The expressibility of artificial neural networks in the quantum many-body physics

Zi Cai

Shanghai Jiaotong University

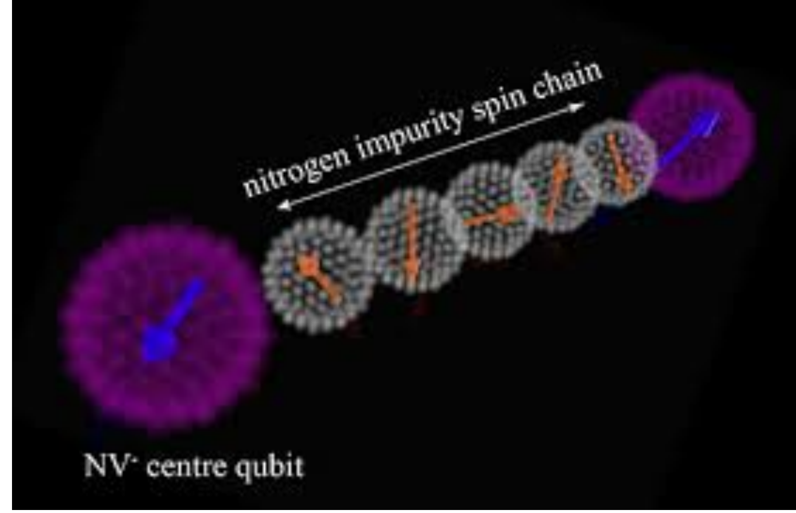
Outline

- Difficulties in strongly correlated systems
- Solutions and limitations: DMRG and QMC
- Artificial neural network: a new solution?
- Some examples: success and failure
- How to improve?

Difficulties: **Exponential explosion of the dimension of Hilbert space of quantum many-body systems !**

$$|\Psi\rangle = \sum_{\{\sigma\}} c_{\{\sigma\}} \,|\sigma_1\sigma_2\cdots\sigma_L\rangle \qquad N \sim \exp(\alpha L)$$

L=60, N~1000000000 TB



Strongly correlated systems:

- Exact solutions in integrable models: Bethe-ansatz
- Bosonization
- Slave boson/fermion Schwinger boson ……

No Universal solutions !

- Exact diagonalization
- Density matrix Renor-malization Group
- Quantum Monte Carlo
- Dynamical mean-field

$$|\Psi\rangle = \sum_{\{\sigma\}} c_{\{\sigma\}} \, |\sigma_1\sigma_2\cdots\sigma_L\rangle \qquad c[\sigma_1\sigma_2\cdots\sigma_L]$$

Two examples:

- **Completely random state (excited states)** $c[\sigma_1\sigma_2\cdots\sigma_L] = random\ number$
- **AKLT state:** $c[\sigma_1\sigma_2\cdots\sigma_L] = Tr[M_1^{\sigma_1} M_2^{\sigma_2} \cdots M_L^{\sigma_L}]$

Purpose: Using a specific form of function f to approximate $c[\sigma_1\sigma_2\cdots\sigma_L]$

$$f[\sigma_1\sigma_2\cdots\sigma_L, \{W\}] \approx c[\sigma_1\sigma_2\cdots\sigma_L]$$

$$\{W\} = \{W_1, W_2, \cdots, W_N\} \;\; N << e^{\alpha L} \;\; (polynomial)$$

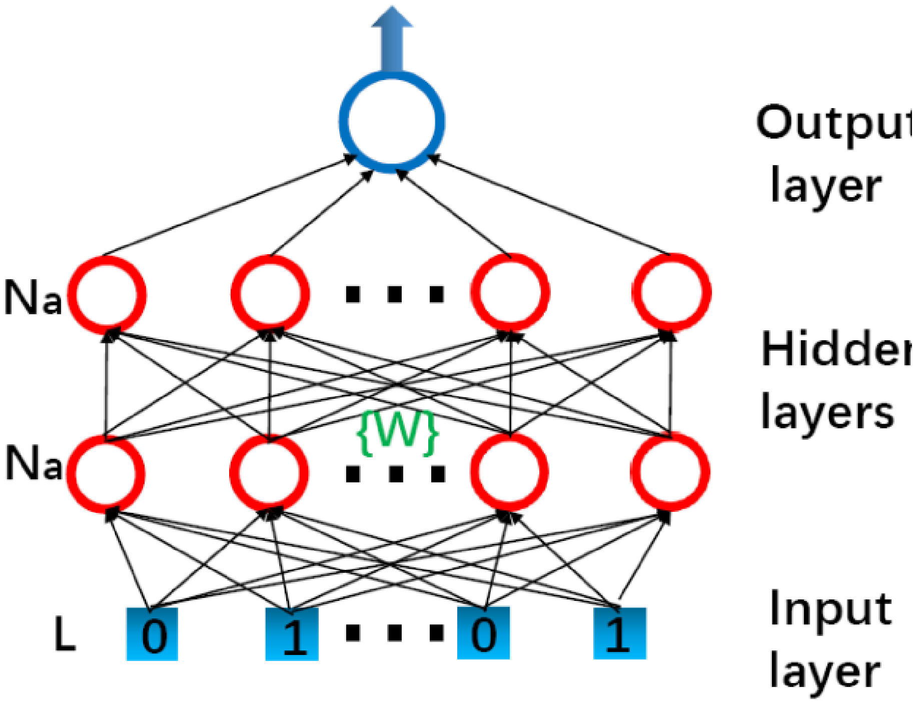
1. MPS or DMRG: $f[\sigma_1\sigma_2\cdots\sigma_L] = Tr[W_1^{\sigma_1} W_2^{\sigma_2} \cdots W_L^{\sigma_L}], \;\; N \sim LD^2 d$

2. Variational QMC:

$$E(\{W\}) = \sum_{sampled\,\{\sigma\}} \frac{1}{Z} f^2[\{\sigma\},\{W\}] \langle\{\sigma\}|H|\{\sigma\}\rangle$$

$$Z = \sum_{Sampled\,\{\sigma\}} f^2[\{\sigma\},\{W\}]$$

Artificial Neural network: a powerful function approximating machine



Extract the rules that may be too complex to be captured by programming or designing explicit algorithms !

The gold is to express the ground states of many-body systems in terms of neural networks with of feasible size and merely a few hidden layers, and most importantly, within learning time polynomial scaled with the system size.

Example I : Ground state of free bosons/fermions

The success of the expressibility relies on the specialty of the ground state compared to a generic eigenstate, where the the information encoded in its wave function is significantly reduced by the physical rules behind it, which on the other hand can be extracted by the neutral networks through big amount of training.

$$P = 1 - |\langle \Psi_{Exact} | \Psi_{ANN} \rangle|$$

1D free bosons: $\Psi_{FB} = \sum_{\boldsymbol{n}} C_{FB}[\boldsymbol{n}] |\boldsymbol{n}\rangle \quad |\boldsymbol{n}\rangle = |n_1 \ldots n_L\rangle$

$$C_{FB}[\boldsymbol{n}] = \sqrt{L!/n_1! \ldots n_L!}/L^{L/2}$$

L=12, N=12: P~10^{-4}

1D free fermion: $C_{FF}[\boldsymbol{n}] = \det[\mathbb{M}] \quad M_{ij} = f_i(x_j) \;\; f_i(x) = \frac{1}{\sqrt{L}} e^{ik_i x}$

L=24, N=11: P~10^{-4}

2D free fermion:

L=24 (4*6), N=5: P~10^{-3}

L=24 (4*6), N=13: P~0.06

Larger systems sizes: Monte Carlo sampling

1D free fermion with L=64, N=31: $O_{nn} = \frac{1}{L}\sum_i \langle n_i n_{i+1} \rangle)$

**we first implement importance sampling to generate millions of “representative“
configurations, and use them to train the ANN and approximate the characteristic
function.**

**After the training is finished, a new set of “representative
“ configurations
are generated according to their weight predicted by the ANN, and we calculate
the value of the physical quantities**

$$f[\sigma_1\sigma_2\cdots\sigma_L,\{W\}] \approx c[\sigma_1\sigma_2\cdots\sigma_L]$$

$$O(\{W\}) = \sum_{sampled\,\{\sigma\}} \frac{1}{Z} f^2[\{\sigma\},\{W\}]\langle\{\sigma\}|O|\{\sigma\}\rangle$$

$$Z = \sum_{Sampled\,\{\sigma\}} f^2[\{\sigma\},\{W\}]$$

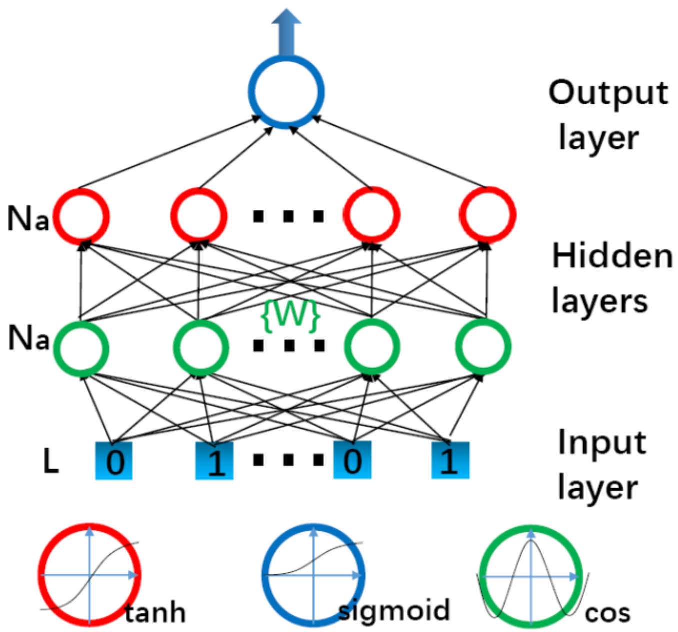
Difference between the exact and ANN values ~10^{-3}

AF ferromagnetism v.s. Frustration

$$H = J\sum \vec{S}_i \cdot \vec{S}_{i+1}$$

$$f[\vec{\sigma}] = |f[\vec{\sigma}]| \times sign\{f[\vec{\sigma}]\}$$

Absolute value **Sign**



Without frustration: Marshall sign rule:
S[σ] = 1/0 if in the total number of down spins in the odd sites is even/odd.

With frustration: Sign rule ???

$$H = J_1\sum_i \vec{S}_i \cdot \vec{S}_{i+1} + J_2\sum_i \vec{S}_i \cdot \vec{S}_{i+2} \quad J_2 = J_1/2,\ \mathrm{dim}er\ phase$$

Practical application: ANN as a variational ground state wavefunction

$$f[\sigma_1\sigma_2\cdots\sigma_L,\{W\}] \approx c[\sigma_1\sigma_2\cdots\sigma_L]$$

$$E(\{W\}) = \sum_{sampled\,\{\sigma\}} \frac{1}{Z} f^2[\{\sigma\},\{W\}]\langle\{\sigma\}\,|\,H\,|\,\{\sigma\}\rangle$$

$$Z = \sum_{Sampled\,\{\sigma\}} f^2[\{\sigma\},\{W\}]$$

$$\frac{\delta E(\{W\})}{\delta\{W\}} = 0$$