Title: Entanglement features of random neural network quantum states

Speakers: Xiaoqi Sun

Series: Machine Learning Initiative

Date: November 18, 2022 - 3:00 PM

URL: https://pirsa.org/22110099

Abstract: Neural networks offer a novel approach to represent wave functions for solving quantum many-body problems. But what kinds of quantum states are efficiently represented by neural networks? In this talk, we will discuss entanglement properties of an ensemble of neural network states represented by random restricted Boltzmann machines. Phases with distinct entanglement features are identified and characterized. In particular, for certain parameters, we will show that these neural network states can look typical in their entanglement profile while still being distinguishable from a typical state by their fractal dimensions. The obtained phase diagrams may help inform the initialization of neural network ansatzes for future computational tasks.

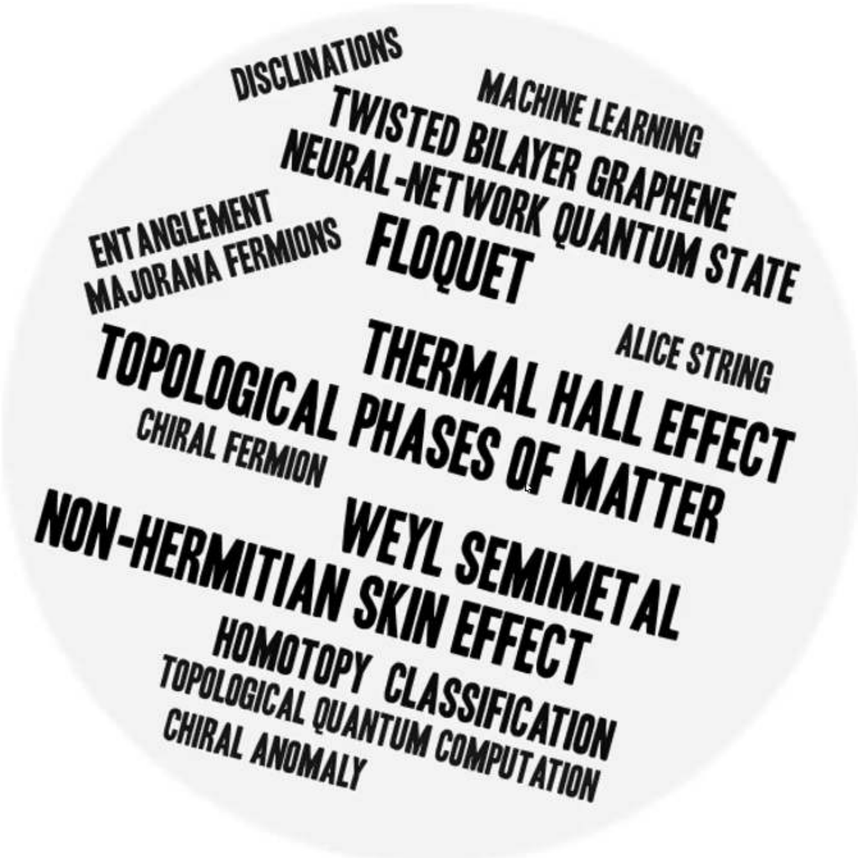
Zoom link: https://pitp.zoom.us/j/94316902357?pwd=RGxWYm9EWGtGYzBvUzM5ZWdwVTB5dz09

Entanglement features of random neural network quantum states

Xiaoqi Sun

XQ Sun, T Nebabu, X Han, MO Flynn, XL Qi,
Phys. Rev. B 106, 115138 (2022)

UNIVERSITY OF ILLINOIS
URBANA-CHAMPAIGN



Theory of quantum matter

Topological phenomena in and out of equilibrium.

Dynamics and transport of quantum materials.

Interdisciplinary studies of machine learning, quantum information science and many-body physics.

......

Many-body physics

Hard problems:

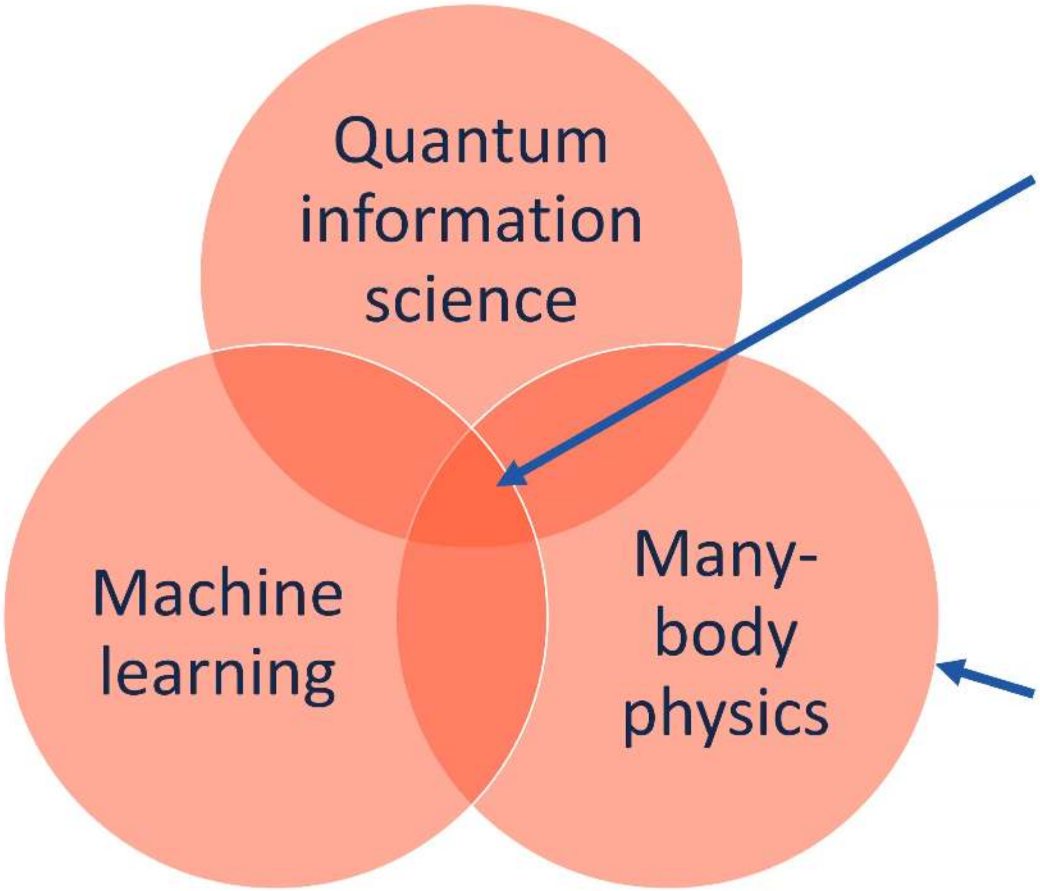
1. Strong interaction
2. Quantum dynamics

"curse of dimensionality"

Modern techniques from quantum information/computation and machine learning greatly enhance our computational power in dealing with problems with high dimensional data.

Many-body physics

Hard problems:
1. Strong interaction
2. Quantum dynamics
"curse of dimensionality"



Neural network quantum states, quantum machine learning, etc.

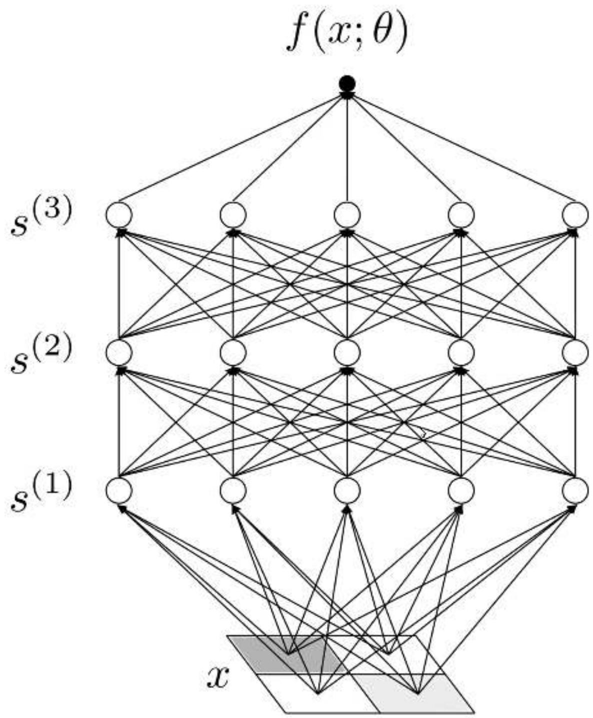
Hard problems:
1. Strong interaction
2. Highly entangled states
3. Quantum dynamics
"curse of dimensionality"

How much can these modern techniques help?

Outline:

1. Introduction to neural network quantum states
2. Thermodynamic limit and infinitely wide neural network
3. Properties of random neural network quantum states:
   - Wavefunction norm and phase transition
   - Second Renyi entanglement entropy and phase transition
   - Fractal dimensions
   - Entanglement level statistics
4. Conclusion and Outlook

Neural networks



Lines represent affine combinations of variables (parameterized by weight w and bias b).
Each circle represents a non-linear neural activation function σ.

Power of representing arbitrarily complex functions

Universal approximation theorems

G Cybenko, Mathematics of Control, Signals, and Systems, 2 (4), 303-314 (1989);K Hornik, et. al. Neural Networks 2, 359 (1989), and several other papers.

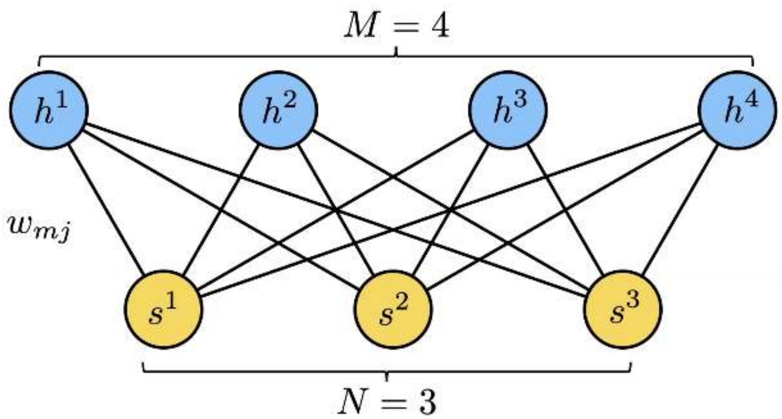
Neural network quantum states

Using neural networks to parameterize many-body quantum wavefunction.

Probability to probability amplitude, real number to complex number

Generalize weight and bias to be complex, G Carleo, M Troyer, Science, 355, 6325 (2017), …
Parametrize magnitude and phase by two neural networks, G Tolai, et. al. Nat. Phys. 14, 447-450 (2018), …
Auxiliary phase node: R Xia & S Kais, Nat. Comm. 9, 4195 (2018), S Kanno & T Tada, Quantum Sci. Technol. 6, 025025 (2021) …;
etc.

Neural network quantum states: simple RBM



A class of generative model:

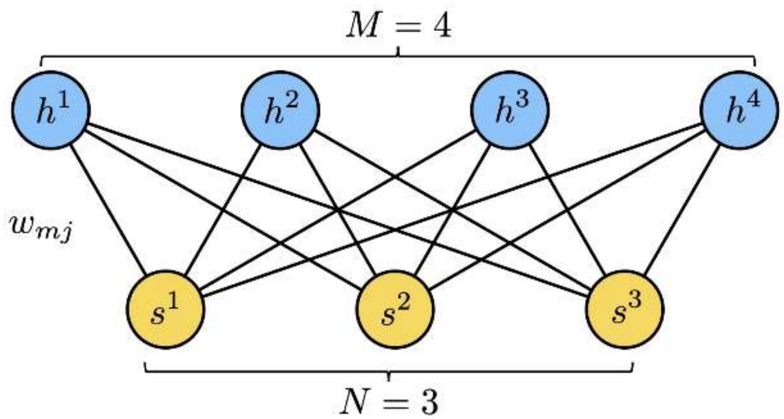
Probability distribution generated by **Boltzmann weight** of an **energy model** on a graph, marginalized over hidden variable.

Restricted meaning no interaction within hidden or visible layer.

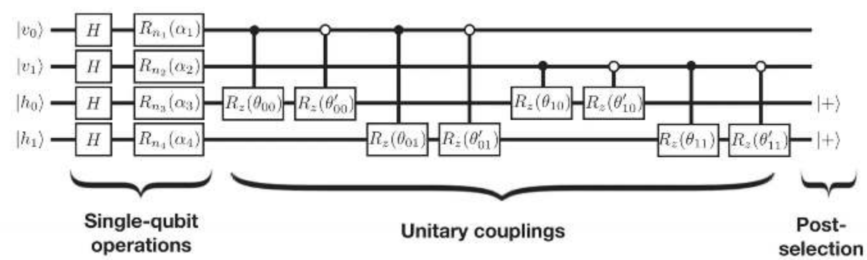
$$|\Psi\rangle \equiv \frac{1}{2^M}\sum_{\boldsymbol{h},\boldsymbol{s}} e^{-\sum_j a_j s^j - \sum_m b_m h^m - \sum_{jm} w_{mj} s^j h^m} |\boldsymbol{s}\rangle$$

Complex w, a, b

Neural network quantum states: simple RBM



Quantum preparation and algorithms:

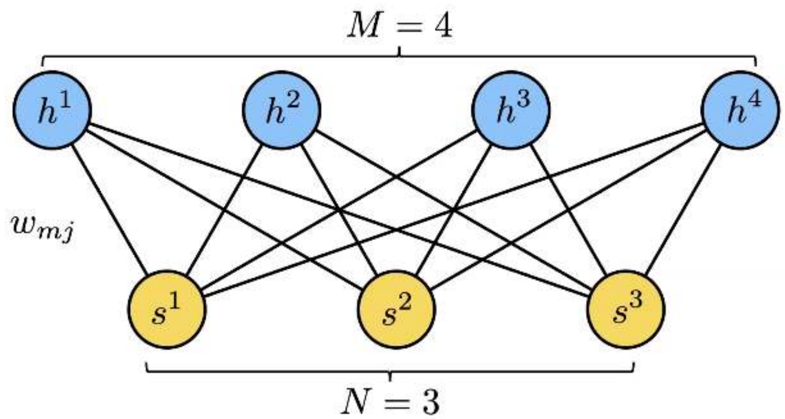


Efficiency in sampling

$$|\Psi\rangle \equiv \frac{1}{2^M} \sum_{\boldsymbol{h},\boldsymbol{s}} e^{-\sum_j a_j s^j - \sum_m b_m h^m - \sum_{jm} w_{mj} s^j h^m} |\boldsymbol{s}\rangle$$

Complex w, a, b

Ability of encoding highly entangled states

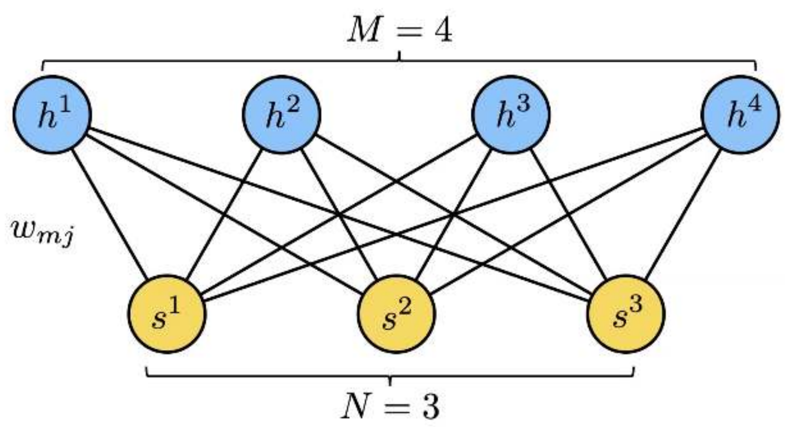


- Polynomial classical resources.
- No known limitations from conventional measures and can represent states that are hard for other ansatz, e.g., volume law entangled states.

DL Deng, X Li, S Das Sarma, PRX, 7, 021021 (2017)

$$|\Psi\rangle \equiv \frac{1}{2^M} \sum_{\boldsymbol{h},\boldsymbol{s}} e^{-\sum_j a_j s^j - \sum_m b_m h^m - \sum_{jm} w_{mj} s^j h^m} |\boldsymbol{s}\rangle$$

Ability of encoding highly entangled states



$$|\Psi\rangle = \frac{1}{2^M} \sum_{\boldsymbol{h}} e^{-\sum_m b_m h_m} \left(|\psi_j^{\boldsymbol{h}}\rangle\right)^{\otimes_j}$$

$$\left|\psi_j^{\boldsymbol{h}}\right\rangle \equiv e^{-a_j - \sum_m h_m w_{mj}} \left|\uparrow\right\rangle_j + e^{a_j + \sum_m h_m w_{mj}} \left|\downarrow\right\rangle_j$$

Understanding:

An entanglement bound: **XQ Sun**, et. al., PRB 106, 115138.

Superposition of 2^M product states. Entanglement bounded by $M \log 2$.

Can produce volume law if $M \propto N$

What kinds of states are efficiently represented by by neural networks?

Can we make some analytic calculation and rigorous statement?

Outline:

1. Introduction to neural network quantum states
2. Thermodynamic limit and infinitely wide neural network
3. Properties of random neural network quantum states:
   - Wavefunction norm and phase transition
   - Second Renyi entanglement entropy and phase transition
   - Fractal dimensions
   - Entanglement level statistics
4. Conclusion and Outlook

Thermodynamic limit & infinite width

Many-body physics naturally focuses on the thermodynamic limit $N \to \infty$.

Likely need the infinite width limit of neural networks. In many situations, it simplifies in the classical theory.

LeCun initialization: $\sigma_w^2 \propto 1/N$, ensemble of neural networks and statistical mechanics.

Thermodynamic limit & infinite width

Example: Neural tangent kernel theory for inifinitely wide neural networks.

Statistical correlations of preactivations of neurons converge to Gaussian (central limit theorem).
Training dynamics is analytically described by differential equations (Neural tangent kernel).

Classical neural network $N \to \infty$ Free statistical field theory
Classical neural network finite N, Interacting statistical field theory with $1/N$ coupling

Similar idea turns out to also work for neural network quantum states in the thermodynamic limit and help understand neural network quantum states and the training dynamics!

Outline:

1. Introduction to neural network quantum states
2. Thermodynamic limit and infinitely wide neural network
3. Properties of random neural network quantum states:
   - Wavefunction norm and phase transition
   - Second Renyi entanglement entropy and phase transition
   - Fractal dimensions
   - Entanglement level statistics if time permits
4. Conclusion and Outlook

Setup of the random ensemble

$$|\Psi\rangle \equiv \frac{1}{2^M} \sum_{\boldsymbol{h},\boldsymbol{s}} e^{-\sum_j a_j s^j - \sum_m b_m h^m - \sum_{jm} w_{mj} s^j h^m} |\boldsymbol{s}\rangle$$

Unnormalized

$$|\Psi\rangle = \sum_{\boldsymbol{s}} \prod_{m=1}^{M} \cosh\left(\sum_j w_{mj} s^j\right) |\boldsymbol{s}\rangle$$

Zero bias, iid Gaussian random weight w_{mj}.
Real part: σ_R, Imaginary part: σ_I.

Proper thermodynamic limit

$$\sigma_R^2 = \frac{u^2}{N}, \qquad \sigma_I^2 = \frac{v^2}{N}, \qquad \lambda = \frac{M}{N}.$$

Setup of the random ensemble

$$|\Psi\rangle = \sum_{\boldsymbol{s}} \prod_{m=1}^{M} \cosh\left(\sum_j w_{mj} s^j\right) |\boldsymbol{s}\rangle$$

Proper thermodynamic limit $$\sigma_R^2 = \frac{u^2}{N}, \qquad \sigma_I^2 = \frac{v^2}{N}, \qquad \lambda = \frac{M}{N}.$$

- Qualitatively different role of u,v: u modulates the magnitude of the wavefunction amplitude much more than v, not good for producing large entanglement.
- Symmetry of the ensemble: The Gaussian iid is symmetric and identical. Ensemble averaged quantity should have symmetry $s^j \to -s^j$.

Outline:

1. Introduction to neural network quantum states
2. Thermodynamic limit and infinitely wide neural network
3. Properties of random neural network quantum states:
   - Wavefunction norm and phase transition (warm-up exercise)
   - Second Renyi entanglement entropy and phase transition
   - Fractal dimensions
   - Entanglement level statistics if time permits
4. Conclusion and Outlook

Norm fluctuations and ferromagnetic transition

Calculation of

$$\overline{\langle\Psi|\Psi\rangle^2} - \overline{\langle\Psi|\Psi\rangle}^2 \equiv \overline{Z_0} - \overline{\langle\Psi|\Psi\rangle}^2$$

$$Z_0 \equiv \mathrm{tr}(\underset{\boldsymbol{s}_1}{\underset{\uparrow}{|\Psi\rangle}}\langle\Psi| \otimes \underset{\boldsymbol{s}_2}{\underset{\uparrow}{|\Psi\rangle}}\langle\Psi|)$$

$$\phi \equiv \sum_i \frac{s_1^i s_2^i}{N}$$

$$\overline{Z_0} = \sum_\phi e^{-N\mathcal{F}(\phi)}$$

Symmetry $\phi \to -\phi$

$$w_{mj} \sim -w_{m'i} \sim w_{m'i}$$

$$\mathcal{F}(\phi) = \mathcal{E}(\phi) - \mathcal{S}(\phi)$$

Favor $\phi \neq 0$ $\phi = 0$

Large N solution = find the free-energy minima

Norm fluctuations and ferromagnetic transition

Calculation of

$$\overline{\langle\Psi|\Psi\rangle^2} - \overline{\langle\Psi|\Psi\rangle}^2 \equiv \overline{Z_0} - \overline{\langle\Psi|\Psi\rangle}^2$$

$$Z_0 \equiv \mathrm{tr}(|\Psi\rangle\langle\Psi| \otimes |\Psi\rangle\langle\Psi|)$$

$\boldsymbol{s}_1$ $\boldsymbol{s}_2$

$$\phi \equiv \sum_i \frac{s_1^i s_2^i}{N}$$

$$\overline{Z_0} = \sum_\phi e^{-N\mathcal{F}(\phi)}$$

Symmetry $\phi \to -\phi$

$$w_{mj} \sim -w_{m'i} \sim w_{m'i}$$

$$\mathcal{F}(\phi) = \mathcal{E}(\phi) - \mathcal{S}(\phi)$$

Favor $\phi \neq 0$ $\phi = 0$

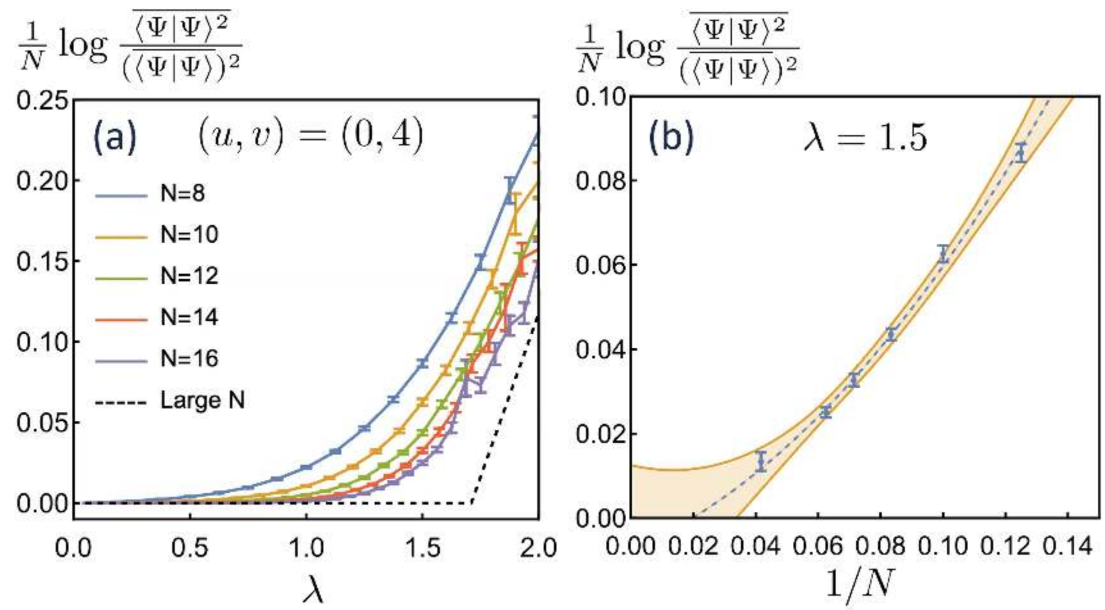
Large N solution = find the free-energy minima

Norm fluctuations and ferromagnetic transition

Calculation of

$$\overline{\langle\Psi|\Psi\rangle^2} - \overline{\langle\Psi|\Psi\rangle}^2 \equiv \overline{Z_0} - \overline{\langle\Psi|\Psi\rangle}^2$$

$$Z_0 \equiv \mathrm{tr}(\underset{\substack{\uparrow \\ \boldsymbol{s}_1}}{|\Psi\rangle}\langle\Psi| \otimes \underset{\substack{\uparrow \\ \boldsymbol{s}_2}}{|\Psi\rangle}\langle\Psi|)$$



Dashed line: symmetry breaking transition.

Outline:

1. Introduction to neural network quantum states
2. Thermodynamic limit and infinitely wide neural network
3. Properties of random neural network quantum states:
   - Wavefunction norm and phase transition (warm-up exercise)
   - Second Renyi entanglement entropy and phase transition
   - Fractal dimensions
   - Entanglement level statistics if time permits
4. Conclusion and Outlook

Second Renyi entropy:

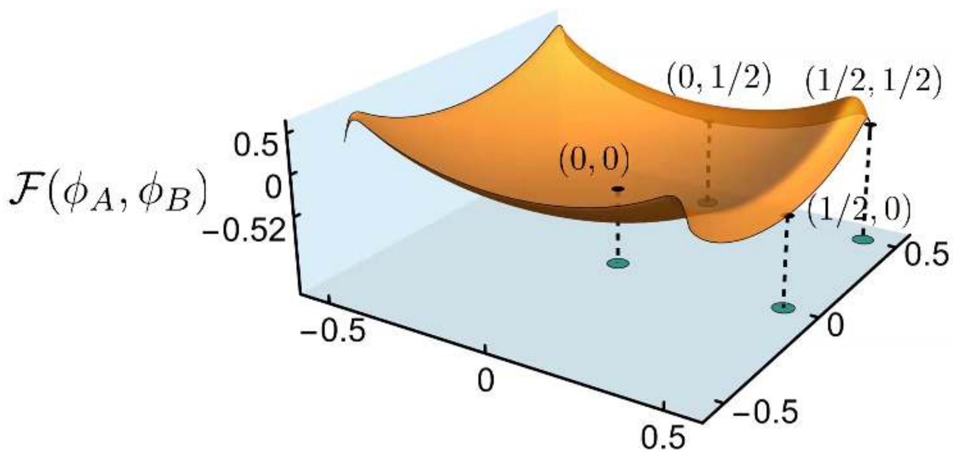
$$\overline{S_2}(A) = -\log \overline{\frac{\mathrm{tr}\left(X_A|\Psi\rangle\langle\Psi|\otimes|\Psi\rangle\langle\Psi|\right)}{\mathrm{tr}\left(|\Psi\rangle\langle\Psi|\otimes|\Psi\rangle\langle\Psi|\right)}}$$
$$\approx -\log\frac{\overline{Z_1}}{\overline{Z_0}} + \ldots$$

Grain of salt:
Needs fluctuation to be small

$$Z_1 \equiv \mathrm{tr}\left(X_A|\Psi\rangle\langle\Psi|\otimes|\Psi\rangle\langle\Psi|\right) \qquad \overline{Z_1} = \sum_{\phi_A,\phi_B} e^{-N\mathcal{F}(\phi_A,\phi_B)}$$

$$\phi_A \equiv \sum_{i\in A}\frac{s_1^i s_2^i}{N}$$

$$\phi_B \equiv \sum_{i\in B}\frac{s_1^i s_2^i}{N}$$



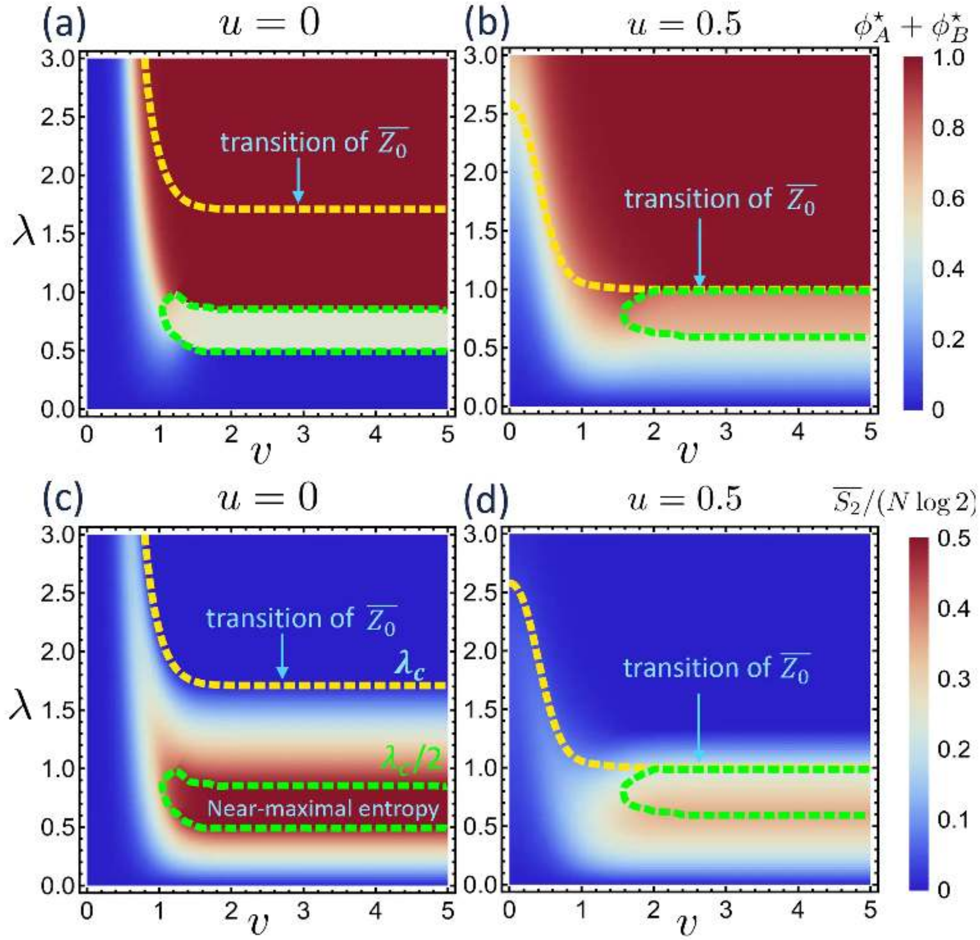
half-system size Renyi entanglement

Symmetry of $\phi_A \leftrightarrow \phi_B$.

Symmetric phase at small λ and large λ.

Symmetry broken phase at intermediate λ.

The symmetry broken phase can have near-maximal entanglement if variance of imaginary weight is large and variance of real weight is 0.



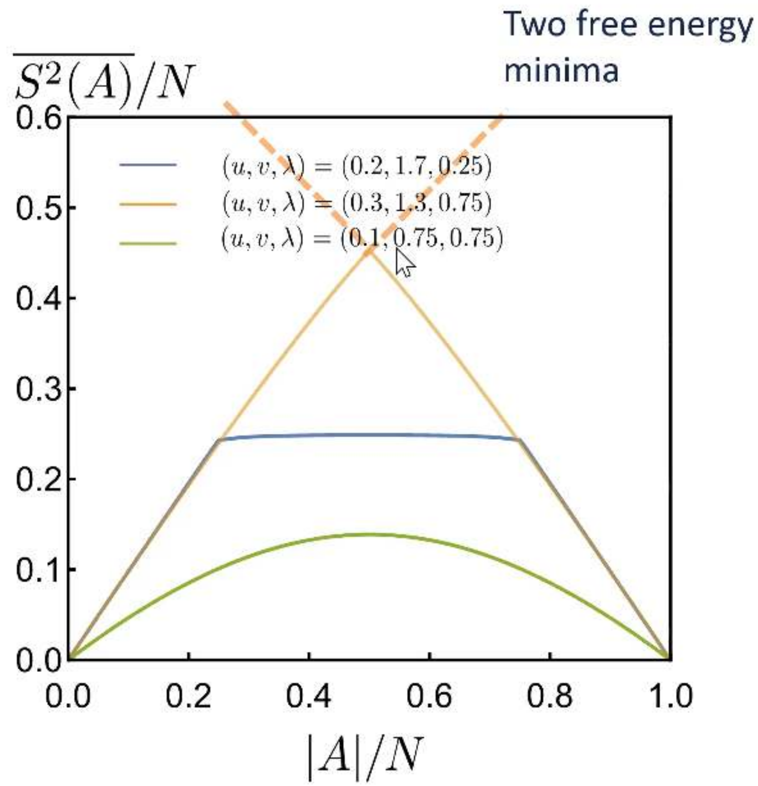
Entanglement phases

Symmetry of $\phi_A \leftrightarrow \phi_B$.

Symmetric phase at small λ and large λ.

Symmetry broken phase at intermediate λ.

The symmetry broken phase can have near-maximal entanglement if variance of imaginary weight is large and variance of real weight is 0.



Entanglement profile

What bounds the entanglement at large λ?
How to characterize these states for random initialization?

Intuition:

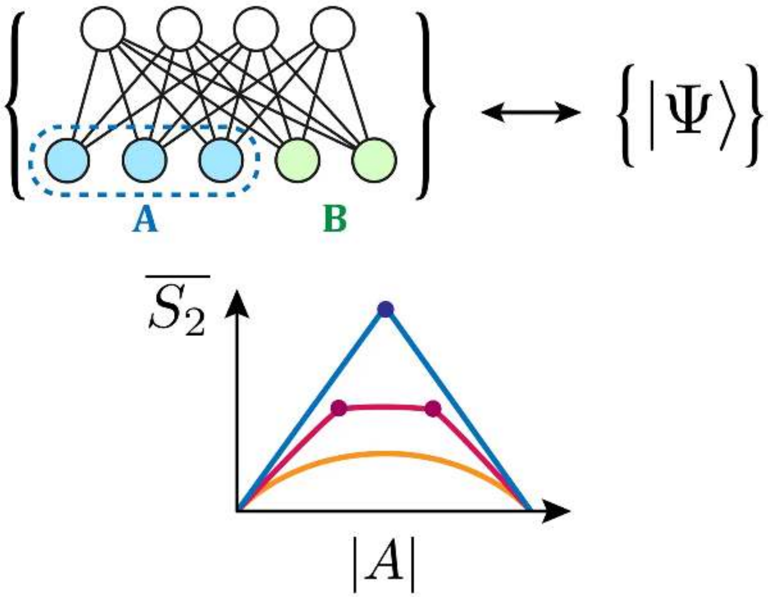
More localized in the Ising spin basis → cannot have large entangled.

$$D_q = \frac{1}{N \log 2} \frac{\log \mathrm{IPR}_q}{1-q}, \quad \mathrm{IPR}_q = \frac{\sum_{\boldsymbol{s}} |\Psi(\boldsymbol{s})|^{2q}}{(\sum_{\boldsymbol{s}} |\Psi(\boldsymbol{s})|^2)^q}$$

Product state in Ising basis: $D_q = 0$, localized states.
Uniform magnitude in Ising basis: $D_q = 1$, extended states.

Other cases: non-ergodic extended states. q-dependence: multifractality.

Conclusion and outlook



Open problems including current projects:

Theory of training dynamics at thermodynamic limit? analogy of neural tangent kernel theory and generalization to neural annealing algorithms?

Understanding the condition for the method to converge to the desired state.

Intuition:

More localized in the Ising spin basis → cannot have large entangled.

$$D_q = \frac{1}{N \log 2} \frac{\log \mathrm{IPR}_q}{1-q}, \quad \mathrm{IPR}_q = \frac{\sum_{\boldsymbol{s}} |\Psi(\boldsymbol{s})|^{2q}}{(\sum_{\boldsymbol{s}} |\Psi(\boldsymbol{s})|^2)^q}$$

Product state in Ising basis: $D_q = 0$, localized states.
Uniform magnitude in Ising basis: $D_q = 1$, extended states.

Other cases: non-ergodic extended states. q-dependence: multifractality.

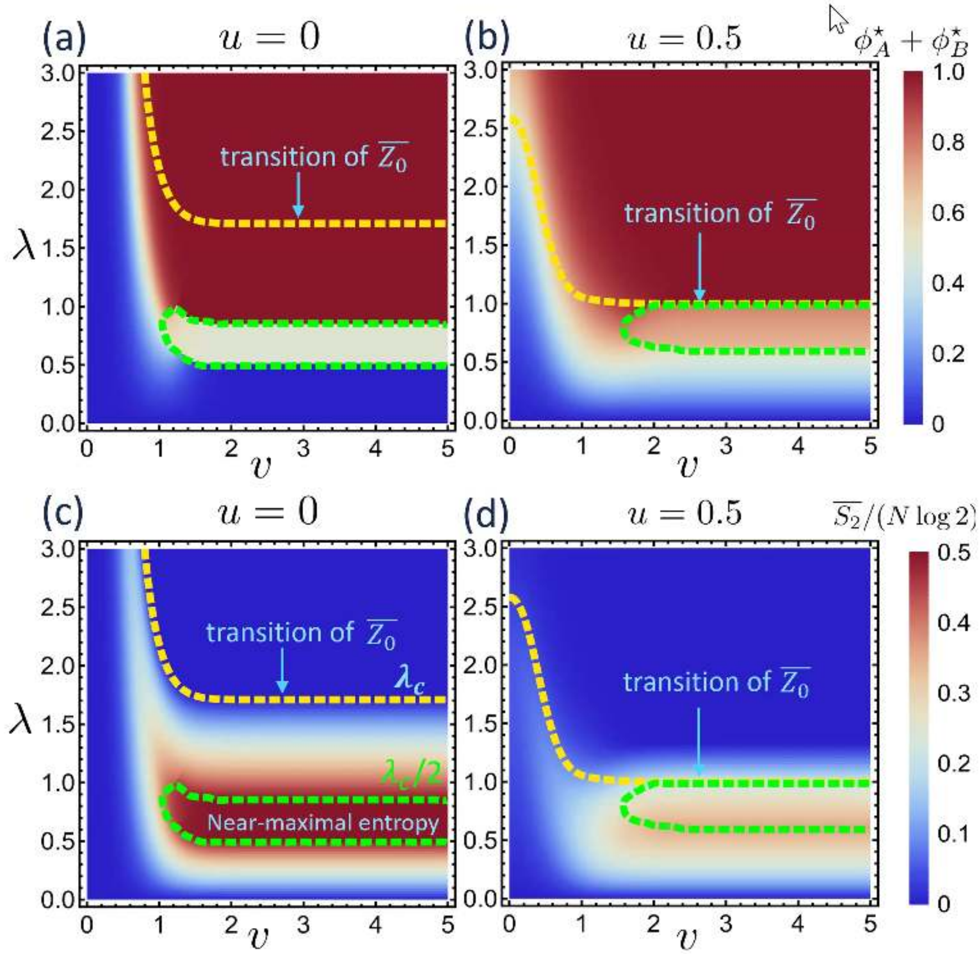
half-system size Renyi entanglement

Symmetry of $\phi_A \leftrightarrow \phi_B$.

Symmetric phase at small λ and large λ.

Symmetry broken phase at intermediate λ.

The symmetry broken phase can have near-maximal entanglement if variance of imaginary weight is large and variance of real weight is 0.



Intuition:

More localized in the Ising spin basis → cannot have large entangled.

$$D_q = \frac{1}{N \log 2} \frac{\log \mathrm{IPR}_q}{1-q}, \quad \mathrm{IPR}_q = \frac{\sum_{\boldsymbol{s}} |\Psi(\boldsymbol{s})|^{2q}}{(\sum_{\boldsymbol{s}} |\Psi(\boldsymbol{s})|^2)^q}$$

Product state in Ising basis: $D_q = 0$, localized states.
Uniform magnitude in Ising basis: $D_q = 1$, extended states.

Other cases: non-ergodic extended states. q-dependence: multifractality.

Moreover, saturation of the bound at half-system

$$\frac{S_q(A)}{N} \leq D_q \log 2$$

G. De Tomasi and I. M. Khaymovich, PRL, 124, 200602 (2020)

Saturation condition is when the reduced density matrix is diagonal in computational basis

$$\rho_A \simeq \frac{1}{\mathcal{N}} \left[\left(e^{-\sum_{ij} J^{ij} \sigma_i^z \sigma_j^z + \ldots} \right) + \delta\rho \right]$$

Like MBL Hamiltonian, many conserved quantity, no level repulsion in the spectrum