“Solving the Quantum Many-Body Problem with Artificial Neural Networks”

Paper by Giuseppe Carleo and Matthias Troyer

https://doi.org/10.1126/science.aag2302

Presentation Group 14: Yizhou, Tian, Xiaocheng, Tai, Prince, Yueying

ILLINOIS

Content

- Many-body problem and existing approaches
- Why neural network
- Terminologies
- Finding the ground state
- Finding the time evolution
- Citation evaluation

Intro

----Prince

Describing The Quantum Many-Body Problem

ψ

- Ψ describes a quantum state, whether a single particle or a complex molecule.
- Encoding a generic many-body quantum state requires exponential amount of information

– Nature Simplifies this complexity

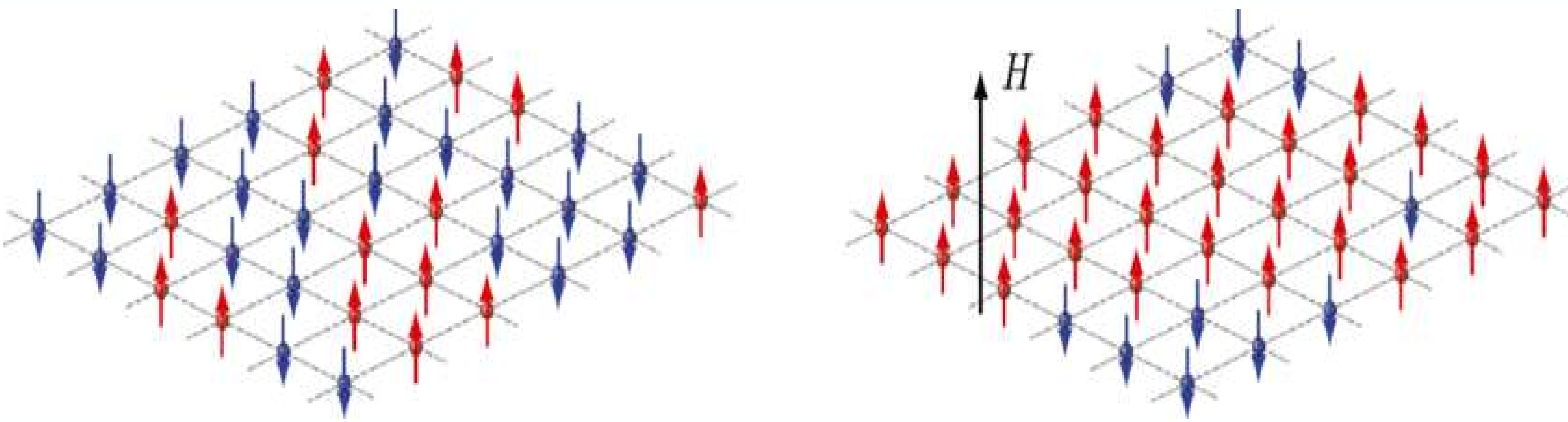


Figure: Parallel cross interpolation for high–precision calculation of high–dimensional integrals - Scientific Figure on ResearchGate

Existing Techniques and Challenges

- Quantum Monte Carlo (QMC) - Samples Finite Relevant Physical Configurations
- General Tensor Networks

- QMC and Sign Problem
- General Tensor Networks: inefficiency of current compression approaches in high-dimensional systems

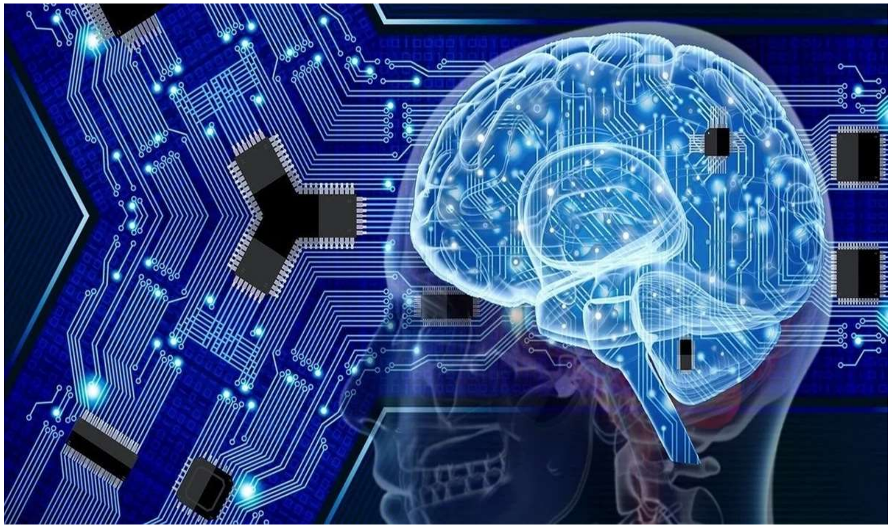
- The challenges in existing techniques makes it complex to analyze dynamical properties of high-dimensional systems and to exact ground-state properties of strongly interacting fermions

Solving the Quantum Many-Body Problem with Artificial Neural Networks

AI and ML has worked well in speech and text recognition, but the benefits of AI in solving many body problem are yet to be explored.

Can artificial neural network modify and adapt itself to describe and analyze a quantum system?

If it can, we could use this ability to solve the quantum many-body problem in those regimes so-far inaccessible by existing exact numerical approaches.



"Artificial Neural Networks Learn Better When They Spend Time Not Learning at All." *Today*, today.ucsd.edu/story/artificial-neural-networks-learn-better-when-they-spend-time-not-learning-at-all. Accessed 11 Dec. 2024.

Why Neural Network for Many-Body Problems?

A Presentation on Quantum Many-Body Problem and Neural Networks --- Yueying Wu

ILLINOIS

Overview

- - Many-body problems involve exponential complexity in describing quantum states.
- - Key Insight: They require dimensional reduction and feature extraction.

Dimensional Reduction and Feature Extraction

- - Dimensional Reduction: Simplifies high-dimensional data (e.g., Hilbert space).
- - Feature Extraction: Identifies and encodes relevant correlations in data.
- - Quantum Context: Wavefunctions are high-dimensional but contain structured correlations.

Why Neural Networks?

- - Efficiently approximate high-dimensional functions.
- - Capture non-local correlations better than tensor networks.
- - Flexible architectures adapt to symmetry and specific problems.
- - Systematic accuracy improvement by increasing hidden neurons.

Previous Applications in Physics

- \- Phase Classification:
- Carrasquilla & Melko (2017): Neural networks classify quantum phases.
- \- Phase Transitions:
- Wang (2016): Detecting phase transitions via unsupervised learning.
- \- This Work: Approximates wavefunctions for ground state and dynamics problems.

Conclusion

- - Neural networks open a new avenue for solving quantum many-body problems.
- - They excel in dimensional reduction, feature extraction, and modeling correlations.

Terminologies

Neural quantum state, variational Monte Carlo and restricted Boltzmann machine

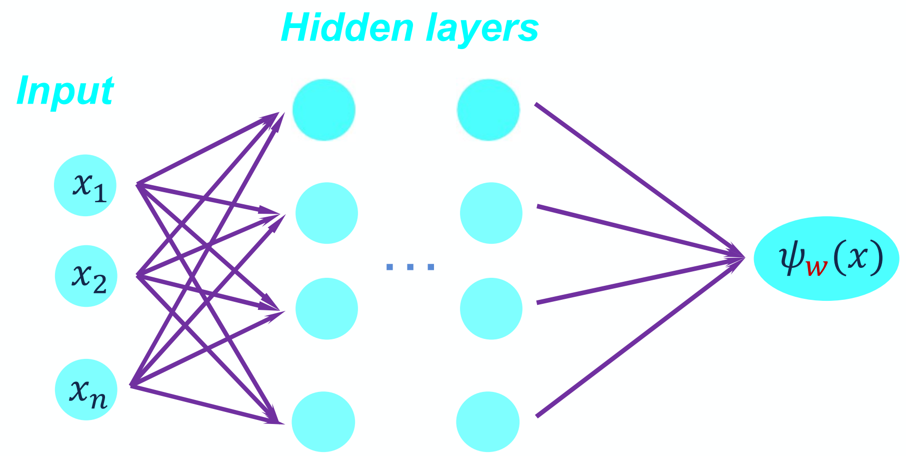
----Tai

ILLINOIS

Neural quantum state

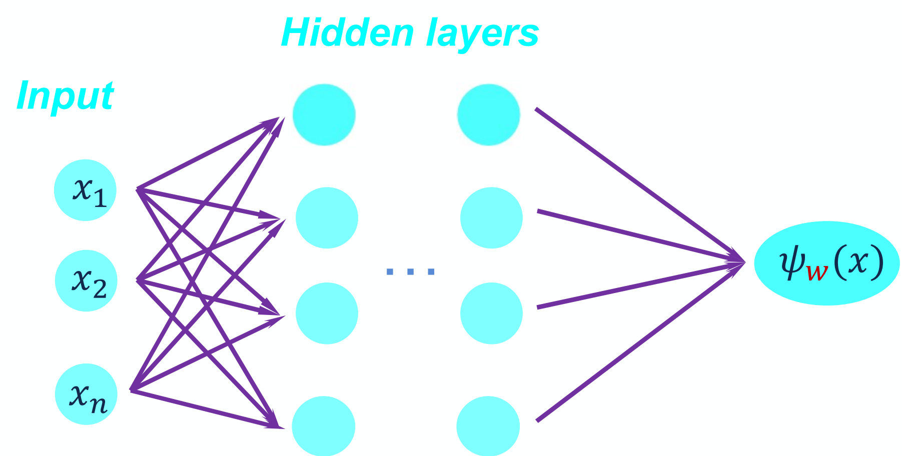
$$|\psi_w(x)\rangle = \sum_{x\in\{0,1\}^n} \psi_w(x)\,|x\rangle$$

Input: $x \in \{0,1\}^n$
Output: $\psi_w(x) \in \mathbb{C}$



Neural quantum state

$|\psi_w(x)\rangle = \sum_{x\in\{0,1\}^n} \psi_w(x)\,|x\rangle$



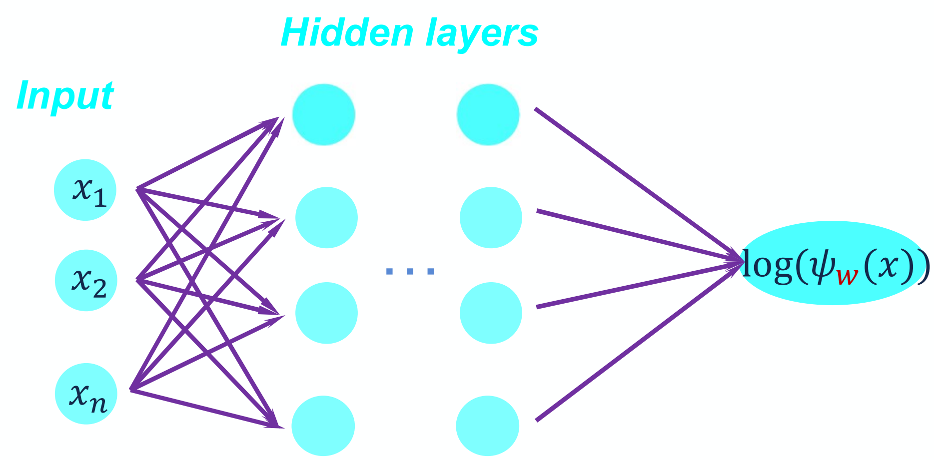
Input: $x \in \{0,1\}^n$
Output: $\psi_w(x) \in \mathbb{C}$

Ground state can be found by minimizing energy with respect to network parameter w

$$E_0 = \min_w \langle \psi_w | H | \psi_w \rangle$$

Neural quantum state

$$|\boldsymbol{\psi_w}(\boldsymbol{x})\rangle = \sum_{x \in \{0,1\}^n} \boldsymbol{\psi_w}(\boldsymbol{x})\,|\boldsymbol{x}\rangle$$



Input: $x \in \{0, 1\}^n$
Output: $\psi_w(x) \in \mathbb{C}$

Ground state can be found by minimizing energy with respect to network parameter w

$$E_0 = \min_w \langle \psi_w | H | \psi_w \rangle$$

Variational Monte Carlo

$$E_0 = \min_w \langle \psi_w | H | \psi_w \rangle$$

Variational Monte Carlo

$$\boldsymbol{E_0} = \min_{w} \langle \psi_w | H | \psi_w \rangle$$
$$= \sum_{x,y} \min_{w} \langle \psi_w | x \rangle \langle x | H | y \rangle \langle y | \psi_w \rangle$$

Variational Monte Carlo

$$\boldsymbol{E_0} = \min_{\boldsymbol{w}} \langle \psi_{\boldsymbol{w}} | H | \psi_{\boldsymbol{w}} \rangle$$

$$= \sum_{x,y} \min_{\boldsymbol{w}} \langle \psi_{\boldsymbol{w}} | x \rangle \langle x | H | y \rangle \langle y | \psi_{\boldsymbol{w}} \rangle$$

$$= \sum_{x,y} \min_{\boldsymbol{w}} \psi_{\boldsymbol{w}}^*(x) \psi_{\boldsymbol{w}}(y) \langle x | H | y \rangle$$

Variational Monte Carlo

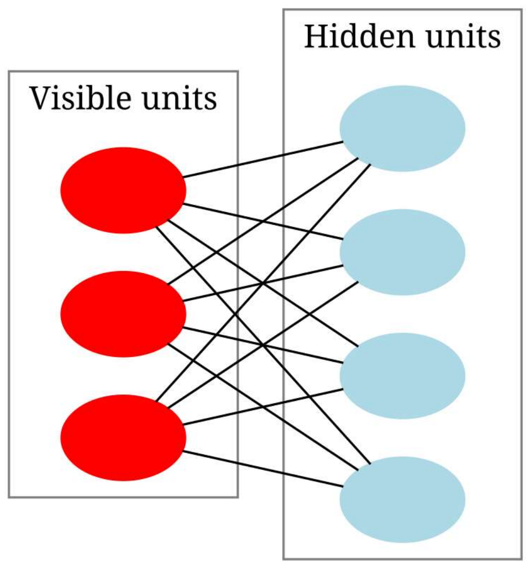
$$\begin{aligned} E_0 &= \min_w \langle \psi_w | H | \psi_w \rangle \\ &= \sum_{x,y} \min_w \langle \psi_w | x \rangle \langle x | H | y \rangle \langle y | \psi_w \rangle \\ &= \sum_{x,y} \min_w \, \psi_w^*(x) \psi_w(y) \, \langle x | H | y \rangle \\ &= \sum_{x,y} \min_w |\psi_w(x)|^2 \frac{\psi_w(y)}{\psi_w(x)} \langle x | H | y \rangle \end{aligned}$$

Variational Monte Carlo

$$
\begin{aligned}
E_0 &= \min_w \langle \psi_w | H | \psi_w \rangle \\
&= \sum_{x,y} \min_w \langle \psi_w | x \rangle \langle x | H | y \rangle \langle y | \psi_w \rangle \\
&= \sum_{x,y} \min_w \, \psi_w^*(x) \psi_w(y) \, \langle x | H | y \rangle \\
&= \sum_{x,y} \min_w |\psi_w(x)|^2 \frac{\psi_w(y)}{\psi_w(x)} \langle x | H | y \rangle \\
&= E_{x \sim p(x)} \left[\sum_y \frac{\psi_w(y)}{\psi_w(x)} \langle x | H | y \rangle \right]
\end{aligned}
$$

Restricted Boltzmann Machine

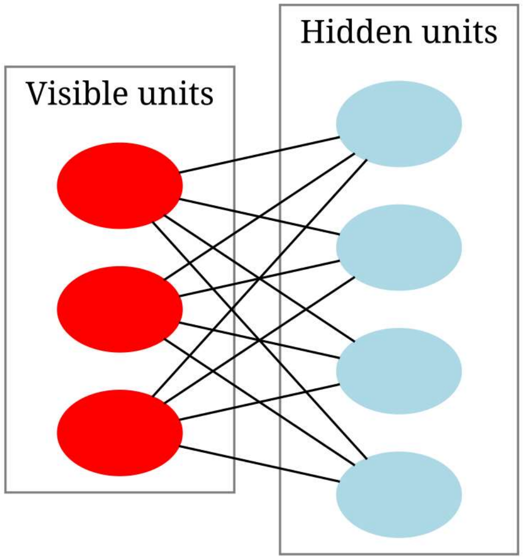
$$E(v,h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i w_{i,j} h_j$$



Restricted Boltzmann Machine

$$E(v,h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i w_{i,j} h_j$$

$$P(v) = \sum_{h \in \{0,1\}^n} \frac{e^{-E(v,h)}}{Z}$$

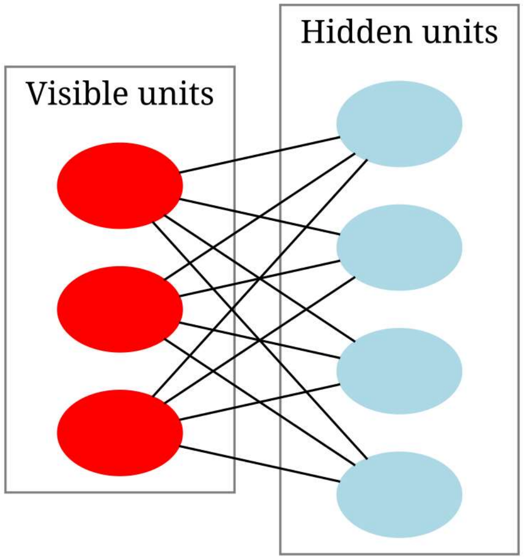


Restricted Boltzmann Machine

$$E(v,h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i w_{i,j} h_j$$

$$P(v) = \sum_{h \in \{0,1\}^n} \frac{e^{-E(v,h)}}{Z}$$

$$= \sum_j \frac{2\cosh\left(\sum_i w_{i,j} v_i + b_j\right) e^{\sum_i a_i v_i}}{Z}$$

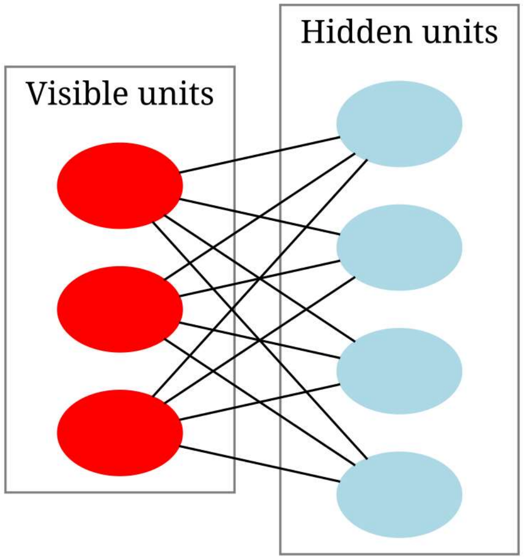


Restricted Boltzmann Machine

$$E(v,h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i w_{i,j} h_j$$

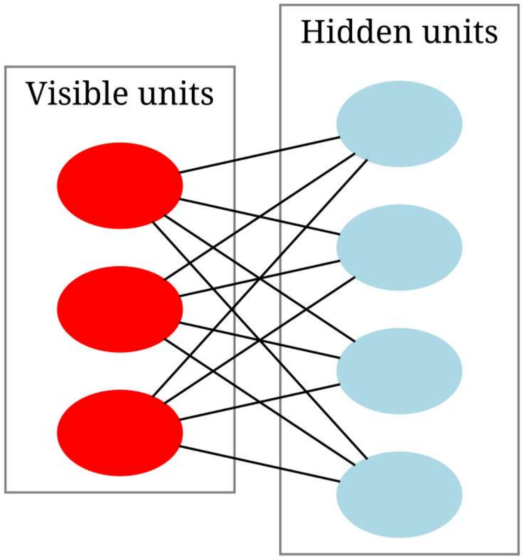
$$P(v) = \sum_{h \in \{0,1\}^n} \frac{e^{-E(v,h)}}{Z}$$

$$= \sum_j \frac{2\cosh\left(\sum_i w_{i,j} v_i + b_j\right) e^{\sum_i a_i v_i}}{Z}$$



Restricted Boltzmann Machine

$$\log(\psi(x)) = 2logcosh\left(\sum_j W_{ij}\, x_j + b_j\right) + \sum_i a_i x_i$$



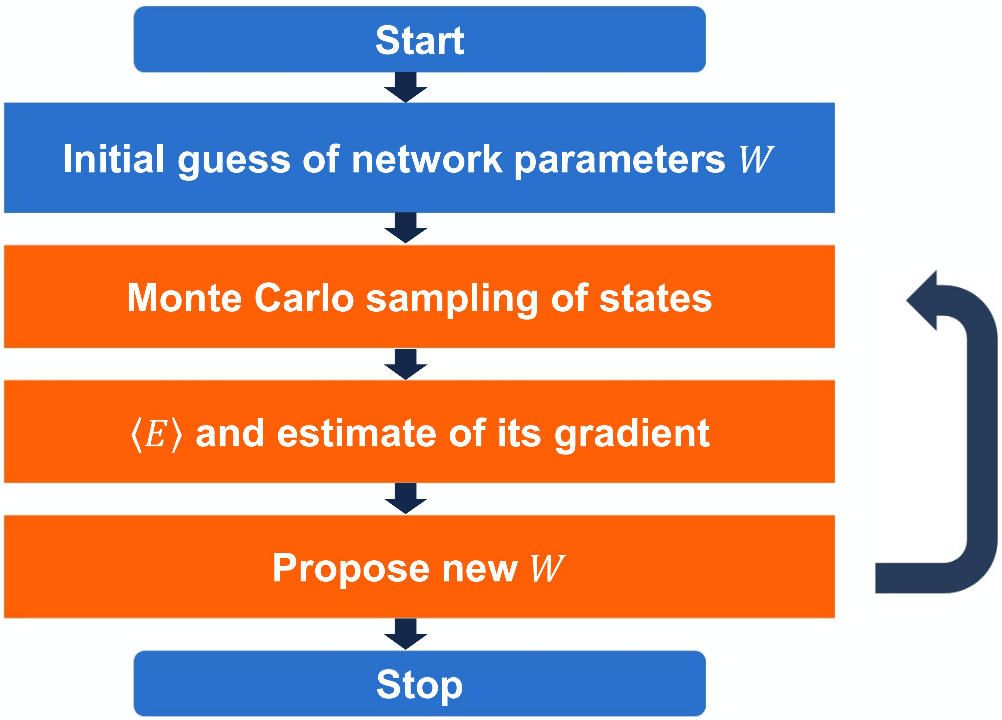
It's actually just single hidden layer neural network

Determining the ground state

----Xiaocheng

ILLINOIS

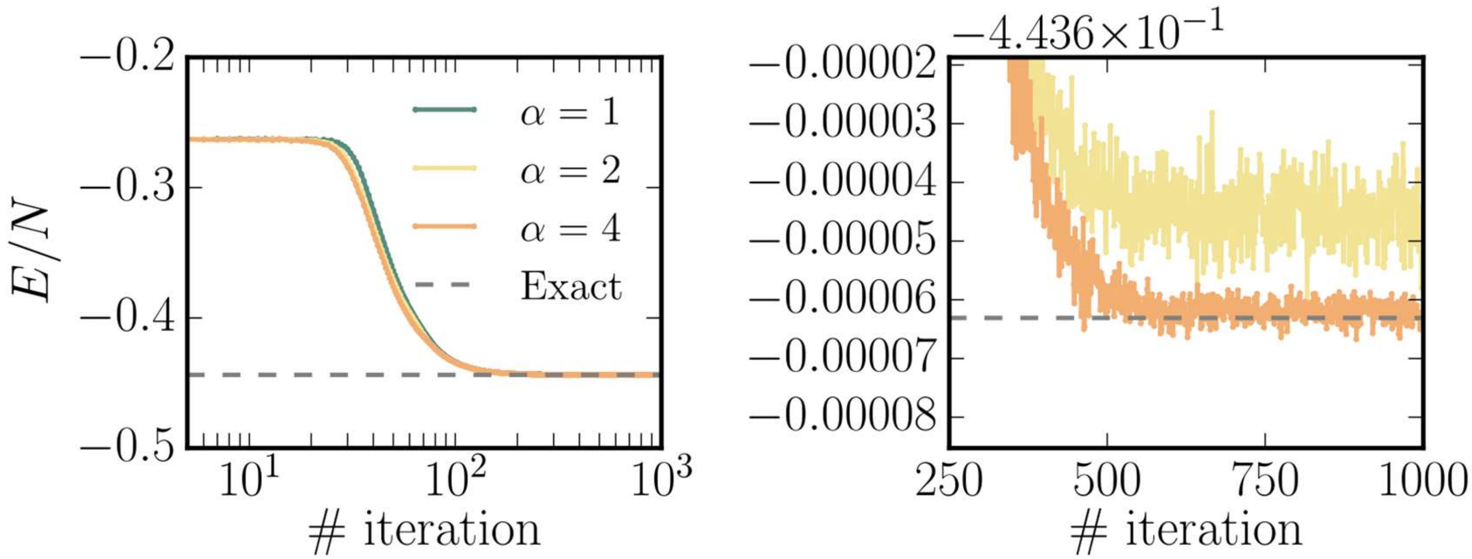
Finding the Ground State: Algorithm



Stochastic Reconfiguration (SR) or Stochastic Gradient Descent (SDG) can be used in new W proposal.

Finding the Ground State: Results

- Even with minimal α the network learns the info of ground state and gives the right result.
- Networks w/ higher value of α converge slightly better



$\alpha = \frac{M}{N}$. Here M, N is number of hidden and visible variables.

Finding the Ground State: Results

Two prototypical spin models were used to validate the scheme

Transverse-field Ising (TFI)

$$H_{TFI} = -h \sum_{i} \sigma_i^x \quad - \quad \sum_{\langle i,j \rangle} \sigma_i^z \, \sigma_j^z$$

Favors alignment

External Transverse Field   Nearest-Neighbor Interaction

Antiferromagnetic Heisenberg (AFH)

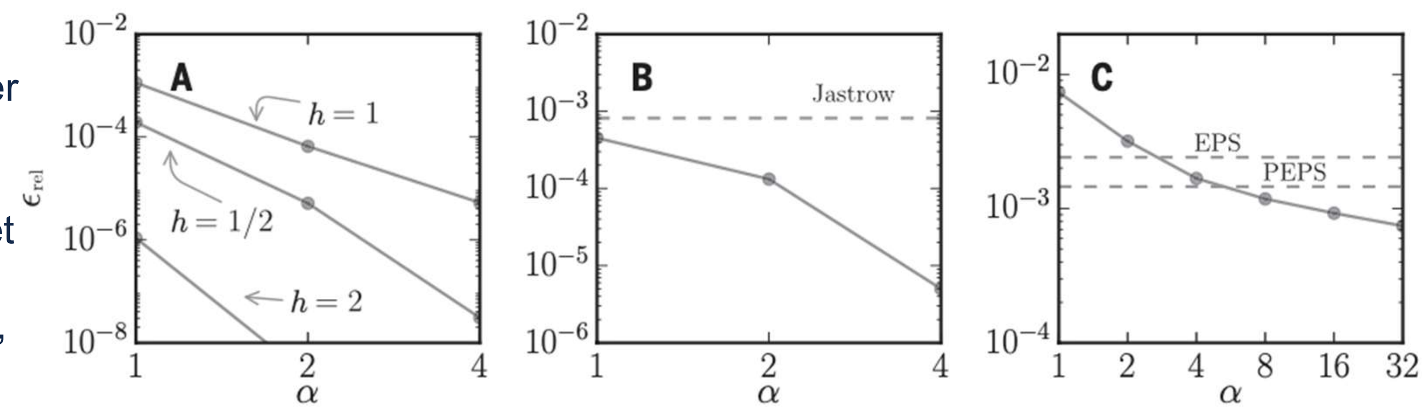
$$H_{AFH} = \sum_{\langle i,j \rangle} \sigma_i^x \sigma_i^x + \sigma_i^y \sigma_i^y + \sigma_i^z \sigma_j^z$$

Favors opposite alignment

Nearest-Neighbor Interaction

Finding the Ground State: Error

- Systematical decrease in error when increase number of hidden variables (α)
- By increasing α one can get higher accuracy than other common methods (Jastrow, EPS, PEPS)
- A network's accuracy depends on model parameters (h in this case)



1D TFI for an 80-spin chain | 1D AFH for an 80-spin chain | 2D AFH on 10-by-10 square lattice

$$\epsilon_{rel} = (E_{NQS} - E_{exact})/|E_{exact}|$$

Time evolution

----Tian

ILLINOIS

Time-dependent Variational Monte Carlo

Static $$\mathcal{H}\Psi = E\Psi$$

minimize $$|\mathcal{H}\Psi - E\Psi|^2$$

Dynamic $$i\hbar\frac{\partial}{\partial t}\Psi(t) = \mathcal{H}\Psi(t)$$

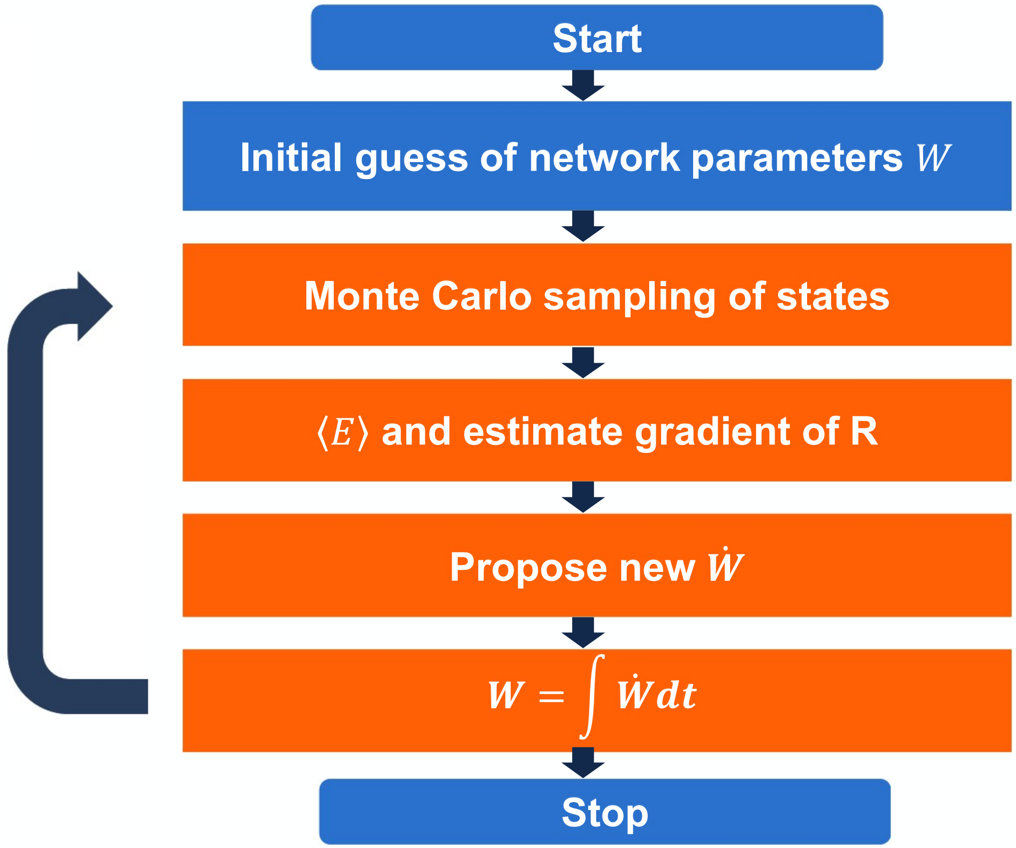
minimize $$|\mathcal{H}\Psi - i\hbar\frac{\partial}{\partial t}\Psi|^2$$

Time-dependent Variational Monte Carlo

$$R(\dot{W}(t)) = \mathrm{dist}(\partial_t \Psi, -i\mathcal{H}\Psi)$$

Weight of the network

Use gradient descendent until the convergence criteria is met

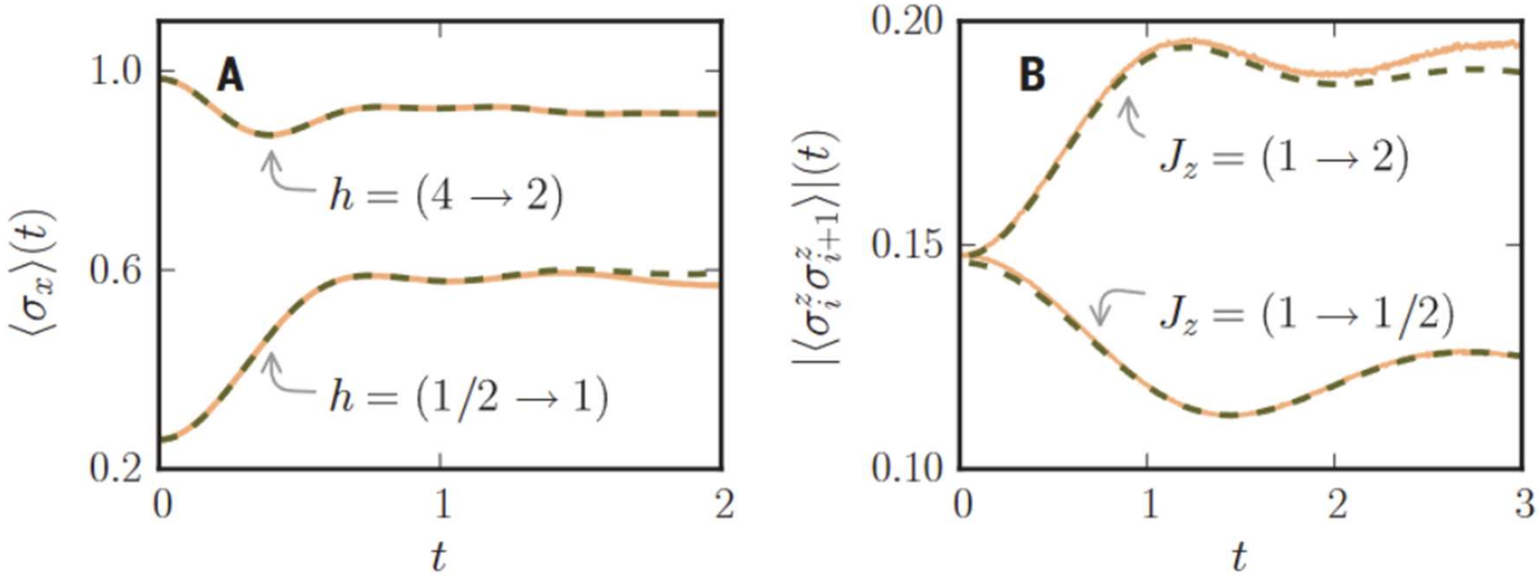


Time-dependent Hamiltonian

$$\mathcal{H}_{TFI} = -h(t)\sum_i \sigma_i^x - \sum_{\langle i,j\rangle} \sigma_i^z \sigma_j^z$$

$$\mathcal{H}_{AFH} = \sum_{\langle i,j\rangle} \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + h(t)\sigma_i^z \sigma_j^z$$

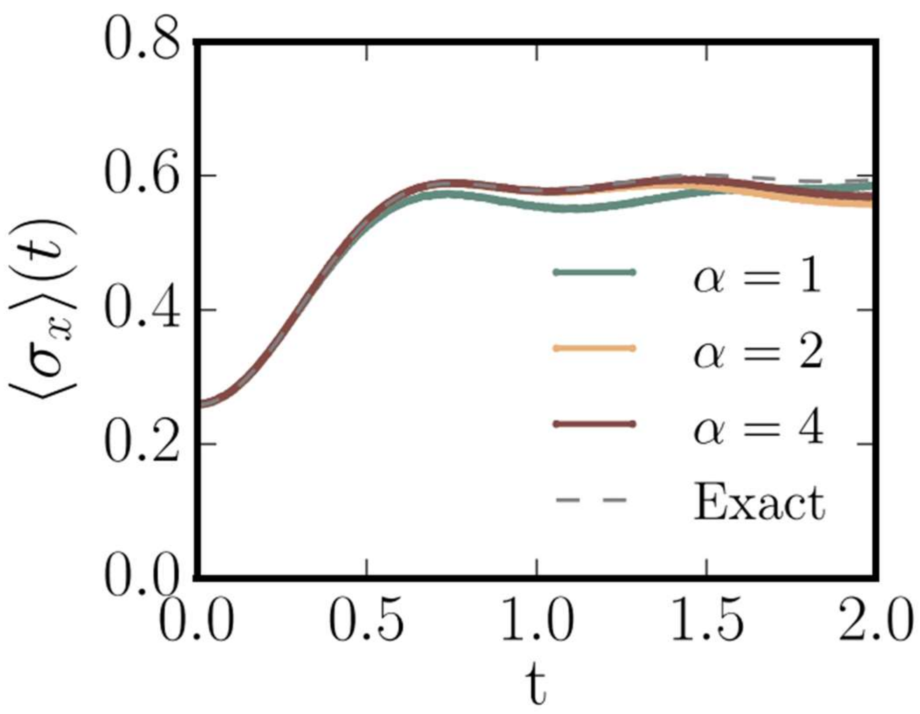
Numerical results of NQS



Numerical result from NQS

Numerical result from t-DMRG

Numerical results of NQS



- Numerical result is close to the exact result
- Numerical result converges to the exact result as the number of hidden variables increases
- All numerical results fails to converge to the exact result at long time

Citation evaluation

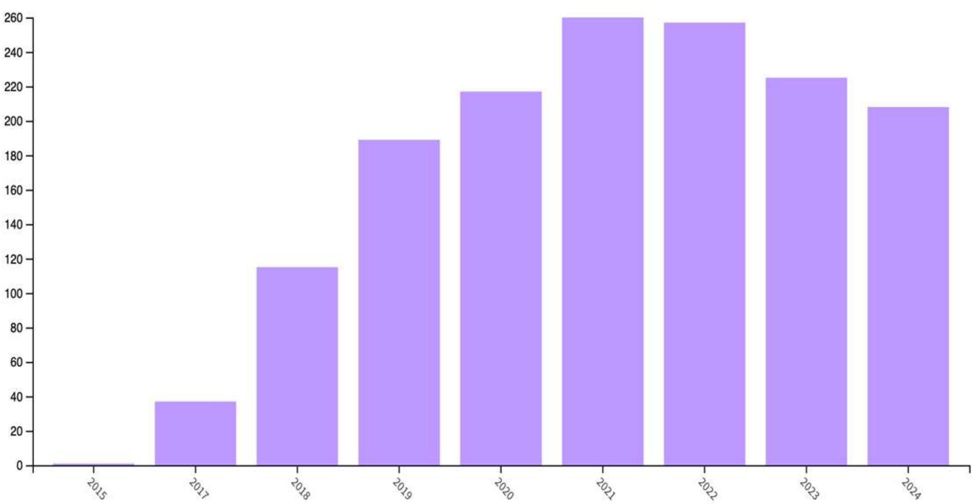
----Yizhou

ILLINOIS

Citation evaluation

According to Web of Science

1549 citations so far



Mainly physics papers.
Also covering math, CS, chemistry…

Field: Research Areas	Record Count	% of 1,509
Physics	1,184	78.463%
Mathematics	827	54.805%
Computer Science	671	44.467%
Chemistry	444	29.423%
Mechanics	407	26.972%

Since then

- Improvements
  - Numerical stability *(M. Schmitt and M. Heyl, Quantum many-body dynamics in two dimensions with artificial neural networks, Phys. Rev. Lett. 125, 100503 (2020))*
  - Noise reduction *(A. Sinibaldi, C. Giuliani, G. Carleo, and F. Vicentini, Unbiasing time-dependent variational Monte Carlo by projected quantum evolution, Quantum 7, 1131 (2023))*
  - Speedup *(Gui, Shaojun, Ho, Tak-San and Rabitz, Herschel, Discrete real-time learning of quantum-state subspace evolution of many-body systems in the presence of time-dependent control fields)*
- Challenge
  - Lack theoretical support *(Hsin-Yuan Huang et al, Provably efficient machine learning for quantum many-body problems.Science**377**,eabk3333(2022))*

In other fields

- Quantum gravity
  - To find Hamiltonian constraints

  *(Hanno Sahlmann and Waleed Sherif 2024 Class. Quantum Grav. **41** 225014)*
- Lattice gauge theory
  - sign problem; dynamics in real time

  (Apte et al, Deep learning lattice gauge theories, Physical Review B)
- ……

Thank you!

Feel free to ask any questions~

Team 14

ILLINOIS