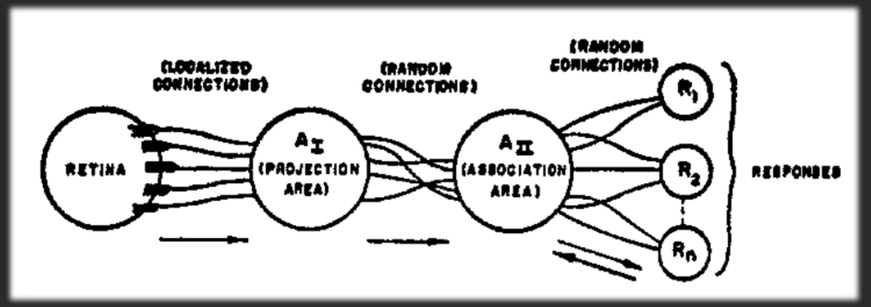
Manual matching?



$$X = \sqrt{\frac{\pi}{4k}} e^{i\pi(2n+1)/4} \sqrt{\frac{k}{\mathcal{H}}} H^{(1)}_{n+3/2}\left(\frac{k}{\mathcal{H}}\right)$$

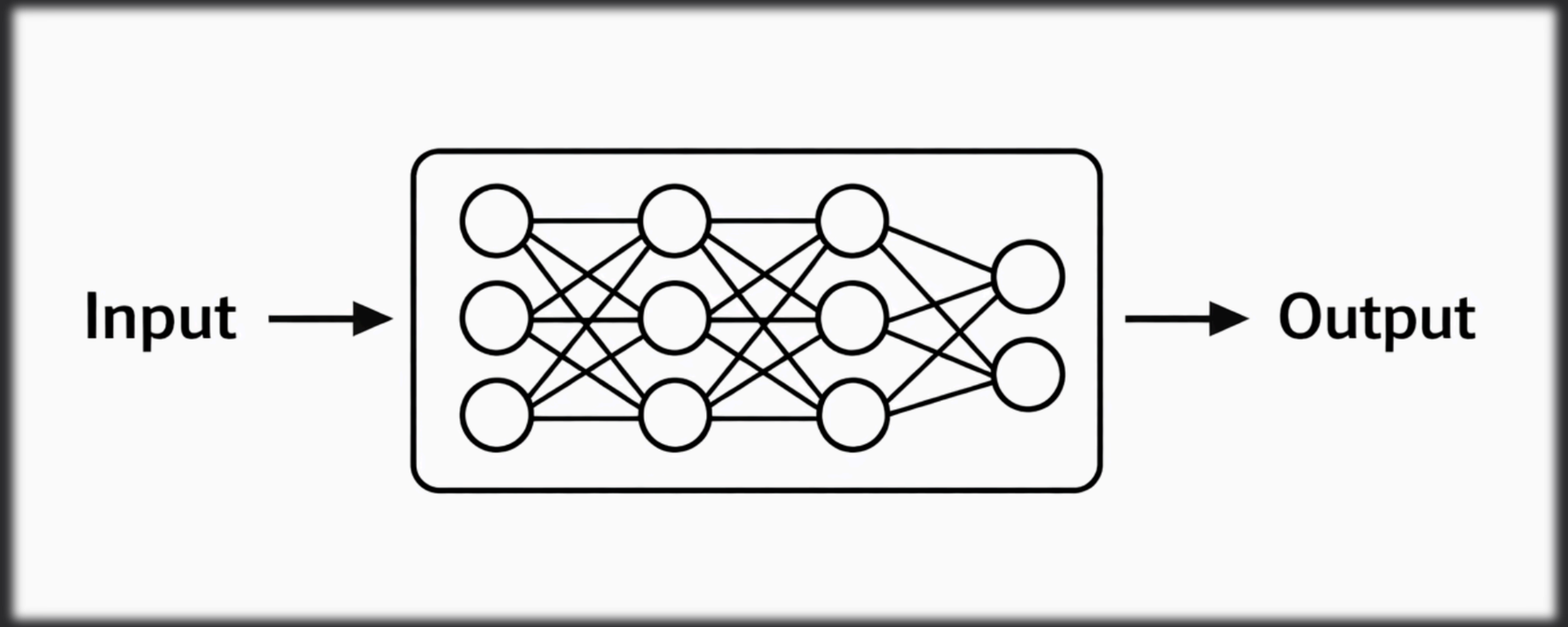
$$X = a^{-1/2}\left[c_1 J_{1/2(n+1)}\left(\sqrt{\frac{2n+1}{n+1}}\frac{k}{am_X}\right) + c_2 J_{-1/2(n+1)}\left(\sqrt{\frac{2n+1}{n+1}}\frac{k}{am_X}\right)\right]$$

Machine Learning



F.Rosenblatt (1958)

ChatGPT

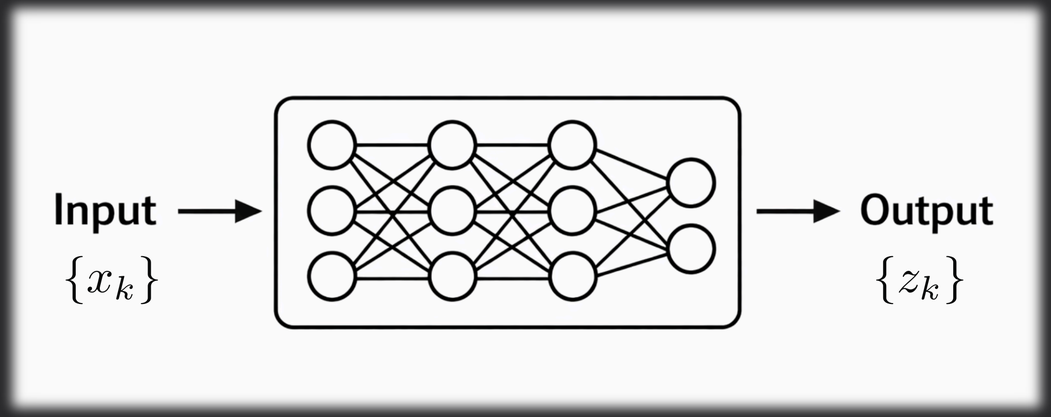


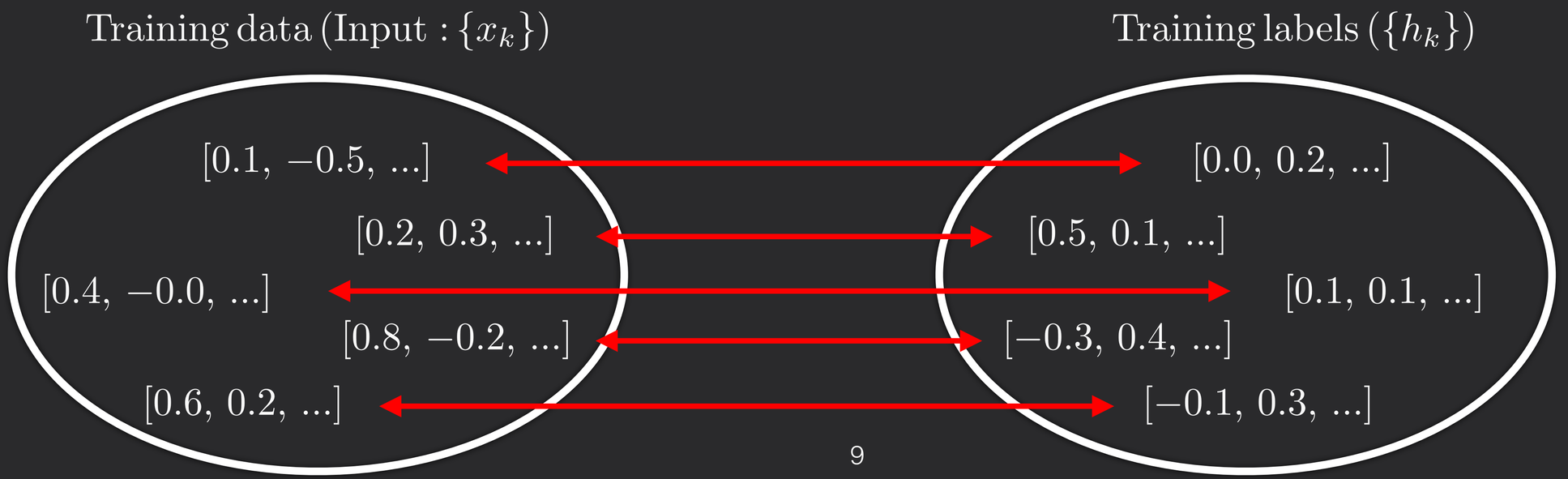
$$\begin{pmatrix} (W_\ell)_{11} & (W_\ell)_{12} & \cdots & (W_\ell)_{1m} \\ (W_\ell)_{21} & (W_\ell)_{22} & \cdots & (W_\ell)_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ (W_\ell)_{n1} & (W_\ell)_{n2} & \cdots & (W_\ell)_{nm} \end{pmatrix} \begin{pmatrix} (y_\ell)_1 \\ (y_\ell)_2 \\ \vdots \\ (y_\ell)_m \end{pmatrix} + \begin{pmatrix} (b_\ell)_1 \\ (b_\ell)_2 \\ \vdots \\ (b_\ell)_n \end{pmatrix} = \begin{pmatrix} (y'_\ell)_1 \\ (y'_\ell)_2 \\ \vdots \\ (y'_\ell)_n \end{pmatrix}$$

$$\begin{pmatrix} \sigma[(y'_\ell)_1] \\ \sigma[(y'_\ell)_2] \\ \vdots \\ \sigma[(y'_\ell)_n] \end{pmatrix} = \begin{pmatrix} (y_{\ell+1})_1 \\ (y_{\ell+1})_2 \\ \vdots \\ (y_{\ell+1})_n \end{pmatrix}$$

(Internal parameters)

Training

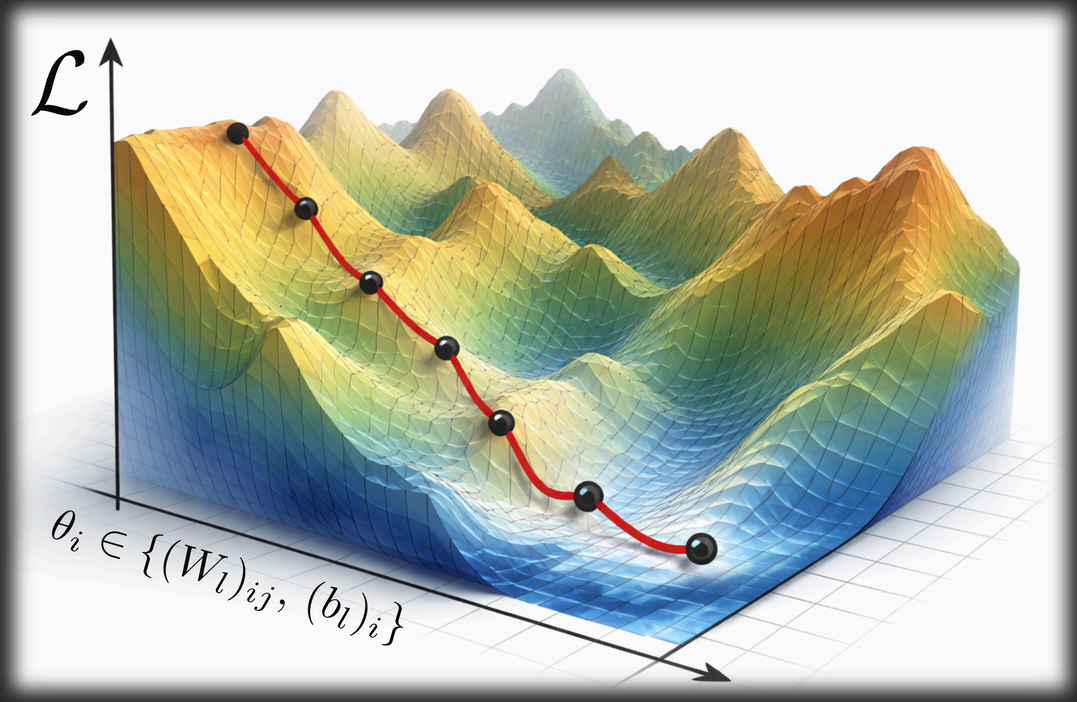

Input
$\{x_k\}$
Output
$\{z_k\}$


Training data (Input : $\{x_k\}$)
Training labels ($\{h_k\}$)
[0.1, −0.5, ...]
[0.2, 0.3, ...]
[0.4, −0.0, ...]
[0.8, −0.2, ...]
[0.6, 0.2, ...]
[0.0, 0.2, ...]
[0.5, 0.1, ...]
[0.1, 0.1, ...]
[−0.3, 0.4, ...]
[−0.1, 0.3, ...]

Minimizing the errors

$$\mathcal{L} = \sum_k ||z_k - h_k||^2$$

$$\{\delta(W_l)_{ij}, \delta(b_l)_i\} = \eta \nabla \mathcal{L}((W_l)_{ij}, b_i)$$



Physics Informed Neural Network

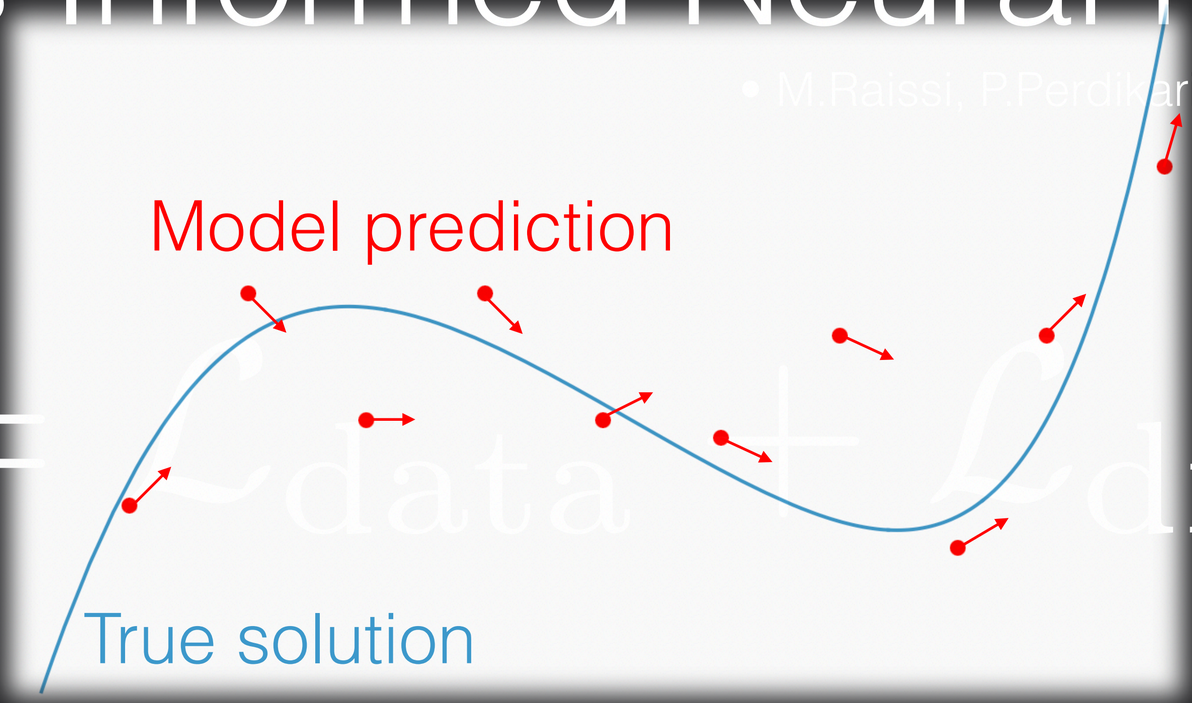
- M.Raissi, P.Perdikaris, G.Karniadakis(2019)

$$\mathcal{L} = \mathcal{L}_{\mathrm{data}} + \mathcal{L}_{\mathrm{diff}}$$

$$D[f(x)] = 0 \qquad \mathcal{L}_{\mathrm{diff}} = \sum_{k} |D[g(x)]|^2$$

Physics Informed Neural Network

- M.Raissi, P.Perdikaris, G.Karniadakis(2019)



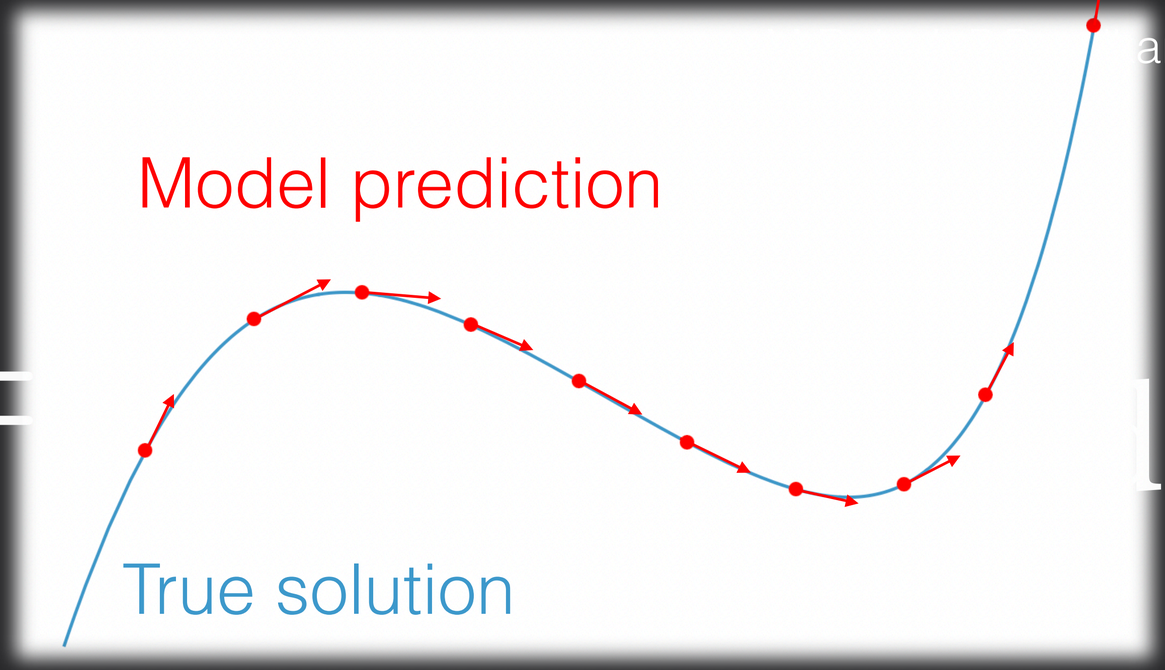
$$\mathcal{L} = \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{diff}}$$

$$\mathcal{L}_{\text{diff}} = \sum_k |D[g(x)]|^2$$

Physics Informed Neural Network

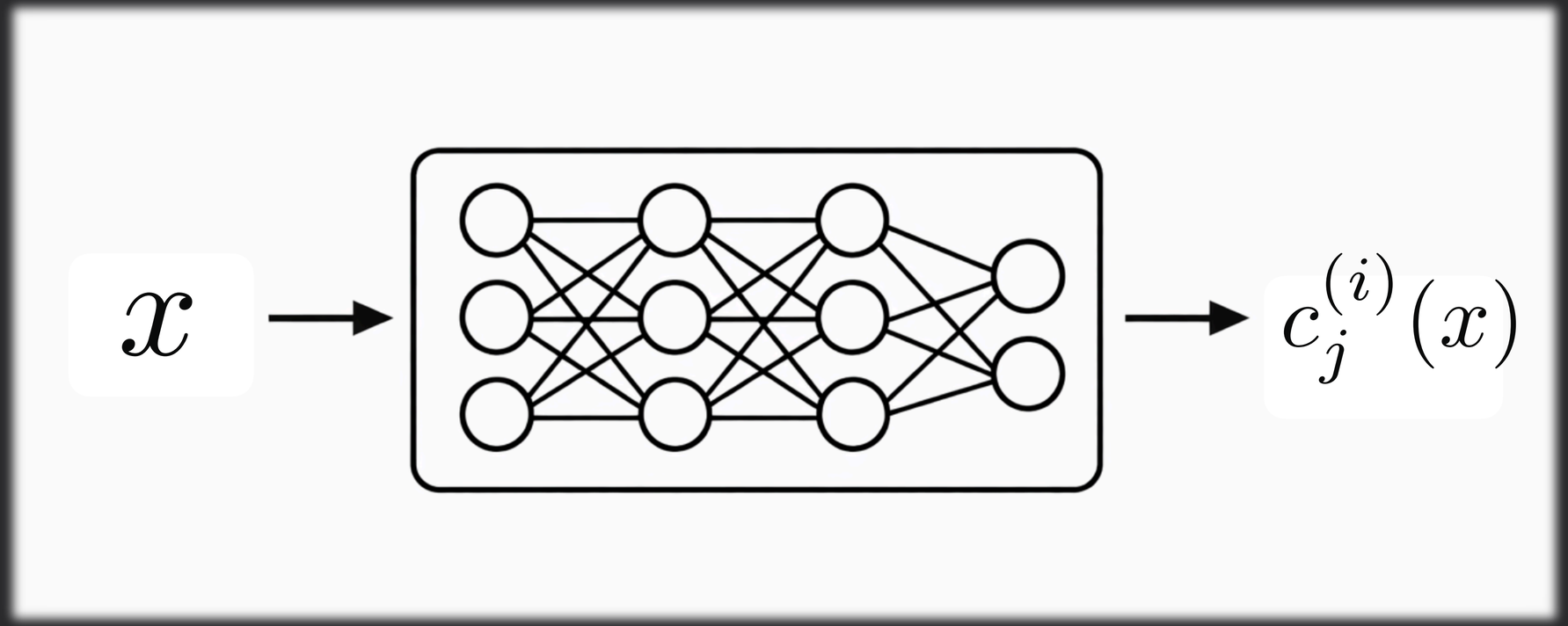
aris, G.Karniadakis(2019)

$$\mathcal{L} = \mathcal{L}_{\text{diff}}$$



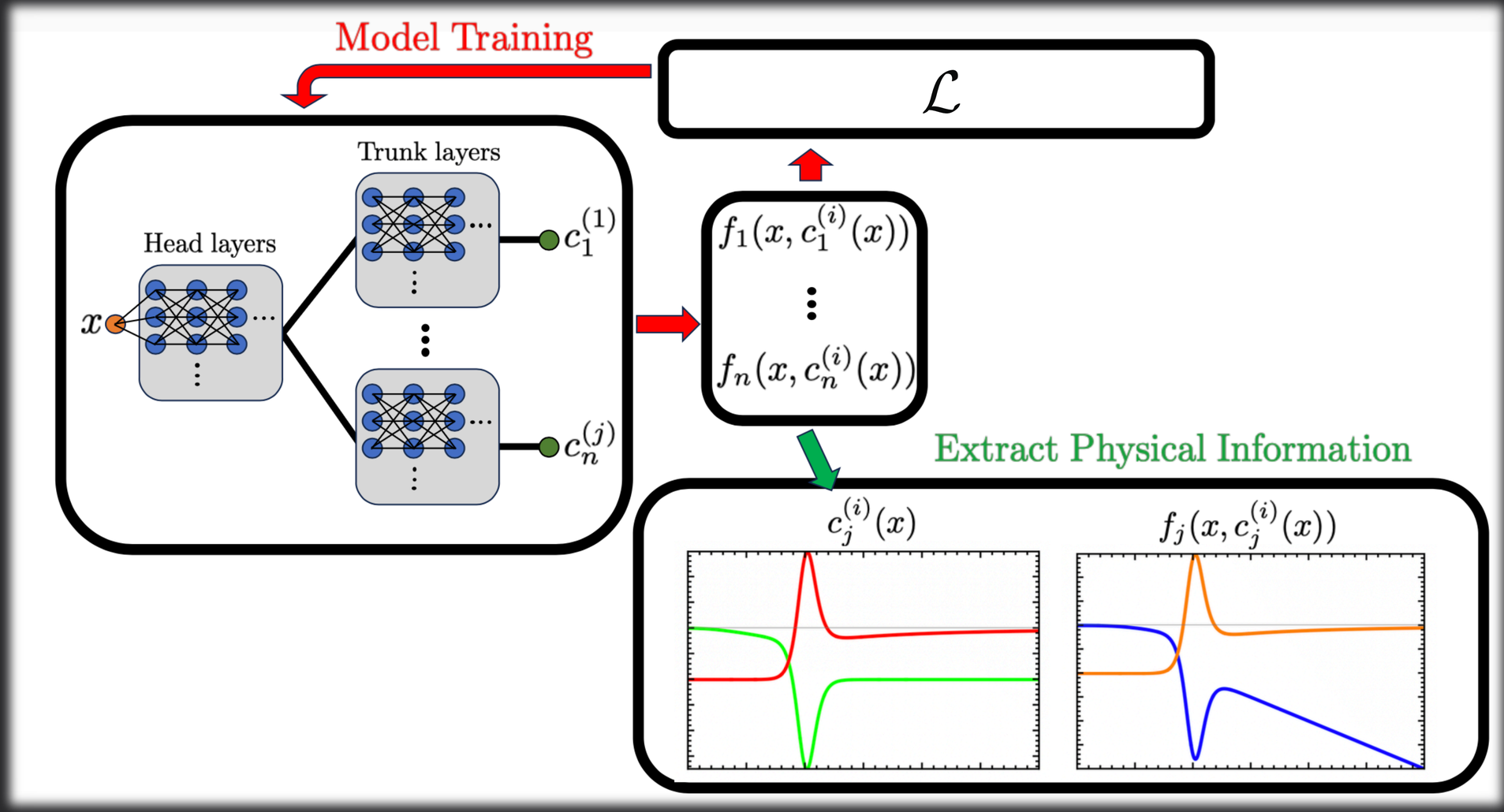
$$\mathcal{L}_{\text{diff}} = \sum_{k} |D[g(x)]|^2$$

GlueNN



$$g(x) \equiv \sum_j f_j(x, c_j^{(i)}(x))$$

GlueNN


Model Training
$\mathcal{L}$
Trunk layers
Head layers
x
$c_1^{(1)}$
$c_n^{(j)}$
$f_1(x, c_1^{(i)}(x))$
$f_n(x, c_n^{(i)}(x))$
Extract Physical Information
$c_j^{(i)}(x)$
$f_j(x, c_j^{(i)}(x))$

GlueNN

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \mathcal{L}_{\text{diff}} + \mathcal{L}_{\text{patch}}$$

$$\mathcal{L}_{\text{data}} = \frac{\lambda_a}{N_a} \sum_{x \in \{x_\alpha\}} \left| \sum_{i,j} f_j\left(x, c_j^{(i)}(x)\right) - h(x) \right|^2$$

$$\mathcal{L}_{\text{diff}} = \frac{\lambda_b}{N_b} \sum_{x \in \{x_\beta\}} \left| \frac{\mathcal{D}\left[\sum_{i,j} f_j\left(x, c_j^{(i)}(x)\right)\right]}{\mathcal{N}\left[\sum_{i,j} f_j\left(x, c_j^{(i)}(x)\right)\right]} \right|^2$$

$$\mathcal{L}_{\text{patch}} = \frac{\lambda_c}{N_c} \sum_{x \in \{x_\gamma\}} \left| \sum_{i} f_j\left(x, c_j^{(i)}(x)\right) \right|^2$$

Ex1. Freeze out

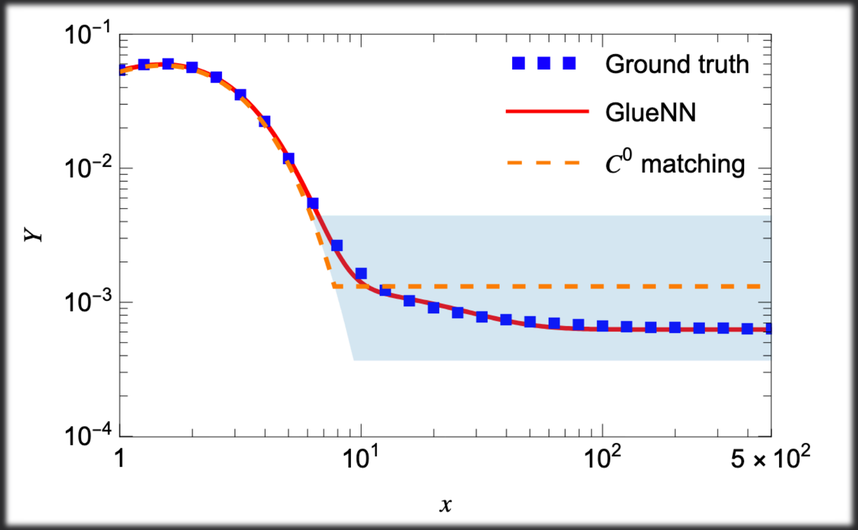
$$\mathcal{D}[Y(x)] = \frac{dY}{dx} + \eta x^{-2}(Y^2 - Y_{\rm eq}^2) = 0 \qquad ; x \equiv m/T \\ Y \equiv n/s$$

Asymptotic solutions :

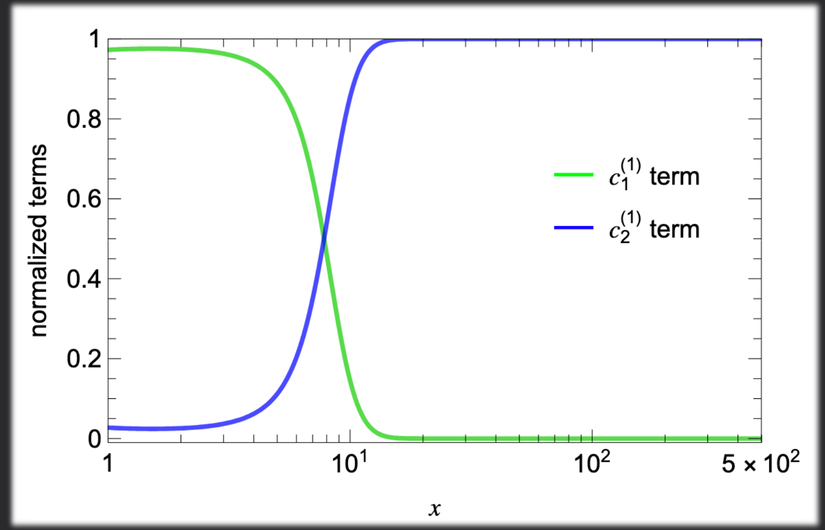
$$Y(x) = \begin{cases} c_1^{(1)} x^{3/2} e^{-x} & \text{for} \quad x < 1 \\ c_2^{(1)} & \text{for} \quad x > 1 \end{cases}$$

Ex1. Freeze out

$$\mathcal{D}[Y(x)] = \frac{dY}{dx} + \eta x^{-2}(Y^2 - Y_{\rm eq}^2) = 0$$



Set	Interval	Points (N)
$\{x_\alpha\}$	$[1.0, 1.9]$	28
$\{x_\beta\}$	$[7.8, 31]$	60



Ex2. Vector production

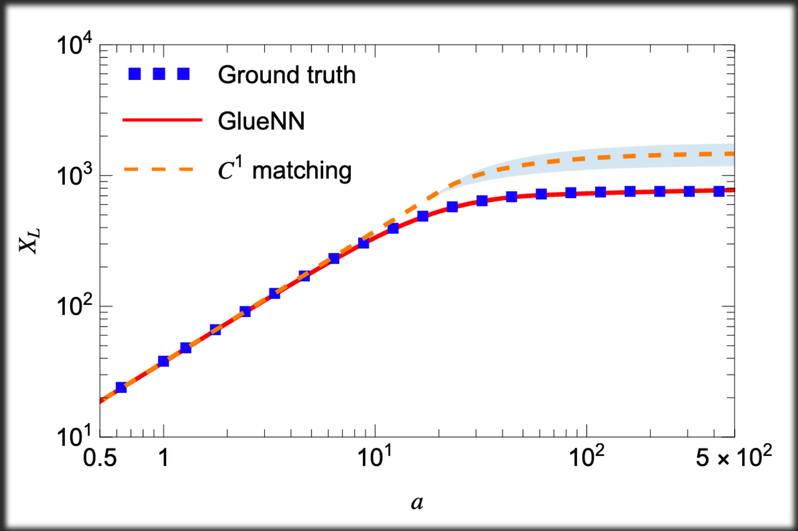
$$\mathcal{D}\left[X_L(a)\right] = \frac{d^2 X_L}{da^2} + \frac{2}{a}\frac{dX_L}{da} + \left\{ \frac{k^2}{H^2 a^4} + \frac{m^2}{H^2 a^2} - \frac{k^2}{k^2 + m^2 a^2}\left(\frac{2}{a^2} - \frac{3m^2}{k^2 + m^2 a^2} \right) \right\} X_L = 0$$

Asymptotic solutions :

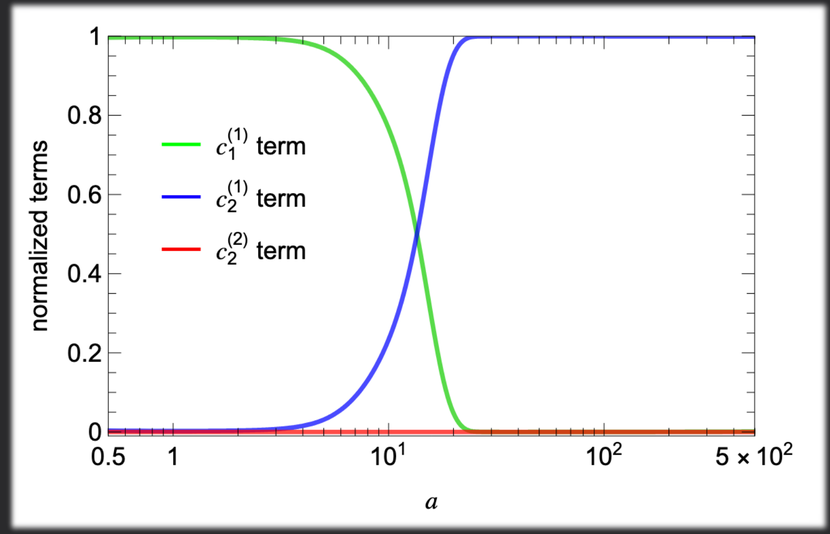
$$X_L(a) = \begin{cases} c_1^{(1)} H_{3/2}^{(1)}\left(\dfrac{k}{aH}\right) + c_1^{(2)} H_{3/2}^{(2)}\left(\dfrac{k}{aH}\right) & \text{for} \quad a \ll k/m, \\ c_2^{(1)} J_{-1/2}\left(\dfrac{k}{ma}\right) + c_2^{(2)} J_{1/2}\left(\dfrac{k}{ma}\right) & \text{for} \quad a \gg k/m, \end{cases}$$

Ex2. Vector production

$$\mathcal{D}\left[X_L(a)\right] = \frac{d^2 X_L}{da^2} + \frac{2}{a}\frac{dX_L}{da} + \left\{ \frac{k^2}{H^2 a^4} + \frac{m^2}{H^2 a^2} - \frac{k^2}{k^2 + m^2 a^2}\left(\frac{2}{a^2} - \frac{3m^2}{k^2 + m^2 a^2} \right) \right\} X_L = 0$$



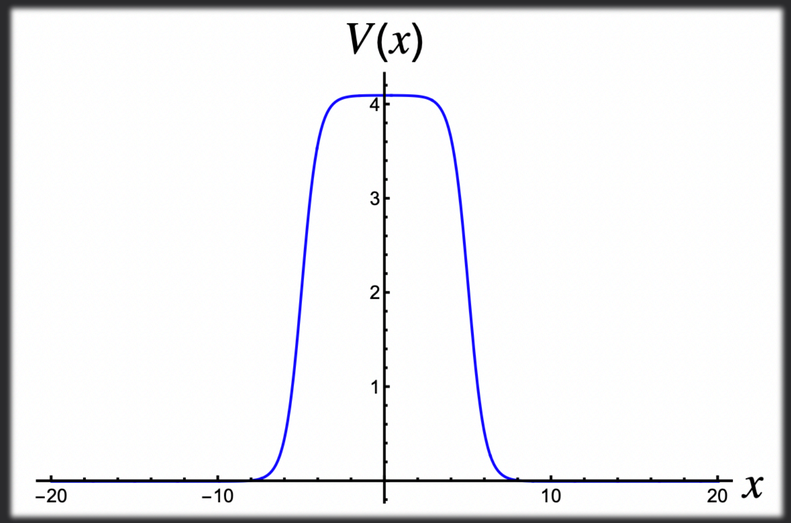
Set	Interval	Points (N)
$\{a_\alpha\}$	$[0.1, 1.0]$	10
$\{a_\beta\}$	$[1.0, 124]$	650
$\{a_\gamma\}$	$[23, 500]$	382



EX3. Quantum tunneling

$$\mathcal{D}[\psi(x)] = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi - E\psi = 0$$

$$V(x) = V_0\left(\frac{1}{e^{-(x+d/2)/\sigma}+1} + \frac{1}{e^{(x-d/2)/\sigma}+1} - 1\right)$$

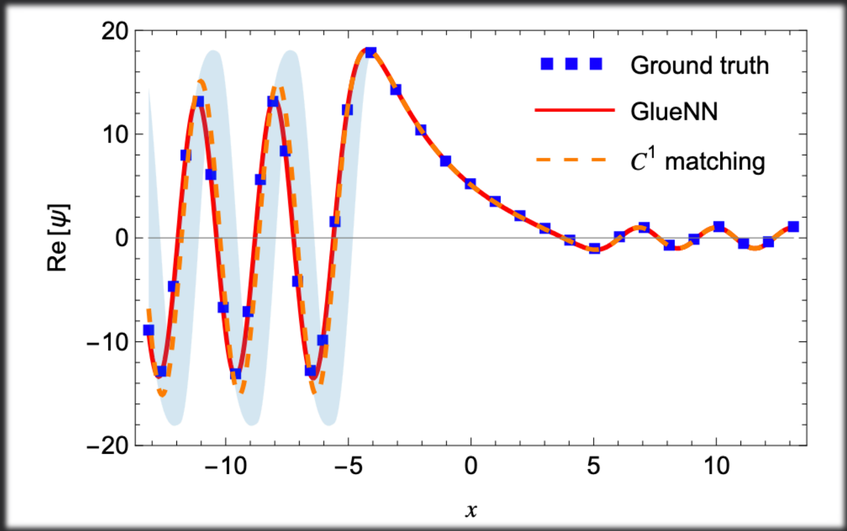


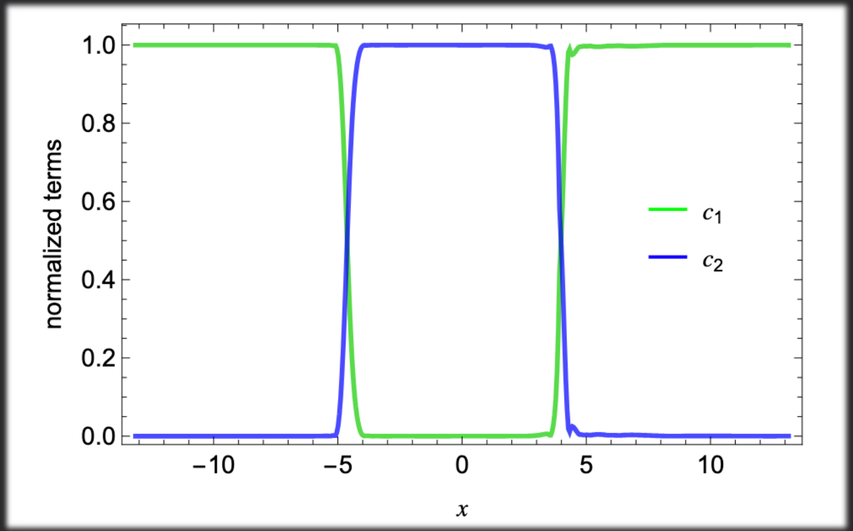
$$\mathrm{Re}[\psi(x)] = \begin{cases} c_1^{(1)}\cos(kx) + c_1^{(2)}\sin(kx) & \text{for} \quad |x| \gg d/2 \\ c_2^{(1)}e^{\kappa x} + c_2^{(2)}e^{-\kappa x} & \text{for} \quad |x| \ll d/2 \end{cases}$$

$$\mathrm{Im}[\psi(x)] = \begin{cases} c_3^{(1)}\cos(kx) + c_3^{(2)}\sin(kx) & \text{for} \quad |x| \gg d/2 \\ c_4^{(1)}e^{\kappa x} + c_4^{(2)}e^{-\kappa x} & \text{for} \quad |x| \ll d/2 \end{cases}$$

EX3. Quantum tunneling

$$\mathcal{D}[\psi(x)] = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V(x)\psi - E\psi = 0$$





Set	Interval	Points (N)
	Real Part Training	
$\{x_\alpha\}$	$[6.05, 13.14]$	140
$\{x_\beta\}$	$[-13.14, 11.62]$	490
$\{x_\gamma\}$	$[-4.08, 4.03]$	160
$\{x_\delta\}$	$[-13.14, -5.09] \cup [4.03, 13.14]$	340

- Pull the physics out of the black box