Journal of Modern Physics, 2025, 16(5), 686-714
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Variational Methods for Solving High-Dimensional Quantum Systems

Daming Li

School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai, China
Email: lidaming@sjtu.edu.cn

How to cite this paper: Li, D.M. (2025) Variational Methods for Solving High-Dimensional Quantum Systems. *Journal of Modern Physics*, **16**, 686-714.
https://doi.org/10.4236/jmp.2025.165038

Received: February 18, 2025
Accepted: May 23, 2025
Published: May 26, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Variational methods are highly valuable computational tools for solving high-dimensional quantum systems. In this paper, we explore the effectiveness of three variational methods: density matrix renormalization group (DMRG), Boltzmann machine learning, and variational quantum eigensolver (VQE). We apply these methods to solve two different quantum systems: the Fermi-Hubbard model in condensed matter physics and the Schwinger model in high energy physics. To facilitate the computations on quantum computers, we map each model to a spin 1/2 system using the Jordan-Wigner transformation. This transformation allows us to take advantage of the capabilities of quantum computing. We calculate the ground state of both quantum systems and compare the results obtained using three variational methods. Our aim is to demonstrate the power and effectiveness of these variational approaches in tackling complex quantum systems.

Keywords

Density Matrix Renormalization Group, Boltzmann Machine Learning, Variational Quantum Eigensolver, Fermi-Hubbard Model, Schwinger Model

1. Introduction

Many exciting phenomena in quantum many-body systems are due to the interplay of quantum fluctuations and correlations. Celebrated examples are the superfluid Helium [1] [2], the fractional quantum Hall effect [3] [4], the Haldane phase in quantum spin chains [5] [6], quantum spin liquids [7], and high-temperature superconductivity [8]. The aim of theoretical physics is to understand the emergent properties for such challenging quantum many-body systems. The main difficulty in investigating quantum many-body problems is due to the fact that the Hilbert space spanned by the possible microstates grows exponentially with the system size. To unravel the physics of microscopic model systems and to study the robustness of quantum phases of matter, large scale numerical simulations are essential. The exact diagonization method is only possible for small many-body systems. For large systems, efficient quantum Monte Carlo (QMC) methods can be applied. In a large class of quantum many-body systems (e.g., fermionic degrees of freedom, geometric frustration), however, these QMC sampling techniques cannot be used effectively due to the sign problem [9]. In this case, the variational methods have been shown to be a powerful tool to efficiently simulate quantum many-body systems.

The first type of variational method is the density matrix renormalization group (DMRG) method which was originally developed to study ground state properties of one-dimensional systems. The success of the DMRG method is based on the area law of the quantum ground state [10], and thus can be represented efficiently using matrix product states (MPS) [11]-[19]. This algorithm has been generalized to study two-dimensional systems [20], Abelian and non-Abelian symmetries [21]-[25], single-site optimization with density matrix perturbation [26] [27], hybrid real-momentum space representation [28] [29], and the development of real-space parallelization [30], continuous matrix product state (cMPS) [31] [32]. The cMPS can also be used to study the Lieb-Liniger model [33] [34], the Gaudin-Yang model [35], periodic bc atomtronics [36] and 1 + 1 relative Bose theories [37].

The machine learning has been extensively used in physical sciences [38]. The second type of variational method is based on machine learning, where neutral network variational ansatz efficiently represents highly correlated quantum states and their parameters are easily optimized by means of the variational Monte Carlo (VMC) method [39], which has powerful expressibility for quantum state [40]-[43]. Alongside with restricted Boltzmann machine (RBM) [44]-[59], other network structures such as a feed-forward [60]-[63], recurrent neural networks [64] [65], autoregressive neural TensorNet [66], and convolutional neural networks [67]-[78], have been adopted. Recurrent neural networks (RNNs) are tremendously powerful tools that have been used in language modeling [79], speech recognition [80], and image generation [81] [82]. RNNs process data streams by maintaining a hidden state, which is updated by applying an identical function to the previous hidden state and the input at the next time step.

At the same time, neural networks using self-attention layers, like the Transformer [83], have had a profound impact on much of machine learning. They have led to breakthroughs in natural language processing [84], language modeling [85], image recognition [86], and protein folding [87]. However, the application of Transformer in this field is still rather limited, with a few results concerning quantum lattice models [88], open systems [89], quantum state tomography [90] and quantum circuit simulation [91]. Therefore, the full potential of the transformer architecture has yet to be explored.

The third type of variational method is called the variational quantum eigensolver (VQE), which is extensively used in quantum calculation. This is a hybrid algorithm, where the quantum state is prepared by a quantum algorithm, which is implemented according to a quantum circuit with many quantum gates. Classical computer is used to optimize the parameters used in quantum circuit. Compared with the matrix product state and machine learning ansatz, the variational state in VQE is realized by quantum algorithm, which can be implemented in the quantum hardware devices, and thus exponential speedup over classical methods becomes possible. The previous quantum algorithms to find the ground state of a given Hamiltonian were based on adiabatic state preparation and quantum phase estimation subroutines [92] [93], both of which have circuit depth requirements beyond those available in the NISQ era. Here we listed both the original VQE architecture [94] and some more advanced methods: orthogonality constrained VQE [95], subspace expansion method [96], subspace VQE [97] [98], multistate contracted VQE [99], adiabatically assisted VQE [100] [101], and accelerated VQE [102]-[104]. To study the dynamics of many-body quantum systems, the conventional quantum Hamiltonian simulation algorithm such as the Trotter-Suzuki product formula [105] was used. However, the circuit depth of standard Trotterization methods can rapidly exceed the coherence time of noisy quantum computers. This has led to recent proposals for variational approaches to dynamical simulation, including iterative variational algorithms [106] [107], imaginary time evolution [108], general first order derivative equations with non-Hermitian Hamiltonians [109], adaptive ansatz to reduce the circuit depth [110] [111] and variational fast forwarding [112]-[115].

In this paper, these three variational methods are discussed in detail and are used to solve the ground state energy of the Fermi-Hubbard model and Schwinger model. The paper is organized as follows: The details of the three variational methods are given in Sections 2, 3, and 4. In Sections 5 and 6, the Fermi-Hubbard model and Schwinger model have been mapped to spin 1/2 systems using the Jordan-Wigner representation. In Section 7, the ground state energy of the Fermi-Hubbard model and Schwinger model are calculated using these three variational methods. The discussion is given in the final Section 8.

2. Density Matrix Renormalization Group

A general discrete Hilbert space of the many-body quantum systems with N sites has the structure

$$\mathcal{H} = \text{span}\left\{ \left| s_0 \right\rangle \otimes \cdots \otimes \left| s_{N-1} \right\rangle \mid s_i \in \mathcal{L}_i, i \in \{0, \cdots, N-1\} \right\} \tag{1}$$

where $\mathcal{L}_i$ is the set of discrete quantum numbers at site i. $\left\{ \left| s_i \right\rangle \right\}_{s_i \in \mathcal{L}_i}$ is the standard orthogonal basis for the local Hilbert space $\mathcal{H}_i$ corresponding to the i th site, $i = 0, \cdots, N-1$. The whole Hilbert space $\mathcal{H}$ is the tensor product of individual local Hilbert space: $\mathcal{H} = \otimes_{i=0}^{N-1} \mathcal{H}_i$. The standard orthogonal basis $\left| s_0 \cdots s_{N-1} \right\rangle \equiv \left| s_0 \right\rangle \otimes \cdots \otimes \left| s_{N-1} \right\rangle$ in $\mathcal{H}$ is also called the computational basis. For the qubit systems, $\mathcal{L}_i = \{0,1\}$. For spin-1/2, $\mathcal{L}_i = \{-1,1\}$, which is the two eigenvalues of Pauli matrix Z. For general spin-S with integral or half integral $S = 1/2, 1, 3/2, 2, \cdots$, $\mathcal{L}_i = \{-2S, -2S+2, \cdots, 2S-2, 2S\}$ and $\mathcal{H}_i$ has the dimension $2S+1$. Sometimes, the dimension of Hilbert space is reduced under some kind of restriction. For example, the spin 1/2 system with the restriction $\sum_{i=0}^{N-1} Z_i = 0$, the dimension 16 of $\mathcal{H}$ is reduced to 6 if $N = 4$.

The state $|\psi\rangle \in \mathcal{H}$ can be represented by

$$|\psi\rangle = \sum_s \psi(s)|s\rangle = \sum_{s_0,\cdots,s_{N-1}} \psi(s_0,\cdots,s_{N-1})|s_0 \cdots s_{N-1}\rangle \tag{2}$$

where the sum s over all quantum numbers is assumed. Since the computational basis is standard orthogonal, the wave function $\psi(s) = \langle s|\psi\rangle$ is the inner product of $|\psi\rangle$ and $|s\rangle$.

The Hamiltonian H of many-body quantum system is a Hermitian operator from d^N dimensional Hilbert space $\mathcal{H}$ into itself. We assume that the dimension d of $\mathcal{H}_i$ are same for $i = 0, \cdots, N-1$. The calculation of the spectra of H is difficult due to the exponentially large dimension over N. For small N, exact diagonalization can be adopted since H can be represented as the Hermitian matrix in the computation basis $|s\rangle$.

A general operator $\hat{O}$ in the basis $|s\rangle$ can be written as

$$\hat{O} = \sum_{s,s'} O_{s'_0\cdots s'_{N-1}}^{s_0\cdots s_{N-1}} |s_0 \cdots s_{N-1}\rangle\langle s'_{N-1} \cdots s'_0| \tag{3}$$

Since the basis $|s\rangle$ is orthogonal to each other,

$$\hat{O}|s'_0 \cdots s'_{N-1}\rangle = \sum_s O_{s'_0\cdots s'_{N-1}}^{s_0\cdots s_{N-1}} |s_0 \cdots s_{N-1}\rangle \tag{4}$$

i.e., $\hat{O}|s'\rangle = \sum_s O_{s'}^s |s\rangle$. Apply this operator $\hat{O}$ to a general state $|\psi\rangle$ in (2), one has

$$\hat{O}|\psi\rangle = \sum_{s,s'} O_{s'}^s \psi(s')|s\rangle \tag{5}$$

The inner product between $\hat{O}|\psi\rangle$ and $|\psi\rangle$ is

$$\langle\psi|\hat{O}|\psi\rangle = \sum_{s,s'} O_{s'}^s \psi(s')\overline{\psi(s)} \tag{6}$$

where $\overline{\psi(s)}$ denotes the complex conjugate of $\psi(s)$.

The wave function $\psi(s_0,\cdots,s_{N-1})$ can be regarded as the tensor $\psi_{s_0\cdots s_{N-1}}$ of order N, and thus the manipulation of the wave function is reduced to be that of tensor $\psi_{s_0\cdots s_{N-1}}$. Similarly, $O_{s'}^s = O_{s'_0\cdots s'_{N-1}}^{s_0\cdots s_{N-1}}$ is the tensor of order $2N$. The computational cost of $\sum_{s'} O_{s'}^s \psi(s') = \sum_{s'_0,\cdots,s'_{N-1}} O_{s'_0\cdots s'_{N-1}}^{s_0\cdots s_{N-1}} \psi_{s'_0,\cdots,s'_{N-1}}$ is $O(d^{2N})$ which is not acceptable. The trick of reducing computational cost is the decomposition of tensor and then truncation. A very important operation for tensor is the decomposition as follows:

$$\psi_{s_0\cdots s_{N-1}} = M^{[0]s_0} M^{[1]s_1} \cdots M^{[N-1]s_{N-1}} \tag{7}$$

where $M^{[i]s_i} = \left(M^{[i]s_i}\right)_{\alpha_i\alpha_{i+1}}$ is the $\chi_i \times \chi_{i+1}$ matrix for fixed (physical) indices s_i, $i = 0, \cdots, N-1$, *i.e.*, $M^{[i]}$ is a tensor of order 3. The index $\{\alpha_i\}_{i=0}^N$ is contracted in the multiplication of all $M^{[i]s_i}$, and $\alpha_0 = \chi_0 = \alpha_N = \chi_N = 1$ since the contracted result $\psi_{s_0\cdots s_{N-1}}$ is a complex number. Since the index $\{\alpha_i\}_{i=0}^{N}$ is not physical, they are called virtual indices.

The decomposition (7) can be realized by $N-1$ steps of singular value decomposition (SVD) of matrices, which is given in detail in Appendix. Inserting (7) into (2), we get the matrix product state

$$|\psi\rangle = \sum_{s_0,\cdots,s_{N-1}} M^{[0]s_0} M^{[1]s_1} \cdots M^{[N-1]s_{N-1}} |s_0 \cdots s_{N-1}\rangle \quad (8)$$

If $s_i = 1,\cdots,d$ for $i = 0,\cdots,N-1$, then $\chi_1 = d$ and $\chi_{i+1} \le d\chi_i$, $i = 1,\cdots,N-2$, and thus $\chi_i \approx d^i$. The exact decomposition of (7) will cause exponential disaster. Fortunately, if the state of interest is not highly entangled, the truncation in SVD will not cause much error for the approximation of this state.

As shown in Appendix, if the tensor $\psi_{s_0\cdots s_{N-1}}$ is normalized

$$\sum_{s_0,\cdots,s_{N-1}} \left|\psi_{s_0\cdots s_{N-1}}\right|^2 = 1 \quad (9)$$

the N tensors $\left\{M^{[i]}\right\}_{i=0}^{N-1}$ of order 3 satisfy the left canonical condition

$$\sum_{s_i} \left(M^{[i]s_i}\right)^\dagger M^{[i]s_i} = \mathbb{I}_{\chi_{i+1}},\ i = 0,\cdots,N-1 \quad (10)$$

where $\mathbb{I}_{\chi_{i+1}}$ is the $\chi_{i+1} \times \chi_{i+1}$ identity matrix. For a general tenor $\psi_{s_0\cdots s_{N-1}}$ of order N, we have the left canonical form

$$\psi_{s_0\cdots s_{N-1}} = A^{[0]s_0} A^{[1]s_1} \cdots A^{[N-1]s_{N-1}} \Lambda^{[N]} \quad (11)$$

with the left canonical condition for the N tensors $\left\{A^{[i]}\right\}_{i=0}^{N-1}$

$$\sum_{s_i} \left(A^{[i]s_i}\right)^\dagger A^{[i]s_i} = \mathbb{I}_{\chi_{i+1}},\ i = 0,\cdots,N-1 \quad (12)$$

Here we intentionally introduce a number $\Lambda^{[N]}$ which becomes 1 if $\psi_{s_0\cdots s_{N-1}}$ satisfies (9). In fact, this left canonical form can also be obtained from the general matrix product form (7) where $M^{[i]}$ does not satisfy (10). This can be realized by $N-1$ steps of SVD. First,

$$\left(M^{[0]s_0}\right)_{\alpha_0,\alpha_1} = \sum_k U_{(\alpha_0 s_0),k} S_k \left(V^\dagger\right)_{k,\alpha_1} \quad (13)$$

where U and V are unitary matrices and $\{S_k\}$ are the singular values. Then $\left(A^{[0]s_0}\right)_{\alpha_0,k} = U_{(\alpha_0 s_0),k}$ satisfies the left canonical condition. The remaining part $S_k\left(V^\dagger\right)_{k,\alpha_1}$ is combined with $\left(M^{[1]s_1}\right)_{\alpha_1,\alpha_2}$ to get $\sum_{\alpha_1} S_k\left(V^\dagger\right)_{k,\alpha_1}\left(M^{[1]s_1}\right)_{\alpha_1,\alpha_2}$ and use SVD for this matrix again, we can find $A^{[1]s_1}$ which satisfies the left canonical condition. Continuing this process, we can get the left canonical form (11).

Similarly, if the SVD is implemented from the right-hand side, the general tensor has the right canonical form

$$\psi_{s_0\cdots s_{N-1}} = \Lambda^{[0]} B^{[0]s_0} B^{[1]s_1} \cdots B^{[N-1]s_{N-1}} \quad (14)$$

with the right canonical condition for the tensor $\left\{B^{[i]}\right\}_{i=0}^{N-1}$ of order 3

$$\sum_{s_i} B^{[i]s_i}\left(B^{[i]s_i}\right)^{\dagger} = \mathbb{I}_{\chi_i},\ i = 0,\cdots,N-1 \tag{15}$$

If the SVD is implemented from both sides, they will meet at some bonds, e.g.,

$$\psi_{s_0\cdots s_{N-1}} = \Lambda^{[0]}A^{[0]s_0}A^{[1]s_1}\Lambda^{[2]}B^{[2]s_2}\cdots B^{[N-1]s_{N-1}}\Lambda^{[N]} \tag{16}$$

We call this the mixed canonical form. Here $\Lambda^{[2]}$ is a $\chi_2 \times \chi_2$ diagonal matrix.

A full decomposition is

$$\psi_{s_0\cdots s_{N-1}} = \Lambda^{[0]}\Gamma^{[0]s_0}\cdots\Lambda^{[N-1]}\Gamma^{[N-1]s_{N-1}}\Lambda^{[N]} \tag{17}$$

where $\Gamma^{[i]s_i}$ is the $\chi_i \times \chi_{i+1}$ matrix and $\Lambda^{[i]}$ is the $\chi_i \times \chi_i$ diagonal matrix. Obviously, these three canonical forms are related to each other

$$A^{[i]s_i} = \Lambda^{[i]}\Gamma^{[i]s_i} = \Lambda^{[i]}B^{[i]s_i}\left(\Lambda^{[i+1]}\right)^{-1},\ B^{[i]s_i} = \Gamma^{[i]s_i}\Lambda^{[i+1]} = \left(\Lambda^{[i]}\right)^{-1}A^{[i]s_i}\Lambda^{[i+1]} \tag{18}$$

for $i = 0,\cdots,N-1$.

Assume that $O_{s'}^{s} = O_{s'_0\cdots s'_{N-1}}^{s_0\cdots s_{N-1}}$ has the decomposition

$$O = v^L W^{[0]}\cdots W^{[N-1]}v^R \tag{19}$$

where $W^{[i]} = \left(W_{s_i s'_i}^{[i]}\right)_{\alpha_i\alpha_{i+1}}$ is the tensor of order 4 with $1 \le s_i, s'_i \le d$, $1 \le \alpha_i \le \chi_i$ and $1 \le \alpha_{i+1} \le \chi_{i+1}$, v^L and v^R are the $1\times\chi_1$ and $\chi_N \times 1$ matrices, respectively. For example, the Hamiltonian

$$H = -\sum_{i=0}^{N-2} JX_iX_{i+1} - \sum_{i=0}^{N-1} gZ_i \tag{20}$$

can be written in the form (19) with

$$v^L = (1,0,0),\ v^R = (0,0,1)^{\mathrm{T}},\ W^{[i]} = \begin{pmatrix} 1 & X & -gZ \\ 0 & 0 & -JX \\ 0 & 0 & 1 \end{pmatrix},\ i = 0,\cdots,N-1$$

Here X, Y and Z are the Pauli x, y and z matrices, respectively.

From the decompositions in (7) and (19), one has

$$\begin{aligned} O_{s'_0\cdots s'_{N-1}}^{s_0\cdots s_{N-1}}\psi_{s'_0,\cdots,s'_{N-1}} = v_{\alpha_0}^L\left(W_{s_0s'_0}^{[0]}\right)_{\alpha_0\alpha_1}\cdots\left(W_{s_{N-1}s'_{N-1}}^{[N-1]}\right)_{\alpha_{N-1}\alpha_N} v_{\alpha_N}^R \\ \left(M^{[0]s'_0}\right)_{\beta_0\beta_1}\cdots\left(M^{[N-1]s'_{N-1}}\right)_{\beta_{N-1}\beta_N} \end{aligned} \tag{21}$$

The inner product $O|\psi\rangle$ with $|\psi\rangle$ in (6) is a scalar

$$\begin{aligned} \langle\psi|\hat{O}|\psi\rangle &= O_{s'_0\cdots s'_{N-1}}^{s_0\cdots s_{N-1}}\psi_{s'_0,\cdots,s'_{N-1}}\overline{\psi_{s_0,\cdots,s_{N-1}}} \\ &= v_{\alpha_0}^L\left(W_{s_0s'_0}^{[0]}\right)_{\alpha_0\alpha_1}\cdots\left(W_{s_{N-1}s'_{N-1}}^{[N-1]}\right)_{\alpha_{N-1}\alpha_N} v_{\alpha_N}^R \\ &\quad \left(M^{[0]s'_0}\right)_{\beta_0\beta_1}\cdots\left(M^{[N-1]s'_{N-1}}\right)_{\beta_{N-1}\beta_N} \\ &\quad \overline{\left(M^{[0]s_0}\right)}_{\gamma_0\gamma_1}\cdots\overline{\left(M^{[N-1]s_{N-1}}\right)}_{\gamma_{N-1}\gamma_N} \end{aligned} \tag{22}$$

The summation over one virtual index α_0 and two physical indices s_0 and s'_0 of $v_{\alpha_0}^L\left(W_{s_0s'_0}^{[0]}\right)_{\alpha_0\alpha_1}$, $\left(M^{[0]s'_0}\right)_{\beta_0\beta_1}$ and $\overline{\left(M^{[0]s_0}\right)}_{\gamma_0\gamma_1}$ is a tensor of order 3 with virtual indices α_1, β_1 and γ_1. Note that $\beta_0=\gamma_0=1$.

If the operator $\hat{O}$ is local and ψ_s is a canonical form, the expression in (22) can be simplified. For example, ψ_s has the mixed canonical form (16), $\hat{O}$ is the operator acting on local site 2 and site 3. For example, $\left(W^{[i]}_{s_i s_i'}\right)_{\alpha_i,\alpha_{i+1}}=\delta_{s_i,s_i'}\delta_{\alpha_i,\alpha_{i+1}}$ for $i\neq 2,3$,

$$
\begin{aligned}
\langle\psi|\hat{O}|\psi\rangle = & v^L_{\alpha_0}\left(W^{[0]}_{s_0s_0'}\right)_{\alpha_0\alpha_1}\left(W^{[1]}_{s_1s_1'}\right)_{\alpha_1\alpha_2}\left(W^{[2]}_{s_2s_2'}\right)_{\alpha_2\alpha_3}\left(W^{[3]}_{s_3s_3'}\right)_{\alpha_3\alpha_4}\cdots\left(W^{[N-1]}_{s_{N-1}s_{N-1}'}\right)_{\alpha_{N-1}\alpha_N}v^R_{\alpha_N}\\
&\left(A^{[0]s_0'}\right)_{\beta_0\beta_1}\left(A^{[1]s_1'}\right)_{\beta_1\beta_2}\left(\Lambda^{[2]}B^{[2]s_2'}\right)_{\beta_2\beta_3}\left(B^{[3]s_3'}\right)_{\beta_3\beta_4}\cdots\left(B^{[N-1]s_{N-1}'}\right)_{\beta_{N-1}\beta_N}\\
&\left(\overline{A^{[0]s_0}}\right)_{\gamma_0\gamma_1}\left(\overline{A^{[1]s_1}}\right)_{\gamma_1\gamma_2}\left(\Lambda^{[2]}\overline{B^{[2]s_2}}\right)_{\gamma_2\gamma_3}\left(\overline{B^{[3]s_3}}\right)_{\gamma_3\gamma_4}\cdots\left(\overline{B^{[N-1]s_{N-1}}}\right)_{\gamma_{N-1}\gamma_N}\\
= & v^L_{\alpha_0}\delta_{\alpha_0,\alpha_1}\delta_{\alpha_1,\alpha_2}\left(W^{[2]}_{s_2s_2'}\right)_{\alpha_2\alpha_3}\left(W^{[3]}_{s_3s_3'}\right)_{\alpha_3\alpha_4}\cdots\delta_{\alpha_{N-1},\alpha_N}v^R_{\alpha_N}\\
&\delta_{\beta_2,\gamma_2}\left(\Lambda^{[2]}B^{[2]s_2'}\right)_{\beta_2\beta_3}\left(B^{[3]s_3'}\right)_{\beta_3\beta_4}\delta_{\beta_4,\gamma_4}\\
&\left(\Lambda^{[2]}\overline{B^{[2]s_2}}\right)_{\gamma_2\gamma_3}\left(\overline{B^{[3]s_3}}\right)_{\gamma_3\gamma_4}\\
= & v^L_{\alpha_2}\left(W^{[2]}_{s_2s_2'}\right)_{\alpha_2\alpha_3}\left(W^{[3]}_{s_3s_3'}\right)_{\alpha_3\alpha_4}v^R_{\alpha_4}\\
&\left(\Lambda^{[2]}B^{[2]s_2'}\right)_{\beta_2\beta_3}\left(B^{[3]s_3'}\right)_{\beta_3\beta_4}\\
&\left(\Lambda^{[2]}\overline{B^{[2]s_2}}\right)_{\beta_2\gamma_3}\left(\overline{B^{[3]s_3}}\right)_{\gamma_3\beta_4}
\end{aligned}
\tag{23}
$$

The ground state of Hamiltonian $\hat{H}$ can be calculated by DMRG algorithm [11], which is a variational method by minimizing $\langle\psi|\hat{H}|\psi\rangle$ over the tensor ψ_s of order N. Here we always require that $\sum_s|\psi_s|^2=1$. Since ψ_s has the structure of matrix product state, this global minimization can be realized by local minimization step by step. For example, $\hat{H}$ and ψ_s have a form (19), and (16), respectively, then $\langle\psi|\hat{H}|\psi\rangle$ has a form (23) can be written as

$$\langle\psi|\hat{H}|\psi\rangle=\langle\Theta|H_{\text{eff}}|\Theta\rangle \tag{24}$$

where $\Theta=\Lambda^{[2]}B^{[2]}B^{[3]}$ is a tensor of order 4. H_{eff}, which is the tensor of order 8, is obtained by the contraction of the left environment $L^{[2]}$, right environment $R^{[3]}$, $W^{[2]}$ and $W^{[3]}$. Here the left environment $L^{[i]}$ is obtained from the contraction of $L^{[i-1]}$, $W^{[i-1]}$, $A^{[i-1]}$ and $\overline{A^{[i-1]}}$, $L^{[0]}_{\beta_0\gamma_0\alpha_0}=\delta_{\beta_0\gamma_0}v^L_{\alpha_0}$. Similarly, $R^{[i]}$ is obtained from the contraction of $R^{[i+1]}$, $W^{[i+1]}$ and $B^{[i+1]}$, $\overline{B^{[i+1]}}$, $R^{[N]}_{\beta_N\gamma_N\alpha_N}=\delta_{\beta_N\gamma_N}v^R_{\alpha_N}$.

To keep the mixed canonical form of tensor of order N, THE SVD for the minimum $\Theta_{\min}$ is

$$\Theta_{\min}=\tilde{A}^{[2]}\tilde{\Lambda}^{[3]}\tilde{B}^{[3]}\Leftrightarrow\Theta_{(\beta_2s_2'),(\beta_4s_3')}=\sum_{\beta_3=1}^{\min(d\chi_2,d\chi_4)}\tilde{A}^{[2]}_{(\beta_2s_2'),\beta_3}\tilde{\Lambda}^{[3]}_{\beta_3\beta_3}\tilde{B}^{[3]}_{\beta_3,(\beta_4s_3')} \tag{25}$$

Here the tensor $\Theta^{s_2's_3'}_{\beta_2\beta_4}$ of order 4 is rewritten as a matrix $\Theta_{(\beta_2s_2'),(\beta_4s_3')}$. The bond dimension becomes d times χ_2 or χ_4 which is almost d times of the bond dimension for the original $\sum_{\beta_3=1}^{\chi_3}\left(A^{[2]s_2'}\right)_{\beta_2,\beta_3}\Lambda^{[3]}_{\beta_3\beta_3}\left(B^{[3]s_3'}\right)_{\beta_3,\beta_4}$. Here we assume that $\chi_2 \sim \chi_3 \sim \chi_4$. To keep the bond dimension under control, the SVD in (25) is truncated for $k = 1, \cdots, \chi_{\max}$. In practice, the largest $\chi_{\max}$ singular values in magnitude $\left\{\Lambda^{[3]}_{\beta_3\beta_3}\right\}_{\beta_3=1}^{\chi_{\max}}$ are kept and then scale these singular values such that $\sum_{k=1}^{\chi_{\max}} \left|\Lambda^{[3]}_{\beta_3\beta_3}\right|^2 = \sum_{k=1}^{\min(d\chi_2, d\chi_4)} \left|\tilde{\Lambda}^{[3]}_{\beta_3\beta_3}\right|^2$ to ensure $\left|\psi_s\right|^2 = 1$. For the ground state with not highly entangled, the singular values in magnitude decays vary fast. This truncation will not cause much error. Replacing $\Theta = \Lambda^{[2]} B^{[2]} B^{[3]}$ by $\Theta_{\min} = \tilde{A}^{[2]} \tilde{\Lambda}^{[3]} \tilde{B}^{[3]}$, we get another better tensor with the less value of $\langle \psi | \hat{O} | \psi \rangle$. This better tensor is also in the mixed canonical form. This is called one step of two-site DMRG algorithm since the local tensors $\Lambda^{[2]} B^{[2]}$ at site 2 and $B^{[3]}$ at site 3 are updated. To update the local tensors $\Lambda^{[3]} B^{[3]}$ at site 3 and $B^{[4]}$ at site 4 in the next step of DMRG, we should prepare the left environment $L^{[3]}$, right environment $R^{[4]}$. $L^{[3]}$ is the contraction of $L^{[2]}$, $W^{[2]}$, $\tilde{A}^{[2]}$ and $\overline{\tilde{A}^{[2]}}$. So we use $N-1$ steps of DMRG to update all N sub-tensors from the left to the right-hand side. This is called one sweep of DMRG. The next sweep is implemented from the right to the left-hand side. This algorithm will stop if $\langle \psi | \hat{H} | \psi \rangle$ will not decrease.

3. Boltzmann Machine Learning

The quantum expectation value of an operator $\hat{O}$ on a non-normalized state $|\psi\rangle$ can be written as a classical expectation value over the distribution $p(s) = |\psi(s)|^2 / \langle \psi | \psi \rangle$ using

$$\langle \hat{O} \rangle \equiv \frac{\langle \psi | \hat{O} | \psi \rangle}{\langle \psi | \psi \rangle} = \mathbb{E}\left[\tilde{O}\right] \equiv \sum_s p(s) \tilde{O}(s) \tag{26}$$

where

$$\tilde{O}(s) = \frac{\langle s | \hat{O} | \psi \rangle}{\langle s | \psi \rangle} = \sum_{s'} \frac{\psi(s')}{\psi(s)} \langle s | \hat{O} | s' \rangle \tag{27}$$

If the matrix presentation O of the operator $\hat{O}$ under the basis $|s'\rangle$ is sparse, there are only $O(n)$ basis $|s\rangle$ such that $\langle s | \hat{O} | s' \rangle$ are nonzero. Equation (26) shows that $\langle \hat{O} \rangle$ can be calculated by Monte Carlo sampling of $\tilde{O}(s)$ with the probability $p(s)$.

We use neutral network to describe the wave function $\psi(s)$, which depends on real parameters $\theta = \{\theta_k\}_{k=1}^{p}$, and thus denoted by $\psi_\theta(s)$. Denote by $\delta\theta_k$ the change of the kth component θ_k. $\psi_{\theta+\delta\theta_k}(s)$ has the expansion

$$\psi_{\theta+\delta\theta_k}(s) = \psi_\theta(s) + \delta\theta_k \frac{\partial \psi_\theta(s)}{\partial \theta_k} + O\left(\delta\theta_k^2\right) \tag{28}$$

Define the local operator $\hat{\Psi}_k$ with matrix elements

$$\langle s | \hat{\Psi}_k | s' \rangle = \delta_{s,s'} \Psi_k(s) \tag{29}$$

where we introduced the log-derivative of $\psi_\theta(s)$

$$\Psi_k(s) = \frac{\partial \ln \psi_\theta(s)}{\partial \theta_k} = \frac{1}{\psi_\theta(s)} \frac{\partial \psi_\theta(s)}{\partial \theta_k} \tag{30}$$

Here $\Psi_k(s)$ depends on θ. We thus have

$$\langle s | \hat{\Psi}_k | \psi_\theta \rangle = \Psi_k(s) \psi_\theta(s) = \frac{\partial \psi_\theta(s)}{\partial \theta_k} \tag{31}$$

From (28), one has

$$\left| \psi_{\theta+\theta_k} \right\rangle = \left(1 + \delta\theta_k \hat{\Psi}_k\right) \left| \psi_\theta \right\rangle \tag{32}$$

where we omitted $O\left(\delta\theta_k^2\right)$.

$|\psi_\theta\rangle$ is normalized to be

$$\left| v_{0,\theta} \right\rangle = \frac{\left| \psi_\theta \right\rangle}{\sqrt{\left\langle \psi_\theta | \psi_\theta \right\rangle}} \tag{33}$$

Define

$$\left| v_{k,\theta} \right\rangle = \left(\hat{\Psi}_k - \left\langle \hat{\Psi}_k \right\rangle \right) \left| v_{0,\theta} \right\rangle, \ k = 1, \cdots, p \tag{34}$$

where

$$\left\langle \hat{\Psi}_k \right\rangle = \left\langle v_{0,\theta} \middle| \hat{\Psi}_k \middle| v_{0,\theta} \right\rangle = \frac{\left\langle \psi_\theta \middle| \hat{\Psi}_k \middle| \psi_\theta \right\rangle}{\left\langle \psi_\theta | \psi_\theta \right\rangle} \tag{35}$$

Thus $\left| v_{k,\theta} \right\rangle$ and $\left| v_{0,\theta} \right\rangle$ are orthogonal, $k = 1, \cdots, p$. But $\left\langle v_{k,\theta} \middle| v_{k',\theta} \right\rangle \neq 0$ for $k, k' = 1, \cdots, p$.

The norm squared of $\left| \psi_{\theta+\delta\theta_k} \right\rangle$ is

$$\begin{aligned} \left\langle \psi_{\theta+\delta\theta_k} \middle| \psi_{\theta+\delta\theta_k} \right\rangle &= \left\langle \psi_\theta \right| \left(1 + \delta\theta_k \hat{\Psi}_k\right)^* \left(1 + \delta\theta_k \hat{\Psi}_k\right) \left| \psi_\theta \right\rangle \\ &= \left\langle \psi_\theta | \psi_\theta \right\rangle \left[1 + 2\,\mathrm{Re}\left(\delta\theta_k \left\langle \hat{\Psi}_k \right\rangle \right) + O\left(\delta\theta_k^2\right) \right] \end{aligned} \tag{36}$$

where Re denotes the real part. One has from (32) (36)

$$\begin{aligned} \left| v_{0,\theta+\delta\theta_k} \right\rangle &= \frac{\left| \psi_{\theta+\delta\theta_k} \right\rangle}{\sqrt{\left\langle \psi_{\theta+\delta\theta_k} \middle| \psi_{\theta+\delta\theta_k} \right\rangle}} \\ &= \left| v_{0,\theta} \right\rangle + \left[\delta\theta_k \hat{\Psi}_k - \mathrm{Re}\left(\delta\theta_k \left\langle \hat{\Psi}_k \right\rangle \right) \right] \left| v_{0,\theta} \right\rangle + O\left(\delta\theta_k^2\right) \\ &= \left[1 + i\,\mathrm{Im}\left(\delta\theta_k \left\langle \hat{\Psi}_k \right\rangle \right) \right] \left| v_{0,\theta} \right\rangle + \delta\theta_k \left| v_{k,\theta} \right\rangle + O\left(\delta\theta_k^2\right) \quad (\text{by } (34)) \\ &= \mathrm{e}^{i\delta\phi} \left[\left| v_{0,\theta} \right\rangle + \delta\theta_k \left| v_{k,\theta} \right\rangle \right] + O\left(\delta\theta_k^2\right) \end{aligned} \tag{37}$$

where $\delta\phi = \mathrm{Im}\left(\delta\theta_k \left\langle \hat{\Psi}_k \right\rangle \right)$ is the imaginary part of $\delta\theta_k \left\langle \hat{\Psi}_k \right\rangle$.

Let $\hat{H}$ be the Hamiltonian operator and using (37), one has

$$\begin{aligned} \frac{\partial}{\partial \theta_k} \frac{\left\langle \psi_\theta \middle| \hat{H} \middle| \psi_\theta \right\rangle}{\left\langle \psi_\theta | \psi_\theta \right\rangle} &= \frac{\partial}{\partial \theta_k} \left\langle v_{0,\theta} \middle| \hat{H} \middle| v_{0,\theta} \right\rangle \\ &= \lim_{\delta\theta_k \to 0} \frac{\left\langle v_{0,\theta+\delta\theta_k} \middle| \hat{H} \middle| v_{0,\theta+\delta\theta_k} \right\rangle - \left\langle v_{0,\theta} \middle| \hat{H} \middle| v_{0,\theta} \right\rangle}{\delta\theta_k} \end{aligned}$$

$$
\begin{aligned}
&= \langle v_{k,\theta} | \hat{H} | v_{0,\theta} \rangle + \langle v_{0,\theta} | \hat{H} | v_{k,\theta} \rangle \\
&= 2\,\mathrm{Re}\left[\frac{\langle \psi_\theta | \hat{H} \left(\hat{\Psi}_k - \langle \hat{\Psi}_k \rangle \right) | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \right] \\
&= 2\,\mathrm{Re}\left[\frac{\langle \psi_\theta | \hat{H} \hat{\Psi}_k | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} - \frac{\langle \psi_\theta | \hat{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \frac{\langle \psi_\theta | \hat{\Psi}_k | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \right]
\end{aligned} \tag{38}
$$

The k th parameter θ_k is updated according to the gradient descent method

$$
\theta_k \leftarrow \theta_k - \eta f_k \tag{39}
$$

where $f_k = \frac{\partial E_\theta}{\partial \theta_k}$, $E_\theta = \frac{\langle \psi_\theta | \hat{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle}$, $\eta > 0$ is the learning rate.

If each component θ_k has a change $\delta\theta_k$, the wave function $\psi_\theta(s)$ has changed to be $\psi_{\theta+\delta\theta}(s)$. Let $|v_{0,\theta+\delta\theta}\rangle = |\psi_{\theta+\delta\theta}\rangle \big/ \sqrt{\langle \psi_{\theta+\delta\theta_k} | \psi_{\theta+\delta\theta_k} \rangle}$. Equation (37) is generalized to be

$$
|v_{0,\theta+\delta\theta}\rangle = \mathrm{e}^{i\delta\phi} \left[|v_{0,\theta}\rangle + \sum_{k=1}^{p} \delta\theta_k |v_{k,\theta}\rangle \right] + O\left(\delta\theta^2\right) \tag{40}
$$

where $\delta\phi = \sum_{k=1}^{p} \mathrm{Im}\left(\delta\theta_k \langle \hat{\Psi}_k \rangle \right)$.

The distance between $|v_{0,\theta}\rangle$ and $|v_{0,\theta+\delta\theta}\rangle$ is

$$
\delta s^2 = \min_{\delta\alpha} \left\| \mathrm{e}^{-i\delta\alpha} |v_{0,\theta+\delta\theta}\rangle - |v_{0,\theta}\rangle \right\|^2 \tag{41}
$$

When $\delta\alpha = \delta\phi$, δs^2 is minimized to be

$$
\delta s^2 = \sum_{1 \le k,k' \le p} \langle v_{k,\theta} | v_{k',\theta} \rangle \delta\theta_k \delta\theta_{k'} + o\left(\delta\theta^2\right) = \sum_{1 \le k,k' \le p} S_{k,k'} \delta\theta_k \delta\theta_{k'} + o\left(\delta\theta^2\right) \tag{42}
$$

with

$$
\begin{aligned}
S_{k,k'} &= \frac{\langle v_{k,\theta} | v_{k',\theta} \rangle + \langle v_{k',\theta} | v_{k,\theta} \rangle}{2} = \mathrm{Re}\left(\langle v_{k,\theta} | v_{k',\theta} \rangle \right) \\
&= \mathrm{Re}\left(\langle \hat{\Psi}_k^* \hat{\Psi}_{k'} \rangle - \langle \hat{\Psi}_k \rangle^* \langle \Psi_{k'} \rangle \right)
\end{aligned} \tag{43}
$$

where we used

$$
\langle v_{k,\theta} | v_{k',\theta} \rangle = \langle v_{0,\theta} | \left(\hat{\Psi}_k - \langle \hat{\Psi}_k \rangle \right)^* \left(\hat{\Psi}_{k'} - \langle \hat{\Psi}_{k'} \rangle \right) | v_{0,\theta} \rangle \tag{44}
$$

Find $\delta\theta$ such that: 1) $\Delta E = E_{\theta+\delta\theta} - E_\theta$ is minimized; 2) the distance δs^2 between $|v_{0,\theta}\rangle$ and $|v_{0,\theta+\delta\theta}\rangle$ is as small as possible. Minimize

$$
\Delta E + \mu \delta s^2 = \sum_k \frac{\partial E_\theta}{\partial \theta_k} \delta\theta_k + \mu \sum_{k,k'} S_{k,k'} \delta\theta_k \delta\theta_{k'} + o\left(\delta\theta^2\right) \tag{45}
$$

to find that $\delta\theta$ satisfies

$$
\sum_{k'} S_{k,k'} \delta\theta_{k'} = -\frac{f_k}{2\mu} \tag{46}
$$

which can be used to update the parameter θ. This algorithm is called the stochastic reconfiguration method.

In the Boltzmann machine ansatz, the wave function is chosen to be

$$\psi_\theta(s)=\exp\left(\sum_{j=0}^{N-1}a_j s_j\right)\prod_{i=0}^{M-1}2\cosh\left(b_i+\sum_{j=0}^{N-1}W_{ij}s_j\right) \tag{47}$$

where $\theta=\{a_j,b_i,W_{ij}\}$ are the real parameters. Here M is a given positive integer number. $\alpha=M/N$ is the ratio of M and N. Compared to the other neutral network ansatz, Ψ_k for the Boltzmann machine ansatz can be calculated analytically [39].

4. Variational Quantum Eigensolver

The variational quantum eigensolver is aimed at finding the ground state energy of a Hamiltonian $\hat{H}$ by minimizing the cost function $L(\theta)=\langle\psi_\theta|\hat{H}|\psi_\theta\rangle$, where he parameter θ is updated by $\theta-\eta\nabla L(\theta)$ by iterations. Here $\eta>0$ is the learning rate. We choose the trial state $|\psi_\theta\rangle=U_\theta|\psi_0\rangle$ for some ansatz U_θ and initial state $|\psi_0\rangle$. The variational Hamiltonian ansatz also aims to prepare a trial ground state for a given Hamiltonian $\hat{H}=\sum_k\hat{H}_k$ (where $\hat{H}_k$ are Hermitian operators, usual Pauli strings) by Trotterizing an adiabatic state preparation process [116]. Here, each Trotter step corresponds to a variational ansatz so that the unitary is given $U_\theta=\prod_l\left(\prod_k e^{-i\theta_{l,k}\hat{H}_k}\right)$.

Consider a general Hamiltonian

$$\begin{aligned}\hat{H}=&\sum_{\sigma=\uparrow,\downarrow}\sum_{j=0}^{N-2}\left(C_X^{\mathrm{h}}X_{j,\sigma}X_{j+1,\sigma}+C_Y^{\mathrm{h}}Y_{j,\sigma}Y_{j+1,\sigma}+C_Z^{\mathrm{h}}Z_{j,\sigma}Z_{j+1,\sigma}\right)\\&+\sum_{j=0}^{N-1}\left(C_X^{\mathrm{v}}X_{j,\uparrow}X_{j,\downarrow}+C_Y^{\mathrm{v}}Y_{j,\uparrow}Y_{j,\downarrow}+C_Z^{\mathrm{v}}Z_{j,\uparrow}Z_{j,\downarrow}\right)\end{aligned} \tag{48}$$

where N is an even integer and the coefficients $C_\alpha^{\mathrm{h}}>0$ and C_α^{v} do not depend on sites. $X_{j,\sigma}$, etc., denotes the Pauli X matrix acting on site j and spin $\sigma=\uparrow,\downarrow$. Splitting the N sites to be even sites and odd sites, this Hamiltonian can be written as

$$\hat{H}=\sum_{\alpha=X,Y,Z}\left(H_\alpha^{\mathrm{h,e}}+H_\alpha^{\mathrm{h,o}}+H_\alpha^{\mathrm{v,e}}+H_\alpha^{\mathrm{v,o}}\right)=H^{\mathrm{e}}+H^{\mathrm{o}} \tag{49}$$

where

$$H^{\mathrm{e}}=\sum_{\alpha=X,Y,Z}\left(H_\alpha^{\mathrm{h,e}}+H_\alpha^{\mathrm{v,e}}\right),\ H^{\mathrm{o}}=\sum_{\alpha=X,Y,Z}\left(H_\alpha^{\mathrm{h,o}}+H_\alpha^{\mathrm{v,o}}\right) \tag{50}$$

$$H_\alpha^{\mathrm{h,e}}=\sum_{\sigma=\uparrow,\downarrow}\sum_{\text{even }j=0}^{N-2}C_\alpha^{\mathrm{h}}\alpha_{j,\sigma}\alpha_{j+1,\sigma},\ H_\alpha^{\mathrm{v,e}}=\sum_{\text{even }j=0}^{N-1}C_\alpha^{\mathrm{v}}\alpha_{j,\sigma}\alpha_{j,\uparrow}\alpha_{j,\downarrow},\ \alpha=X,Y,Z \tag{51}$$

$H_\alpha^{\mathrm{h,o}}$ and $H_\alpha^{\mathrm{v,o}}$ are defined similarly by replacing summation over even site j with odd site j.

Inspired by Ref. [117] [118], the quantum circuit is $U_\theta=\prod_{k=1}^{p}U^{\mathrm{h}}\left(\theta_k^{\mathrm{h}}\right)U^{\mathrm{v}}\left(\theta_k^{\mathrm{v}}\right)$, where

$$\begin{aligned}U^{\mathrm{h}}\left(\theta_k^{\mathrm{h}}\right)=&G\left(\theta_{k,X}^{\mathrm{h,e}},H_X^{\mathrm{h,e}}\right)G\left(\theta_{k,Y}^{\mathrm{h,e}},H_Y^{\mathrm{h,e}}\right)G\left(\theta_{k,Z}^{\mathrm{h,e}},H_Z^{\mathrm{h,e}}\right)\\&G\left(\theta_{k,Z}^{\mathrm{h,o}},H_Z^{\mathrm{h,o}}\right)G\left(\theta_{k,Y}^{\mathrm{h,o}},H_Y^{\mathrm{h,o}}\right)G\left(\theta_{k,X}^{\mathrm{h,o}},H_X^{\mathrm{h,o}}\right)\end{aligned} \tag{52}$$

where $G(x,H) = \mathrm{e}^{-ixH}$. $U^{\mathrm{v}}(\theta_k^{\mathrm{v}})$ is defined similarly, with the subscription h replaced by v. Since the terms in $H_\alpha^{\mathrm{h,o}}$, etc., commute with each other,

$$G\left(\theta_{k,X}^{\mathrm{h,e}}, H_X^{\mathrm{h,e}}\right) = \prod_{\sigma=\uparrow,\downarrow} \prod_{\text{even } j=0}^{N-2} \mathrm{e}^{-iC_\alpha^{\mathrm{h}} \alpha_{j,\sigma}\alpha_{j+1,\sigma}}$$

The initial state $|\psi_0\rangle$ is chosen to be the ground state $|\psi_0\rangle = \otimes_{\sigma=\uparrow,\downarrow} \otimes_{\text{even } j=0}^{N-2} \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right)$ of $\sum_{\alpha=X,Y,Z} H_\alpha^{\mathrm{h,e}}$ with the ground state energy $-3NC_\alpha^{\mathrm{h}}$. This is because

$$\alpha_{j,\sigma}\alpha_{j+1,\sigma}\left(|01\rangle - |10\rangle\right) = -\left(|01\rangle - |10\rangle\right) \tag{53}$$

for $\alpha = X,Y,Z$, $\sigma = \uparrow,\downarrow$ and even site $j = 0,2,\cdots,N-2$.

5. Fermi-Hubbard Model

The Hamiltonian for Fermi-Hubbard model is

$$\begin{aligned}\hat{H} &= -t\sum_{\sigma=\uparrow,\downarrow}\sum_{j=0}^{N-2}\left(c_{j,\sigma}^\dagger c_{j+1,\sigma} + \text{h.c.}\right) + U\sum_{j=0}^{N-1} n_{j,\uparrow}n_{j,\downarrow} - \mu\sum_{\sigma=\uparrow,\downarrow}\sum_{j=0}^{N-1} n_{j,\sigma} \\ &= -t\sum_{\sigma=\uparrow,\downarrow}\sum_{j=0}^{N-2}\left(c_{j,\sigma}^\dagger c_{j+1,\sigma} + \text{h.c.}\right) + U\sum_{j=0}^{N-1}\left(n_{j,\uparrow} - \frac{1}{2}\right)\left(n_{j,\downarrow} - \frac{1}{2}\right) \\ &\quad -\left(\mu - \frac{U}{2}\right)\sum_{\sigma=\uparrow,\downarrow}\sum_{j=0}^{N-1} n_{j,\sigma} - \frac{NU}{4}\end{aligned} \tag{54}$$

where open boundary condition is assumed and $n_{j,\sigma} = c_{j,\sigma}^\dagger c_{j,\sigma}$. The creation operator $c_{j,\sigma}^\dagger$ and the annihilate operators $c_{j,\sigma}$ satisfy

$$\left\{c_{j,\sigma}, c_{k,\tau}^\dagger\right\} = \delta_{jk}\delta_{\sigma\tau},\ \left\{c_{j,\sigma}, c_{k,\tau}\right\} = 0,\ \left\{c_{j,\sigma}^\dagger, c_{k,\tau}^\dagger\right\} = 0 \tag{55}$$

For the half filling case: $\mu = U/2$, the Hamiltonian is reduced to

$$\hat{H} = -t\sum_{\sigma=\uparrow,\downarrow}\sum_{j=0}^{N-2}\left(c_{j,\sigma}^\dagger c_{j+1,\sigma} + \text{h.c.}\right) + U\sum_{j=0}^{N-1}\left(n_{j,\uparrow} - \frac{1}{2}\right)\left(n_{j,\downarrow} - \frac{1}{2}\right) - \frac{NU}{4} \tag{56}$$

Combining the index $\sigma = \uparrow(0),\downarrow(1)$ and $j = 0,\cdots,N-1$ to be $k = N\sigma + j$, we have

$$\hat{H} = -\sum_{k=0}^{2N-2} t_k\left(c_k^\dagger c_{k+1} + \text{h.c.}\right) + U\sum_{k=0}^{N-1}\left(n_k - \frac{1}{2}\right)\left(n_{k+N} - \frac{1}{2}\right) - \frac{NU}{4} \tag{57}$$

where $t_k = t$ if $k \neq N-1$ and 0 if $k = N-1$. The Jordan-Wigner representation shows that

$$n_j = \frac{1}{2} + \frac{1}{2}Z_j \Leftrightarrow Z_j = 2n_j - 1$$

$$2\left(c_{k+1}^\dagger c_k + c_k^\dagger c_{k+1}\right) = X_k X_{k+1} + Y_k Y_{k+1}$$

Here X,Y,Z are Pauli operators. In the Jordan-Wigner representation, the Hamiltonian for half-filling case, is written as

$$\frac{1}{tN}\hat{H} = -\frac{1}{2N}\sum_{k=0}^{2N-2} t_k\left(X_k X_{k+1} + Y_k Y_{k+1}\right) + \frac{U}{4tN}\sum_{k=0}^{N-1} Z_k Z_{k+N} - \frac{U}{4t} \tag{58}$$

which is the special case for the Hamiltonian in (48) except the constant $-\frac{U}{4t}$.

If one uses the index $k=2j+\sigma$ for $j=0,\cdots,N-1$ and $\sigma=\uparrow(0),\downarrow(1)$, the Hamiltonian in (56) becomes

$$\hat{H}=-t\sum_{k=0}^{2N-3}\left(c_k^{\dagger}c_{k+2}+\text{h.c.}\right)+U\sum_{k=0}^{N-1}\left(n_{2k}-\frac{1}{2}\right)\left(n_{2k+1}-\frac{1}{2}\right)-\frac{NU}{4} \tag{59}$$

Under the Jordan-Wigner representation

$$2\left(c_{k+2}^{\dagger}c_k+c_k^{\dagger}c_{k+2}\right)=X_k\left(-Z_{k+1}\right)X_{k+2}+Y_k\left(-Z_{k+1}\right)Y_{k+2}$$

it becomes

$$\frac{1}{tN}\hat{H}=-\frac{1}{2N}\sum_{k=0}^{2N-3}\left(X_k\left(-Z_{k+1}\right)X_{k+2}+Y_k\left(-Z_{k+1}\right)Y_{k+2}\right)+\frac{U}{4t}\sum_{k=0}^{N-1}Z_{2k}Z_{2k+1}-\frac{U}{4t} \tag{60}$$

6. Schwinger Model

The massive Schwinger model, or QED in two space-time dimensions, is a gauge theory describing the interaction of a single fermionic species with the photon field. In the temporal gauge, the Hamiltonian density is

$$\mathcal{H}=-i\psi^{\dagger}\gamma^0\gamma^1\left(\partial_x+igA\right)\psi+m\psi^{\dagger}\left(x\right)\gamma^0\psi+\frac{1}{2}E^2 \tag{61}$$

where

$$\gamma^0=\begin{pmatrix}1&0\\0&-1\end{pmatrix},\ \gamma^1=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$$

are the γ matrices, m and g are two parameters. The electric field, $E=\partial_t A$ satisfies the Gauss's law

$$\partial_x E\left(x\right)-g\psi^{\dagger}\left(x\right)\psi\left(x\right)=0$$

Here ∂_x and ∂_t denotes the partial derivative along one-dimensional space direction x and time direction t.

The theory is quantized using canonical quantization by imposing anticommutation relations on the fermion fields

$$\left\{\psi^{\dagger}\left(x,t\right),\psi\left(y,t\right)\right\}=\delta\left(x-y\right) \tag{62}$$

and by imposing commutation relations on the gauge fields

$$\left[E\left(x,t\right),A\left(y,t\right)\right]=i\delta\left(x-y\right) \tag{63}$$

The model can now be formulated on a "staggered" spatial lattice. Let the lattice spacing be a, and label the sites with an integer n. Define a single-component fermion field $\phi(n)$ at each site n, obeying anticommutation relations

$$\left\{\phi^{\dagger}\left(n\right),\phi\left(m\right)\right\}=\delta_{mn} \tag{64}$$

The gauge field is defined on the links $(n,n+1)$ connecting each pair of sites by

$$U\left(n,n+1\right)=\mathrm{e}^{i\theta(n)}=\mathrm{e}^{iagA(n)} \tag{65}$$

Then the lattice Hamiltonian equivalent to Equation (61) is

$$\hat{H} = \frac{-i}{2a}\sum_{n=0}^{N-1}\left[\phi^{\dagger}(n)\mathrm{e}^{i\theta(n)}\phi(n+1) - \text{H.c.}\right] + m\sum_{n=0}^{N-1}(-1)^n \phi^{\dagger}(n)\phi(n) + \frac{g^2 a}{2}\sum_{n=0}^{N-1} L^2(n) \quad (66)$$

where the number of lattice sites N is even, and the correspondence between lattice and continuum fields is

$$\frac{\phi(n)}{\sqrt{a}} \to \begin{cases} \psi_e(x), & n \text{ even,} \\ \psi_o(x), & n \text{ odd} \end{cases} \quad (67)$$

ψ_e and ψ_o are the two components of the continuum spinor $\psi = (\psi_e, \psi_o)$,

$$\frac{1}{ag}\theta(n) \to A(x),\ gL(n) \to E(x) \quad (68)$$

From the commutation relations (63) and the above correspondence, one has

$$\left[\theta(n), L(m)\right] = i\delta_{nm} \quad (69)$$

The dimensionless Hamiltonian is $\frac{2}{ag^2}\hat{H}$, which is also denoted by $\hat{H}$

$$\hat{H} = -ix\sum_{n=0}^{N-1}\left[\phi^{\dagger}(n)\mathrm{e}^{i\theta(n)}\phi(n+1) - \text{H.c.}\right] + \mu\sum_{n=0}^{N-1}(-1)^n \phi^{\dagger}(n)\phi(n) + \sum_{n=0}^{N-1} L^2(n) \quad (70)$$

with the dimensionless parameters

$$x = \frac{1}{g^2 a^2},\ \mu = \frac{2m}{g^2 a} \quad (71)$$

The Gauss law $\partial_x E = g\psi^{\dagger}\psi$ is discretized to be

$$L(n) - L(n-1) = \phi(n)^{\dagger}\phi(n) - \frac{1}{2}\left[1 - (-1)^n\right] \quad (72)$$

The elimination of fermion fields can be realized by using a Jordan-Wigner transformation

$$\phi(n) = \prod_{k<n}\left(-\sigma_k^z\right)\sigma_n^-,\ \phi(n)^{\dagger} = \prod_{k<n}\left(-\sigma_k^z\right)\sigma_n^+$$

where $\sigma_n^{\pm} = \left(X_n \pm iY_n\right)/2$ denotes the matrix $|0\rangle\langle 1|$ ($|1\rangle\langle 0|$) acting on the *j*th qubit.

The Gauss law in (72) is reduced to be

$$L(n) - L(n-1) = \frac{1}{2}\left[\sigma_n^z + (-1)^n\right] \quad (73)$$

and the dimensionless Hamiltonian in (70) is

$$H = -ix\sum_{n=0}^{N-1}\left[\sigma_n^+ \mathrm{e}^{i\theta(n)}\sigma_{n+1}^- - \text{H.c.}\right] + \mu\sum_{n=0}^{N-1}(-1)^n \frac{1}{2}\left(\sigma_n^z + 1\right) + \sum_{n=0}^{N-1} L^2(n) \quad (74)$$

The factor $e^{i\theta(n)}$ has been eliminated by a residual gauge transformation [119]

$$\sigma_n^- \to \prod_{l<n}\left\{\mathrm{e}^{-i\theta(l)}\right\}\sigma_n^- \quad (75)$$

Finally, the model (70) can be mapped to a spin 1/2 Hamiltonian [119] [120]

$$
\begin{aligned}
H &= x\sum_{n=0}^{N-2}\left[\sigma_n^+\sigma_{n+1}^- + \sigma_n^-\sigma_{n+1}^+\right] + \frac{\mu}{2}\sum_{n=0}^{N-1}\left[1+(-1)^n\sigma_n^z\right] + \sum_{n=0}^{N-2}\left[\ell + \frac{1}{2}\sum_{k=0}^{n}\left((-1)^k+\sigma_k^z\right)\right]^2 \\
&= \frac{x}{2}\sum_{n=0}^{N-2}\left[X_nX_{n+1}+Y_nY_{n+1}\right] + \frac{\mu}{2}\sum_{n=0}^{N-1}\left[1+(-1)^n Z_n\right] + \sum_{n=0}^{N-2}\left[\ell + \frac{1}{2}\sum_{k=0}^{n}\left((-1)^k+Z_k\right)\right]^2
\end{aligned}
\tag{76}
$$

where ℓ is the boundary electric field (on the leftmost link), which can describe the background field. In the second equality, we used

$$
\sigma_n^+\sigma_{n+1}^- + \sigma_n^-\sigma_{n+1}^+ = \frac{X_nX_{n+1}+Y_nY_{n+1}}{2},\quad \sigma_n^z = Z_n
$$

7. Simulation Results

We investigate the Fermi-Hubbard model at half filling, where the Hamiltonian is given by (56) or (58). Table 1 presents the ground state energy per site E_0/N obtained by DMRG algorithm for different values of N and U/t. It is evident that the thermodynamic limit ($N\to\infty$) is achieved in each column. We also check these results by exact diagonalization and find that the error between the DMRG and exact ground state energies is less than 1.42×10^{-9} for all $(U/t,N)$ in Table 1, demonstrating the high accuracy of the DMRG method. Figure 1 illustrates the energy $\langle\hat{H}\rangle=\langle\psi|\hat{H}|\psi\rangle$ in Equation (24) as a function of iteration for different values of U/t with $N=8$. We perform 8 sweeps, with 14 ($=2(N-1)$) local minimizations in each sweep (multiplied by 2 due to spin up and spin down states). It is noteworthy that the ground state energy is obtained after a single sweep (14 local minimizations). Furthermore, the energy $\langle\hat{H}\rangle$ decreases with each step of local minimization.

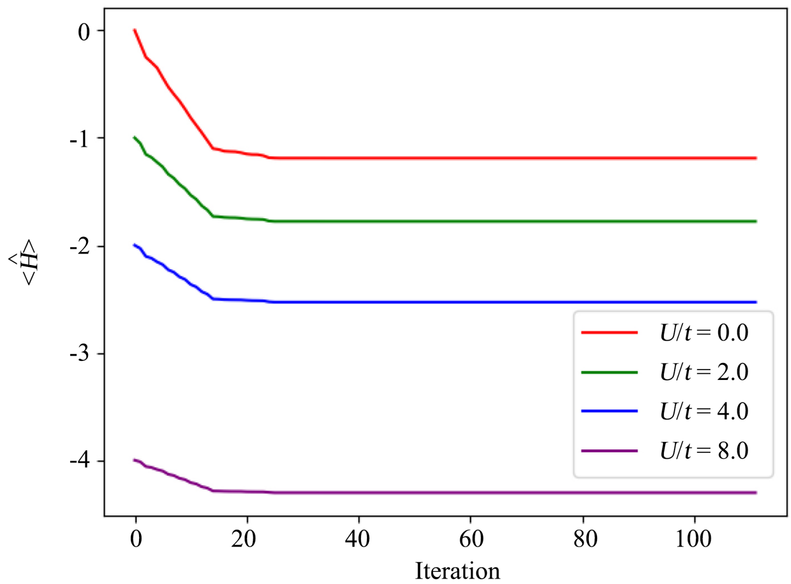

Figure 1. The dependence of $\langle\hat{H}\rangle$ of the Fermi-Hubbard model for half-filling case on the iteration for different U/t, $N=8$ by DMRG algorithm.

Table 1. The ground state energy E_0/N of the Fermi-Hubbard model for half-filling case obtained by DMRG algorithm.

U/t	0	2	4	8
$N=4$	−1.118033	−1.718985	−2.488286	−4.279293
$N=6$	−1.164653	−1.757718	−2.515427	−4.294683
$N=8$	−1.189692	−1.778204	−2.529475	−4.302603

We display the ground state energy per site, E_0/N, obtained using the restricted Boltzmann machine with $\alpha=4$ in Table 2. To calculate the ground state energy E_0/N, we use the last 10,000 sampling data out of a total of 40,000 data points. Comparing these results with those obtained using the DMRG algorithm, we find that the results from the restricted Boltzmann machine are reliable when U/t and N are not too large. However, for $U/t=8$ and $N=8$, the value $E_0/N=-4.211473$ in Table 2 is significantly higher than $E_0/N=-4.302603$ in Table 1. Figure 2 illustrates the dependence of $\langle\hat{H}\rangle$ on the iteration number for the restricted Boltzmann machine with $\alpha=4$ and $N=8$. For the case $U/t=8.0$, there are large fluctuations of $\langle\hat{H}\rangle$ with each iteration.

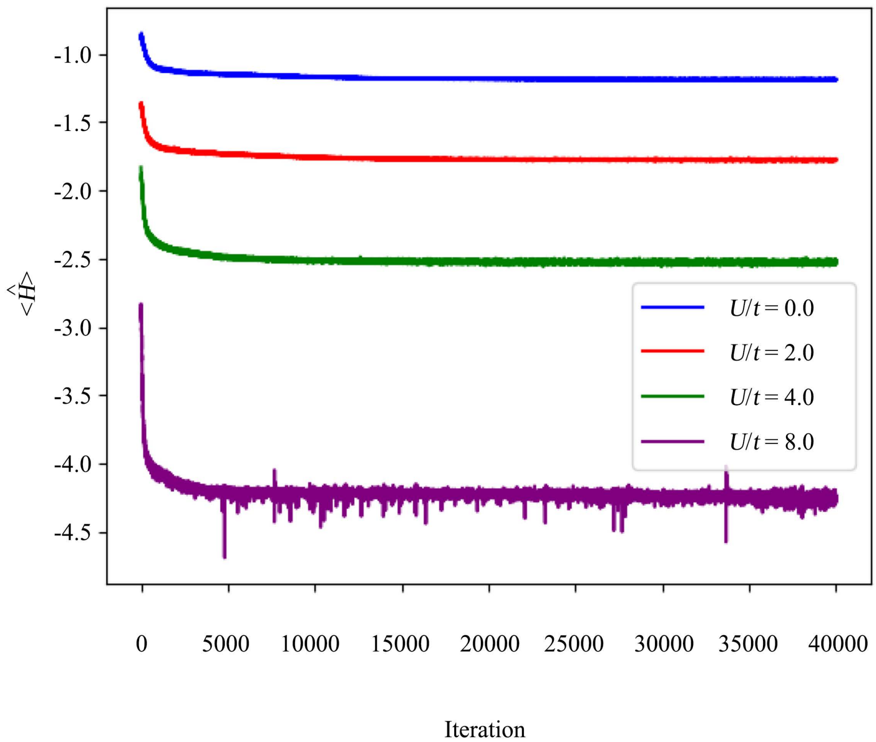

Figure 2. The dependence of $\langle\hat{H}\rangle$ of the Fermi-Hubbard model for half-filling case on the iteration for different U/t by restricted Boltzmann machine with $\alpha=4$ and $N=8$.

Table 2. The ground state energy E_0/N of the Fermi-Hubbard model for half-filling case obtained from the last 10,000 sampling data (total 40,000 data). Using restricted Boltzmann machine with $\alpha = 4$.

U/t	0	2	4	8
$N=4$	−1.112495 ± 0.00458	−1.713604 ± 0.00430	−2.482209 ± 0.00468	−4.270301 ± 0.00639
$N=6$	−1.154108 ± 0.00698	−1.747218 ± 0.00706	−2.502630 ± 0.00775	−4.272329 ± 0.01133
$N=8$	−1.172597 ± 0.00820	−1.759270 ± 0.00934	−2.507544 ± 0.01090	−4.211473 ± 0.01207

Table 3 displays the ground state energy per site, E_0/N, obtained using the variational quantum eigensolver (VQE) with $p = 6$. The results from VQE show good agreement with those obtained using the DMRG algorithm. The largest deviation in E_0/N between VQE and DMRG is 0.017611, which occurs for the case $U/t = 8$ and $N = 8$. Figure 3 illustrates the dependence of $\langle \hat{H} \rangle$ on the iteration number for VQE with different values of U/t and $N = 8$. When compared with the fluctuation in $\langle \hat{H} \rangle$ shown in Figure 2 for the restricted Boltzmann machine, the fluctuation observed in Figure 3 is very small.

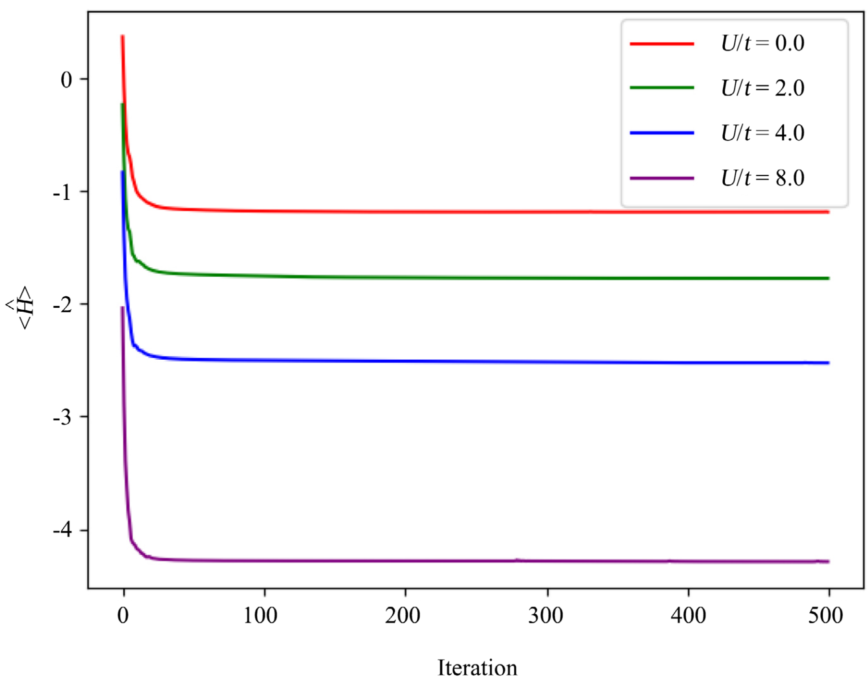

Figure 3. The dependence of $\langle \hat{H} \rangle$ of the Fermi-Hubbard model for half-filling case on the iteration by VQE with $p = 6$ for different U/t, $N = 8$.

Table 3. The ground state energy E_0/N of the Fermi-Hubbard model for half-filling case obtained by VQE after 2000 iterations, $p=6$ in quantum circuit $U(\theta)$, the learning rate $\eta=0.01$.

U/t	0	2	4	8
$N=4$	−1.11802 ± 0.000003	−1.718857 ± 0.000036	−2.488273 ± 0.000009	−4.274382 ± 0.002363
$N=6$	−1.164649 ± 0.000001	−1.757449 ± 0.000014	−2.514657 ± 0.000070	−4.284413 ± 0.000717
$N=8$	−1.189618 ± 0.000004	−1.776534 ± 0.000159	−2.520144 ± 0.001265	−4.284992 ± 0.001337

The ground state energy per site, E_0/N, of the Schwinger model (76) has been calculated using both the density matrix renormalization group (DMRG) algorithm and the variational quantum eigensolver (VQE). However, the restricted Boltzmann machine is not suitable for solving the Schwinger model as it fails to approach the ground state energy accurately. In this study, we have chosen $x=100$ and $\ell=0$. Table 4 presents the ground state energy per site obtained by the DMRG algorithm for different values of μ and N. Comparing the results to the exact values, we observe that E_0/N for $N=4$ and 8 has an error less than $O(10^{-10})$, while for $N=16$ the error is on the order of $O(10^{-4})$. For fixed values of $\mu=0$ and 0.5, E_0/N tends to the thermodynamic limit as $N\to\infty$. Table 5 displays the ground state energy per site obtained by VQE for different values of μ and $N=4,8$, and 16 due to computational limitations. It is worth noting that E_0/N obtained by VQE is slightly larger than those obtained by the DMRG algorithm.

Table 4. The ground state energy E_0/N of the Schwinger model with $x=100$ obtained by DMRG algorithm.

μ	0.0	2.5
$N=4$	−55.627908	−54.403737
$N=8$	−59.184968	−57.966911
$N=16$	−61.162808	−59.954053
$N=32$	−62.209047	−61.012423
$N=64$	−62.752515	−61.564360

Table 5. The ground state energy E_0/N of the Schwinger model with $x=100$ obtained by VQE. The last 1000 iterations are used (total 3000 iterations), $p=6$ in quantum circuit $U(\theta)$, and the learning rate $\eta=0.01$.

μ	0.0	2.5
$N=4$	−55.513396 ± 0.000435	−54.263672 ± 0.000295

Continued

$N = 8$	−58.831446 ± 0.000582	−57.581606 ± 0.000463
$N = 16$	−60.318314 ± 0.000503	−59.068237 ± 0.000498

8. Discussion

We have used three variational methods to calculate the ground state energy for Fermi-Hubbard model at half filling and Schwinger model. The DMRG algorithm has proven to be successful for these two models, while the VQE algorithm, although less accurate, is also effective for calculating ground state energies. However, it seems that the restricted Boltzmann learning method is not as accurate as the DMRG algorithm for the Fermi-Hubbard model. It's unfortunate that the restricted Boltzmann does not work for solving the ground state energy of Schwinger model. We also used mean field ansatz, Jastrow ansatz, multi-layer dense network, group convolutional neural networks [75] [121] and autoregressive networks [62], etc., but it also does not work. The authors in [122] used the transformer quantum state (TQS) to study the ground state of 1D transverse field Ising model and showed the power of TQS in expressing the quantum states for quantum many-body systems. We used TQS to calculate the ground state energy of Schwinger model with $N = 20$ for different μ in **Figure 4**. TQS is trained for many randomly chosen $\mu \in [1.5, 3.5]$ and then the TQS model is extrapolated to the other values of $\mu = [0, 1.5] \cup [3.5, 5]$. The trained TQS model can also be fine-tuned. The calculated result in **Figure 4** should be comparable with the ground state energy obtained by DMRG algorithm.

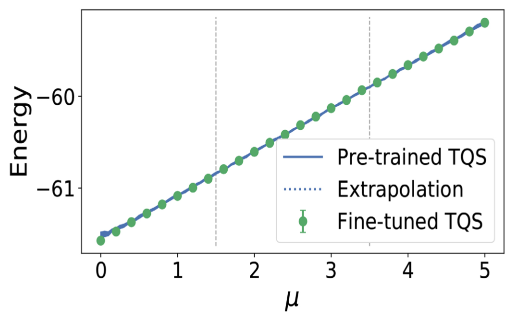

Figure 4. The ground state energy E_0/N of Schwinger model with $N = 20$ obtained by transformer quantum state (TQS).

Variational methods can indeed be applied to various other aspects of quantum systems, such as excited states, thermal states, real-time dynamics, dissipative dynamics, and quantum control. It's beneficial to learn from algorithms and ideas used in one variational method and apply them to other variational methods. This cross-pollination of ideas can lead to further advancements in the field.

Acknowledgements

Daming Li was supported by National Key Research and Development Program of China (No. 2024YFA1014104), and Shanghai Science and Technology Innovation Action Plan Project (No. 22JC1401500).

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Feynman, R.P. (1954) Atomic Theory of the Two-Fluid Model of Liquid Helium. *Physical Review*, **94**, 262-277. https://doi.org/10.1103/physrev.94.262

[2] Feynman, R.P. and Cohen, M. (1956) Energy Spectrum of the Excitations in Liquid Helium. *Physical Review*, **102**, 1189-1204. https://doi.org/10.1103/physrev.102.1189

[3] Tsui, D.C., Stormer, H.L. and Gossard, A.C. (1982) Two-Dimensional Magnetotransport in the Extreme Quantum Limit. *Physical Review Letters*, **48**, 1559-1562. https://doi.org/10.1103/physrevlett.48.1559

[4] Laughlin, R.B. (1983) Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations. *Physical Review Letters*, **50**, 1395-1398. https://doi.org/10.1103/physrevlett.50.1395

[5] Haldane, F.D.M. (1983) Continuum Dynamics of the 1-D Heisenberg Antiferromagnet: Identification with the O(3) Nonlinear Sigma Model. *Physics Letters A*, **93**, 464-468. https://doi.org/10.1016/0375-9601(83)90631-x

[6] Haldane, F.D.M. (1983) Nonlinear Field Theory of Large-Spin Heisenberg Antiferromagnets: Semiclassically Quantized Solitons of the One-Dimensional Easy-Axis Néel State. *Physical Review Letters*, **50**, 1153-1156. https://doi.org/10.1103/physrevlett.50.1153

[7] Anderson, P.W. (1973) Resonating Valence Bonds: A New Kind of Insulator? *Materials Research Bulletin*, **8**, 153-160. https://doi.org/10.1016/0025-5408(73)90167-0

[8] Bednorz, J.G. and Muller, K.A. (1986) Possible Hight C Superconductivity in the Ba-La-Cu-O System. *Zeitschrift für Physik B Condensed Matter*, **64**, 189-193. https://doi.org/10.1007/bf01303701

[9] Li, D. (2016) Fermion Bag Approach for the Massive Thirring Model at Finite Density. *Physical Review D*, **94**, Article 114501. https://doi.org/10.1103/physrevd.94.114501

[10] Hastings, M.B. (2007) An Area Law for One-Dimensional Quantum Systems. *Journal of Statistical Mechanics: Theory and Experiment*, **2007**, P08024. https://doi.org/10.1088/1742-5468/2007/08/p08024

[11] White, S.R. (1992) Density Matrix Formulation for Quantum Renormalization Groups. *Physical Review Letters*, **69**, 2863-2866. https://doi.org/10.1103/physrevlett.69.2863

[12] Schollwöck, U. (2011) The Density-Matrix Renormalization Group in the Age of Matrix Product States. *Annals of Physics*, **326**, 96-192. https://doi.org/10.1016/j.aop.2010.09.012

[13] White, S.R. and Feiguin, A.E. (2004) Real-Time Evolution Using the Density Matrix Renormalization Group. *Physical Review Letters*, **93**, Article 076401. https://doi.org/10.1103/physrevlett.93.076401

[14] Zaletel, M.P., Mong, R.S.K., Karrasch, C., Moore, J.E. and Pollmann, F. (2015) Time-Evolving a Matrix Product State with Long-Ranged Interactions. *Physical Review B*, **91**, Article 165112. https://doi.org/10.1103/physrevb.91.165112

[15] Haegeman, J., Cirac, J.I., Osborne, T.J., Pižorn, I., Verschelde, H. and Verstraete, F. (2011) Time-Dependent Variational Principle for Quantum Lattices. *Physical Review Letters*, **107**, Article 070601. https://doi.org/10.1103/physrevlett.107.070601

[16] Haegeman, J., Lubich, C., Oseledets, I., Vandereycken, B. and Verstraete, F. (2016) Unifying Time Evolution and Optimization with Matrix Product States. *Physical Review B*, **94**, Article 165116. https://doi.org/10.1103/physrevb.94.165116

[17] Pollmann, F., Khemani, V., Cirac, J.I. and Sondhi, S.L. (2016) Efficient Variational Diagonalization of Fully Many-Body Localized Hamiltonians. *Physical Review B*, **94**, Article 041116. https://doi.org/10.1103/physrevb.94.041116

[18] Khemani, V., Pollmann, F. and Sondhi, S.L. (2016) Obtaining Highly Excited Eigenstates of Many-Body Localized Hamiltonians by the Density Matrix Renormalization Group Approach. *Physical Review Letters*, **116**, Article 247204. https://doi.org/10.1103/physrevlett.116.247204

[19] Yu, X., Pekker, D. and Clark, B.K. (2017) Finding Matrix Product State Representations of Highly Excited Eigenstates of Many-Body Localized Hamiltonians. *Physical Review Letters*, **118**, Article 017201. https://doi.org/10.1103/physrevlett.118.017201

[20] Liang, S. and Pang, H. (1994) Approximate Diagonalization Using the Density Matrix Renormalization-Group Method: A Two-Dimensional-Systems Perspective. *Physical Review B*, **49**, 9214-9217. https://doi.org/10.1103/physrevb.49.9214

[21] McCulloch, I.P. and Gulácsi, M. (2002) The Non-Abelian Density Matrix Renormalization Group Algorithm. *Europhysics Letters* (*EPL*), **57**, 852-858. https://doi.org/10.1209/epl/i2002-00393-0

[22] Singh, S., Pfeifer, R.N.C. and Vidal, G. (2010) Tensor Network Decompositions in the Presence of a Global Symmetry. *Physical Review A*, **82**, Article 050301. https://doi.org/10.1103/physreva.82.050301

[23] Singh, S., Pfeifer, R.N.C. and Vidal, G. (2011) Tensor Network States and Algorithms in the Presence of a Global U(1) Symmetry. *Physical Review B*, **83**, Article 115125. https://doi.org/10.1103/physrevb.83.115125

[24] Singh, S. and Vidal, G. (2012) Tensor Network States and Algorithms in the Presence of a Global SU(2) Symmetry. *Physical Review B*, **86**, Article 195114. https://doi.org/10.1103/physrevb.86.195114

[25] Weichselbaum, A. (2012) Non-abelian Symmetries in Tensor Networks: A Quantum Symmetry Space Approach. *Annals of Physics*, **327**, 2972-3047. https://doi.org/10.1016/j.aop.2012.07.009

[26] White, S.R. (2005) Density Matrix Renormalization Group Algorithms with a Single Center Site. *Physical Review B*, **72**, Article 180403. https://doi.org/10.1103/physrevb.72.180403

[27] Hubig, C., McCulloch, I.P., Schollwöck, U. and Wolf, F.A. (2015) Strictly Single-Site DMRG Algorithm with Subspace Expansion. *Physical Review B*, **91**, Article 155115. https://doi.org/10.1103/physrevb.91.155115

[28] Motruk, J., Zaletel, M.P., Mong, R.S.K. and Pollmann, F. (2016) Density Matrix Renormalization Group on a Cylinder in Mixed Real and Momentum Space. *Physical Review B*, **93**, Article 155139. https://doi.org/10.1103/physrevb.93.155139

[29] Ehlers, G., White, S.R. and Noack, R.M. (2017) Hybrid-Space Density Matrix Renormalization Group Study of the Doped Two-Dimensional Hubbard Model. *Physical Review B*, **95**, Article 125125. https://doi.org/10.1103/physrevb.95.125125

[30] Stoudenmire, E.M. and White, S.R. (2013) Real-Space Parallel Density Matrix Renormalization Group. *Physical Review B*, **87**, Article 155137.

https://doi.org/10.1103/physrevb.87.155137

[31] Verstraete, F. and Cirac, J.I. (2010) Continuous Matrix Product States for Quantum Fields. *Physical Review Letters*, **104**, Article 190405. https://doi.org/10.1103/physrevlett.104.190405

[32] Haegeman, J., Cirac, J.I., Osborne, T.J. and Verstraete, F. (2013) Calculus of Continuous Matrix Product States. *Physical Review B*, **88**, Article 085118. https://doi.org/10.1103/physrevb.88.085118

[33] Rincón, J., Ganahl, M. and Vidal, G. (2015) Lieb-Liniger Model with Exponentially Decaying Interactions: A Continuous Matrix Product State Study. *Physical Review B*, **92**, Article 115107. https://doi.org/10.1103/physrevb.92.115107

[34] Ganahl, M., Rincón, J. and Vidal, G. (2017) Continuous Matrix Product States for Quantum Fields: An Energy Minimization Algorithm. *Physical Review Letters*, **118**, Article 220402. https://doi.org/10.1103/physrevlett.118.220402

[35] Chung, S.S., Bauman, S., Sun, K. and Bolech, C.J. (2016) On the New Continuous Matrix Product Ansatz. *Journal of Physics*: *Conference Series*, **702**, Article 012004. https://doi.org/10.1088/1742-6596/702/1/012004

[36] Draxler, D., Haegeman, J., Verstraete, F. and Rizzi, M. (2017) Continuous Matrix Product States with Periodic Boundary Conditions and an Application to Atomtronics. *Physical Review B*, **95**, Article 045145. https://doi.org/10.1103/physrevb.95.045145

[37] Tilloy, A. (2021) Relativistic Continuous Matrix Product States for Quantum Fields without Cutoff. *Physical Review D*, **104**, Article 096007. https://doi.org/10.1103/physrevd.104.096007

[38] Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., *et al*. (2019) Machine Learning and the Physical Sciences. *Reviews of Modern Physics*, **91**, Article 045002. https://doi.org/10.1103/revmodphys.91.045002

[39] Carleo, G. and Troyer, M. (2017) Solving the Quantum Many-Body Problem with Artificial Neural Networks. *Science*, **355**, 602-606. https://doi.org/10.1126/science.aag2302

[40] Gao, X. and Duan, L. (2017) Efficient Representation of Quantum Many-Body States with Deep Neural Networks. *Nature Communications*, **8**, Article No. 662. https://doi.org/10.1038/s41467-017-00705-2

[41] Carleo, G., Nomura, Y. and Imada, M. (2018) Constructing Exact Representations of Quantum Many-Body Systems with Deep Neural Networks. *Nature Communications*, **9**, Article No. 5322. https://doi.org/10.1038/s41467-018-07520-3

[42] Freitas, N., Morigi, G. and Dunjko, V. (2018) Neural Network Operations and Susuki-Trotter Evolution of Neural Network States. *International Journal of Quantum Information*, **16**, Article 1840008. https://doi.org/10.1142/s0219749918400087

[43] Glasser, I., Pancotti, N., August, M., Rodriguez, I.D. and Cirac, J.I. (2018) Neural-Network Quantum States, String-Bond States, and Chiral Topological States. *Physical Review X*, **8**, Article 011006. https://doi.org/10.1103/physrevx.8.011006

[44] Nomura, Y., Darmawan, A.S., Yamaji, Y. and Imada, M. (2017) Restricted Boltzmann Machine Learning for Solving Strongly Correlated Quantum Systems. *Physical Review B*, **96**, Article 205152. https://doi.org/10.1103/physrevb.96.205152

[45] Vieijra, T., Casert, C., Nys, J., De Neve, W., Haegeman, J., Ryckebusch, J., *et al*. (2020) Restricted Boltzmann Machines for Quantum States with Non-Abelian or Anyonic Symmetries. *Physical Review Letters*, **124**, Article 097201. https://doi.org/10.1103/physrevlett.124.097201

[46] Golubeva, A. and Melko, R.G. (2022) Pruning a Restricted Boltzmann Machine for Quantum State Reconstruction. *Physical Review B*, **105**, Article 125124. https://doi.org/10.1103/physrevb.105.125124

[47] Pilati, S. and Pieri, P. (2020) Simulating Disordered Quantum Ising Chains via Dense and Sparse Restricted Boltzmann Machines. *Physical Review E*, **101**, Article 063308. https://doi.org/10.1103/physreve.101.063308

[48] Lu, S., Gao, X. and Duan, L.-. (2019) Efficient Representation of Topologically Ordered States with Restricted Boltzmann Machines. *Physical Review B*, **99**, Article 155136. https://doi.org/10.1103/physrevb.99.155136

[49] Nomura, Y., Yoshioka, N. and Nori, F. (2021) Purifying Deep Boltzmann Machines for Thermal Quantum States. *Physical Review Letters*, **127**, Article 060601. https://doi.org/10.1103/physrevlett.127.060601

[50] Vicentini, F., Biella, A., Regnault, N. and Ciuti, C. (2019) Variational Neural-Network Ansatz for Steady States in Open Quantum Systems. *Physical Review Letters*, **122**, Article 250503. https://doi.org/10.1103/physrevlett.122.250503

[51] Yoshioka, N. and Hamazaki, R. (2019) Constructing Neural Stationary States for Open Quantum Many-Body Systems. *Physical Review B*, **99**, Article 214306. https://doi.org/10.1103/physrevb.99.214306

[52] Nagy, A. and Savona, V. (2019) Variational Quantum Monte Carlo Method with a Neural-Network Ansatz for Open Quantum Systems. *Physical Review Letters*, **122**, Article 250501. https://doi.org/10.1103/physrevlett.122.250501

[53] Hartmann, M.J. and Carleo, G. (2019) Neural-Network Approach to Dissipative Quantum Many-Body Dynamics. *Physical Review Letters*, **122**, Article 250502. https://doi.org/10.1103/physrevlett.122.250502

[54] Nomura, Y. (2021) Helping Restricted Boltzmann Machines with Quantum-State Representation by Restoring Symmetry. *Journal of Physics: Condensed Matter*, **33**, Article 174003. https://doi.org/10.1088/1361-648x/abe268

[55] Park, C. and Kastoryano, M.J. (2022) Expressive Power of Complex-Valued Restricted Boltzmann Machines for Solving Nonstoquastic Hamiltonians. *Physical Review B*, **106**, Article 134437. https://doi.org/10.1103/physrevb.106.134437

[56] Park, C. and Kastoryano, M.J. (2020) Geometry of Learning Neural Quantum States. *Physical Review Research*, **2**, Article 023232. https://doi.org/10.1103/physrevresearch.2.023232

[57] Choo, K., Carleo, G., Regnault, N. and Neupert, T. (2018) Symmetries and Many-Body Excitations with Neural-Network Quantum States. *Physical Review Letters*, **121**, Article 167204. https://doi.org/10.1103/physrevlett.121.167204

[58] Torlai, G. and Melko, R.G. (2018) Latent Space Purification via Neural Density Operators. *Physical Review Letters*, **120**, Article 240503. https://doi.org/10.1103/physrevlett.120.240503

[59] Zen, R., My, L., Tan, R., Hébert, F., Gattobigio, M., Miniatura, C., *et al.* (2020) Transfer Learning for Scalability of Neural-Network Quantum States. *Physical Review E*, **101**, Article 053301. https://doi.org/10.1103/physreve.101.053301

[60] Cai, Z. and Liu, J. (2018) Approximating Quantum Many-Body Wave Functions Using Artificial Neural Networks. *Physical Review B*, **97**, Article 035116. https://doi.org/10.1103/physrevb.97.035116

[61] Luo, D. and Clark, B.K. (2019) Backflow Transformations via Neural Networks for Quantum Many-Body Wave Functions. *Physical Review Letters*, **122**, Article 226401. https://doi.org/10.1103/physrevlett.122.226401

[62] Sharir, O., Levine, Y., Wies, N., Carleo, G. and Shashua, A. (2020) Deep Autoregressive Models for the Efficient Variational Simulation of Many-Body Quantum Systems. *Physical Review Letters*, **124**, Article 020503. https://doi.org/10.1103/physrevlett.124.020503

[63] Kessler, J., Calcavecchia, F. and Kühne, T.D. (2021) Artificial Neural Networks as Trial Wave Functions for Quantum Monte Carlo. *Advanced Theory and Simulations*, **4**, Article 200069. https://doi.org/10.1002/adts.202000269

[64] Hibat-Allah, M., Ganahl, M., Hayward, L.E., Melko, R.G. and Carrasquilla, J. (2020) Recurrent Neural Network Wave Functions. *Physical Review Research*, **2**, Article 023358. https://doi.org/10.1103/physrevresearch.2.023358

[65] Roth, C. (2020) Iterative Retraining of Quantum Spin Models Using Recurrent Neural Networks.

[66] Chen, Z., Newhouse, L., Chen, E., Luo, D. and Soljačić, M. (2023) ANTN: Bridging Autoregressive Neural Networks and Tensor Networks for Quantum Many-Body Simulation.

[67] Schmitt, M. and Heyl, M. (2020) Quantum Many-Body Dynamics in Two Dimensions with Artificial Neural Networks. *Physical Review Letters*, **125**, Article 100503. https://doi.org/10.1103/physrevlett.125.100503

[68] Gutiérrez, I.L. and Mendl, C.B. (2022) Real Time Evolution with Neural-Network Quantum States. *Quantum*, **6**, 627. https://doi.org/10.22331/q-2022-01-20-627

[69] Yang, L., Leng, Z., Yu, G., Patel, A., Hu, W. and Pu, H. (2020) Deep Learning-Enhanced Variational Monte Carlo Method for Quantum Many-Body Physics. *Physical Review Research*, **2**, Article 012039. https://doi.org/10.1103/physrevresearch.2.012039

[70] Irikura, N. and Saito, H. (2020) Neural-Network Quantum States at Finite Temperature. *Physical Review Research*, **2**, Article 013284. https://doi.org/10.1103/physrevresearch.2.013284

[71] Vieijra, T. and Nys, J. (2021) Many-body Quantum States with Exact Conservation of Non-Abelian and Lattice Symmetries through Variational Monte Carlo. *Physical Review B*, **104**, Article 045123. https://doi.org/10.1103/physrevb.104.045123

[72] Liang, X., Liu, W., Lin, P., Guo, G., Zhang, Y. and He, L. (2018) Solving Frustrated Quantum Many-Particle Models with Convolutional Neural Networks. *Physical Review B*, **98**, Article 104426. https://doi.org/10.1103/physrevb.98.104426

[73] Liu, C. and Wang, D. (2021) Random Sampling Neural Network for Quantum Many-Body Problems. *Physical Review B*, **103**, Article 205107. https://doi.org/10.1103/physrevb.103.205107

[74] Saito, H. and Kato, M. (2018) Machine Learning Technique to Find Quantum Many-Body Ground States of Bosons on a Lattice. *Journal of the Physical Society of Japan*, **87**, Article 014001. https://doi.org/10.7566/jpsj.87.014001

[75] Roth, C. and MacDonald, A.H. (2021) Group Convolutional Neural Networks Improve Quantum State Accuracy.

[76] Fu, C., Zhang, X., Zhang, H., Ling, H., Xu, S. and Ji, S. (2021) Lattice Convolutional Networks for Learning Ground States of Quantum Many-Body Systems.

[77] Liang, X., Li, M., Xiao, Q., An, H., *et al.* (2022) 2^{1296} Exponentially Complex Quantum Many-Body Simulation via Scalable Deep Learning Method.

[78] Choo, K., Neupert, T. and Carleo, G. (2019) Two-Dimensional Frustrated J1 - J2 Model Studied with Neural Network Quantum States. *Physical Review B*, **100**, Article 125124. https://doi.org/10.1103/physrevb.100.125124

[79] Mikolov, T., Kombrink, S., Burget, L., Cernocky, J. and Khudanpur, S. (2011) Extensions of Recurrent Neural Network Language Model. 2011 *IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*), Prague, 22-27 May 2011, 5528-5531. https://doi.org/10.1109/icassp.2011.5947611

[80] Graves, A., Mohamed, A. and Hinton, G. (2013) Speech Recognition with Deep Recurrent Neural Networks. 2013 *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, 26-31 May 2013, 6645-6649. https://doi.org/10.1109/icassp.2013.6638947

[81] Gregor, K., Danihelka, I., Graves, A., Rezende, D.J. and Wierstra, D. (2015) Draw: A Recurrent Neural Network for Image Generation.

[82] Oord, A.V.D., Kalchbrenner, N. and Kavukcuoglu, K. (2016) Pixel Recurrent Neural Networks.

[83] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., *et al.* (2017) Attention Is All You Need. 2017 31*st Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 1-11.

[84] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K. (2018) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.

[85] Brown, T.B., Mann, B., Ryder, N., *et al.* (2020) Language Models Are Few-Shot Learners. *Advances in Neural Information Processing Systems*, **33**, 1877-1901.

[86] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale.

[87] Jumper, J., Evans, R., Pritzel, A., *et al.* (2021) Highly Accurate Protein Structure Prediction with AlphaFold. *Nature*, **596**, Article 583589.

[88] Luo, D., Chen, Z., Hu, K., Zhao, Z., *et al.* (2021) Gauge Invariant and Anyonic Symmetric Transformer and RNN Quantum States for Quantum Lattice Models.

[89] Luo, D., Chen, Z., Carrasquilla, J. and Clark, B.K. (2022) Autoregressive Neural Network for Simulating Open Quantum Systems via a Probabilistic Formulation. *Physical Review Letters*, **128**, Article 090501. https://doi.org/10.1103/physrevlett.128.090501

[90] Cha, P., Ginsparg, P., Wu, F., Carrasquilla, J., McMahon, P.L. and Kim, E. (2021) Attention-Based Quantum Tomography. *Machine Learning*: *Science and Technology*, **3**, 01LT01. https://doi.org/10.1088/2632-2153/ac362b

[91] Carrasquilla, J., Luo, D., Pérez, F., Milsted, A., Clark, B.K., Volkovs, M., *et al.* (2021) Probabilistic Simulation of Quantum Circuits Using a Deep-Learning Architecture. *Physical Review A*, **104**, Article 032610. https://doi.org/10.1103/physreva.104.032610

[92] Abrams, D.S. and Lloyd, S. (1999) Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors. *Physical Review Letters*, **83**, 5162-5165. https://doi.org/10.1103/physrevlett.83.5162

[93] Aspuru-Guzik, A., Dutoi, A.D., Love, P.J. and Head-Gordon, M. (2005) Simulated Quantum Computation of Molecular Energies. *Science*, **309**, 1704-1707. https://doi.org/10.1126/science.1113479

[94] Kandala, A., Temme, K., Córcoles, A.D., Mezzacapo, A., Chow, J.M. and Gambetta, J.M. (2019) Error Mitigation Extends the Computational Reach of a Noisy Quantum Processor. *Nature*, **567**, 491-495. https://doi.org/10.1038/s41586-019-1040-7

[95] Higgott, O., Wang, D. and Brierley, S. (2019) Variational Quantum Computation of Excited States. *Quantum*, **3**, 156. https://doi.org/10.22331/q-2019-07-01-156

[96] McClean, J.R., Kimchi-Schwartz, M.E., Carter, J. and de Jong, W.A. (2017) Hybrid

Quantum-Classical Hierarchy for Mitigation of Decoherence and Determination of Excited States. *Physical Review A*, **95**, Article 042308.
https://doi.org/10.1103/physreva.95.042308

[97] Heya, K., Nakanishi, K.M., Mitarai, K. and Fujii, K. (2019) Subspace Variational Quantum Simulator.

[98] Nakanishi, K.M., Mitarai, K. and Fujii, K. (2019) Subspace-Search Variational Quantum Eigensolver for Excited States. *Physical Review Research*, **1**, Article 033062.
https://doi.org/10.1103/physrevresearch.1.033062

[99] Parrish, R.M., Hohenstein, E.G., McMahon, P.L. and Martínez, T.J. (2019) Quantum Computation of Electronic Transitions Using a Variational Quantum Eigensolver. *Physical Review Letters*, **122**, Article 230401.
https://doi.org/10.1103/physrevlett.122.230401

[100] Garcia-Saez, A. and Latorre, J. (2018) Addressing Hard Classical Problems with Adiabatically Assisted Variational Quantum Eigensolvers.

[101] Cerezo, M., Sharma, K., Arrasmith, A. and Coles, P.J. (2020) Variational Quantum State Eigensolver.

[102] Wang, D., Higgott, O. and Brierley, S. (2019) Accelerated Variational Quantum Eigensolver. *Physical Review Letters*, **122**, Article 140504.
https://doi.org/10.1103/physrevlett.122.140504

[103] Wang, G., Koh, D.E., Johnson, P.D. and Cao, Y. (2021) Minimizing Estimation Runtime on Noisy Quantum Computers. *PRX Quantum*, **2**, Article 010346.
https://doi.org/10.1103/prxquantum.2.010346

[104] Wang, G., Koh, D.E., Johnson, P.D. and Cao, Y. (2006) Bayesian Inference with Engineered Likelihood Functions for Robust Amplitude Estimation.

[105] Wang, Y. (2012) Quantum Computation and Quantum Information. *Statistical Science*, **27**, 373-394. https://doi.org/10.1214/11-sts378

[106] Li, Y. and Benjamin, S.C. (2017) Efficient Variational Quantum Simulator Incorporating Active Error Minimization. *Physical Review X*, **7**, Article 021050.
https://doi.org/10.1103/physrevx.7.021050

[107] Yuan, X., Endo, S., Zhao, Q., Li, Y. and Benjamin, S.C. (2019) Theory of Variational Quantum Simulation. *Quantum*, **3**, 191. https://doi.org/10.22331/q-2019-10-07-191

[108] McArdle, S., Jones, T., Endo, S., Li, Y., Benjamin, S.C. and Yuan, X. (2019) Variational Ansatz-Based Quantum Simulation of Imaginary Time Evolution. *NPJ Quantum Information*, **5**, Article No. 75. https://doi.org/10.1038/s41534-019-0187-2

[109] Endo, S., Sun, J., Li, Y., Benjamin, S.C. and Yuan, X. (2020) Variational Quantum Simulation of General Processes. *Physical Review Letters*, **125**, Article 010501.
https://doi.org/10.1103/physrevlett.125.010501

[110] Yao, Y., Gomes, N., Zhang, F., Wang, C., Ho, K., Iadecola, T., *et al.* (2021) Adaptive Variational Quantum Dynamics Simulations. *PRX Quantum*, **2**, Article 030307.
https://doi.org/10.1103/prxquantum.2.030307

[111] Zhang, Z.J., Sun, J., Yuan, X. and Yung, M.H. (2020) Low-Depth Hamiltonian Simulation by Adaptive Product Formula.

[112] Cîrstoiu, C., Holmes, Z., Iosue, J., Cincio, L., Coles, P.J. and Sornborger, A. (2020) Variational Fast Forwarding for Quantum Simulation beyond the Coherence Time. *npj Quantum Information*, **6**, Article No. 82.
https://doi.org/10.1038/s41534-020-00302-0

[113] Gibbs, J., Gili, K., Holmes, Z., Commeau, B., Arrasmith, A., Cincio, L., *et al.* (2022)

Long-Time Simulations for Fixed Input States on Quantum Hardware. *npj Quantum Information*, **8**, Article No. 135. https://doi.org/10.1038/s41534-022-00625-0

[114] Khatri, S., LaRose, R., Poremba, A., Cincio, L., Sornborger, A.T. and Coles, P.J. (2019) Quantum-Assisted Quantum Compiling. *Quantum*, **3**, 140. https://doi.org/10.22331/q-2019-05-13-140

[115] Commeau, B., Cerezo, M., Holmes, Z., Cincio, L., *et al.* (2020) Variational Hamiltonian Diagonalization for Dynamical Quantum Simulation.

[116] Wecker, D., Hastings, M.B. and Troyer, M. (2015) Progress towards Practical Quantum Variational Algorithms. *Physical Review A*, **92**, Article 042303. https://doi.org/10.1103/physreva.92.042303

[117] Wiersema, R., Zhou, C., de Sereville, Y., Carrasquilla, J.F., Kim, Y.B. and Yuen, H. (2020) Exploring Entanglement and Optimization within the Hamiltonian Variational Ansatz. *PRX Quantum*, **1**, Article 020319. https://doi.org/10.1103/prxquantum.1.020319

[118] Ho, W.W. and Hsieh, T.H. (2019) Efficient Variational Simulation of Non-Trivial Quantum States. *Sci Post Physics*, **6**, 10-19. https://doi.org/10.21468/scipostphys.6.3.029

[119] Hamer, C.J., Weihong, Z. and Oitmaa, J. (1997) Series Expansions for the Massive Schwinger Model in Hamiltonian Lattice Theory. *Physical Review D*, **56**, 55-67. https://doi.org/10.1103/physrevd.56.55

[120] Bañuls, M.C., Cichy, K., Cirac, J.I. and Jansen, K. (2013) The Mass Spectrum of the Schwinger Model with Matrix Product States. *Journal of High Energy Physics*, **2013**, Article No. 158. https://doi.org/10.1007/jhep11(2013)158

[121] Cohen, T. and Welling, M. (2016) Group Equivariant Convolutional Networks.

[122] Zhang, Y. and Di Ventra, M. (2023) Transformer Quantum State: A Multipurpose Model for Quantum Many-Body Problems. *Physical Review B*, **107**, Article 075147. https://doi.org/10.1103/physrevb.107.075147

Appendix

The decomposition in (7)

For the sake of notation convenience, $s_i = 1,\cdots,d$ for $i = 0,\cdots,N-1$. Combining the last $N-1$ indices $\{s_i\}_{i=1}^{N-1}$ together: $(s_1\cdots s_{N-1}) = (s_1-1)d^{N-2} + \cdots + (s_{N-1}-1)d^0 + 1 \le d^{N-1}$, the tensor $\psi_{s_0,\cdots,s_{N-1}}$ can be regarded as $d\times d^{N-1}$ matrix

$$\psi_{s_0,\cdots,s_{N-1}} \to \psi_{s_0,(s_1\cdots s_{N-1})} \tag{77}$$

The SVD of $d\times d^{N-1}$ matrix $\psi_{s_0,(s_1\cdots s_{N-1})}$ is

$$\psi_{s_0,(s_1\cdots s_{N-1})} = \sum_{\alpha_1=1}^{\chi_1} U_{s_0,\alpha_1} S_{\alpha_1} V^{\dagger}_{\alpha_1,(s_1\cdots s_{N-1})} \tag{78}$$

with $\chi_1 \le d$, S_{α_1} is the singular value, U and V are $d\times\chi_1$ and $d^{N-1}\times\chi_1$ unitary matrix (the column vectors with unit length are orthogonal to each other), respectively. Let $\left(M^{[0]s_0}\right)_{\alpha_0,\alpha_1} = U_{s_0,\alpha_1}$, which is a $1\times\chi_1$ matrix ($\alpha_0 = 1$ and s_0 fixed). The left two parts in SVD are combined to be $(\chi_1 d)\times d^{N-2}$ matrix

$$S_{\alpha_1} V^{\dagger}_{\alpha_1,(s_1\cdots s_{N-1})} \to \psi_{(\alpha_1 s_1),(s_2\cdots s_{N-1})} \tag{79}$$

where $(\alpha_1 s_1) = (\alpha_1 - 1)d + s_1 \le (\chi_1 - 1)d + d = \chi_1 d$, $(s_2\cdots s_{N-1}) = (s_2-1)d^{N-3} + \cdots + (s_{N-1}-1)d^0 + 1 \le d^{N-2}$. So, we have

$$\psi_{s_0,(s_1\cdots s_{N-1})} = \sum_{\alpha_1=1}^{\chi_1} \left(M^{[0]s_0}\right)_{\alpha_0,\alpha_1} \psi_{(\alpha_1 s_1),(s_2\cdots s_{N-1})}$$

The SVD of $\psi_{(\alpha_1 s_1),(s_2\cdots s_{N-1})}$ is

$$\psi_{(\alpha_1 s_1),(s_2\cdots s_{N-1})} = \sum_{\alpha_2=1}^{\chi_2} U_{(\alpha_1 s_1),\alpha_2} S_{\alpha_2} V^{\dagger}_{\alpha_2,(s_2\cdots s_{N-1})} \tag{80}$$

Let $\left(M^{[1]s_1}\right)_{\alpha_1,\alpha_2} = U_{(\alpha_1 s_1),\alpha_2}$, which is a $\chi_1\times\chi_2$ matrix (s_1 fixed). The remaining two parts are combined as $(\chi_2 d)\times d^{N-3}$ matrix

$$S_{\alpha_2} V^{\dagger}_{\alpha_2,(s_2\cdots s_{N-1})} \to \psi_{(\alpha_2 s_2),(s_3\cdots s_{N-1})} \tag{81}$$

After $N-2$ steps of SVD, we get $(\chi_{N-2} d)\times d$ matrix $\psi_{(\alpha_{N-2}s_{N-2}),s_{N-1}}$, and use SVD again

$$\psi_{(\alpha_{N-2}s_{N-2}),s_{N-1}} = \sum_{\alpha_{N-1}=1}^{\chi_{N-1}} U_{(\alpha_{N-2}s_{N-2}),\alpha_{N-1}} S_{\alpha_{N-1}} V^{\dagger}_{\alpha_{N-1},s_{N-1}} \tag{82}$$

Let $\left(M^{[N-2]s_{N-2}}\right)_{\alpha_{N-2},\alpha_{N-1}} = U_{(\alpha_{N-2}s_{N-2}),\alpha_{N-1}}$, which is a $\chi_{N-2}\times\chi_{N-1}$ matrix. Let $\left(M^{[N-1]s_{N-1}}\right)_{\alpha_{N-1},\alpha_N} = S_{\alpha_{N-1}} V^{\dagger}_{\alpha_{N-1},s_{N-1}}$, which is a $\chi_{N-1}\times 1$ matrix ($\alpha_N = 1$ and $s_{N-1} = 1$ fixed). Thus after $N-1$ steps of SVD, the decomposition in (7) is obtained.

From the above SVD processes, $\left\{M^{[i]}\right\}_{i=0}^{N-2}$ satisfy the left canonical condition. For example,

$$\sum_{s_0}\left(M^{[0]s_0}\right)^\dagger M^{[0]s_0}=\mathbb{I}_{\chi_1}$$
$$\Leftrightarrow\sum_{s_0,k}\overline{\left(M^{[0]s_0}\right)_{ki}}\left(M^{[0]s_0}\right)_{kj}=\sum_{s_0}\overline{U_{s_0 i}}U_{s_0 j}=\delta_{ij},\ 1\le i,j\le\chi_1 \tag{83}$$

due to the unitary matrix U where $\mathbb{I}_{\chi_1}$ is the $\chi_1\times\chi_1$ identity matrix. Here the summation over $k=1$ can be ignored. Since $\left\{M^{[i]}\right\}_{i=0}^{N-2}$ satisfy the left canonical condition, one has

$$\begin{aligned}\langle\psi|\psi\rangle&=\sum_s\left|\psi_{s_0,\cdots,s_{N-1}}\right|^2\\&=\sum_s\left(M^{[0]s_0}\right)_{1,\beta_1}\left(M^{[1]s_1}\right)_{\beta_1\beta_2}\cdots\left(M^{[N-1]s_{N-1}}\right)_{\beta_{N-1},1}\\&\quad\overline{\left(M^{[0]s_0}\right)_{1,\gamma_1}}\,\overline{\left(M^{[1]s_1}\right)_{\gamma_1\gamma_2}}\cdots\overline{\left(M^{[N-1]s_{N-1}}\right)_{\gamma_{N-1},1}}\\&=\sum_{s_{N-1},\alpha_{N-1}}\overline{\left(M^{[N-1]s_{N-1}}\right)_{\alpha_{N-1},1}}\left(M^{[N-1]s_{N-1}}\right)_{\alpha_{N-1},1}\\&=\sum_{s_{N-1},\alpha_{N-1}}\overline{S_{\alpha_{N-1}}V^\dagger_{\alpha_{N-1},s_{N-1}}}S_{\alpha_{N-1}}V^\dagger_{\alpha_{N-1},s_{N-1}}\\&=\sum_{s_{N-1},\alpha_{N-1}}S^2_{\alpha_{N-1}}V_{s_{N-1},\alpha_{N-1}}\overline{V_{s_{N-1},\alpha_{N-1}}}=\sum_{\alpha_{N-1}=1}^{\chi_{N-1}}S^2_{\alpha_{N-1}}\end{aligned}\tag{84}$$

If $\sum_{s_0,\cdots,s_{N-1}}\left|\psi_{s_0,\cdots,s_{N-1}}\right|^2=1$, the left canonical condition for $M^{[N-1]}$ is also satisfied.