Tensor Weights-NN and Fermionic QFT

Guojun Huang

The Chinese University of Hong Kong, Shenzhen,

August 22, 2025

arXiv: 2507.05303 [hep-ph]

Outline

Outline

A Brief Retrospective on NNQFT:
Basic conception

The output of the NN for α-th sampling

$$f_\alpha(\boldsymbol{x}) = \sum_{h=1}^{H} \sigma(\boldsymbol{x} W_{\mathrm{in,h}}^{(\alpha)} + b_{\mathrm{h}}^{(\alpha)}) W_{\mathrm{h,out}}^{(\alpha)}, \tag{1}$$

with the parameters sampling from the Gaussian distribution ($\mu_{\mathrm{w}} = \mu_{\mathrm{b}} = 0$)

$$W_{\mathrm{in,h}}^{(\alpha)} \sim \mathcal{N}(\mu_{\mathrm{w}}, \sigma_{\mathrm{w}}/\sqrt{D}),\ W_{\mathrm{h,out}}^{(\alpha)} \sim \mathcal{N}(\mu_{\mathrm{w}}, \sigma_{\mathrm{w}}/\sqrt{H}),\ b_{\mathrm{h}}^{(\alpha)} \sim \mathcal{N}(\mu_{\mathrm{b}}, \sigma_{\mathrm{b}}), \tag{2}$$

the correlators

$$G_{\mathrm{NN}}^{(n)}(\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n) \equiv \frac{1}{n_{\mathrm{nets}}} \sum_{\alpha=1}^{n_{\mathrm{nets}}} f_\alpha(\boldsymbol{x}_1) f_\alpha(\boldsymbol{x}_2) \cdots f_\alpha(\boldsymbol{x}_n). \tag{3}$$

Generating functional understanding

Generating functional (denote $\boldsymbol{Q}_h \equiv (W_{1,h}, W_{2,h}, ..., W_{D,h})$, $V_h \equiv b_h$, $d_{\rm out} = 1$ and $\varphi_h \equiv H * W_{\rm h,out=1}$).

$$Z_H[J] = \prod_{h=1}^{H} \left[\int \mathrm{d}\boldsymbol{Q}_h \mathrm{d}V_h \mathrm{d}\varphi_h \; P(\boldsymbol{Q}_h) P(V_h) P(\varphi_h) \right] \mathrm{e}^{\mathrm{i} \int \mathrm{d}^D \boldsymbol{x} J(\boldsymbol{x}) f(\boldsymbol{Q}, V | \boldsymbol{x})}, \tag{4}$$

$$= \left[1 - \frac{\sigma_{\rm w}^2}{2H} \int \mathrm{d}^D \boldsymbol{x} \mathrm{d}^D \boldsymbol{y} J(\boldsymbol{x}) \mathbb{E}_{\boldsymbol{Q}, V} [\mathcal{M}_1(\boldsymbol{Q}, V | \boldsymbol{x}, \boldsymbol{y})] J(\boldsymbol{y}) + \mathcal{O}\left(\frac{1}{H^2} \right) \right]^H,$$

with $\mathcal{M}_h(\boldsymbol{Q}, V | \boldsymbol{x}, \boldsymbol{y}) \equiv \sigma(\boldsymbol{x} \cdot \boldsymbol{Q}_h + V_h) \sigma(\boldsymbol{y} \cdot \boldsymbol{Q}_h + V_h)$.

$$\lim_{H \to \infty} Z_H[J] = \mathrm{e}^{-\frac{\sigma_{\rm w}^2}{2} \int \mathrm{d}^D \boldsymbol{x} \mathrm{d}^D \boldsymbol{y} J(\boldsymbol{x}) \mathbb{E}_{\boldsymbol{Q}, V} [\mathcal{M}_1(\boldsymbol{Q}, V | \boldsymbol{x}, \boldsymbol{y})] J(\boldsymbol{y})}, \tag{5}$$

the correlators can be calculated via the free bosonic QFT Feynman diagrams,

$$G_{\rm NN}^{(n)}(\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n) \Rightarrow \text{n-points Feynman diagrams.} \tag{6}$$

Outline

Complex-valued NN and Corresponding Bosonic Complex Scalar Field

Both the real and imaginary parts of the complex-valued parameters in NN satisfy the same Gaussian distribution

$$\mathrm{Re}(W^{(\alpha)}_{\mathrm{in,h}}),\ \mathrm{Im}(W^{(\alpha)}_{\mathrm{in,h}}) \sim \mathcal{N}(\mu_{\mathrm{w}}, \sigma_{\mathrm{w}}/\sqrt{D}), \tag{7}$$

$$\mathrm{Re}(W^{(\alpha)}_{\mathrm{h,out}}),\ \mathrm{Im}(W^{(\alpha)}_{\mathrm{h,out}}) \sim \mathcal{N}(\mu_{\mathrm{w}}, \sigma_{\mathrm{w}}/\sqrt{H}), \tag{8}$$

$$\mathrm{Re}(b^{(\alpha)}_{\mathrm{h}}),\ \mathrm{Im}(b^{(\alpha)}_{\mathrm{h}}) \sim \mathcal{N}(\mu_{\mathrm{b}}, \sigma_{\mathrm{b}}). \tag{9}$$

The output funciton of the NN is

$$f(\boldsymbol{Q}, V|\boldsymbol{x}) = \frac{1}{H}\sum_{h=1}^{H} \lambda_h(\boldsymbol{Q}, V|\boldsymbol{x})\varphi_h, \tag{10}$$

$$\lambda_h(\boldsymbol{Q}, V|\boldsymbol{x}) \equiv \frac{\exp\left(\boldsymbol{x}\cdot\boldsymbol{Q}_h + V_h\right)}{\exp\left(\sigma_{\mathrm{b}}^2 + \sigma_{\mathrm{w}}^2\boldsymbol{x}^2/D\right)}, \tag{11}$$

The corresponding field theory is *free Bosonic Complex Scalar Field*

$$\lim_{H\to\infty} Z_H[J] = \exp\left\{ -2\sigma_{\mathrm{w}}^2 \int \mathrm{d}^D\boldsymbol{x}\mathrm{d}^D\boldsymbol{y}\, J(\boldsymbol{x})\mathbb{E}_{\boldsymbol{Q},V}[\mathcal{M}_1(\boldsymbol{Q}, V|\boldsymbol{x}, \boldsymbol{y})]J^*(\boldsymbol{y})\right\}, \tag{12}$$

with $\mathcal{M}_h(\boldsymbol{Q}, V|\boldsymbol{x}, \boldsymbol{y}) \equiv \lambda_h^*(\boldsymbol{Q}, V|\boldsymbol{x})\lambda_h(\boldsymbol{Q}, V|\boldsymbol{y})$.

Renormalization and divergence of finite width hidden layer CVNN

The generating functional for finite width hidden layer CVNN is

$$Z[J^*, J] = \left\{1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \frac{(2\sigma_{\mathrm{w}}^2)^m}{H^m} \int \prod_{i=1}^{m} [\mathrm{d}^D \boldsymbol{x}_i \mathrm{d}^D \boldsymbol{y}_i] \left[\prod_{j=1}^{m} J(\boldsymbol{x}_j)\right] \mathbb{E}_{\boldsymbol{Q},V} \left[\prod_{k=1}^{m} \mathcal{M}_1(\boldsymbol{Q}, V | \boldsymbol{x}_k, \boldsymbol{y}_k)\right] \left[\prod_{l=1}^{m} J^*(\boldsymbol{y}_l)\right]\right\}^H . \quad (13)$$

To analyse ultraviolet (UV) effects we switch to momentum space and impose a hard cutoff,

$$f_<(\boldsymbol{Q}, V|\boldsymbol{x}) \equiv \int \frac{\mathrm{d}^D \boldsymbol{p}}{(2\pi)^{D/2}} \mathrm{e}^{-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{p}} f_<(\boldsymbol{Q}, V|\boldsymbol{p}), \quad (14)$$

$$f_<(\boldsymbol{Q}, V|\boldsymbol{p}) \equiv \begin{cases} f(\boldsymbol{Q}, V|\boldsymbol{p}), & \boldsymbol{p} \in \mathcal{V}, \\ 0, & \text{others}, \end{cases} \quad (15)$$

where $\mathcal{V}$ is the area of D-dimensional cube of side length p_L (thus $p_i \in [-p_L/2, p_L/2]$, $(i = 1, 2, ..., D)$) and $f(\boldsymbol{Q}, V|\boldsymbol{p})$ is the momentum representation of $f(\boldsymbol{Q}, V|\boldsymbol{x})$,

$$f(\boldsymbol{Q}, V|\boldsymbol{p}) = \frac{1}{H} \sum_{h=1}^{H} \lambda_h(\boldsymbol{Q}, V|\boldsymbol{p}) \varphi_h, \quad (16)$$

$$\lambda_h(\boldsymbol{Q}, V|\boldsymbol{p}) \equiv \left(\frac{D}{2\sigma_{\mathrm{w}}^2}\right)^{D/2} \mathrm{e}^{-\frac{(\boldsymbol{p}-\mathrm{i}\boldsymbol{Q}_h)^2}{4\sigma_{\mathrm{w}}^2/D} + V_h - \sigma_{\mathrm{b}}^2}. \quad (17)$$

It can be noticed that the UV divergence arises when we calculate the Feynman diagrams in momentum representation. This divergence is caused by exchanging the integral order of $\boldsymbol{Q}$ (especially the real part of $\boldsymbol{Q}$) and $\boldsymbol{p}$, mathematically.

A single Wilsonian RG step enlarges the cutoff, $p_L \mapsto s\, p_L$ with $s > 1$. Under this rescaling the connected m-point coefficients transform as

$$\mathbb{E}_{\boldsymbol{Q},V}\left[\prod_{k=1}^{m}\mathcal{M}_1^{<}(\boldsymbol{Q},V|\boldsymbol{x}_k,\boldsymbol{y}_k)\right] \Rightarrow \mathbb{E}_{s\boldsymbol{Q},V}\left[\prod_{k=1}^{m}\mathcal{M}_1^{<}(\boldsymbol{Q},V|s\boldsymbol{x}_k,s\boldsymbol{y}_k)\right], \tag{18}$$

where the notation $\mathbb{E}_{s\boldsymbol{Q},V}$ means that the variance of each real component of $\mathbf{Q}_h$ is rescaled, $\sigma_{\mathrm{w}}^2/D \mapsto \sigma_{\mathrm{w}}^2/(s^2 D)$. Taking $s \to \infty$ restores the original uncut theory:

$$\lim_{s\to+\infty}\mathbb{E}_{s\boldsymbol{Q},V}\left[\prod_{k=1}^{m}\mathcal{M}_1^{<}(\boldsymbol{Q},V|s\boldsymbol{x}_k,s\boldsymbol{y}_k)\right] = \mathbb{E}_{\boldsymbol{Q},V}\left[\prod_{k=1}^{m}\mathcal{M}_1(\boldsymbol{Q},V|\boldsymbol{x}_k,\boldsymbol{y}_k)\right], \tag{19}$$

hence the coefficients of the full CVNN-QFT are invariant under the scaling $\boldsymbol{x} \to s\boldsymbol{x}$ once the cutoff has been removed, confirming that they define a fixed-point theory, which is more important for Deep Neural Networks.

Quantum State in the ∞-Width Limit:

Degeneracy of Interaction Terms Arising from Measure Zero

Denote $W_{\mathrm{in,h}}$, b_{h} and $W_{\mathrm{h,out}}$ with $\boldsymbol{Q}_h$, V_h and ψ_h as

$$\boldsymbol{Q}_h \equiv W_{\mathrm{in,h}},\ V_h \equiv b_{\mathrm{h}},\ \psi_h \equiv \sqrt{H} * W_{\mathrm{h,out}}, \tag{20}$$

then

$$W^{(\alpha)}_{\mathrm{h,out}} \sim \mathcal{N}(\mu_{\mathrm{w}}, \sigma_{\mathrm{w}}/\sqrt{H}) \Rightarrow \psi_h \sim \mathcal{N}(\mu_{\mathrm{w}}, \sigma_{\mathrm{w}}) \tag{21}$$

the vertexes higher than 4-points (including) vanish for measuring zero,

$$f(\boldsymbol{Q}, V|\boldsymbol{x}) = \frac{1}{\sqrt{H}} \sum_{h=1}^{H} \frac{\exp\left(\boldsymbol{x}\cdot\boldsymbol{Q}_h + V_h\right)}{\exp(\sigma_{\mathrm{b}}^2 + \sigma_{\mathrm{w}}^2 \boldsymbol{x}^2/D)} \psi_h \tag{22}$$

the vertex terms higher than 4-points will reduce the number of h-summation

$$\langle f^*_\infty(\boldsymbol{Q}, V|\boldsymbol{x}_1)\cdots f^*_\infty(\boldsymbol{Q}, V|\boldsymbol{x}_N) f_\infty(\boldsymbol{Q}, V|\boldsymbol{y}_1)\cdots f_\infty(\boldsymbol{Q}, V|\boldsymbol{y}_N)\rangle, \tag{23}$$

only the all free progator connection will survive $(H^N$, for sum of $N)$ in the $H\to\infty$ condition for measuring zero $(1/H^N)$.

Quantum fluctuation

Denote $\hat{O}_h(\boldsymbol{x}) \equiv |\lambda_h(\boldsymbol{Q}, V|\boldsymbol{x})|^2 = \lambda_h^*(\boldsymbol{Q}, V|\boldsymbol{x})\lambda_h(\boldsymbol{Q}, V|\boldsymbol{x})$, then

$$\langle |f(\boldsymbol{Q}, V|\boldsymbol{x})|^4\rangle = 8\sigma_{\mathrm{w}}^4 \left\{\mathbb{E}_{\boldsymbol{Q},V}\left[\hat{O}_1(\boldsymbol{x})\right]\right\}^2 + \frac{8\sigma_{\mathrm{w}}^4}{H}\mathbb{E}_{\boldsymbol{Q},V}\left[(\Delta\hat{O}_1(\boldsymbol{x}))^2\right], \tag{24}$$

the quantum fluctuation term is depressed for infinite H, $H\to\infty$.

Quantum State in the ∞-Width Limit: Equivalent representation

Denote $W_{\rm in,h}$, $b_{\rm h}$ and $W_{\rm h,out}$ with $\boldsymbol{Q}_h$, V_h and φ_h as

$$\boldsymbol{Q}_h \equiv W_{\rm in,h},\ V_h \equiv b_{\rm h},\ \varphi_h \equiv H * W_{\rm h,out}, \tag{25}$$

infinite-width last hidden layer condition: h-summation$\rightarrow$ integration

$$f(\boldsymbol{Q}, V|\boldsymbol{x}) = \frac{1}{H}\sum_{h=1}^{H} \frac{\exp\left(\boldsymbol{x}\cdot\boldsymbol{Q}_h + V_h\right)}{\exp(\sigma_{\rm b}^2 + \sigma_{\rm w}^2 \boldsymbol{x}^2/D)}\varphi_h \Rightarrow f_\infty(\boldsymbol{Q}, V|\boldsymbol{x}) = \int_0^1 {\rm d}\xi \frac{\exp\left(\boldsymbol{x}\cdot\boldsymbol{Q}(\xi) + V(\xi)\right)}{\exp(\sigma_{\rm b}^2 + \sigma_{\rm w}^2 \boldsymbol{x}^2/D)}\varphi(\xi), \tag{26}$$

since the higher order correlation functions vanish in the $H\rightarrow\infty$ limit, we can absorb the arithmetic square root of $\boldsymbol{Q}(\xi)$ and $V(\xi)$'s distributions into the definition of the field to reproduce the 2-pt correlation function of f equivalently reproduces the expectation of $\langle f^*(\boldsymbol{Q}, V|\boldsymbol{x}) f(\boldsymbol{Q}, V|\boldsymbol{y})\rangle$,

$$f_{\rm eff}(\boldsymbol{Q}, V|\boldsymbol{x}) = \int_0^1 {\rm d}\xi \sqrt{P(V(\xi))P(\boldsymbol{Q}(\xi))}\frac{\exp\left(\boldsymbol{x}\cdot\boldsymbol{Q}(\xi) + V(\xi)\right)}{\exp\left(\sigma_{\rm b}^2 + \sigma_{\rm w}^2 \boldsymbol{x}^2/D\right)}\varphi(\xi), \tag{27}$$

with $\langle\varphi^*(\xi)\varphi(\xi')\rangle = 2\sigma_{\rm w}^2\delta(\xi-\xi')$.

Quantum State in the ∞-Width Limit: Simplification

It can be noticed that $V(\xi)$ does not connect to the position $\boldsymbol{x}$, and $f_{\text{eff}}(\boldsymbol{Q}, V|\boldsymbol{x})$ can be divided into two parts

$$f_{\text{eff}}(\boldsymbol{Q}, V|\boldsymbol{x}) \equiv \int_0^1 \mathrm{d}\xi\; g_1(\boldsymbol{Q}(\xi)|\boldsymbol{x}) g_2(V(\xi))\varphi(\xi), \tag{28}$$

$g_2(V(\xi))$ will not influence $\langle f^*(\boldsymbol{Q}, V|\boldsymbol{x}) f(\boldsymbol{Q}, V|\boldsymbol{y})\rangle$ for given $\boldsymbol{x}$ and $\boldsymbol{y}$,

$$f_{\text{eff}}(\boldsymbol{Q}, V|\boldsymbol{x}) = \int_0^1 \mathrm{d}\xi \frac{\mathrm{e}^{-|V(\xi)-2\sigma_{\mathrm{b}}^2|^2/(4\sigma_{\mathrm{b}}^2)-\frac{1}{2}[V^*(\xi)-V(\xi)]}}{\sqrt{2\pi\sigma_{\mathrm{b}}^2}} \times \frac{\mathrm{e}^{-||\boldsymbol{Q}(\xi)-2\sigma_{\mathrm{w}}^2\boldsymbol{x}/D||^2/(4\sigma_{\mathrm{w}}^2/D)-\frac{1}{2}\boldsymbol{x}\cdot[\boldsymbol{Q}^*(\xi)-\boldsymbol{Q}(\xi)]}}{(2\pi\sigma_{\mathrm{w}}^2/D)^{D/2}}\varphi(\xi). \tag{29}$$

Quantum State in the ∞-Width Limit: Simplification

Then the quantum state can effectively be

$$f_{\text{eff}}(\boldsymbol{Q}|\boldsymbol{x}) = \frac{1}{(2\pi\sigma_{\text{w}}^2/D)^{D/2}} \int_0^1 \mathrm{d}\xi \mathrm{e}^{-\frac{[\boldsymbol{Q}_{\text{R}}(\xi) - 2\sigma_{\text{w}}^2 \boldsymbol{x}/D]^2 + \boldsymbol{Q}_{\text{I}}(\xi)^2}{4\sigma_{\text{w}}^2/D} + \mathrm{i}\boldsymbol{x}\cdot\boldsymbol{Q}_{\text{I}}(\xi)} \varphi(\xi). \tag{30}$$

The Wick contraction of $\varphi(\xi)$ actually controls the equivalence of $\boldsymbol{Q}(\xi)$,

$$\langle \varphi^*(\xi)\varphi(\xi')\rangle = 2\sigma_{\text{w}}^2 \delta(\xi - \xi') \quad \Rightarrow \quad \xi = \xi' \quad \Rightarrow \quad \boldsymbol{Q}(\xi) = \boldsymbol{Q}(\xi') \tag{31}$$

and one can transform φ to the function of $\boldsymbol{Q}$ with its 2-pt correlation function gives the Dirac delta function of $\boldsymbol{Q}$, namely,

$$\varphi(\xi) \to \tilde{\varphi}(\boldsymbol{Q}), \tag{32}$$

$$\left\langle \tilde{\varphi}^*(\boldsymbol{Q})\tilde{\varphi}(\boldsymbol{Q}')\right\rangle = 2\sigma_{\text{w}}^2 \delta^{(D)}(\boldsymbol{Q} - \boldsymbol{Q}'), \tag{33}$$

the corresponding ξ integral becomes $\boldsymbol{Q}$ integral,

$$f_{\text{eff}}(\boldsymbol{x}) = \int \frac{\mathrm{d}^D\boldsymbol{Q}_{\text{R}}\mathrm{d}^D\boldsymbol{Q}_{\text{I}}}{(2\pi\sigma_{\text{w}}^2/D)^{D/2}} \mathrm{e}^{-\frac{(\boldsymbol{Q}_{\text{R}} - 2\sigma_{\text{w}}^2\boldsymbol{x}/D)^2}{4\sigma_{\text{w}}^2/D}} \mathrm{e}^{-\frac{\boldsymbol{Q}_{\text{I}}^2}{4\sigma_{\text{w}}^2/D} + \mathrm{i}\boldsymbol{x}\cdot\boldsymbol{Q}_{\text{I}}} \tilde{\varphi}(\boldsymbol{Q}). \tag{34}$$

Quantum State in the ∞-Width Limit:

Tensor-Product State of Dual Gaussian Wave Packets

$$f_{\rm eff}(\boldsymbol{x}) = \int \frac{\mathrm{d}^D\boldsymbol{Q}_{\rm R}\mathrm{d}^D\boldsymbol{Q}_{\rm I}}{(2\pi\sigma_{\rm w}^2/D)^{D/2}} \mathrm{e}^{-\frac{(\boldsymbol{Q}_{\rm R}-2\sigma_{\rm w}^2\boldsymbol{x}/D)^2}{4\sigma_{\rm w}^2/D}} \mathrm{e}^{-\frac{\boldsymbol{Q}_{\rm I}^2}{4\sigma_{\rm w}^2/D}+\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{Q}_{\rm I}} \tilde{\varphi}(\boldsymbol{Q}).$$

Considering the variables' transformations $\boldsymbol{Q}_{\rm R} \to 2\sigma_{\rm w}^2\boldsymbol{y}/D$ and $\boldsymbol{Q}_{\rm I} \to \boldsymbol{q}$, the quantum field $f_{\rm eff}$ can be written as the direct product of the two Gaussian wave packets as

$$f_{\rm eff}(\boldsymbol{x}) = \sqrt{2}\sigma_{\rm w} \int \mathrm{d}^D\boldsymbol{y}\ \psi_{\boldsymbol{p}=\boldsymbol{0}}(2^{-1}\sqrt{D}/\sigma_{\rm w}, \boldsymbol{x}, \boldsymbol{y})\hat{\kappa}(\boldsymbol{y}) \otimes \int \mathrm{d}^D\boldsymbol{q}\ \psi_{\boldsymbol{x}}(\sigma_{\rm w}/\sqrt{D}, \boldsymbol{q}, \boldsymbol{0})\hat{\eta}(\boldsymbol{q}), \tag{35}$$

with and the wave function of a Gaussian wave packet being

$$\psi_{\boldsymbol{x}}(\sigma, \boldsymbol{q}, \boldsymbol{q}_0) \equiv (2\pi\sigma^2)^{-D/4}\mathrm{e}^{-\frac{(\boldsymbol{q}-\boldsymbol{q}_0)^2}{4\sigma^2}+\mathrm{i}\boldsymbol{x}\cdot(\boldsymbol{q}-\boldsymbol{q}_0)}, \tag{36}$$

$$\langle \hat{\eta}(\boldsymbol{p})^\dagger \hat{\eta}(\boldsymbol{q})\rangle = \delta^{(D)}(\boldsymbol{p}-\boldsymbol{q}), \tag{37}$$

and the Fourier transformation relation that

$$\hat{\kappa}(\boldsymbol{y}) = \int \frac{\mathrm{d}^D\boldsymbol{q}}{(2\pi)^{D/2}} \mathrm{e}^{-\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{q}}\hat{\eta}(\boldsymbol{q}), \tag{38}$$

$$\psi_{\boldsymbol{p}=\boldsymbol{0}}((2\sigma)^{-1}, \boldsymbol{x}, \boldsymbol{y}) = \int \frac{\mathrm{d}^D\boldsymbol{q}}{(2\pi)^{D/2}} \mathrm{e}^{-\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{q}}\psi_{\boldsymbol{x}}(\sigma, \boldsymbol{q}, \boldsymbol{0}). \tag{39}$$

Quantum State in the ∞-Width Limit:

Key understanding of the quantum state

The replacement that $\boldsymbol{Q}_{\rm R} \to 2\sigma_{\rm w}^2 \boldsymbol{y}/D$, $\boldsymbol{Q}_{\rm I} \to \boldsymbol{q}$ and the neglect of V ($\boldsymbol{Q}_h \equiv W_{\rm in,h},\ V_h \equiv b_{\rm h},\ \varphi_h \equiv H * W_{\rm h,out}$) tell that

$$\boldsymbol{Q}_h, V_h \Rightarrow \text{eigenvalues}, \tag{40}$$

$$\varphi_h \Rightarrow \text{bosonic field}. \tag{41}$$

- For NN with multi-hidden layers, only the weights between the last hidden layer and the output layer connect to the quantum fields, the parameters between the input layer and the last hidden layer should be the eigenvalues.

Here we tackle the complementary questions posed by NN-QFT: **what neural architecture gives rise *intrinsically* to a fermionic quantum field theory?**

$$\varphi_h \to \text{fermionic quantum field?} \tag{42}$$

Outline

Fermion Field Representation via Tensors Weights

$\varphi_h \to$ Clifford algebra

Construct the anticommute matrices γ_h via Pauli matrices. For $H = 3$, set $\gamma_h^{(3)} \equiv \sigma_h$ $(h = 1, 2, 3)$. For H being odd numbers and $H \geq 5$, one can construct the gamma matrices by recursive definition as

$$\gamma_h^{(H)} \equiv -\sigma_2 \otimes \gamma_h^{(H-2)}, \;\; h = 1, 2, ..., H-2, \tag{43}$$

$$\gamma_{H-1}^{(H)} \equiv \sigma_1 \otimes I_{\frac{d}{2}\times\frac{d}{2}}, \tag{44}$$

$$\gamma_H^{(H)} \equiv \sigma_3 \otimes I_{\frac{d}{2}\times\frac{d}{2}}, \tag{45}$$

with the matrices $\gamma_h^{(H)}$ meet the Clifford algebra $\{\gamma_h^{(H)}, \gamma_{h'}^{(H)}\} = 2\delta_{hh'} I_{d\times d}$ and $(\gamma_h^{(H)})^\dagger = \gamma_h^{(H)}$, the dimensions of $\gamma_h^{(H)}$ should be $d = 2^{(H-1)/2}$.

Transform the weighting parameters between last hidden layer and the output layer to the tensor-formatted coefficients by multiplying the gamma matrices as

$$\varphi_h \to \varphi_h \gamma_h, \;\; \varphi_h^* \to \varphi_h^* \gamma_h^\dagger, \;\; h = 1, 2, ..., H, \tag{46}$$

redefine the correlator

$$G_{\mathrm{NN}}^{(n)}(\boldsymbol{x}_1, ..., \boldsymbol{x}_n, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_{n'}) \equiv \frac{1}{d}\mathrm{Tr}\left(\frac{1}{n_{\mathrm{nets}}}\sum_{\alpha=1}^{n_{\mathrm{nets}}} f_\alpha(\boldsymbol{x}_1)^\dagger \cdots f_\alpha(\boldsymbol{x}_n)^\dagger f_\alpha(\boldsymbol{y}_1)\cdots f(\boldsymbol{y}_{n'})\right). \tag{47}$$

CVNN-fermionic QFT: Feynman Rules for ∞-Width NN

In the infinite width hidden layer condition, $H \to \infty$, the vertexes higher than 4-point vertex (including) vanish, and the Clifford algebra of $\gamma_h^{(H)}$ transforms to the Grassman algebra, the first term in Eq. (48) being depressed, for example.

Take the 4-pt correlation function as a example, there exist the measure zero that

$$\frac{1}{d}\langle \varphi^*_{h_1}\varphi^*_{h_2}\varphi_{h_1'}\varphi_{h_2'}\rangle \mathrm{tr}\left[\gamma_{h_1}\gamma_{h_2}\gamma_{h_1'}\gamma_{h_2'}\right] = \frac{4\sigma_{\mathrm{w}}^4}{H^2}\left(2\delta_{h_1h_2}\delta_{h_1h_1'}\delta_{h_1h_2'} - \delta_{h_1h_1'}\delta_{h_2h_2'} + \delta_{h_1h_2'}\delta_{h_2h_1'}\right),$$

with the last two terms inside the bracket give the anticommute characters of φ_h.

$$\lim_{H\to\infty}\langle \hat{f}^*(\mathbf{x}_1)\hat{f}^*(\mathbf{x}_2)\hat{f}(\mathbf{x}_3)\hat{f}(\mathbf{x}_4)\rangle = G^{(2)}_{\mathrm{NN}}(\mathbf{x}_1,\mathbf{x}_4)G^{(2)}_{\mathrm{NN}}(\mathbf{x}_2,\mathbf{x}_3) - G^{(2)}_{\mathrm{NN}}(\mathbf{x}_1,\mathbf{x}_3)G^{(2)}_{\mathrm{NN}}(\mathbf{x}_2,\mathbf{x}_4), \tag{48}$$

the fermionic Feynman rules holds for arbitrary n-points correlators,

$$\lim_{H\to\infty}\langle \hat{f}^*(\mathbf{x}_1)\cdots\hat{f}^*(\mathbf{x}_n)\hat{f}(\mathbf{y}_1)\cdots\hat{f}(\mathbf{y}_{n'})\rangle = \delta_{nn'} \times (\text{free fermion Feynman diagrams}). \tag{49}$$

Necessity of Grassmann Algebra in Generating Functionals

The generating functional can not be transformed to the fermionic condition using the traditional auxiliary field method, for the derivatives of the auxiliary field will not give the extra tensor-formatted field $f(\boldsymbol{Q}, V|\boldsymbol{x})$, considering the inequation that

$$\frac{\delta}{\delta J(\boldsymbol{y})} \mathrm{e}^{\int \mathrm{d}^D \boldsymbol{x} J(\boldsymbol{x}) f_\gamma(\boldsymbol{Q}, V|\boldsymbol{x})} \Big|_{J=0} \neq f_\gamma(\boldsymbol{Q}, V|\boldsymbol{y}). \tag{50}$$

The Grassmann algebra is necessary to introduce the generating functional.

Outlook

- Finite width hidden layer condition (fermi ψ^4?)
- What the tensor-weights NN can do in daily life? image identification? ......
- How to use the tensor-weights NN in data simulation?
- Is it possible for tensor-weights NN to solve the sign problem?

Backup

Tensor weights 4-point correlation function

The 4-points correlation function of tensor weights last hidden layer condition is

$$\langle \ \hat{f}^*(\mathbf{x}_1)\hat{f}^*(\mathbf{x}_2)\hat{f}(\mathbf{x}_3)\hat{f}(\mathbf{x}_4) \ \rangle \equiv \frac{1}{d} \ \mathrm{tr}\left(\frac{1}{n_{\mathrm{nets}}} \sum_{\alpha=1}^{n_{\mathrm{nets}}} f_\alpha^*(\mathbf{x}_1) f_\alpha^*(\mathbf{x}_2) f_\alpha(\mathbf{x}_3) f_\alpha(\mathbf{x}_4) \right), \tag{51}$$

which can be expended as

$$\begin{aligned} &\langle \ \hat{f}^*(\mathbf{x}_1)\hat{f}^*(\mathbf{x}_2)\hat{f}(\mathbf{x}_3)\hat{f}(\mathbf{x}_4) \ \rangle \\ = &\langle \sigma\left(\mathbf{x}_1 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right)^* W_{\mathrm{h,out}}^* \sigma\left(\mathbf{x}_2 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right)^* W_{\mathrm{h,out}}^* \sigma\left(\mathbf{x}_3 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right) W_{\mathrm{h,out}} \sigma\left(\mathbf{x}_4 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right) W_{\mathrm{h,out}} \rangle, \end{aligned} \tag{52}$$

considering that the dimension of the output function f is $d_{\mathrm{out}} = 1$ and the relation that

$$\frac{1}{d} \left\langle W_i^* W_j^* W_k W_l \right\rangle \mathrm{tr}[\gamma_i \gamma_j \gamma_k \gamma_l] = \frac{4\sigma_{\mathrm{w}}^4}{H^2} (2\delta_{ij}\delta_{ik}\delta_{il} - \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), \tag{53}$$

the 4-points correlation function can be calculated as

$$\begin{aligned} &\langle \ \hat{f}^*(\mathbf{x}_1)\hat{f}(\mathbf{x}_2)\hat{f}^*(\mathbf{x}_3)\hat{f}(\mathbf{x}_4) \ \rangle \\ = &\frac{4\sigma_{\mathrm{w}}^4}{H^2} \left\langle \sigma\left(\mathbf{x}_4 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right) \sigma\left(W_{\mathrm{in,h}}^\dagger \mathbf{x}_1^\dagger + \mathbf{b}_{\mathrm{h}}^\dagger \right) \sigma\left(\mathbf{x}_3 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right) \sigma\left(W_{\mathrm{in,h}}^\dagger \mathbf{x}_2^\dagger + \mathbf{b}_{\mathrm{h}}^\dagger \right) \right\rangle \\ &- \frac{4\sigma_{\mathrm{w}}^4}{H^2} \left\langle \sigma\left(\mathbf{x}_3 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right) \sigma\left(W_{\mathrm{in,h}}^\dagger \mathbf{x}_1^\dagger + \mathbf{b}_{\mathrm{h}}^\dagger \right) \sigma\left(\mathbf{x}_4 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}\right) \sigma\left(W_{\mathrm{in,h}}^\dagger \mathbf{x}_2^\dagger + \mathbf{b}_{\mathrm{h}}^\dagger \right) \right\rangle \\ &+ \frac{8\sigma_{\mathrm{w}}^4}{H^2} \sum_{j=1}^{H} \left\langle \exp\left\{ \sum_{i=1}^{d_{\mathrm{in}}} [x_1^{(i)*} w_{ij}^* + x_2^{(i)*} w_{ij}^* + x_3^{(i)} w_{ij} + x_4^{(i)} w_{ij}] + 2b_j + 2b_j^* \right\} \right\rangle. \end{aligned} \tag{54}$$

According to the matrix multiplication relation that

$$
\begin{aligned}
&\exp(\mathbf{x}_4 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}) \exp(W_{\mathrm{in,h}}^{\dagger} \mathbf{x}_1^{\dagger} + \mathbf{b}_{\mathrm{h}}^{\dagger}) \exp(\mathbf{x}_3 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}) \exp(W_{\mathrm{in,h}}^{\dagger} \mathbf{x}_2^{\dagger} + \mathbf{b}_{\mathrm{h}}^{\dagger}) \\
&= \sum_{j=1}^{H} \exp\left\{ \sum_{i=1}^{d_{\mathrm{in}}} [x_4^{(i)} w_{ij} + x_1^{(i)*} w_{ij}^*] + b_j + b_j^* \right\} \sum_{j'=1}^{H} \exp\left\{ \sum_{i'=1}^{d_{\mathrm{in}}} [x_3^{(i')} w_{i'j'} + x_2^{(i')*} w_{i'j'}^*] + b_{j'} + b_{j'}^* \right\}
\end{aligned} \tag{55}
$$

the mathematical expectation of the above equation gives

$$
\begin{aligned}
&\langle \exp(\mathbf{x}_4 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}) \exp(W_{\mathrm{in,h}}^{\dagger} \mathbf{x}_1^{\dagger} + \mathbf{b}_{\mathrm{h}}^{\dagger}) \exp(\mathbf{x}_3 W_{\mathrm{in,h}} + \mathbf{b}_{\mathrm{h}}) \exp(W_{\mathrm{in,h}}^{\dagger} \mathbf{x}_2^{\dagger} + \mathbf{b}_{\mathrm{h}}^{\dagger}) \rangle \\
&= H(H-1) \exp\left\{ \frac{2\sigma_{\mathrm{w}}^2}{d_{\mathrm{in}}} (\mathbf{x}_1 \cdot \mathbf{x}_4 + \mathbf{x}_2 \cdot \mathbf{x}_3) + 4\sigma_{\mathrm{b}}^2 \right\} + H \times \exp\left\{ \frac{2\sigma_{\mathrm{w}}^2}{d_{\mathrm{in}}} (\mathbf{x}_1 + \mathbf{x}_2) \cdot (\mathbf{x}_3 + \mathbf{x}_4) + 8\sigma_{\mathrm{b}}^2 \right\},
\end{aligned} \tag{56}
$$

considering $\mathbf{x}_i$ $(i = 1, 2, 3, 4)$ being real valued vectors, with the first term comes from the $j \neq j'$ case, and the second term comes from the $j = j'$ case.
As for the third term in equation (54), the mathematical expectation gives the value of the second term in equation (56), thus, in the large H limit, this term vanishes, and the 4-points correlation function gives

$$
\lim_{H \to \infty} \langle \hat{f}^*(\mathbf{x}_1) \hat{f}^*(\mathbf{x}_2) \hat{f}(\mathbf{x}_3) \hat{f}(\mathbf{x}_4) \rangle = G_{\mathrm{NN}}^{(2)}(\mathbf{x}_1, \mathbf{x}_4) G_{\mathrm{NN}}^{(2)}(\mathbf{x}_2, \mathbf{x}_3) - G_{\mathrm{NN}}^{(2)}(\mathbf{x}_1, \mathbf{x}_3) G_{\mathrm{NN}}^{(2)}(\mathbf{x}_2, \mathbf{x}_4). \tag{57}
$$

Thank you!