Gradient optimization of fermionic projected entangled pair states on directed lattices

Shao-Jun Dong, Chao Wang, Yongjian Han, Guang-can Guo, and Lixin He

1 Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei, Anhui, 230026, People’s Republic of China
2 Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, 230026, China

The recently developed stochastic gradient method combined with Monte Carlo sampling techniques [PRB 95, 195154 (2017)] offers a low scaling and accurate method to optimize the projected entangled pair states (PEPS). We extended this method to the fermionic PEPS (fPEPS). To simplify the implementation, we introduce a fermi arrow notation to specify the order of the fermion operators in the virtual entangled EPR pairs. By defining some local operation rules associated with the fermi arrows, one can implement fPEPS algorithms very similar to that of standard PEPS. We benchmark the method for the interacting spinless fermion models, and the t-J models. The numerical calculations show that the gradient optimization greatly improves the results of simple update method. Furthermore, much larger virtual bond dimensions (D) and truncation dimensions (Dc) than those of boson and spin systems are necessary to converge the results. The method therefore offer a powerful tool to simulate fermion systems because it has much lower scaling than the direct contraction methods.

I. INTRODUCTION

Interacting quantum many-body systems pose some of the most exciting open problems in physics. Particularly, fermion systems are central to many of the most fascinating effects in condensed matter physics, such as high-temperature superconductivity,\(^1\) the fractional quantum Hall effect,\(^2\) and Mott insulator transitions.\(^3,4\) The simulation of the strongly correlated fermion system plays the critical role to understand these system and is also one of the most challenging problems in condensed matter physics.

The Quantum Monte Carlo (QMC)\(^5\) method as one of the leading methods in studying many-body physics has achieved great success in Bosonic and spin systems since its first proposed. However, except in some special cases,\(^6\) the fermion systems are extremely difficult to treat using QMC simulations\(^7,8\) because of the notorious “sign problems”.

Recently, the methods based on tensor network states (TNS), especially the projected entangled states (PEPS)\(^9,10\) have shown their power on simulation of the strongly correlated many-particle systems. The PEPS is sign-problem free and has achieved great successes in studying the frustrated spin models.\(^17,18\) The PEPS method has been extended to study fermion models (namely fPEPS) by different approaches.\(^20,21\) Apparently, the fPEPS are more complicated than PEPS because of the anti-communication properties of the fermion operators. In addition, fermion systems are highly frustrated. It has been proven that the entanglements of the ground states of some fermion systems are beyond the area law.\(^25,26\) Therefore, to faithfully simulate such models, it usually requires much larger bond dimensions (D) than those for boson and spin models. Furthermore, the highly entangled property makes the environment of the PEPS more subtle in the fermion system. To exactly consider the environment, the traditional contraction methods suffer from extremely high computational scaling to the bond dimensions. This problem is more serious for fermions than bosons because much larger bond dimensions are needed in the simulations. The recently developed Monte Carlo sampling techniques for PEPS can greatly reduce the computational scaling.\(^15,27\) By combining with stochastic gradient method, one can achieve great precision in obtaining the ground states.\(^31,32\) In this work, we extended the stochastic gradient method to optimize the fPEPS wave functions for fermion systems.

To simplify the implementation of the fPEPS algorithms, we introduce a “fermi arrow” notation to specify the order of the fermion operators in the entangled EPR pairs. With this notation and some local operation rules associated with the fermi arrows, we can greatly simplify the implementation of the stochastic gradient optimization method (and other methods) for fPEPS. We implement this local operation rules for fPEPS in our recently developed TNSpack,\(^33\) in which the anti-communication properties of the fermi operators are automatically taken account of. Therefore, one can implement fPEPS algorithm very similar to that of the standard PEPS without worry too much about the details of the anti-communication between the fermion operators.

We benchmark the stochastic gradient method for fPEPS on the interacting spinless fermion models, and the t-J models. The numerical calculations show that the gradient optimization greatly improves the results of simple update method. Furthermore, much larger virtual bond dimensions (D) and truncation dimensions (Dc)
than those of Bosonic systems are necessary to converge the results which is the dominant difficulty to simulate the fermion systems. Therefore the present method is advantageous because it has much lower scaling than the traditional direct contraction method.

II. DEFINITION OF FPEPS BASED ON DIRECTED NETWORK

The definition of the fPEPS\textsuperscript{24} on a lattice is similar to that of the standard PEPS.\textsuperscript{14,16} Without lose of generality, we take a fermion system on a $L_1 \times L_2$ square lattice as an example, where the physical dimension of each site is $d$. For each horizontal bond connecting sites $(i,j)$ and $(i,j+1)$, there is an EPR pair, i.e., a Bell type entangled state,

\[ \hat{I}_k(i,j)|0\rangle = \sum_{k=0}^{D-1} |k\rangle_{(i,j)\rightarrow k}|k\rangle_{(i,j+1)\rightarrow k}, \]

where $|k\rangle_{(i,j)\rightarrow k}$ and $|k\rangle_{(i,j+1)\rightarrow k}$ are the fermion states on site $(i,j)$ and site $(i,j+1)$. States $|k\rangle$ are generated as, $|k\rangle = |k_1 k_2 \cdots k_p\rangle = a_1^{k_1} a_2^{k_2} \cdots a_p^{k_p}|0\rangle$, where $(k_1 k_2 \cdots k_p)$ is the binary representation of $k$ and $|0\rangle$ is the vacuum state. $\hat{a}_s$ and $\hat{a}_d^\dagger$ are the fermion operators that satisfy $\{\hat{a}_s, \hat{a}_d^\dagger\} = \delta_{sd}$. For convenience, we denote the state $|k_{(i,j)\rightarrow k}\rangle = a_{(i,j)\rightarrow k}^{\dagger}|0\rangle$. Similarly, for each vertical bond connecting site $(i,j)$ and $(i+1,j)$, there is also a Bell type entangled state, (short) $\hat{I}_k(i,j)|0\rangle = \sum_{k=0}^{D-1} a_{(i,j)\rightarrow k}^{\dagger} a_{(i+1,j)\rightarrow k}|0\rangle$. Therefore, a standard virtual mother state of a fPEPS can be defined as,

\[ |\Phi_0\rangle = \prod_{i=1}^{L_1-1} \prod_{j=1}^{L_2-1} \hat{I}_k(i,j)\hat{I}_k(i,j)|0\rangle. \]

To define a quantum state in the real physical space, we project $|\Phi_0\rangle$ to the physical space. The projector on site $(i,j)$ is defined as:

\[ \hat{P}_k[i,j] = \sum_{\beta=0}^{d-1} \sum_{\beta_1, \beta_2, \beta_3, \beta_4} T_{\beta, \beta_1, \beta_2, \beta_3, \beta_4} [i,j] a_{(i,j)\rightarrow k}^{\beta_1} a_{(i,j)\rightarrow k}^{\beta_2} a_{(i,j)\rightarrow k}^{\beta_3} a_{(i,j)\rightarrow k}^{\beta_4}. \]

Here, $a_{(i,j)\rightarrow k}^{\beta}$ is the creation operator of the physical particle on site $(i,j)$ whereas $a_{(i,j)\rightarrow k}^{\beta}$ are the annihilation operators of the state $|n\rangle_{(i,j),m}$. The fPEPS is then defined as,

\[ |\Phi_{\text{fPEPS}}\rangle = \prod_{i,j} \hat{P}_k[i,j]|\Phi_0\rangle. \]

To make the fPEPS well defined, the state $|\Phi_{\text{fPEPS}}\rangle$ should be independent of the order of the projectors up to a global phase, i.e., the parity of all elements in a projector should be the same. The parity of the element $T_{\beta, \beta_1, \beta_2, \beta_3, \beta_4} [i,j] a_{(i,j)\rightarrow k}^{\beta_1} a_{(i,j)\rightarrow k}^{\beta_2} a_{(i,j)\rightarrow k}^{\beta_3} a_{(i,j)\rightarrow k}^{\beta_4}$ of the projector $\hat{P}_k[i,j]$ is obtained by $\tilde{p}(\beta_3, \beta_4) \tilde{p}(\beta_2, \beta_3) \tilde{p}(\beta_1, \beta_2)$, where $\tilde{p}(x) = -1$, if the parity of $x$ is odd, and $\tilde{p}(x) = +1$ if the parity of $x$ is even. Therefore, the parity of all elements can be obtained by the lower indices of tensor $T_{\beta, \beta_1, \beta_2, \beta_3, \beta_4} [i,j]$. Without lose of generality, we assume all nonzero projector elements have even parity in this paper. As a consequence, the elements with odd parity vanish, i.e., $T_{\beta, \beta_1, \beta_2, \beta_3, \beta_4} [i,j] = 0$, if $\beta + \beta_1 + \beta_2 + \beta_3 + \beta_4$ is odd. In this definition of fPEPS, we may interchange the positions of any two projectors and EPR pairs, because they all have even parity.

One of the key issues in the fPEPS is the order of the fermion operators, including the operators in the projectors and in EPRs. We define the standard order of the fermion operators in each projector operators on the square lattice as follows: physical creation operator, left, down, right, and up virtual operators (i.e., anti-clockwise order), which is the same as the order of the lower indices in the tensor $T_{\beta, \beta_1, \beta_2, \beta_3, \beta_4}$ (see Eq. 3). When changing the order of fermion operators, a sign which is determined by the parity of the indices will appear. For example, if we exchange the two adjacent fermion operators $a_{(i,j)\rightarrow k}^{\beta}$ and $a_{(i,j)\rightarrow k}^{\delta}$, there will be an extra phase, i.e., $T_{\beta, \beta_1, \beta_2, \beta_3, \beta_4} [i,j] \rightarrow \tilde{p}(\beta_3, \beta_4) T_{\beta, \beta_1, \beta_2, \beta_3, \beta_4} [i,j]$, where

\[ \tilde{p}(\beta_3, \beta_4) = \begin{cases} -1, & \text{both } \tilde{p}(\beta_3), \tilde{p}(\beta_4) = -1, \\ 1, & \text{otherwise}. \end{cases} \]

Beside the fermion operators appeared in projector $\hat{P}$, we also need to specify the operators’ order in the EPR pairs, which is not given in the tensors explicitly. In previous methods,\textsuperscript{20,22,24} the operator orders in EPS pairs are usually assumed beforehand and fixed during calculations. However, sometimes the fermion operators’ order in an EPR may change during the calculations, which bring extra complexity to the algorithms. In this work, we introduce a fermi arrow notation to specify the order of the EPR pairs. For example, as shown in Fig. 1(a), the arrow points from site A to site B, and the corresponding EPR state is $I_A \rightarrow B |0\rangle = \sum_{k=0}^{D-1} a_{A}^{k+1} a_{B}^{k+1}|0\rangle$, whereas in Fig. 1(b), the arrow points from B to A, and the corresponding EPR state is $I_B \rightarrow A |0\rangle = \sum_{k=0}^{D-1} a_{B}^{k} a_{A}^{k+1}|0\rangle$. The two states can be transformed to each other as follows,

\[ \sum_{k=0}^{D-1} a_{A}^{k} a_{B}^{k} |0\rangle = \sum_{k=0}^{D-1} (-1)^{\tilde{p}(k)} a_{B}^{k} a_{A}^{k+1} |0\rangle. \]

We may also assign fermi arrows to the physical indices: the fermi arrows point into the sites for
FIG. 2: A schematic example of a fPEPS on the 4×4 lattice. The circles are the tensors on the lattice whereas the black solid lines are the virtual EPR pairs and the arrows on the bonds specify the order of the fermi operators in the EPR pairs. The red solid lines are the physical indices associated with creation operators.

The physical creation operators, and pointing out of the sites for the annihilation operators. This definition is equivalent to insert EPR pairs between the physical operators when contracting the physical indices e.g., \[ \langle 0 | \sum_{\beta_1} A_1^\dagger A_1 B_2 a_{\beta_2} a_{\beta_2}^\dagger | 0 \rangle = \langle 0 | \sum_{\beta_1} A_1^\dagger A_1 B_2 a_{\beta_2} a_{\beta_2}^\dagger \sum_{\beta_3} B_3 a_{\beta_3}^\dagger a_{\beta_3}^\dagger | 0 \rangle. \] With this definition, we can treat the physical indices and the virtual indices in the same manner, and do not need to distinguish the real fermions and virtual fermions during operations. We can now uniquely define a fPEPS on a directed lattice, as shown for example in Fig. 2, on a 4×4 lattice.

By defining some calculation rules associated with fermi arrows, we are able to perform fPEPS calculations. Contraction is one of the most important operations in PEPS algorithms. When we contract the tensors on two sites in a fPEPS, we need to consider the fermi arrow direction. We take the two situations in Fig. 1 as an example, which gives two different contraction formula, for Fig. 1(a) and,

\[ \sum_{\beta,A,B} A\beta A_{\beta}^\dagger B_{\beta} a_{\beta} B_{\beta}^\dagger a_{\beta}^\dagger | 0 \rangle = \sum_{\beta} A\beta B_{\beta} | 0 \rangle \]

for Fig. 1(b). The anti-commutation relation of fermions has been used. The fermi arrow helps to distinguish the two situations of contraction in the graphical notions of Fig. 1. Using the graphic representation may greatly simplify the notation.

More generally, giving two tensors A and B, connected via multi virtual bonds (EPRs), where \{i_1, i_2, \ldots, i_k, i_{k+1}, \ldots, i_N\} are the joint bonds to be contracted. Assume that bonds \{i_1, i_2, \ldots, i_k\} have fermi arrows pointing from B to A, and the rest bonds \{i_{k+1}, \ldots, i_N\} have fermi arrows pointing from A to B. We first reshape A to \(A_{I,iN,iN-1, \ldots, i_1}\), and reshape B to \(B_{i_1,i_2, \ldots, i_N,J}\), where \(\{I\}, \{J\}\) are the bonds that are not to be contracted in A and B respectively. For the convenience of notation, we assume the signs due to the change of bond order in the tensors according to Eq. 5 have been absorbed into the tensors, then the resulting tensor is,

\[ R_{I,J} = \sum_{i_1,i_2, \ldots, i_N} \prod_{l=1,k} \tilde{p}(i_l) A_{I,iN,iN-1, \ldots, i_1} B_{i_1,i_2, \ldots, i_N,J} \]

III. STOCHASTIC GRADIENT OPTIMIZATION OF FPEPS

In order to find the ground state of a given Hamiltonian using fPEPS, different methods have been introduced. The leading method is the imaginary time evolution (ITE) method. However, due to the high computation complexity to obtain the exact environment during the time evolution, some kinds of approximations are necessary. The simple update method has been widely used, which however may have large errors because the environment is over simplified. Several methods have been developed to treat the environment more rigorously, such as the full update method, and the gradient method, which may significantly improve the results, but the scaling to \(D\) of these methods is rather high.

We recently developed stochastic gradient optimization method for PEPS, combined with Monte Carlo sampling techniques, which gives remarkable accuracy of the results. The method has two advantages. First, the environments of tensors are treated rigorously, and therefore, the results are more accurate than SU and even SU methods. Secondly, the Monte Carlo sampling technique may greatly reduce the scaling of the method to the virtual bond dimension \(D\) from \(D^{10}\) to \(D^6\) for OBC, which is even more crucial for fPEPS, where larger \(D\) is often needed to converge the results. In this work, we extended this method to fPEPS.

The fPEPS wave functions of a many-particle system.
in Eq. 4 can be rewritten as,

$$|\Psi\rangle = \sum_\{n\} C \left[ \prod_i T[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4} \right]_{n_1, n_2, \ldots, n_N}$$

$$\equiv \sum_{\{n\}} W(n)|n\rangle,$$  \hspace{1cm} (10)

where $i=(i,j)$ is the site index of the lattice, and $C$ means to contract all the entangled virtual fermions according to the rules defined in Sec. II. $W(n)$ is the coefficient of the physical state $|n\rangle = |n_1, n_2, \ldots, n_N\rangle$ in the particle number representation. The energy of the system can be written as,

$$E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{1}{\sum_{n'}|W(n')|^2} \sum_n |W(n)|^2 E(n)$$  \hspace{1cm} (11)

where,

$$E(n) = \sum_{n'} \frac{W(n')}{W(n)} \langle n' | H | n \rangle.$$  \hspace{1cm} (12)

The total energy of the system for a given fPEPS can be evaluated via Monte Carlo sampling over the physical configurations space.\textsuperscript{15,27-29,31}

To optimize the energy function, we need the derivatives of the energy with respect to the tensor elements,

$$\frac{\partial E}{\partial T[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4}} = 2\langle \Delta[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4} (n) E(n) \rangle$$  \hspace{1cm} (13)

$$-2\langle \Delta[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4} (n) \rangle \langle E(n) \rangle,$$

where $\langle \cdots \rangle$ denotes the MC average. $\Delta[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4}$ is defined as

$$\Delta[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4} (n) = \frac{1}{W(n)} \frac{\partial W(n)}{\partial T[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4}},$$  \hspace{1cm} (14)

and the derivative of $W(n)$ is

$$\frac{\partial W(n)}{\partial T[i]_{n,\beta_1,\beta_2,\beta_3,\beta_4}} = C \left[ T[1] \cdots T[i-1] T[i+1] \cdots T[N] \right].$$  \hspace{1cm} (15)

The derivatives can be also evaluated by the MC samplings.\textsuperscript{31}

Once we have the energy and its gradients, we can optimize the system energy using stochastic gradient method\textsuperscript{27,31} which has been successfully applied to the standard PEPS method.

The overall algorithm for fPEPS is similar to that of PEPS. We need to contract the fPEPS tensors at given particle configuration to obtain $W(n)$ and the gradients. However, contracting a fPEPS is much more complicated than contacting a standard PEPS, because of the anti-commutation relation of fermions. One must be very careful about the contraction order and underlying fermions’ order in EPR pairs. We show here that with the help of fermi arrows and the operation rules associated with them, the contraction can be done easily as in the standard PEPS algorithms.

To obtain $W(n)$, we need to contact a single layer of fPEPS with fixed particle configuration $|n\rangle$. We adopt the boundary-MPO method\textsuperscript{34,35} where we need to find a fermionic matrix product operator (fMPO) denoted as $|X\rangle$ with bond dimension $D_c$ [see Fig.3(b)] to approximate the two rows of fPEPS with bond dimension $D$, denoted as $|\Psi\rangle$ [see Fig.3(a)]. To find such $|X\rangle$, we minimize

$$\delta = \frac{|X\rangle - |\Psi\rangle|^2}{\langle \Psi | \Psi \rangle},$$  \hspace{1cm} (16)

which lead to the linear equation for each tensor $T^{i,j}$ on site $(i,j)$,

$$\langle X^{i,j}|X \rangle = \langle X^{i,j}|\Psi \rangle,$$  \hspace{1cm} (17)

where $|X^{i,j}\rangle$ is obtained by taking the tensor $T^{i,j}$ out of $|X\rangle$, as graphically shown in Fig. 4.
To solve the equation, we first contract the tensors on the left side of Fig. 4. We change the fermi arrow directions from Fig. 3(b) to Fig. 5(a), i.e., all arrows are pointing into site \((i,j)\). The rule of changing the directions of fermi arrows are given in the Appendix. As will be seen in the following text, the change of fermi arrow directions is to take the advantages of the canonic form of MPO\textsuperscript{10}.

We next do QR decomposition to the tensor on the first site of \(|X\rangle\), resulting in two tensors, \(Q_1\) and \(R\) as shown in Fig. 5(b). The rules for QR (and other decompositions) in the presence of fermi arrows are also given in the Appendix. We then contract the \(R\) tensor with the second tensor on the right side, and perform QR decomposition on the second site again to obtain the \(Q_2\) tensor. We repeat this process until reach the tensor \(T_n\). We contract the last \(R\) tensor with the \(T_n\), resulting in a new tenor \(T_R^{n,j}\). Similarly, we perform the LQ decomposition on the right side of \(|X\rangle\), starting from the last site to the site \((i,j)\), and contract the last \(L\) tensor with \(T_L^{n,j}\) to get \(T_R^{n,j}\). During the LQ (QR) decompositions, new fermi arrows have been inserted between \(L\) (Q) tensors and \(Q\) (R) tensor. After these processes, we obtain \(|X\rangle\) in Fig. 5(c).

We perform the same operations for \(|X^{i,j}\rangle\). After these operations, the left side of Fig. 6 become that of Fig. 6. By using the orthogonality of \(Q\) and \(Q^\dagger\), i.e., \(QQ^\dagger=I\), which is discussed in the Appendix for the fPEPS with fermi arrows, we obtain the right side of Fig. 6. The original equation Fig. 4 become of Fig. 7, which can be solved iteratively as in standard boundary MPO method\textsuperscript{34,35} which usually converges in a few sweeps.

The contraction in Eq. 15 can be calculated in the same procedures. Once we have \(W(n)\), and \(\Delta_n\beta_1\beta_2\beta_3\beta_4\), the energies and their gradients can be easily calculated.

In our calculations, we first perform ITE with simple update method to obtain a approximate ground state\textsuperscript{31} which usually have energy errors around \(10^{-2}\). We further optimize the fPEPS via gradient decent method to obtain the highly accurate ground state.

**IV. BENCHMARK RESULTS**

We benchmark our method for two typical fermion models on finite size square lattices, including the interacting spinless fermions model and the t-J model. We demonstrate that our method can give highly accurate results compared with the exact results.

**A. spinless fermions model**

The interacting spinless fermions model reads,

\[
H = -t \sum_{\langle i,j \rangle} (c_i^\dagger c_j + H.c.) + V \sum_{\langle i,j \rangle} n_i n_j ,
\]

where \(c_i^\dagger\) and \(c_j\) are the creation and the annihilation operators, and \(\langle i,j \rangle\) denotes the nearest neighbor pairs. The chemical potential \(\mu\) is set to zero here. We set hopping parameter \(t=1\), and the interaction strength \(V\geq 0\). In general this model is not exactly solvable, and has been numerically investigated in Ref.\textsuperscript{37}. For the parameters we used, the ground state is in a uniform phase when \(V\) is small and moves towards the phase boundary between the uniform phase and the phase separation with the increasing of \(V\).

We firstly calculate this model on a 4\times4 square lattice so we can compare the fPEPS results with those

\[\text{FIG. 5: Simplification of } |X\rangle: \text{(a) To make use of the orthogonality conditions, we first reverse the directions some fermi arrows. (b) We perform QR decompositions starting form the first tensor. (c) The tensor state }|X\rangle \text{ after simplification.}\]

\[\text{FIG. 6: By applying the orthogonality conditions } Q_i Q_i^\dagger = I, \text{ the left side of Fig. 4 reduces to a single tensor } T_{RL}^{n,ij}.\]

\[\text{FIG. 7: The equation in Fig. 4 after the LQ and the QR decomposition on the } |X^{i,j}\rangle.\]
obtained from the exact diagonalization method. In the calculations, we take \( D = 10 \), and \( D_c = 30 \sim 40 \). The convergence of these parameters will be discussed in details in Sec. V. The results are presented in Table I for various \( V \). As seen from Table I, the SU method may give the results with errors around \( 5 \times 10^{-3} \) when \( V \) is small, but the errors increase for larger \( V \). When \( V = 2 \), the error of SU is about \( 10^{-2} \). The GO may significantly improve the ground state energies. By using the given \( D \) and \( D_c \), we are able to obtain an impressively highly accurate ground state with relative error near \( \sim 10^{-5} \).

We now consider a special case of \( V = 0 \), where the model reduces to the free fermion model. Although in this case, the model is exactly solvable, it is a challenging model for the PEP method because the free fermions have strong entanglement in real space \( S \sim L^d-1 \log L \) that violates the area law.\(^{26} \) Especially at \( \mu = 0 \), the fermi surface is very large, making the problem more difficult. One may expect that to obtain the high accuracy results it requires very large \( D \) and \( D_c \). Furthermore, the required parameters \( D \) and \( D_c \) will generally increase rapidly with the increasing of the size of the system to keep the given accuracy. In Table II, we list the ground states of the free fermion model on the square lattice with different sizes obtained from the SU and GO methods, compared with the exact results \( E_{ex} \). We see that the relative errors of the SU are usually about \( 10^{-2} \) for \( D = 8 \), but sometimes the SU method may have numerical instability in some small systems. The performance of GO is much better, and we always get stable results. Even with a small \( D = 6 \), the relative errors are about \( 10^{-3} \sim 10^{-4} \), and reduce to \( \sim 10^{-5} \) when \( D = 8 \) is used. On the other hand, the violation of the area law is also showed in this table, that the accuracy gets lower in larger systems for a given \( D \).

From the above tests, we find that the SU method sometimes may give rather accurate results (\( \sim 10^{-2} \sim 10^{-3} \)), but the situation may change from case to case. On the other hand the GO always gives reliable and highly accurate results (\( \sim 10^{-5} \)).

| \( V \) | SU | GO | \( E_{ex} \) | relative error |
|-------|----|----|------------|----------------|
| 0.1   | -0.66590 | -0.67124 | -0.67125 | 1 \times 10^{-5} |
| 0.8   | -0.59255 | -0.59309 | -0.59312 | 5 \times 10^{-5} |
| 2     | -0.48136 | -0.50643 | -0.50646 | 5 \times 10^{-5} |

| size   | SU(D=8) | GO(D=6) | GO(D=8) | \( E_{ex} \) |
|--------|---------|---------|---------|-------------|
| 4 \times 4 | -0.68308 | -0.68401 | -0.68401 |
| 6 \times 6 | -0.73269 | -0.73305 | -0.73309 |
| 8 \times 8 | -0.75414 | -0.75492 | -0.75510 |
| 10 \times 10 | -0.76619 | -0.76705 | -0.76748 |
| 12 \times 12 | -0.77094 | -0.77538 |          |

**TABLE II: Compare the ground state energies of the free fermion model of SU (\( D = 8 \)), and GO (\( D = 6, 8 \)) with the exact results. For the \( 4 \times 4 \) lattice, the SU result is numerically instable for \( D = 8 \).**

**TABLE III: Compare the ground state energies of t-J model with hole filling \( \bar{n}_h = 0.125 \) calculated by SU and GO methods. A virtual bond dimension \( D = 12 \) is used.**

| size   | SU | GO |                |
|--------|----|----|----------------|
| 4 \times 4 | -0.55108 | -0.56420 |               |
| 4 \times 8 | -0.57994 | -0.59055 |               |
| 6 \times 8 | -0.59431 | -0.60349 |               |
| 8 \times 8 | -0.60849 | -0.61184 |               |
| 8 \times 10 | -0.61068 | -0.61738 |               |
| 8 \times 12 | -0.61707 | -0.62164 |               |
| 12 \times 12 | -0.62307 | -0.62973 |               |

**B. t-J model**

In this section, we benchmark our method on the t-J model,

\[
H = -t \sum_{\langle i,j \rangle, \sigma} (c_i^\dagger \sigma c_j \sigma + H.c.) + J \sum_{\langle i,j \rangle} (\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j)
\]  

(19)

where \( \sigma = \uparrow, \downarrow \) is the spin index and \( \vec{S}_i \) is the spin 1/2 operator on site \( i \). \( n_i = \sum_{\sigma} c_i^\dagger \sigma c_i \sigma \) is the number of electrons on site \( i \). In t-J model, the electron double occupancy is forbidden. The t-J model is one of the key models to understand many important physical phenomena,\(^{28} \) such as high \( T_c \) superconductivity.\(^1 \) Here, we calculate the model with \( J/t = 0.4 \) and hole filling of \( \bar{n}_h = 0.125 \). The U(1) symmetry is adopted to enforce the particle number conservation. The true physics of the system at this point, whether the ground state is a stripe state\(^{21,39-41} \) or an uniform phase,\(^{42,43} \) is still under debate. Without doubt, the energy is one of the critical criterions to determine the ground state of the system. We calculate the ground energies of different system sizes, using \( D = 12, D_c = 36 \sim 50 \), and the results are shown in Table III for both SU and GO methods. Again we see GO method greatly improves the energies obtained from SU method.
V. CONVERGENCE OF FPEPS

Fermion systems may have large entanglement that beyond the area law and therefore it may need large $D$ to represent the many-particle state. One may also expect that the $D$ and $D_c$ will increase with the size of the system. The speed of the increasing of $D$ and $D_c$ along with the size of the system indicates the efficiency of the simulation methods. It is important to understand how the fPEPS calculations converge respect to $D$ and $D_c$. For finite systems, we can explicitly exam what $D$ and $D_c$ are needed to converge the results as the size of the systems grow up. In this section, we will discuss the convergence of the parameters $D$ and $D_c$ respectively. We show that the behavior depends strongly on the models.

We first investigate the convergency of the ground state energies to $D_c$ in a given system with fixed parameter $D$. We calculate the error of energy defined as,

$$\Delta E = E(D_c) - E(D_c^{\text{max}}),$$  

(20)

where $E(D_c)$ is the energy with a given truncation parameter $D_c$, and $E(D_c^{\text{max}})$ is the converged energy where the maximal $D_c$ is used.

Figure 8(a) depicts the results of the spinless fermion model with $V = 0$ (free fermion) and $V = 2$; and different system sizes, the $4 \times 4$ and $10 \times 10$ lattices. We fix the bond dimension at $D = 8$. We use $D_c^{\text{max}} = 48$ for the $4 \times 4$ system and $D_c^{\text{max}} = 64$ for the $10 \times 10$ system. We first note that $\Delta E$s approach 0 in a non-trivial way, which are not always from above (i.e., $\Delta E > 0$). This means that the ground state energy is not variational to $D_c$, and therefore one must be very careful to extrapolate $D_c$ to infinite. The convergence of energy is model dependent. As shown in the figure, $D_c$ converge much faster for $V = 2$ (correlated electrons) than for $V = 0$ (free electrons). In both cases, the convergency of energy strongly depend on the size of the systems. In the cases of small sizes $L=4$, the energies converge rather fast with $D_c$. However, for the $10 \times 10$ system, $\Delta E$ converge much slower as functions of $D_c$. For $V = 2$, the energy is well converged at $D_c = 26$ (about 3D), whereas the energy of free fermions is not well converged even at $D_c = 48$.

We investigate the convergence of the t-J model at hole doping $\bar{n}_h = 0.125$, and the results are shown in Fig. 8(b). In the calculations, $D = 12$ is used, and the result of $D_c^{\text{max}} = 50$ is used as a reference. Interestingly, we find the ground state energies converge rather fast with $D_c$. The errors reduce to $3 \times 10^{-4}$ for $D_c = 2D$, and the errors reduce to $1 \times 10^{-5}$ for $D_c = 3D$. More importantly, unlike the interacting fermion model, $D_c$ is only slightly dependent on the size of the system.

The energy errors in the calculations are induced by the contraction errors due to bond dimension truncation. We further test the relationship between the convergent truncation dimension $D_c$ and the size of the system, i.e. we examine the minimal $D_c$ needed to ensure the relative contracting error $\delta < 10^{-6}$ (see Eq. 16) along with the increasing of the system size. The bond dimension used here is fixed to a relatively small one $D = 6$. We compare the truncation errors for the spinless interacting electron model at $V=0$ and $V=2$. For the t-J model, we compare two situations, the hole doping $\bar{n}_h = 0.125$, and the $\bar{n}_h = 0$, and the later one reduces to the Heisenberg model. The results are shown in Fig. 9(a). We find that the required $D_c$ is almost independent of the size of the system for the Heisenberg model, and for the t-J model with hole filling $\bar{n}_h = 0.125$. However, the required $D_c$ increases rapidly with the size of the system, especially for the free electrons. At $L=12$, $D_c=80$ is required to ensure the desired contraction accuracy for the free electron model and $D_c=40$ for the $V=2$ model.

We also examined the relationship between $D_c$ and the bond dimension $D$. In this test, we fixed the size of the system to $L=10$. The results are shown in Fig. 9(b). We see that the required $D_c$ increase roughly linearly with $D$ for these models. For the Heisenberg model and the t-J model with hole filling $\bar{n}_h = 0.125$, $D_c \sim 2D - 3D$ is enough to ensure the accuracy of contraction, whereas for the interacting electron model, $D_c \sim 9D - 15D$ are required to ensure the desired contraction accuracy, which becomes the major difficult to simulate these models. We note that in the standard contraction method, the bond truncation dimension $D_{\text{trunc}}$ for a double layer tensor network should scale as $D^2$ making the simulation of fermions with large $D$ even more difficult.

With the convergent $D_c$ for each $D$, we can analyze the convergence of the energy against $D$ for a given system.
FIG. 9: The bond truncation dimensions $D_c$ needed to ensure the contraction error $\delta < 10^{-6}$ (see Eq. 16) as functions of (a) the lattice size $L$ and (b) the virtual bond dimension $D$ for various models, including the spinless interacting fermion models, and the t-J model. The “t-J model” in the figure is calculated with parameters $J = 0.4$, $\bar{n}_h = 0.125$, whereas the “Heisenberg” model is calculated using t-J model in the limit $\bar{n}_h = 0$.

FIG. 10: The convergence of ground state energy as functions of $D$ for (a) the free fermion model, (b) the interaction electron model with $V = 2$, and (c) the t-J model with $\bar{n}_h = 0.125$, on the lattices of different sizes.

VI. SUMMARY

In this work, we extend the stochastic gradient optimization method combined with Monte Carlo sampling techniques to optimize the fPEPS wave functions for fermion systems. The Monte Carlo sampling techniques may greatly reduce the scaling of the calculation, and therefore allow using larger bond dimensions ($D$) and bond truncation dimensions ($D_c$) in the calculations, which is important for the faithful simulations of fermion systems.

We benchmark the method on the interacting spinless fermion models, and the t-J models. The numerical calculations show that the gradient optimization may greatly enhance the accuracy of the results over the simple update method. We further investigate the convergence of fPEPS calculation with respect to $D$ and $D_c$ for the models. The free fermion model is most challenging to simulate with fPEPS, because the $D_c$ increase very rapidly with $D$ and the size of the system. For t-J models, we find that large $D$s are needed to converge the results. Our method therefore offer a powerful tool to simulate fermion systems because it has much lower scaling in both computational time and memory than direct contraction methods.

VII. ACKNOWLEDGE

This work was funded by the National Key Research and Development Program of China (Grant No. 2016YFB0201202), the Chinese National Science Foun-
Appendix A: Rules for fermi arrows

In this Appendix, we give the rules of operations associated with the fermi arrows in fPEPS. These rules are straightforward to prove.

a. Reversing fermi arrows and the hermitian conjugate

Sometimes, we need to reverse the direction of a fermi arrow. The rule of reversing fermi arrows is giving as follows. Suppose,

\[ \hat{A} = \sum_{\beta_1,\beta_2} A_{\beta_1,\beta_2} a_{1}^{\beta_1} a_{2}^{\beta_2}, \]
\[ \hat{B} = \sum_{\beta_3,\beta_4} B_{\beta_3,\beta_4} a_{3}^{\beta_3} a_{4}^{\beta_4}, \]

are two projectors in a fPEPS that are connected by a fermi arrow pointing from \( A \) to \( B \), as shown on the left side of Fig. 11. We may reverse the fermi arrow, pointing from \( B \) to \( A \), and resulting in two possible (but equivalent) forms that are given on the right side of Fig. 11. It is easy to prove that,

\[ \hat{A}' = \sum_{\beta_1,\beta_2} \hat{p}(\beta_2) A_{\beta_1,\beta_2} a_{1}^{\beta_1} a_{2}^{\beta_2}, \]
\[ \hat{B}' = \sum_{\beta_3,\beta_4} \hat{p}(\beta_3) B_{\beta_3,\beta_4} a_{3}^{\beta_3} a_{4}^{\beta_4}. \] (A1)

When we calculate the expectation value of a physical quantity, \( \langle \Psi_{\text{fPEPS}} | O | \Psi_{\text{fPEPS}} \rangle \), we need take the hermitian conjugate of ket state \( | \Psi_{\text{fPEPS}} \rangle \) to get the bar state \( \langle \Psi_{\text{fPEPS}} | \). When taking the hermitian conjugate of the projectors in a fPEPS, we need to (i) reserve the orders of the indices of the tensor associated with the projectors, e.g., change tensor \( A_{i,d,r,u} \) to \( A_{i,d,r,u}^{\dagger} \), as shown in Fig. 12; and (ii) reverse all the fermi arrows associated with the projectors. Note that here the reversion of the fermi arrows is required by the hermitian conjugate, and no change is needed for the tensors during the process.

b. Matrix decompositions and contractions

The operations such as tensor decompositions also have close relation to the fermi arrows. For example, in the standard PEPS, when we do SVD to a matrix \( C \), we have \( C = U S V \). However, in fPEPS, two fermi arrows should be inserted into the inner bonds after the decomposition, i.e., the fermi arrow pointing from \( U \) to \( S \), and the one pointing from \( S \) to \( V \) as follows, and schematically shown in Fig. 13(a),

\[ \hat{C} = \hat{U} S \hat{V} I_{U \rightarrow S} \hat{I}_{S \rightarrow V}, \] (A2)

where

\[ \hat{C} = \sum_{\alpha,\beta} C_{\alpha,\beta} a_{\alpha}^{\dagger} a_{\beta}, \]
\[ \hat{U} = \sum_{\alpha,\delta} U_{\alpha,\delta} a_{\alpha}^{\dagger} a_{\delta}, \]
\[ \hat{S} = \sum_{\delta_1,\delta_2} S_{\delta_1,\delta_2} a_{\delta_1}^{\dagger} a_{\delta_2}, \]
\[ \hat{V} = \sum_{\gamma,\delta} V_{\gamma,\delta} a_{\gamma}^{\dagger} a_{\delta} \] (A3)

Other matrix decompositions such as LQ/QR decompositions follow the similar rules, i.e., one need to insert fermi arrows (i.e., directed EPR pairs) between the decomposed matrices, as shown in Fig. 13(b),(c).
to contract the

i.e.,

helx@ustc.edu.cn

form of MPS in the MPO algorithm\textsuperscript{34,35} to contract the PEPS, taking the advantage of the orthogonality of the \(Q\) tensors obtained from LQ/QR decompositions (or the \(U\) and \(V\) matrices from SVD decompositions),\textsuperscript{9} i.e., \(Q_i Q_i^\dagger = I\), where \(I\) is a unit matrix. However, this relation can’t be directly used in the PEPS, where we need to take the fermi arrows into consideration during the contractions. It is easily prove that only when the fermi arrows have “consistent directions”, i.e., all fermi arrows point from \(Q\) to \(Q_i\), or from \(Q_i^\dagger\) to \(Q\), we can use the orthogonality condition for \(Q\) matrix. The results after contraction are fermi arrows pointing to the right or to the left, as schematically shown in Fig. 14. If the fermi arrows are not “consistent”, we need to rearrange the directions of the fermi arrows first to make them “consistent”, before we can use the orthogonality condition. This is done in Sec.\textsuperscript{III}, when we contract two rows of PEPS via a MPO scheme.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig14.png}
\caption{The rule of contacting \(Q_i Q_i^\dagger\). To use the orthogonality \(Q_i Q_i^\dagger = I\), the fermi arrows must have “consistent directions” as shown above.}
\end{figure}

In standard PEPS, we often use so called canonical form of MPS in the MPO algorithm\textsuperscript{34,35} to contract the PEPS, taking the advantage of the orthogonality of the \(Q_i\) tensors obtained from LQ/QR decompositions (or the \(U\) and \(V\) matrices from SVD decompositions),\textsuperscript{9} i.e., \(Q_i Q_i^\dagger = I\), where \(I\) is a unit matrix. However, this relation can’t be directly used in the PEPS, where we need to take the fermi arrows into consideration during the contractions. It is easily prove that only when the fermi arrows have “consistent directions”, i.e., all fermi arrows point from \(Q\) to \(Q_i\), or from \(Q_i^\dagger\) to \(Q\), we can use the orthogonality condition for \(Q\) matrix. The results after contraction are fermi arrows pointing to the right or to the left, as schematically shown in Fig. 14. If the fermi arrows are not “consistent”, we need to rearrange the directions of the fermi arrows first to make them “consistent”, before we can use the orthogonality condition. This is done in Sec.\textsuperscript{III}, when we contract two rows of PEPS via a MPO scheme.

* Electronic address: smhan@ustc.edu.cn
† Electronic address: helx@ustc.edu.cn
1 P. A. Lee, N. Nagaosa, and X.-G. Wen, Rev. Mod. Phys. \textbf{78}, 17 (2006).
2 H. L. Stormer, D. C. Tsui, and A. C. Gossard, Rev. Mod. Phys. \textbf{71}, S298 (1999).
3 D. M. Edwards and A. C. Hewson, Rev. Mod. Phys. \textbf{40}, 810 (1968).
4 M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. \textbf{70}, 1039 (1998).
5 W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. \textbf{73}, 33 (2001).
6 Z.-X. Li and H. Yao, arXiv:1805:08219 (2018).
7 E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B \textbf{41}, 9301 (1990).
8 M. Troyer and U.-J. Wiese, Phys. Rev. Lett. \textbf{94}, 170201 (2005).
9 U. Schollwck, Annals of Physics \textbf{326}, 96 (2011).
10 D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Info. Comput. \textbf{7}, 401 (2007).
11 F. Verstraete, V. Murg, and J. Cirac, Advances in Physics \textbf{57}, 143 (2008).
12 H. C. Jiang, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. \textbf{101}, 090603 (2008).
13 G. Vidal, Phys. Rev. Lett. \textbf{101}, 110501 (2008).
14 F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066 (2004).
15 A. Sfondrini, J. Cerrillo, N. Schuch, and J. I. Cirac, Phys. Rev. B \textbf{81}, 214426 (2010).
16 F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. \textbf{96}, 220601 (2006).
17 L. Wang, Z.-C. Gu, F. Verstraete, and X.-G. Wen, Phys. Rev. B \textbf{94}, 075143 (2016).
18 G. Vidal, Phys. Rev. Lett. \textbf{98}, 070201 (2007).
19 A. M.-H., J. I. Cirac, and M. C. Bauls, New Journal of Physics \textbf{14}, 075003 (2012).
20 P. Corboz, R. Orús, B. Bauer, and G. Vidal, Phys. Rev. B \textbf{81}, 165104 (2010).
21 P. Corboz, S. R. White, G. Vidal, and M. Troyer, Phys. Rev. B \textbf{84}, 041108 (2011).
22 Z.-C. Gu, F. Verstraete, and X.-G. Wen, arXiv:1004.2563 (2010).
23 Z.-C. Gu, H.-C. Jiang, D. N. Sheng, H. Yao, L. Balents, and X.-G. Wen, Phys. Rev. B \textbf{88}, 155112 (2013).
24 C. V. Kraus, N. Schuch, F. Verstraete, and J. I. Cirac, Phys. Rev. A \textbf{81}, 052338 (2010).
25 M. B. Hastings, Journal of Statistical Mechanics: Theory and Experiment \textbf{2007}, P08024 (2007).
26 M. M. Wolf, Phys. Rev. Lett. \textbf{96}, 010404 (2006).
27 A. W. Sandvik and G. Vidal, Phys. Rev. Lett. \textbf{99}, 220602 (2007).
28 N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. \textbf{100}, 040501 (2008).
29 M. C. Baurils, K. Cichy, J. Ignacio Cirac, K. Jansen, S. Kühn, and H. Saito, in \textit{European Physical Journal Web of Conferences} (2017), vol. 137, p. 04001.
30 S.-J. Dong, W. Liu, X.-F. Zhou, G.-C. Guo, Z.-W. Zhou, Y.-J. Han, and L. He, Phys. Rev. B \textbf{96}, 045119 (2017).
31 W.-Y. Liu, S.-J. Dong, Y.-J. Han, G.-C. Guo, and L. He, Phys. Rev. B \textbf{95}, 195154 (2017).
32 L. He, H. An, C. Yang, F. Wang, J. Chen, C. Wang, W. Liang, S. Dong, Q. Sun, W. Han, et al., IEEE Transactions on Parallel and Distributed Systems \textbf{29}, 2838 (2018).
33 S.-J. Dong, W.-Y. Liu, C. Wang, Y. Han, G.-C. Guo, and L. He, Computer Physics Communications \textbf{228}, 163 (2018).
34 M. Lubasch, J. I. Cirac, and M.-C. Baurils, Phys. Rev. B \textbf{90}, 064425 (2014).
35 M. Lubasch, J. I. Cirac, and M.-C. Bauls, New Journal of Physics \textbf{16}, 033014 (2014).
36 L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, Phys. Rev. B \textbf{94}, 155123 (2016).
37 J. de Woul and E. Langmann, Journal of Statistical Physics \textbf{139}, 1033 (2010).
38 F. C. Zhang and T. M. Rice, Phys. Rev. B \textbf{37}, 3759 (1988).
39 C. S. Hellberg and E. Manousakis, Phys. Rev. Lett. \textbf{83}, 132 (1999).
40 S. R. White and D. J. Scalapino, Phys. Rev. Lett. \textbf{80}, 1272 (1998).
41 P. Corboz, T. M. Rice, and M. Troyer, Phys. Rev. Lett. \textbf{113}, 046402 (2014).
42 A. Sherman and M. Schreiber, European Physical Journal B \textbf{32}, 203 (2003).
A. Vineet Mallik, G. K. Gupta, V. B. Shenoy, and H. R. Krishnamurthy, arXiv:1805.02429 (2018).