Journal of Physics:
Condensed Matter

PAPER

On the fly swapping algorithm for ordering of degrees of freedom in density matrix renormalization group

To cite this article: Weitang Li *et al* 2022 *J. Phys.: Condens. Matter* **34** 254003

IOP Publishing Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **34** (2022) 254003 (15pp) https://doi.org/10.1088/1361-648X/ac640e

On the fly swapping algorithm for ordering of degrees of freedom in density matrix renormalization group

Weitang Li, Jiajun Ren, Hengrui Yang and Zhigang Shuai*

MOE Key Laboratory of Organic OptoElectronics and Molecular Engineering, Department of Chemistry, Tsinghua University, Beijing, People's Republic of China

E-mail: zgshuai@tsinghua.edu.cn

Received 23 January 2022, revised 22 March 2022
Accepted for publication 4 April 2022
Published 22 April 2022

Abstract

Density matrix renormalization group (DMRG) and its time-dependent variants have found widespread applications in quantum chemistry, including *ab initio* electronic structure of complex bio-molecules, spectroscopy for molecular aggregates, and charge transport in bulk organic semiconductors. The underlying wavefunction ansatz for DMRG, matrix product state (MPS), requires mapping degrees of freedom (DOF) into a one-dimensional topology. DOF ordering becomes a crucial factor for DMRG accuracy. In this work, we propose swapping neighboring DOFs during the DMRG sweeps for DOF ordering, which we term 'on the fly swapping' (OFS) algorithm. We show that OFS is universal for both static and time-dependent DMRG with minimum computational overhead. Examples are given for one dimensional antiferromagnetic Heisenberg model, *ab initio* electronic structure of N_2 molecule, and the S_1/S_2 internal conversion dynamics of pyrazine molecule. It is found that OFS can indeed improve accuracy by finding better DOF ordering in all cases.

Keywords: density matrix renormalization group, ordering, degree of freedom, quantum chemistry, quantum dynamics, vibronic model

(Some figures may appear in colour only in the online journal)

1. Introduction

Density matrix renormalization group (DMRG) is originally proposed for strongly correlated one dimensional lattice model by White [1, 2]. It was quickly generalized to quantum chemistry with semiempirical long-range Coulomb potential demonstrating surprisingly high accuracy by Shuai, Ramasesha and Fano *et al* [3–5]. Then DMRG for *ab initio* quantum chemistry was later explored in 1999 by White and Martin [6]. After two decades of development, DMRG has become one of the mainstream methods in quantum chemistry [7–9] thanks to the efforts made by Chan, Reiher, Kurashige and Ma *et al* [10–13]. The success of DMRG can be ascribed to its underlying ansatz, namely matrix product state (MPS), which satisfies the area law [14]. Apart from static DMRG that targets the ground state or low-lying excited states, time-dependent DMRG (TD-DMRG) [15–19] has also received much attention in recent years [20–23]. The projector-splitting (PS) time evolution scheme [24], based on time-dependent variational principle (TDVP) [25, 26], has enabled accurate simulation of a variety of chemical properties, such as the spectra of molecular aggregates [27], singlet fission dynamics in molecular dimers [28], and charge mobility of organic semiconductors [29, 30]. However, numerically exact DMRG simulation is still quite time-consuming, even with massive parallelism over CPU and GPU [31–34]. How to simulate larger systems with limited computational resources is one of the central topics of DMRG studies [35, 36].

The MPS ansatz requires an ordering of the degree of freedoms (DOFs) of the system and the DOF ordering is a determining factor for the efficiency of DMRG. For one dimensional lattice models, the optimal DOF ordering for MPS

* Author to whom any correspondence should be addressed.

is according to the geometry of the chain structure. If the bipartition of the MPS chain corresponds to a bipartition of the one dimensional lattice, then according to the area law the MPS ansatz is able to express the ground state of any gapped lattice Hamiltonian with polynomial number of parameters. Even for the long-range Coulomb potential with density-density type terms such as Pariser–Parr–Pople Hamiltonian with only nearest-neighbor hopping integrals, Shuai *et al* demonstrated that DMRG is almost as accurate as the short-ranged Hubbard model [3]. Nevertheless, the application of DMRG to general models with nonlocal interaction, such as momentum-space DMRG [37] or DMRG in two-dimensional lattice [38], brings about the problem of DOF ordering. Recently, tree tensor networks appear as useful extensions of DMRG, and the problem of finding an optimal tree topology is analog to the optimal DOF ordering problem in DMRG [39–42]. Intuitively, the strongly entangled DOFs should be arranged as close as possible. In the context of quantum chemistry, the DOF ordering problem is especially prominent with the molecular orbitals as DOFs. There have been extensive investigations along this direction [43–49]. Roughly speaking, the existing strategies can be classified into two categories. The first category is to derive DOF ordering according to the single-electron or two-electron integrals that appear in the *ab initio* electronic Hamiltonian. The second category is to order DOFs according to the information extracted from trial DMRG calculations. In both categories, a common strategy is to firstly establish the entanglement between DOFs and then use reversed Cuthill–Mckee (RCM) algorithm [43], the Fiedler vector [47, 48], genetic algorithm (GA) [48] or simulated annealing [46] to group strongly entangled DOFs in the DOF ordering. For the two categories, the exchange integral and the mutual information are the most frequently used metrics to characterize the entanglement respectively. For the second category, it is also possible to directly optimize DOF ordering using GAs taking DMRG optimized energy as the evolution criteria [45]. The computational overhead for the first category is negligibly small and they are used routinely in DMRG calculations [7, 9, 50, 51]. Yet the resulting orderings are usually not as good as the orderings by the second category methods [45]. Consensus has been reached in the community that finding the globally optimal ordering for a general Hamiltonian is quite difficult. The search space for the DOF ordering scales as $\mathcal{O}(N!)$ where N is the total number of DOFs and in general cases it is highly nontrivial to judge whether a given DOF ordering is better than another one.

To date, there is hardly any effort on the DOF ordering for TD-DMRG quantum dynamics, in particular the vibronic problem. In vibronic problem the size of the basis set varies with DOFs. Phonon DOFs that have low frequency and are strongly coupled with other DOFs require larger basis set. It is recently suggested by Ma *et al* that the size of the local basis set should also be taken into account for optimal DOF ordering, along with the entanglement considerations [28]. For efficient two-site TDVP-PS time evolution, it should be avoided to arrange phonon DOFs with large basis set next to each other. Such complication makes finding an optimal DOF ordering even more challenging.

In this work, we propose an algorithm, on the fly swapping (OFS), for DOF ordering optimization. OFS complements the existing two categories of DOF ordering strategies mentioned above because it can exploit the information retrieved during DMRG optimization and at the same time spares the computational overhead. OFS can also take the effect of the basis set sizes into consideration. The key idea behind OFS is to swap neighboring DOFs during the DMRG sweeping process according to a certain loss function, such as the bipartite von Neumann entanglement entropy. OFS can be applied to any models and is compatible with both static DMRG and TD-DMRG. We demonstrate OFS can help find optimized DOF ordering in a variety of examples, including spin-$\frac{1}{2}$ one dimensional antiferromagnetic Heisenberg model, *ab initio* electronic structure of N_2 at dissociation bond length, and the S_1/S_2 internal conversion dynamics for the pyrazine molecule. The examples are versatile in nature and demonstrate the proposed OFS is a general method for DOF ordering optimization.

2. The on the fly swapping (OFS) algorithm

2.1. Brief overview of matrix product states

The MPSs and matrix product operators represent the wavefunction of many-body system $|\Psi\rangle$ with overall N DOFs as well as corresponding operators $\hat{O}$ as the product of a series of matrices [1, 14]

$$|\Psi\rangle = \sum_{\{a\},\{\sigma\}} A[1]^{\sigma_1}_{a_1} A[2]^{\sigma_2}_{a_1 a_2} \dots A[N]^{\sigma_N}_{a_{N-1}} |\sigma_1 \sigma_2 \dots \sigma_N\rangle, \quad (1)$$

$$O = \sum_{\{w\},\{\sigma\},\{\sigma'\}} W[1]^{\sigma'_1,\sigma_1}_{w_1} W[2]^{\sigma'_2,\sigma_2}_{w_1 w_2} \dots \times W[N]^{\sigma'_N,\sigma_N}_{w_{N-1}} |\sigma'_1 \sigma'_2 \dots \sigma'_N\rangle\langle\sigma_N \sigma_{N-1} \dots \sigma_1|. \quad (2)$$

$|\sigma_i\rangle$ is the basis for each DOF. $A[i]^{\sigma_i}_{a_{i-1}a_i}$ and $W[i]^{\sigma'_i,\sigma_i}_{w_{i-1}w_i}$ are numerical matrices in the chain. Here i in the square bracket is the label for the DOF, and the ordering of the A and W in the MPS/MPO representation is implied by the sequence of multiplication. The dimension of a_i and w_i is called (virtual) bond dimension, while the dimension of σ_i is called the size of the physical index. For MPS and MPO, the virtual bond dimension is denoted as M_S and M_O respectively, and the the size of the physical index is denoted as d. For certain applications of MPS such as *ab initio* electronic structure it is essential to exploit the sparsity of W. In this work, we do not consider such complication for a unified treatment of all models.

The MPS parametrization for a given state is not unique, and in practice mixed/left/right-canonical MPS are commonly used. A mixed-canonical MPS with the gauge center at the nth site can be written as:

$$|\Psi\rangle = \sum_{\{l\},\{r\}\{\sigma\}} L[1]^{\sigma_1}_{l_1} \dots L[n-1]^{\sigma_{n-1}}_{l_{n-2}l_{n-1}} C[n]^{\sigma_n}_{l_{n-1}r_n} \times R[n+1]^{\sigma_{n+1}}_{r_n r_{n+1}} \dots R[N]^{\sigma_N}_{r_{N-1}} |\sigma_1\sigma_2\dots\sigma_N\rangle \quad (3)$$

where $L^{\sigma_i}_{l_{i-1}l_i}$ and $R^{\sigma_j}_{r_{j-1}r_j}$ satisfy orthonormal condition:

$$\sum_{\sigma_i,l_{i-1}} L^{\sigma_i *}_{l_{i-1}l'_i} L^{\sigma_i}_{l_{i-1}l_i} = \delta_{l'_i l_i}, \quad (4)$$

$$\sum_{\sigma_j,r_j} R^{\sigma_j *}_{r'_{j-1}r_j} R^{\sigma_j}_{r_{j-1}r_j} = \delta_{r'_{j-1}r_{j-1}}. \quad (5)$$

Here the superscript stars represent complex conjugation.

For brevity, in this article we will not outline the detailed algorithm for static ground state DMRG and TD-DMRG. Rather, we only focus on a common sub-routine of the algorithms that are relevant to OFS, which is the decomposition of a two-site tensor. Readers are referred to several excellent reviews [14, 20, 21, 52] for more context. We note that the OFS framework can also be generalized to dynamical DMRG in frequency domain [53, 54] quite straightforwardly, which we will not discuss in detail in this work.

In a left-to-right sweep of two-site static ground state DMRG, the gauge center is moved from 1 to $N-1$, and the local tensors are updated along the way. Suppose during one sweep the gauge center is at the ith site, in order to update $C[i]$ and $R[i+1]$, the computationally heavy step is to solve the eigenvalue problem:

$$\sum_{l_{i-1},\sigma_i,\sigma_{i+1},r_{i+1}} H[i,i+1]^{\sigma'_i\sigma_i,\sigma'_{i+1}\sigma_{i+1}}_{l'_{i-1}l_{i-1},r'_{i+1}r_{i+1}} C[i,i+1]^{\sigma_i,\sigma_{i+1}}_{l_{i-1}r_{i+1}} = EC[i,i+1]^{\sigma'_i,\sigma'_{i+1}}_{l'_{i-1}r'_{i+1}} \quad (6)$$

where $H[i,i+1]$, the effective Hamiltonian, is defined as:

$$H[i,i+1]^{\sigma'_i\sigma_i,\sigma'_{i+1}\sigma_{i+1}}_{l'_{i-1}l_{i-1},r'_{i+1}r_{i+1}} = \sum_{\{w\}} h[1:i-1]_{\{l',w,l\}_{i-1}} W[i]^{\sigma'_i,\sigma_i}_{w_{i-1}w_i} \times W[i+1]^{\sigma'_{i+1},\sigma_{i+1}}_{w_i w_{i+1}} h[i+2:N]_{\{r',w,r\}_{i+1}}, \quad (7)$$

where

$$h[1:i-1]_{\{l',w,l\}_{i-1}} = \sum_{\{l'\},\{w\},\{l\}} h[1]_{\{l',w,l\}_1} \dots \times h[i-1]_{\{l',w,l\}_{i-2},\{l',w,l\}_{i-1}}, \quad (8)$$

$$h[i+2:N]_{\{r',w,r\}_{i+1}} = \sum_{\{r'\},\{w\},\{r\}} \times h[i+2]_{\{r',w,r\}_{i+1},\{r',w,r\}_{i+2}} \dots \times h[N]_{\{r',w,r\}_{N-1}}, \quad (9)$$

$$h[i]_{\{a',w,a\}_{i-1},\{a',w,a\}_i} = \sum_{\sigma_i,\sigma'_i} A[i]^{\sigma'_i *}_{a'_{i-1}a'_i} W[i]^{\sigma'_i,\sigma_i}_{w_{i-1}w_i} A[i]^{\sigma_i}_{a_{i-1}a_i}$$

$$(A = L \text{ or } R,\ a = l \text{ or } r).$$

Then, the resulting four-legged tensor $C[i,i+1]$ is decomposed by SVD

$$C[i,i+1]^{\sigma_i,\sigma_{i+1}}_{l_{i-1}r_{i+1}} = \sum_{j,j'} U_{l_{i-1}\sigma_i,j} s_{jj'} V_{j',\sigma_{i+1}r_{i+1}} \quad (11)$$

where $s_{jj'} = s_j \delta_{jj'}$ are the singular values and satisfy $\sum_j s_j^2 = 1$ for normalized $C[i,i+1]$. We assume s_j is sorted such that $s_j \geqslant s_{j+1}$. Generally speaking, the SVD decomposition has two purposes. The first is to move the gauge center from the ith site to the $(i+1)$th site, by setting $L[i]^{\sigma_i}_{l_{i-1}l_i} = U_{l_{i-1}\sigma_i,j}$ and $C[i+1]^{\sigma_{i+1}}_{l_i r_{i+1}} = \sum_{j'} s_{jj'} V_{j',\sigma_{i+1}r_{i+1}}$ with appropriate reshaping. And the second purpose is to compress the MPS wavefunction by truncating the singular values according to certain criteria. Suppose the compression strategy is to retain only the largest M_S singular values, then after moving the gauge center to the $(i+1)$th site the bond dimension between the ith site and the $(i+1)$th site is simply M_S.

The algorithm for two-site TDVP-PS time evolution shares a lot in common with the two-site static DMRG algorithm. In a left-to-right sweep, suppose the gauge center is at the ith site, the forward time evolution for $C[i]$ and $R[i+1]$ is carried out according to

$$i\frac{\partial C[i,i+1]^{\sigma'_i,\sigma'_{i+1}}_{l'_{i-1}r'_{i+1}}}{\partial t} = \sum_{l_{i-1},\sigma_i,\sigma_{i+1},r_{i+1}} H[i,i+1]^{\sigma'_i\sigma_i,\sigma'_{i+1}\sigma_{i+1}}_{l'_{i-1}l_{i-1},r'_{i+1}r_{i+1}} \times C[i,i+1]^{\sigma_i,\sigma_{i+1}}_{l_{i-1}r_{i+1}} \quad (12)$$

where the initial value for $C[i,i+1]$ is $\sum_{r_i} C[i]^{\sigma_i}_{l_{i-1}r_i} R[i+1]^{\sigma_{i+1}}_{r_i r_{i+1}}$. Then, similar to static DMRG algorithm, the updated $C[i,i+1]$ is decomposed by SVD to move the gauge center and compress the wavefunction. We note that different from the static DMRG algorithm, in TDVP-PS it is necessary to carry out 'backward time evolution', however, this step of the TDVP-PS algorithm is irrelevant to OFS, at least in the two-site case. We would like to mention here that there exists a number of different TD-DMRG time evolution schemes and some of them are quite different from the two-site TDVP-PS algorithm [21], yet the TDVP-PS algorithm seems to be the most popular choice among recent studies [27–31, 55, 56].

In both static ground state two-site DMRG algorithm and the two-site TDVP-PS time evolution algorithm, one of the key steps is to perform SVD over the updated two-site tensor $C[i,i+1]$, specified in equation (11). In order to prevent the

computational cost of the DMRG algorithms from exponential explosion, it is compulsory to compress the MPS wavefunction according to the singular values s_j, based on either predefined bond dimension or truncation threshold, so that only the largest singular values are reserved. In the following we assume s_j is truncated to the largest M_S singular values, while generalization to other truncation schemes is straightforward. If the MPS wavefunction before the truncation is $|\Psi\rangle$ and the MPS wavefunction after the truncation is $|\Phi\rangle$, then the error induced by the truncation can be measured by

$$\|\Phi - \Psi\|^2 = 1 - \sum_{j=1}^{M_\mathrm{S}} s_j^2 = \sum_{j=M_\mathrm{S}+1}^{N_S} s_j^2 \tag{13}$$

where $N_S = M_\mathrm{S} d$ is the number of singular values. The error introduced in this step, albeit controllable, is arguably the most important source of error for two-site static DMRG algorithm and TDVP-PS time evolution algorithm. The efficiency for the approximation is closely related to the bipartite von Neumann entanglement entropy

$$S = -\sum_j s_j^2 \ln s_j^2. \tag{14}$$

If S is large, s_j tend to be uniformly distributed, resulting in relatively large truncation error. If S is small, then s_j tends to have an unbalanced distribution, resulting in a relatively small truncation error. In general, a good DOF ordering that minimizes S reduces the truncation error, while a bad DOF ordering with large S is harmful to accuracy. In order to reduce S the most entangled DOFs should be arranged closely in the linear chain, however, it is not straightforward to implement this philosophy for general models, where there's little physical intuition on which DOFs are most strongly entangled.

2.2. On the fly swapping (OFS) for optimized DOF ordering

The key idea behind OFS is the observation that the bipartite entanglement entropy between the ith site and the $(i+1)$th site usually changes upon swapping the DOF for the ith site and the DOF for the $(i+1)$th site. For example, consider the following state in the basis of $|l_{i-1}\sigma_i\sigma_{i+1}r_{i+1}\rangle$ for the $C[i,i+1]$ tensor:

$$|\Psi\rangle = \frac{1}{2}\left(|0000\rangle + |1010\rangle + |0101\rangle + |1111\rangle\right). \tag{15}$$

In this case, we have $s_j = \frac{1}{2}$ for $1 \leqslant j \leqslant 4$ and $s_j = 0$ otherwise, and thus $S = 2\ln 2$. If the DOFs on the ith site and the $(i+1)$th site are allowed to be swapped, such that $|\Psi\rangle$ is expressed in the $|l_{i-1}\sigma_{i+1}\sigma_i r_{i+1}\rangle$ basis, then

$$|\Psi\rangle = \frac{1}{2}\left(|00\rangle + |11\rangle\right) \otimes \left(|00\rangle + |11\rangle\right) \tag{16}$$

and the corresponding bipartite entanglement entropy S between the ith site and the $(i+1)$th site is zero. To exploit this possibility of reducing entanglement entropy, we propose to firstly permute the indices of $C[i,i+1]$

$$C[i+1,i]^{\sigma_{i+1},\sigma_i}_{l_{i-1}r_{i+1}} \equiv \mathrm{PermuteIndex}\left(C[i,i+1]^{\sigma_i,\sigma_{i+1}}_{l_{i-1}r_{i+1}}\right) \tag{17}$$

and then carry out SVD for $C[i+1,i]$

$$C[i+1,i]^{\sigma_{i+1},\sigma_i}_{l_{i-1}r_{i+1}} = \sum_{j,j'} U'_{l_{i-1}\sigma_{i+1},j} s'_{jj'} V'_{j',\sigma_i r_{i+1}} \tag{18}$$

in addition to the decomposition in the traditional workflow equation (11). The bipartite entanglement entropy S for the two sets of singular values s_j and s'_j are denoted as S and S' respectively. If $S' < S$, then swapping the DOFs results in a lower entanglement entropy, and the DOF ordering after the swapping is more favoured compared to the original ordering. The tensor at the ith site $L[i]$ is then set to $U'_{l_{i-1}\sigma_{i+1},j}$ (again taking left-to-right sweep as an example), and now it represents the DOF originally at the $(i+1)$th site. Meanwhile, the tensor at the $(i+1)$th site $C[i+1]$ is set to $\sum_{j'} s'_{jj'} V'_{j',\sigma_i r_{i+1}}$. The resulting MPS has the DOFs at the ith site and the $(i+1)$th site swapped, with improved accuracy after truncated to M_S largest singular values compared to the original DOF ordering. The same swapping should be performed to Hamiltonian MPO before the iterative MPS algorithm proceeds to the next site. An efficient and accurate DOF swapping algorithm for MPO will be introduced in section 2.3. If $S \leqslant S'$, then the original DOF ordering is fine and nothing needs to be done. If the SVD truncation is performed according to singular value threshold instead of pre-defined M_S, so that M_S is adjusted dynamically, then OFS should be able to find the DOF ordering that results in a lower bond dimension. One of the main features of the proposed algorithm is that the DOF ordering optimization is taken place during the sweeps, thus we term this algorithm as OFS. Apparently, the computational scaling for OFS does not exceed the overall scaling for vanilla static DMRG and TD-DMRG. Moreover, the overhead caused by an additional SVD decomposition is small for reasonably large M_O, since the computational bottleneck resides in the evaluation equations (6) and (12), which scale as $\mathcal{O}(M_\mathrm{S}^3 M_\mathrm{O} d^2 + M_\mathrm{S}^2 M_\mathrm{O}^2 d^3)$ with a prefactor for the number of iterations, while the SVD of $C[i,i+1]$ scales at $\mathcal{O}(M_\mathrm{S}^3 d^3)$. Furthermore, the two SVD decompositions over $C[i,i+1]$ and $C[i+1,i]$ can be performed in perfect parallel. In this regard, the computational overhead is negligible. We note that it is only meaningful to carry out the swapping on the fly, because after the truncation the system wavefunction is actually biased toward the current DOF ordering, and at this stage the relative entanglement before and after swapping is not reliable.

The swapping criteria based on bipartite entanglement entropy S is straight-forward, however, S is only 'closely related to' the error induced by the truncation. The exact measurement for the error is the sum of the discarded singular values D, as indicated by equation (13):

$$D \equiv \sum_{j=M_\mathrm{S}+1}^{N_S} s_j^2. \tag{19}$$

We note that $S < S'$ does not guarantee $D < D'$ for all possible values of M_S. The bipartite entanglement S before and after the swapping is not a function of the truncation bond dimension M_S. Meanwhile, the truncation error D before and after the swapping is dominated by M_S. From this point of view,

D can be regarded as a variant of S that is tailored for the current truncation bond dimension M_S. Thus, both S and D can serve as indicators to decide if it is worthwhile to swap the DOFs at the ith site and the $(i+1)$th site. Other possible alternatives include Rényi entropies [57]. More generally, we can design loss function $\mathcal{L}$ as the swapping criteria, and swap the DOFs when $\mathcal{L}' < \mathcal{L}$. $\mathcal{L}$ may not only include information from the entanglement point of view, but also include other aspects of DOF ordering, such as the distribution of the sizes of the physical index d, the MPO bond dimension M_O, or even computational overhead. For example, static DMRG for vibrational models [58] may employ the following strategy to take the effect of physical basis size into consideration. During the first a few rounds of sweeps, when the bond dimension M_S is typically small, use the entanglement entropy as the swapping criteria. And before moving on to the large M_S limit, add the product of the sizes of the physical index of adjacent sites into the loss function such that the large M_S limit calculation is affordable. In this work, we focus on reducing entanglement entropy. In figure 1 we present a schematic illustration of the OFS approach.

If $\mathcal{L}$ is set to S, we denote this approach as OFS-S:

$$\mathcal{L}_{\mathrm{OFS-S}} = S. \tag{20}$$

If $\mathcal{L}$ is set to D, we denote this approach as OFS-D:

$$\mathcal{L}_{\mathrm{OFS-D}} = D. \tag{21}$$

Although $\mathcal{L}_{\mathrm{OFS-D}}$ seems to measure the truncation error accurately, it does not work well at the sites that are close to the boundary of the MPS. At these sites, because the Hilbert space of one side of the environment is actually rather small, there's virtually no truncation error regardless of how local DOFs are ordered. On the contrary, S is still able to correctly predict the ordering with smaller entanglement entropy in this case. To deal with this situation, we propose a hybrid loss function:

$$\mathcal{L}_{\mathrm{OFS-D/S}} = \begin{matrix} S & \text{if } D < \xi \text{ and } D' < \xi \\ D & \text{otherwise} \end{matrix} \tag{22}$$

that tries to combine the advantages of both $\mathcal{L}_{\mathrm{OFS-D}}$ and $\mathcal{L}_{\mathrm{OFS-S}}$. Here ξ is a pre-defined small threshold and for the results reported in this article we set $\xi = 10^{-10}$.

It is instructive to compare the OFS approach with on the fly local basis optimization [35]. The authors propose to apply a unitary transformation U over the physical indices to $C[i, i+1]$ during the static DMRG sweeping process in order to reduce entanglement between DOFs:

$$C[i,i+1]^{\sigma_i',\sigma_{i+1}'}_{l_{i-1}r_{i+1}} = \sum_{\sigma_i,\sigma_{i+1}} U^{\sigma_i',\sigma_{i+1}'}_{\sigma_i,\sigma_{i+1}} C[i,i+1]^{\sigma_i,\sigma_{i+1}}_{l_{i-1}r_{i+1}}. \tag{23}$$

The unitary transformation U acts as a 'disentangler'. Formally, the OFS approach introduced here can be viewed as a special case for the local basis optimization algorithm, where U is restricted to be a swap gate U_{swap}. However, the simplicity of OFS has earned it at least three advantages against the local basis optimization approach: (1) swapping the MPS and the MPO has a lower scaling and smaller computational cost than finding and applying an optimal U; (2) the overall ordering after several iterations is more interpretable than the optimized basis encoded in a set of U; (3) the design of $\mathcal{L}$ is more flexible.

Generally, $\mathcal{L}$ has a lot of local minimums, and a similar problem is also reported for the local basis optimization approach [35]. Thus, it is highly unlikely that OFS is able to find out the globally best DOF ordering given a moderate initial guess. Yet, the OFS approach is orthogonal to existing DOF ordering schemes and has minimal computational overhead. Therefore, we can always expect OFS to improve the DOF ordering, in combination with other DOF ordering methods or simply physical intuition, almost for free. One of the most promising combinations is to integrate OFS with the GA that takes DMRG optimized energy as the evolution criteria [45]. OFS should speed up the convergence of GAs and the crossovers/mutations in GAs can help to avoid local minimums in OFS.

2.3. Symbolic MPO DOF swapping

Swapping the DOF ordering in MPS requires the DOF ordering in the corresponding Hamiltonian MPO to be swapped accordingly:

$$W'[i+1]W'[i] = W[i+1,i] = U_{\mathrm{swap}}\,(W[i]W[i+1])\,U_{\mathrm{swap}}. \tag{24}$$

Note that we use the index in the square brackets to label DOFs rather than the order of the DOFs in the MPO representation, while the latter (the order of the DOFs) is implied by the order of the multiplication. For general MPO, the swap gates will increase the MPO bond dimension, and computationally expensive numerical compression is required to maintain the computational scaling. Besides, as we shall discuss in section 3.2, sometimes it is required to change the 'Hamiltonian' on the fly, which cannot be accomplished straightforwardly by equation (24).

Another naïve solution is to reconstruct the whole MPO once the DOF ordering is updated. Because we do not pose any constraints to the model, to the best of our knowledge the automatic MPO construction algorithm [59] is the optimal choice for MPO construction under arbitrary DOF ordering. This approach is inefficient in two ways. First, reconstructing MPO takes time, especially when the MPO is rather complicated, such as *ab initio* Hamiltonian. Second, the local W tensors might have changed for the sites that are not swapped, since the optimal MPO representation for any operator is usually not unique. The change of W means previously calculated $h[1:i]$ and $h[i:N]$ become out-dated and should also be re-calculated, which is unfavorable in DMRG algorithms.

Here we propose an efficient and accurate approach to the MPO DOF swapping problem using the idea of symbolic MPO. Following the notation in reference [59], the symbolic MPO representation of any operators in a sum-of-products (SOP) form can be written as

$$\hat{O} = \sum_{\{z\}} \gamma_{z_1 z_2 \ldots z_N} \hat{z}_1 \hat{z}_2 \ldots \hat{z}_N \tag{25}$$

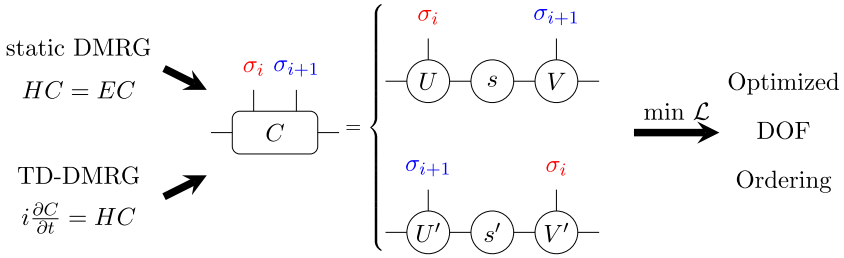

Figure 1. Schematic illustration of the OFS approach. During the sweeping iteration in both static DMRG and TD-DMRG it is necessary to decompose the coefficient tensor $C[i, i+1]$ by SVD. We propose to perform an additional SVD decomposition after swapping the DOFs then decide if such swapping should be adopted by minimizing the loss function $\mathcal{L}$. Then the DMRG algorithm sweeps to the next site for another iteration.

$$= \sum_{\{w\},\{z\}} W[1]^{z_1}_{w_1} W[2]^{z_2}_{w_1 w_2} \dots W[N]^{z_N}_{w_{N-1}} \hat{z}_1 \hat{z}_2 \dots \hat{z}_N \quad (26)$$

$$= \sum_{\{w\}} \hat{W}[1]_{w_1} \hat{W}[2]_{w_1 w_2} \dots \hat{W}[N]_{w_{N-1}}, \quad (27)$$

where $\{\hat{z}_i\}$ represents the elementary operators of each local DOF and $\gamma_{z_1 z_2 \dots z_N}$ is the prefactor for each term, commonly very sparse. This form of operator is similar to equation (2) except that the matrix elements for $\hat{W}[i]_{w_{i-1} w_i}$ are symbolic operators while $W[i]^{\sigma_i', \sigma_i}_{w_{i-1} w_i}$ is a numeric tensor.

Imagine the MPO chain is cut between the $(i-1)$th and the ith site, or equivalently at bond w_{i-1}, the operator $\hat{O}$ can be represented by operators assembled at the left block and the right block of the MPO after the cutting

$$\hat{O} = \sum_{w_{i-1}} \hat{W}[1 : i-1]_{w_{i-1}} \hat{W}[i : N]_{w_{i-1}} \quad (28)$$

with the left/right block operator

$$\hat{W}[1 : i-1]_{w_{i-1}} = \sum_{\{w_j | 1 \leqslant j < i-1\}} \hat{W}[1]_{w_1} \dots \hat{W}[i-1]_{w_{i-2} w_{i-1}} \quad (29)$$

$$\hat{W}[i : N]_{w_{i-1}} = \sum_{\{w_j | i \leqslant j < N\}} \hat{W}[i]_{w_{i-1} w_i} \dots \hat{W}[N]_{w_{N-1}}. \quad (30)$$

It is evident from the above definition that $\hat{O}$ can also be written as:

$$\hat{O} = \sum_{w_{i-1}, w_i, w_{i+1}} \hat{W}[1 : i-1]_{w_{i-1}} \hat{W}[i]_{w_{i-1} w_i} \times \hat{W}[i+1]_{w_i w_{i+1}} \hat{W}[i+2 : N]_{w_{i+1}}. \quad (31)$$

For brevity, we denote the set of operators in the left/right block when cutting between the $(i-1)$th site and the ith site as

$$\mathcal{W}_{i-1} = \{\hat{W}[1 : i-1]_{w_{i-1}}\}, \quad (32)$$

$${}_{i-1}\mathcal{W} = \{\hat{W}[i : N]_{w_{i-1}}\}. \quad (33)$$

Suppose we wish to swap DOF i and $i+1$, located at site i and $i+1$, we actually aim to determine the local symbolic matrices at site i and site $i+1$ after the swapping, denoted as $\hat{Q}[i+1]_{w_{i-1} w_i}$ and $\hat{Q}[i]_{w_i w_{i+1}}$ respectively, with $\mathcal{W}_{i-1}$ and ${}_{i+1}\mathcal{W}$ already known from the initial optimal MPO construction:

$$\hat{O} = \sum_{w_{i-1}, w_i, w_{i+1}} \hat{W}[1 : i-1]_{w_{i-1}} \hat{Q}[i+1]_{w_{i-1} w_i} \times \hat{Q}[i]_{w_i w_{i+1}} \hat{W}[i+2 : N]_{w_{i+1}}. \quad (34)$$

$\hat{Q}[i+1]$ and $\hat{Q}[i]$ are required to possess minimal $|w_i|$, the dimension of w_i.

There are many possible approaches to this constrained optimization problem, and perhaps the most efficient and easy-to-implement algorithm is to interface with the existing optimal MPO construction algorithm [59] and reuse the code. Toward this goal, equation (31) is partially expanded into the SOP form:

$$\hat{O} = \sum_{z_L, z_i, z_{i+1}, z_R} \gamma_{z_L z_i z_{i+1} z_R} \hat{z}_L \hat{z}_i \hat{z}_{i+1} \hat{z}_R \quad (35)$$

with $\{\hat{z}_L\} = \mathcal{W}_{i-1}$ and $\{\hat{z}_R\} = {}_{i+1}\mathcal{W}$. In other words, the assembled operators $\mathcal{W}_{i-1}$ and ${}_{i+1}\mathcal{W}$ are recognized as elementary operators for imaginary sites positioned at $i-1$ and $i+2$. Then, the ith DOF and the $(i+1)$th DOF in the SOP representation are swapped

$$\hat{O} = \sum_{z_L, z_i, z_{i+1}, z_R} \gamma_{z_L z_i z_{i+1} z_R} \hat{z}_L \hat{z}_{i+1} \hat{z}_i \hat{z}_R \quad (36)$$

where we have assumed, as in the case of MPOs, the operators for different DOFs commute, and we will discuss the non-commuting case for fermions in section 3.2. Equation (36) can be efficiently transformed into a four-site MPO with minimal bond dimension in exactly the same form as equation (34) using the existing optimal MPO construction algorithm. $\hat{Q}[i+1]$ and $\hat{Q}[i]$ correspond to the two sites at the middle of the four-site MPO. Additional trivial matrix permutation might be required to ensure the indices w_{i-1} and w_{i+1} for $\hat{Q}[i+1]_{w_{i-1} w_i}$ and $\hat{Q}[i]_{w_i w_{i+1}}$ are in the desired order. In the worst case scenario, this symbolic MPO DOF swapping algorithm scales at $\mathcal{O}(M_O^{\frac{5}{2}})$ using the Hopcroft–Karp algorithm [60], negligible for DMRG algorithms.

3. Experiments and discussion

In this section, we present our numerical simulation results to evaluate the performance of the OFS schemes. All simulations are performed using the Renormalizer package [61]. In our implementation particle number conservation is explicitly enforced. More specifically, we track the quantum number of each renormalized basis as well as the physical basis and SVD is performed within each symmetry block. Meanwhile, the MPS and MPO are stored in the dense form. The OFS algorithm proposed in this work fully respects the particle number conservation symmetry. The SVD decompositions before and after the swapping are carried out in the same fashion. The symbolic MPO swapping algorithm ensures that good quantum number can be assigned to each normal and complementary operator based on the quantum number of each elementary operator. For non-Abelian groups such as SU(2) how to perform OFS deserves further investigation.

3.1. One dimensional Heisenberg model

We first test the OFS algorithm in a spin-$\frac{1}{2}$ one dimensional antiferromagnetic Heisenberg model:

$$\hat{H}/J = \frac{1}{4}\sum_i [\hat{X}_i\hat{X}_{i+1} + \hat{Y}_i\hat{Y}_{i+1} + \hat{Z}_i\hat{Z}_{i+1}]. \qquad (37)$$

Here $\hat{X}$, $\hat{Y}$ and $\hat{Z}$ are Pauli matrices and $J > 0$ is the coupling constant. The one dimensional Heisenberg model is one of the initial applications of DMRG [2]. With moderate computational cost the ground state energy can be obtained at machine precision. The ordering problem for this model is trivial, in that the optimal ordering is equal to the spin index or its reversion. This kind of one dimensional spin lattice model is suitable for evaluating DOF ordering algorithms since the globally optimal solution is known. In the following we simulate the Heisenberg model with 32 spin sites and open boundary condition, for which the exact E/J with Bethe ansatz is $-13.997\,3156$. The DMRG sweeps are performed for 25 rounds starting from random initial guesses with the warm-up algorithm [62].

For a given bond dimension M_S, we randomly generate 64 different DOF orderings, and then compare the lowest energy during the sweeps with or without OFS. Note that when comparing different ordering schemes both the initial DOF ordering and the wavefunction initial guess are the same. The results are summarized in figure 2. The error bar indicates the standard deviation of the E/J distribution. For reference, converged DMRG energy with optimal ordering, which can be considered as the numerically exact result, are also included in figure 2. For example, when $M_S = 32$, E/J is determined to be $-13.997\,3153$, in close agreement with the exact solution. With M_S from 16 to 64, randomly ordered DOF yields far inferior E/J than the optimal ordering, and convergence with respect to M_S is slow, implying the importance of DOF ordering in these systems. The OFS schemes are able to ameliorate the problem. Notably, E/J obtained with OFS schemes at $M_S = 16$ is already more accurate than the $M_S = 64$ results without OFS. Nevertheless, E/J obtained with OFS schemes is still nowhere close to the exact solution even at $M_S = 64$.

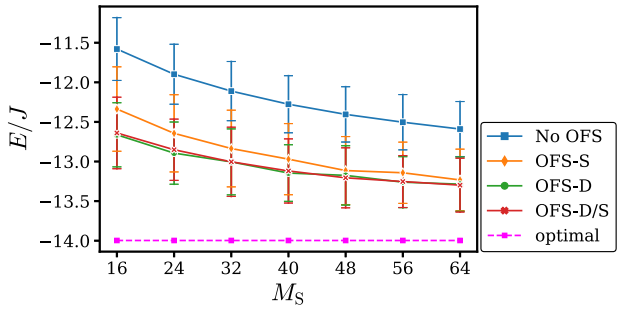

Figure 2. DMRG optimized energy E/J starting from random orderings with or without OFS ordering optimization at different bond dimension M_S. For reference E/J obtained with optimal ordering is also shown, which is very close to the exact result. The error bar is the standard deviation of E/J distribution.

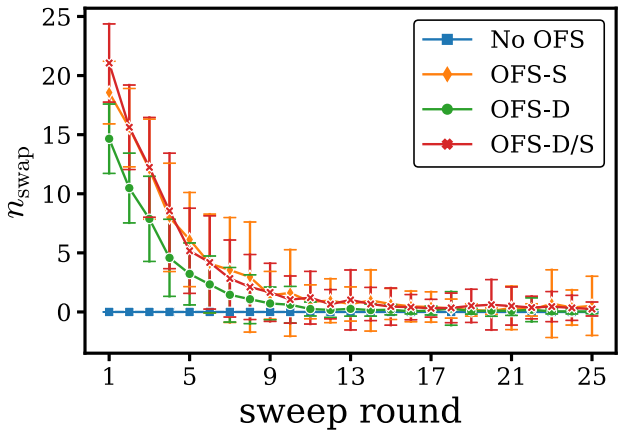

Figure 3. The number of OFS swaps n_{swap} performed in each DMRG sweep for the OFS schemes. The error bar is the standard deviation of the number of OFS swaps.

All three OFS schemes are sensitive to the initial ordering guess, leading to quite a significant standard deviation. Manual inspection of the ordering optimized by OFS shows that the arrange of the DOFs is somewhat 'clustered' by the algorithm yet the overall ordering is far from optimal. Several examples of the optimized ordering are included in appendix A. Within the OFS schemes, it is found that OFS-D and OFS-D/S perform better than OFS-S by a moderate margin. In figure 3 we show the number of OFS swaps (n_{swap}) performed in each DMRG sweep. It is found that after around 10 sweeps the number of OFS swaps is approaching zero. So we can conclude that the OFS orderings are trapped in local minimums. Within OFS schemes OFS-D performs less swap, as OFS-D is not able to optimize DOF ordering at the boundary of the MPS chain.

3.2. Ab initio electronic structure of N_2

The second-quantized Hamiltonian for general chemical system reads:

$$\hat{H} = \sum_{ij} t_{ij}\hat{c}_i^\dagger\hat{c}_j + \sum_{ijkl} v_{ijkl}\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l. \qquad (38)$$

i, j, k, l are indices for spin-orbital. t_{ij} is the single-electron integral and v_{ijkl} is the two-electron integral. In contrast to the

Heisenberg model discussed in section 3.1, the optimal DOF ordering for the *ab initio* Hamiltonian equation (38) cannot be read out directly. One of the feasible approaches in determining a good ordering is to arrange strongly correlated DOFs as close as possible, which is determined semi-empirically by the integrals t_{ij} and v_{ijkl}. In this regard, t_{ij} is a suitable metric since it is a symmetric matrix with exactly two indices. The four-indexed two-body integrals need to be somewhat compressed before the interaction pattern can be extracted. Two possible candidates for the compression are the Coulomb integral $J_{ij} = (ii|jj)$ and exchange integral $K_{ij} = (ij|ji)$. Here i and j are indices for spatial-orbital. Suppose the matrix chosen as the interaction metric is M_{ij}, a lot of methods can be used to derive a DOF ordering from M_{ij}. The RCM approach firstly converts M_{ij} to binary matrices according to a certain threshold and then solves the matrix bandwidth reduction problem using the reversed Cuthill–Mckee algorithm [43, 63]. Another approach is to use the idea of Fiedler vector [64] to avoid information loss during the thresholding process [47, 48]. The spirit is to recognize orbital ordering or ranking as continuous one dimensional variables x_i and use the Fiedler vector to minimize loss function $\sum_{ij}(x_i - x_j)^2|M_{ij}|$. Loss functions in more general form can be optimized using GA. In practice, one popular choice for M_{ij} is K_{ij} [45, 48, 65]. Another matrix that is commonly chosen as M_{ij} is the mutual information matrix I_{ij} [44, 46, 47]:

$$I_{ij} = \frac{1}{2}\left(S^i + S^j - S^{ij}\right)\left(1 - \delta_{ij}\right) \tag{39}$$

where $S^{\mathrm{sys}} = -\mathrm{Tr}\,\rho^{\mathrm{sys}} \ln \rho^{\mathrm{sys}}$ and ρ^{sys} is the reduced density matrix of the system. The calculation of I_{ij} requires the knowledge of the system wavefunction, which is typically obtained by trial DMRG calculation.

In order to maintain the anti-commutation relation of fermionic creation and annihilation operators, $c_i^\dagger$ and c_i are transformed to Pauli matrices via Jordan–Wigner transformation:

$$\begin{aligned}\hat{c}_i^\dagger &= \hat{+}_i \prod_{j<i} \hat{Z}_j \\ \hat{c}_i &= \hat{-}_i \prod_{j<i} \hat{Z}_j\end{aligned} \tag{40}$$

where we have used $\hat{+}_i = \frac{1}{2}(\hat{X}_i - i\hat{Y}_i)$ and $\hat{-}_i = \frac{1}{2}(\hat{X}_i + i\hat{Y}_i)$ to represent spin creation/annihilation operators. During the OFS sweeping, it is necessary to re-perform the Jordan–Wigner transformation according to DOF ordering in order to keep M_{O} of the MPO constant. The reason is that equation (40) assumes a natural ordering of the DOFs, and the MPO representation is more compact if the DOF ordering is the same as the ordering implied by the transformation. An illustrative example of this point is shown in appendix B. Thus, if the ordering of the Jordan–Wigner transformation remains unchanged during OFS iteration, the bond dimension of the MPO obtained by symbolic MPO DOF swapping algorithm presented in section 2.3 will continue to grow. To address this issue, it is mandatory to update the Jordan–Wigner mapping once DOF swapping takes place. Suppose the ith DOF and the $(i+1)$th DOF are to be swapped, and the Jordan–Wigner transformation is carried out according to equation (40), then the updated Jordan–Wigner transformation for $\hat{c}_i^\dagger(\hat{c}_i)$ and $\hat{c}_{i+1}^\dagger(\hat{c}_{i+1})$ reads:

$$\begin{aligned}\hat{c}_i^\dagger &= \hat{+}_i \hat{Z}_{i+1} \prod_{j<i} \hat{Z}_j \\ \hat{c}_{i+1}^\dagger &= \hat{+}_{i+1} \prod_{j<i} \hat{Z}_j\end{aligned} \tag{41}$$

while other operators remain unchanged. Equivalently, the following mapping (together with its Hermite conjugation) should be applied to the previously Jordan–Wigner transformed Hamiltonian:

$$\begin{aligned}\hat{+}_i &\to \hat{+}_i \hat{Z}_{i+1} \\ \hat{+}_{i+1} &\to \hat{Z}_i \hat{+}_{i+1}.\end{aligned} \tag{42}$$

Before updating the Jordan–Wigner transformation, the system wavefunction stored in the MPS should also be updated accordingly. More specifically, the definition for the states $|\ldots 11 \ldots\rangle = \hat{c}_i^\dagger \hat{c}_{i+1}^\dagger |\ldots 00 \ldots\rangle$ has changed due to the DOF swapping for Jordan–Wigner transformation. Thus, the coefficient matrix $C[i+1, i]$ should be updated before the decomposition according to

$$C[i+1, i]^{1,1}_{l_{i-1} r_{i+1}} \to -C[i+1, i]^{1,1}_{l_{i-1} r_{i+1}} \tag{43}$$

to keep the Jordan–Wigner DOF ordering for the state and for the operator synchronized.

To test the performance of OFS for the *ab initio* Hamiltonian we choose N_2 with cc-pVDZ basis set [66] as benchmark platform. The bond length is set to 1.905 Å, which is close to dissociation and induces strong correlation. In our calculations the 1*s* orbitals are frozen. The integrals are calculated using the PySCF package [67]. FCI energy with 1*s* orbital frozen, corresponding to a (10e, 26o) active space, is available for this particular system [68], which is $E_{\mathrm{FCI}} = -108.9948$ Hartree. Meanwhile, the RHF energy is $E_{\mathrm{RHF}} = -108.38478$ Hartree. Highly accurate DMRG with bond dimension M_{S} up to 4000 yields an error of 0.02 mH [69]. We compare OFS with six different traditional ordering strategies, labeled as energy, symmetry, RCM-*K*, Fiedler-*K*, GA-*K* and GA-*I*. The details of the strategies are listed below.

(a) Energy. Sort orbitals based on their RHF energy;
(b) Symmetry. Sort orbitals based on their symmetry and energy. The irreducible representations are sorted as $A_g, B_{1u}, B_{2u}, B_{3u}, B_{2g}, B_{3g}, B_{1g}, A_u$. Within the same irreducible representation, the orbitals are sorted according to energy;
(c) RCM-*K*. The ordering obtained by solving the band reduction problem of the K matrix using the RCM algorithm;
(d) Fiedler-*K*. The ordering given by the solution of minimizing $\sum_{ij}(x_i - x_j)^2|K_{ij}|$ using the Fiedler vector approach;
(e) GA-*K*. The ordering given by the solution of minimizing $\sum_{ij}(i-j)^2|K_{ij}|$ using GA [70];

Table 1. The effect of OFS on the DMRG calculated correlation energy $E - E_{\mathrm{RHF}}$ starting from different initial orderings. The unit is Hartree. The most accurate result for each bond dimension is shown in bold. The exact correlation energy by $E_{\mathrm{FCI}} - E_{\mathrm{RHF}}$ is −0.6100.

		M_{S}			
Initial ordering	OFS type	64	128	256	512
Energy	No OFS	−0.495	−0.543	−0.5804	−0.599 57
	OFS-S	−0.522	−0.549	−0.5892	−0.600 10
	OFS-D	−0.525	−0.554	−0.5828	−0.600 40
	OFS-D/S	−0.526	−0.546	−0.5876	−0.600 43
Symmetry	No OFS	−0.541	−0.582	−0.6003	−0.606 86
	OFS-S	−0.554	−0.586	−0.6003	−0.607 29
	OFS-D	−0.564	−0.589	−0.6028	−0.607 62
	OFS-D/S	−0.565	−0.587	**−0.6035**	−0.607 00
RCM-*K*	No OFS	−0.543	−0.584	−0.6005	−0.606 89
	OFS-S	−0.562	−0.591	−0.6017	−0.606 84
	OFS-D	−0.546	**−0.593**	−0.6030	**−0.608 44**
	OFS-D/S	**−0.585**	−0.579	**−0.6035**	−0.607 88
Fiedler-*K*	No OFS	−0.435	−0.557	−0.5854	−0.600 16
	OFS-S	−0.473	−0.488	−0.4937	−0.601 54
	OFS-D	−0.435	−0.569	−0.5925	−0.600 30
	OFS-D/S	−0.436	−0.555	−0.5852	−0.603 59
GA-*K*	No OFS	−0.520	−0.548	−0.5905	−0.601 88
	OFS-S	−0.523	−0.576	−0.5919	−0.602 34
	OFS-D	−0.520	−0.549	−0.5893	−0.602 12
	OFS-D/S	−0.521	−0.549	−0.5905	−0.601 75
GA-*I*	No OFS	−0.556	−0.584	−0.5983	−0.604 96
	OFS-S	−0.546	−0.587	−0.5989	−0.604 84
	OFS-D	−0.550	−0.586	−0.5984	−0.605 08
	OFS-D/S	−0.556	−0.586	−0.5995	−0.605 06

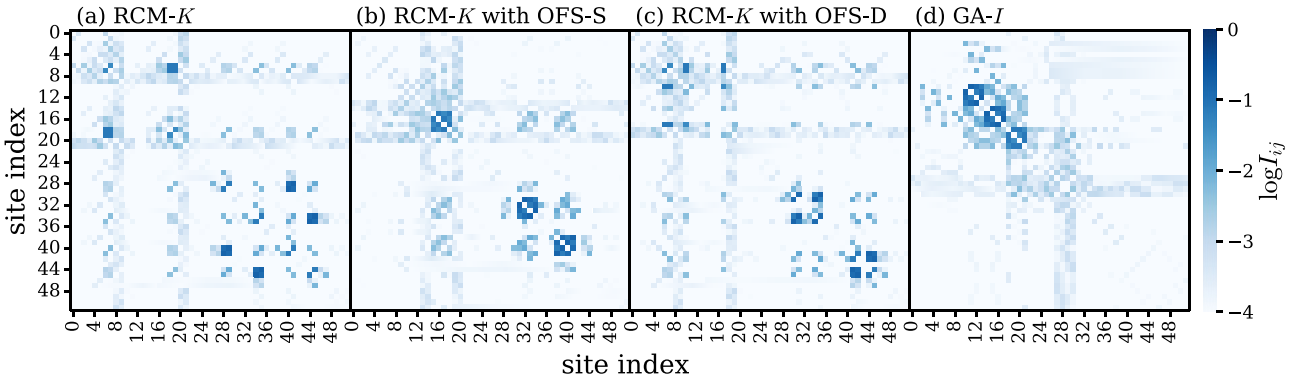

Figure 4. Mutual information I_{ij} from the DMRG optimized wavefunction. The DOF orderings are (a) RCM-*K*, (b) OFS-S optimized ordering starting from RCM-*K* initial ordering, (c) OFS-D optimized ordering starting from RCM-*K* initial ordering and (d) GA-*I*. The specific DOF orderings are included in appendix A.

(f) GA-*I*. The ordering given by the solution of minimizing $\sum_{ij}(i-j)^2|I_{ij}|$ using GA [70]. I_{ij} is the mutual information matrix calculated using DMRG with energy ordering and $M_{\mathrm{S}} = 512$.

The initial ordering of OFS is set to the orderings derived from the strategies, and we test if OFS is able to improve the ordering. All the initial orderings are included in appendix B. The largest bond dimension M_{S} employed in our calculation is 512, which is typically smaller than practical DMRG calculations for *ab initio* quantum chemistry. The reason we adopt such a small bond dimension is two-fold. The first is reducing computational cost and the second is making the effect of different orderings more prominent. The DMRG sweeps are performed for 80 rounds with the warm-up algorithm [62]. The wavefunction initial guess is set to the RHF determinant in all cases.

In table 1 we show the DMRG calculated correlation energy $E - E_{\mathrm{RHF}}$ with or without OFS. The exact correlation energy by $E_{\mathrm{FCI}} - E_{\mathrm{RHF}}$ is −0.6100. In most cases enabling OFS will lead to lower energy and better DMRG accuracy. The lowest energy is obtained by OFS optimized ordering using

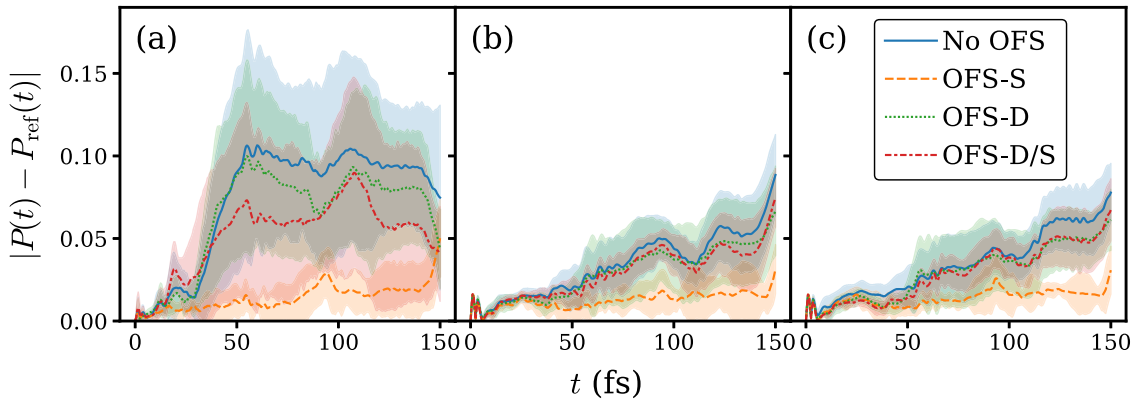

Figure 5. The time evolution error of the S_2 diabatic state population $P(t)$ with or without the OFS algorithm for the S_1/S_2 internal conversion of pyrazine, starting from (a) random ordering, (b) $\{v9a, v6a, |S_1, S_2\rangle, v10a, v1\}$ plus random bath mode ordering and (c) $\{|S_1, S_2\rangle, v6a, v10a, v1, v9a\}$ plus random bath mode ordering. The shades represent the standard deviation of the error $|P(t) - P_{\mathrm{ref}}(t)|$. The bond dimension M_S is set to 16 and the reference population $P_{\mathrm{ref}}(t)$ is obtained with bond dimension 128.

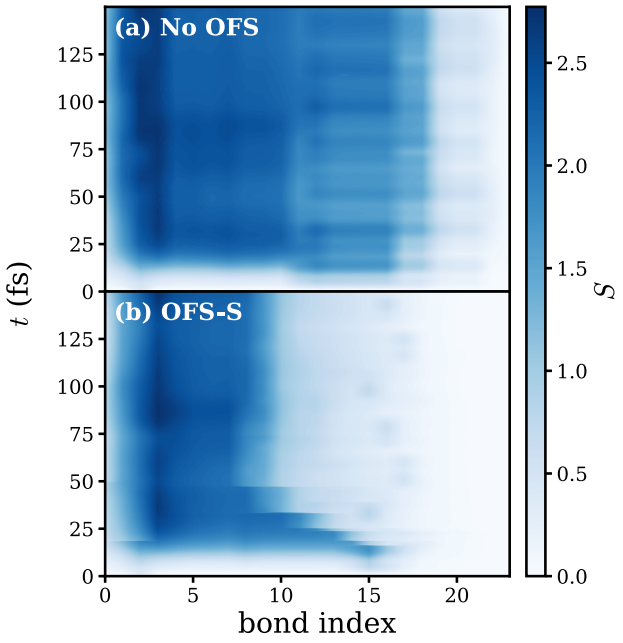

Figure 6. Time evolution of the bipartite entanglement entropy S at each virtual bond for the S_1/S_2 internal conversion of pyrazine with (a) OFS disabled and (b) OFS-S enabled. The initial DOF ordering groups strongly correlated DOFs at the left side of the MPS chain in the same pattern as figure 5(b). Details of the initial DOF ordering as well as the optimized DOF ordering are presented in appendix A. The bond dimension M_S is set to 16.

RCM-*K* as the initial ordering. In particular, when OFS-D is used in combination with the RCM-*K* initial ordering, more than 99.7% of the correlation energy can be captured. Generally speaking, OFS-D and OFS-D/S yield slightly better DOF ordering than OFS-S, in agreement with our observation in figure 2. It is a surprising result that the GA-*I* ordering is not as accurate as the RCM-*K* ordering or the symmetry ordering. Three possible reasons are the inaccuracy of *I*, the local minimums in DMRG optimization and the defects in the loss function used in the GA (for example, the punishment over the distance should not be quadratic). For the GA-*I* initial ordering the improvement by OFS is not significant. The reason is probably that the initial ordering given by GA-*I* is close to an OFS local minimum. To prove this, we calculated the Kendall rank correlation coefficient τ between the initial ordering and the OFS optimized ordering. τ measures the similarity between two rankings. $\tau = 0$ for uncorrelated DOF orderings and $\tau = 1$ for identical DOF orderings. If GA-*I* is used as the initial ordering, the correlation coefficient averaged over the bond dimension is $\tau_{\mathrm{GA}-I} = 0.91$. For all other initial DOF orderings, the averaged correlation coefficient is $\tau_{\mathrm{others}} = 0.82$. $\tau_{\mathrm{GA}-I} > \tau_{\mathrm{others}}$ means that the similarity between the GA-*I* ordering and its OFS optimized ordering is bigger than the similarities between other initial orderings and their OFS optimized orderings. Particularly inaccurate DMRG energies are obtained for Fiedler-*K*/OFS-S with $M_S = 128$ or $M_S = 256$ and the culprit is probably the local minimums present in DMRG optimizations.

In figure 4 we illustrate I_{ij} of the optimized wavefunction by several different orderings. The bond dimension M_S is set to 512. By comparing figures 4(b) and (c) with figure 4(a) we can see that OFS indeed reduces the distance between strongly entangled DOFs. Although figure 4(b) seems to be a better ordering than figure 4(c), in table 1 the energy calculated using the ordering shown in figure 4(c) is more accurate than that of figure 4(b). For comparison we also include the I_{ij} by the GA-*I* ordering in figure 4(d). As expected, GA-*I* puts large I_{ij} elements around the diagonal of the matrix most effectively. However, from table 1 its DMRG optimized energy is higher than the other three cases. Details of the optimized ordering are included in appendix A.

3.3. *Pyrazine* S_1/S_2 *internal conversion dynamics*

Next, we test the OFS algorithm in the time evolution problem and choose the S_1/S_2 internal conversion dynamics of pyrazine after UV photoexcitation to the S_2 state as the benchmark platform. This model is firstly studied in detail by multiconfiguration time-dependent Hartree (MCTDH) [71–74]. Pyrazine features a conical intersection between the S_1 and the S_2 states as well as strong vibronic linear and quadratic couplings with

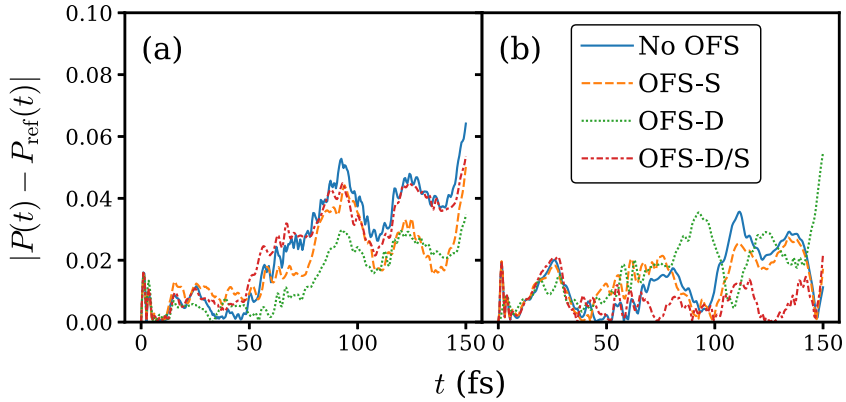

Figure 7. The time evolution error of the S_2 diabatic state population $P(t)$ for the S_1/S_2 internal conversion of pyrazine, with initial ordering derived from (a) ML-MCTDH tree and (b) OFS-S optimized ordering. The specific DOF orderings are included in appendix A.

24 vibration modes:

$$\hat{H} = \begin{pmatrix} -\Delta & 0 \\ 0 & \Delta \end{pmatrix} + \sum_{i=1}^{24} \frac{\omega_i}{2} \left(-\frac{\partial^2}{\partial q_i^2} + q_i^2 \right) + \sum_{i=1}^{24} q_i \begin{pmatrix} g_{i,11} & g_{i,12} \\ g_{i,12} & g_{i,22} \end{pmatrix} + \sum_{i,j=1}^{24} q_i q_j \begin{pmatrix} g'_{ij,11} & g'_{ij,12} \\ g'_{ij,12} & g'_{ij,22} \end{pmatrix} \tag{44}$$

where the matrix element indices refer to the diabatic potential energy surfaces for S_1 or S_2 state, 2Δ is the energy difference between S_1/S_2 at the ground state equilibrium geometry, q_i is the coordinate of the ith ground state normal mode, ω_i is the vibration frequency of the ith vibration mode, and $g(g')$ is linear(quadratic) coupling constants. All parameters are adopted directly from the references [73, 74].

The absorption spectra $I(\omega)$ can be calculated via the Fourier transformation of dipole–dipole correlation function:

$$I(\omega) \propto \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\omega t - t/\Gamma} \langle \hat{\mu}(t)\hat{\mu}(0) \rangle_{\mathrm{ground}} \, \mathrm{d}t \tag{45}$$

Here $\hat{\mu}$ is the dipole operator and the subscript 'ground' denotes the ground state. Γ accounts for a homogeneous broadening of the spectrum. At zero temperature, the calculation of the correlation function reduces to the calculation of the overlap between the initial vertical excitation S_2 state wavefunction $\Psi(0)$ and the time evolved wavefunction $\Psi(t)$:

$$\langle \hat{\mu}(t)\hat{\mu}(0) \rangle_{\mathrm{ground}} \propto \langle \Psi(0) | \Psi(t) \rangle. \tag{46}$$

Various TD-DMRG algorithms have been applied to the system [28, 75, 76]. Four core vibration modes that exhibit strong vibronic coupling can be identified and they should be arranged close to the electronic DOF [28]. Nevertheless, it is difficult to further optimize the DOF ordering from physical intuition because of the complicated interaction pattern.

Figure 5(a) shows the effect of OFS for the internal conversion dynamics of pyrazine starting from random initial DOF ordering by comparing the time evolution error on the population of the S_2 diabatic state $P(t)$. The bond dimension M_S is set to 16 and the time evolution step is set to 0.5 fs. Again, the M_S used here is much smaller than what is required for numerically exact DMRG calculation. The reference data $P_{\mathrm{ref}}(t)$ is obtained with TD-DMRG and the bond dimension M_S is set to 128. The shades represent the standard deviation of the $|P(t) - P_{\mathrm{ref}}(t)|$ for seven different initial DOF orderings. Figure 5(a) indicates that the OFS-S scheme significantly reduces the time evolution error. The OFS-D and the OFS-D/S scheme are able to reduce the error, but only marginally. In figures 5(b) and (c) we use a different strategy to generate the initial ordering. More specifically, following the suggestion by Ma *et al* [28], we put the electronic DOF and the vibrational DOFs that are most strongly coupled to the electronic DOF at one side of the chain, with the order of $\{v9a, v6a, |S_1, S_2\rangle, v10a, v1\}$ for (b) and $\{|S_1, S_2\rangle, v6a, v10a, v1, v9a\}$ for (c). To test OFS, the 20 bath modes are put at the other side of the chain randomly instead of by ω_i in the literature [28]. The symbols of the modes are the same as literature [28, 73]. The general trends reflected by figures 5(b) and (c) are quite similar. When OFS is disabled, we find that this ordering indeed leads to a smaller overall error compared to random DOF ordering. Meanwhile, the OFS algorithms are able to further reduce the error and the lowest error is observed for the OFS-S scheme. We note that here the effect of the size of the physical indices d is not taken into account, and whether DOFs with large d are arranged at neighboring sites can be a crucial factor for the overall computational cost, [28] in addition to the entanglement considerations.

In figure 6(a) we show one example of the evolution of bipartite entanglement entropy S without OFS together with S with OFS in figure 6(b). The initial ordering is generated in the same way as in figure 5(b) and the specific form of the initial ordering as well as the optimized ordering are listed in appendix A. In both figures 6(a) and (b) the DOFs that are most strongly entangled reside in the left end of the chain, which correspond to $\{v9a, v6a, |S_1, S_2\rangle, v10a, v1\}$. In figure 6(a), the randomly generated ordering for the bath modes spreads the entanglement over a large fraction of the MPS chain. If the ordering is allowed to be optimized, then the spreading of the entanglement over the MPS chain is effectively reduced over the time evolution iteration, as shown in figure 6(b). From physical intuition, the best DOF ordering should place the most strongly entangled DOFs at the center of the chain. Such global optimum is not found by OFS because moving DOFs at the

right side of the chain to the left side one site at a time will temporarily separate the most entangled DOFs and increase entanglement. We note that the maximum entanglement entropy S in both figures 6(a) and (b) is approaching $\ln M_S$, so a larger bond dimension is required for practical computation.

Lastly we compare the DOF ordering used in the ML-MCTDH literature [77] with the OFS-S optimized ordering used in figure 6(b). The DOF mapping from a ML-MCTDH tree to a MPS chain is not unique, here we take a simple approach which is to read out the vibrational DOFs in the ML-MCTDH tree shown in figure 9 of the reference [77] from left to right. The specific form of the ordering is included in appendix A. In the figure 7 we compare the time evolution error starting from the ordering derived from ML-MCTDH tree and an ordering optimized by OFS-S. When OFS is disabled, by comparing figures 7(a) and (b) we can see that using OFS-S optimized ordering as the fixed ordering yields more accurate population than the population calculated from the ML-MCTDH ordering. Furthermore, figure 7(a) shows that OFS is able to optimize the ordering derived from the ML-MCTDH tree and increase accuracy. For the ordering that has already been optimized by OFS-S, as the entanglement entropy evolves with time, OFS still tries to find better ordering during the time evolution. Yet, as shown in figure 7(b) the overall improvement is not particularly visible especially for OFS-S.

4. Conclusion

In this work, we have proposed a general framework termed OFS to optimize the DOF ordering in MPS, which features on the fly swapping of neighboring MPS sites during DMRG sweeps. The OFS framework is agnostic to the underlying model Hamiltonian and can be applied in static ground state DMRG, TD-DMRG or even dynamical DMRG in the frequency domain. Besides, OFS can be straightforwardly generalized to other tensor network structures [41, 78]. The criteria for whether to perform the swapping in each iteration is flexible and three schemes are suggested. The OFS-S scheme minimizes the bipartite entanglement entropy S, and the OFS-D scheme minimizes the truncation error D. A hybrid scheme named OFS-D/S is also designed as an attempt to combine the advantage of both schemes. We propose an efficient DOF swapping algorithm for MPO in order to keep the DOF ordering of MPO adapted to that of MPS during the OFS iterations. We argue that the overall computational cost for the OFS scheme is negligible for practical models.

The OFS schemes are tested in the one dimensional Heisenberg model, the *ab initio* electronic structure of the N_2 molecule at the dissociation bond length and the internal conversion dynamics of pyrazine. The three cases demonstrate that the OFS schemes are able to find a better DOF ordering compared to the initial ordering and improve the accuracy of DMRG simulation. For the N_2 case and the pyrazine case, we present evidence that OFS effectively moves DOFs that are strongly entangled toward each other. Nevertheless, simulation results based on the Heisenberg model indicate that the DOF ordering optimized by OFS may still lay quite far from the globally optimal ordering. The convergence to the locally optimized ordering is achieved in approximately 10 rounds of sweeps and we expect for larger systems OFS will also quickly converges to a local minimum. For the pyrazine case, the OFS-S scheme yields better overall DOF ordering than the other two schemes. For the Heisenberg model and the N_2 case the DOF ordering by the OFS-S scheme is not as good as the other two schemes, but still acceptable. Thus in general we recommend the OFS-S scheme for practical calculation. Whether it is possible to find a loss function $\mathcal{L}$ that constantly produces better DOF ordering than other schemes requires additional investigation.

We note that OFS is not designed to 'outperform' any existing DOF ordering methods. Rather, we suggest closely integrating OFS with other DOF ordering schemes so that OFS can further improve the DOF ordering, with almost no additional computational cost. A particularly interesting combination is OFS and the GA that takes DMRG optimized energy as the evolution criteria [45]. We believe that OFS can help speed up the convergence of the GA and in the context of GA the presence of local minimums in OFS is not a significant drawback.

Conflict of interest

The authors declare that there are no competing interests.

Acknowledgments

This work is dedicated to Prof. Enrico Clementi (1931–2021), a Giant in computational chemistry, best known for his 'global simulation approach' for natural system through Dirac – Schr ö dinger – Newton – Langevin – Navier – Stokes equations starting with his IBMOL/MOTECC open source program package, dating back to the sixties at IBM-San Jose. ZS met Enrico already in early 90's in Belgium during a symposium dedicated for his 60th birthday. ZS has been deeply influenced by Enrico's fascinating personality and warm encouragement. We will all miss this great mind. This work is supported by the National Natural Science Foundation of China (NSFC) through the project 'Science CEnter for Luminescence from Molecular Aggregates (SCELMA)' Grant Number 21788102, as well as by the Ministry of Science and Technology of China through the National Key R & D Plan Grant Number 2017YFA0204501. JR is also supported by the NSFC via Grant Number 22003029. We thank Prof. Tao Xiang for stimulating discussions.

Data availability

The data that support the findings of this study are available upon reasonable request from the authors.

Appendix A. DOF orderings in the main text

In the following we list five of the OFS-D optimized spin orderings for the Heisenberg model:

- 12, 19, 6, 26, 5, 4, 3, 11, 22, 21, 23, 24, 13, 10, 9, 16, 31, 30, 28, 0, 2, 1, 14, 15, 29, 8, 7, 18, 17, 20, 27, 25
- 21, 7, 8, 30, 31, 28, 19, 20, 6, 5, 4, 3, 29, 27, 10, 12, 16, 17, 14, 2, 0, 1, 11, 13, 15, 26, 25, 23, 18, 24, 22, 9
- 15, 9, 14, 1, 10, 16, 22, 30, 21, 19, 20, 23, 31, 29, 0, 2, 24, 25, 26, 3, 5, 6, 12, 13, 11, 27, 28, 18, 8, 4, 7, 17
- 9, 7, 28, 29, 10, 12, 13, 8, 24, 22, 23, 18, 4, 3, 5, 17, 21, 20, 15, 27, 25, 0, 30, 31, 26, 6, 11, 1, 19, 14, 2, 16
- 29, 10, 21, 15, 25, 26, 31, 30, 14, 16, 0, 1, 2, 13, 11, 12, 17, 18, 19, 20, 24, 23, 8, 7, 6, 3, 4, 27, 22, 5, 9, 28

For the N_2 molecule, the initial orderings in terms of spin-orbital are listed below. The orbital index is Hartree–Fock energy, from low to high.

- Energy: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52
- Symmetry: 1, 2, 5, 6, 23, 24, 29, 30, 33, 34, 39, 40, 3, 4, 15, 16, 21, 22, 31, 32, 45, 46, 51, 52, 7, 8, 17, 18, 35, 36, 9, 10, 19, 20, 37, 38, 11, 12, 25, 26, 47, 48, 13, 14, 27, 28, 49, 50, 41, 42, 43, 44
- RCM-*K*: 33, 34, 29, 30, 23, 24, 5, 6, 1, 2, 39, 40, 51, 52, 31, 32, 21, 22, 15, 16, 3, 4, 45, 46, 49, 50, 27, 28, 13, 14, 37, 38, 19, 20, 9, 10, 35, 36, 17, 18, 7, 8, 47, 48, 11, 12, 25, 26, 41, 42, 43, 44
- Fiedler-*K*: 29, 30, 1, 2, 5, 6, 23, 24, 43, 44, 33, 34, 41, 42, 39, 40, 49, 50, 13, 14, 27, 28, 15, 16, 31, 32, 21, 22, 3, 4, 51, 52, 45, 46, 47, 48, 11, 12, 25, 26, 19, 20, 37, 38, 9, 10, 35, 36, 7, 8, 17, 18
- GA-*K*: 37, 38, 9, 10, 19, 20, 47, 48, 11, 12, 25, 26, 41, 42, 39, 40, 29, 30, 5, 6, 23, 24, 1, 2, 33, 34, 17, 18, 7, 8, 35, 36, 49, 50, 13, 14, 27, 28, 43, 44, 15, 16, 21, 22, 3, 4, 31, 32, 51, 52, 45, 46
- GA-*I*: 52, 51, 26, 27, 19, 25, 28, 20, 18, 17, 12, 9, 10, 11, 13, 8, 7, 14, 6, 15, 16, 5, 22, 21, 30, 24, 23, 1, 2, 4, 3, 29, 32, 31, 34, 33, 38, 37, 36, 35, 47, 39, 41, 42, 40, 48, 50, 46, 44, 45, 43, 49

The optimized ordering by OFS-S from the RCM-*K* initial ordering shown in figure 4(b) is

- 51, 39, 40, 31, 32, 34, 33, 29, 30, 23, 24, 21, 22, 2, 1, 6, 15, 5, 16, 4, 3, 52, 46, 45, 49, 50, 38, 37, 28, 27, 19, 13, 8, 7, 14, 18, 36, 20, 10, 12, 9, 11, 17, 26, 25, 35, 41, 42, 43, 44, 48, 47

and the optimized ordering by OFS-D from the RCM-*K* initial ordering shown in figure 4(c) is

- 33, 34, 29, 30, 23, 24, 5, 6, 1, 2, 15, 21, 31, 32, 22, 51, 52, 16, 3, 4, 39, 46, 40, 45, 49, 50, 36, 35, 17, 27, 13, 7, 18, 28, 14, 8, 37, 47, 38, 20, 19, 10, 9, 48, 11, 12, 25, 26, 41, 42, 43, 44

The initial ordering shown in figure 6(a) for the pyrazine model is

- $v9a$, $v6a$, $|S_1, S_2\rangle$, $v10a$, $v1$, $v17a$, $v11$, $v5$, $v12$, $v13$, $v4$, $v18b$, $v8b$, $v8a$, $v7b$, $v20b$, $v19a$, $v19b$, $v18a$, $v14$, $v6b$, $v16b$, $v3$, $v16a$, $v2$

and the optimized ordering shown in figure 6(b) is

- $|S_1, S_2\rangle$, $v9a$, $v1$, $v6a$, $v10a$, $v8a$, $v12$, $v11$, $v18b$, $v19b$, $v14$, $v16a$, $v4$, $v17a$, $v5$, $v6b$, $v8b$, $v3$, $v19a$, $v20b$, $v18a$, $v7b$, $v2$, $v16b$, $v13$

The ordering derived from ML-MCTDH tree used in figure 7(a) is

- $|S_1, S_2\rangle$, $v10a$, $v6a$, $v1$, $v9a$, $v8a$, $v2$, $v6b$, $v8b$, $v4$, $v5$, $v3$, $v16a$, $v12$, $v13$, $v19b$, $v18b$, $v18a$, $v14$, $v19a$, $v17a$, $v20b$, $v16b$, $v11$, $v7b$

Appendix B. MPO representation of Jordan–Wigner transformed Hamiltonian

In this section, we illustrate the importance of keeping consistent DOF ordering in Jordan–Wigner transformation and in MPO. We take the following toy Hamiltonian with totally 12 terms as an example:

$$\hat{H} = \sum_{i,j=1,i\neq j}^{4} \hat{c}_i^\dagger \hat{c}_j. \tag{B1}$$

The Hamiltonian is designed to be invariant with respect to DOF index permutation. Jordan–Wigner transformation with operator order $\{1, 2, 3, 4\}$, can be written as:

$$\begin{aligned}
\hat{c}_1^\dagger &= \hat{+}_1, & \hat{c}_1 &= \hat{-}_1 \\
\hat{c}_2^\dagger &= \hat{Z}_1\hat{+}_2, & \hat{c}_2 &= \hat{Z}_1\hat{-}_2 \\
\hat{c}_3^\dagger &= \hat{Z}_1\hat{Z}_2\hat{+}_3, & \hat{c}_3 &= \hat{Z}_1\hat{Z}_2\hat{-}_3 \\
\hat{c}_4^\dagger &= \hat{Z}_1\hat{Z}_2\hat{Z}_3\hat{+}_4, & \hat{c}_4 &= \hat{Z}_1\hat{Z}_2\hat{Z}_3\hat{-}_4.
\end{aligned} \tag{B2}$$

For DOF order $\{1, 2, 3, 4\}$, optimal MPO in the symbolic form reads:

$$\hat{H} = \begin{bmatrix} \hat{I}_1 & \hat{+}_1 & \hat{-}_1 \end{bmatrix} \begin{bmatrix} \hat{I}_2 & \hat{+}_2 & \hat{-}_2 & 0 \\ 0 & \hat{Z}_2 & 0 & \hat{-}_2 \\ 0 & 0 & \hat{Z}_2 & \hat{+}_2 \end{bmatrix} \begin{bmatrix} \hat{-}_3 & \hat{+}_3 & 0 \\ 0 & \hat{Z}_3 & \hat{-}_3 \\ \hat{Z}_3 & 0 & \hat{+}_3 \\ 0 & 0 & \hat{I}_3 \end{bmatrix} \begin{bmatrix} \hat{+}_4 \\ \hat{-}_4 \\ \hat{I}_4 \end{bmatrix} \tag{B3}$$

which means

$$\begin{aligned}
\hat{W}[1:2] &= \begin{bmatrix} \hat{I}_1\hat{I}_2, & \hat{+}_1\hat{Z}_2 + \hat{+}_2, & \hat{-}_1\hat{Z}_2 + \hat{-}_2, & \hat{+}_1\hat{-}_2 + \hat{-}_1\hat{+}_2 \end{bmatrix} \\
\hat{W}[3:4] &= \begin{bmatrix} \hat{+}_3\hat{-}_4 + \hat{-}_3\hat{+}_4, & \hat{-}_3 + \hat{Z}_3\hat{-}_4, & \hat{+}_3 + \hat{Z}_3\hat{+}_4, & \hat{I}_3\hat{I}_4 \end{bmatrix}.
\end{aligned} \tag{B4}$$

The maximum MPO bond dimension $M_{\mathrm{O}} = 4$.

On the other hand, if the DOF ordering is set to be $\{1, 3, 2, 4\}$, with the same Jordan Wigner transformation equation (B2), the optimal MPO in the symbolic form reads:

$$\hat{H} = \begin{bmatrix} \hat{I}_1 & \hat{+}_1 & \hat{-}_1 \end{bmatrix} \begin{bmatrix} \hat{Z}_3 & 0 & \hat{+}_3 & \hat{-}_3 & \hat{+}_3 & \hat{-}_3 & 0 & 0 \\ 0 & \hat{-}_3 & \hat{I}_3 & 0 & 0 & 0 & \hat{Z}_3 & 0 \\ 0 & \hat{+}_3 & 0 & \hat{I}_3 & 0 & 0 & 0 & \hat{Z}_3 \end{bmatrix} \times \begin{bmatrix} \hat{-}_2 & \hat{+}_2 & 0 \\ 0 & 0 & \hat{Z}_2 \\ 0 & 0 & \hat{-}_2 \\ 0 & 0 & \hat{+}_2 \\ 0 & \hat{I}_2 & 0 \\ \hat{I}_2 & 0 & 0 \\ 0 & \hat{Z}_2 & 0 \\ \hat{Z}_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{+}_4 \\ \hat{-}_4 \\ \hat{I} \end{bmatrix} \tag{B5}$$

which means

$$\begin{aligned} \hat{W}[1:2] &= \begin{bmatrix} \hat{Z}_3, & \hat{+}_1\hat{-}_3 + \hat{-}_1\hat{+}_3, & \hat{+}_1 + \hat{+}_3, & \hat{-}_1 + \hat{-}_3, & \hat{+}_3\ , & \hat{-}_3, & \hat{+}_1\hat{Z}_3, & \hat{-}_1\hat{Z}_3 \end{bmatrix} \\ \hat{W}[3:4] &= \begin{bmatrix} \hat{+}_2\hat{-}_4 + \hat{-}_2\hat{+}_4, & \hat{Z}_2, & \hat{-}_2, & \hat{+}_2, & \hat{-}_4, & \hat{+}_4, & \hat{Z}_2\hat{-}_4, & \hat{Z}_2\hat{+}_4 \end{bmatrix}. \end{aligned} \tag{B6}$$

The maximum MPO bond dimension $M_{\mathrm{O}} = 8$. It is clear that equation (B3) is a much more compact representation than equation (B5) for the same Hamiltonian. Thus, if the DOF is ordered to be $\{1, 3, 2, 4\}$, the fermion operators should be retransformed to spin operators according to the ordering for minimal MPO bond dimension.

ORCID iDs

Weitang Li https://orcid.org/0000-0002-8739-641X
Jiajun Ren https://orcid.org/0000-0002-1508-4943
Hengrui Yang https://orcid.org/0000-0001-8525-9772
Zhigang Shuai https://orcid.org/0000-0003-3867-2331

References

[1] White S R 1992 *Phys. Rev. Lett.* **69** 2863
[2] White S R 1993 *Phys. Rev.* B **48** 10345
[3] Shuai Z, Bredas J-L, Pati S K and Ramasesha S 1997 *Optical Probes of Conjugated Polymers* vol 3145 ed Z V Vardeny and L J Rothberg (Washington: International Society for Optics and Photonics (SPIE)) pp 293–302
[4] Shuai Z, Brédas J L, Saxena A and Bishop A R 1998 *J. Chem. Phys.* **109** 2549
[5] Fano G, Ortolani F and Ziosi L 1998 *J. Chem. Phys.* **108** 9246
[6] White S R and Martin R L 1999 *J. Chem. Phys.* **110** 4127
[7] Li Z, Guo S, Sun Q and Chan G K-L 2019 *Nat. Chem.* **11** 1026
[8] Taffet E J, Beljonne D and Scholes G D 2020 *J. Am. Chem. Soc.* **142** 20040
[9] Phung Q M, Muchammad Y, Yanai T and Ghosh A 2021 *JACS Au* **1** 2303
[10] Chan G K-L and Sharma S 2011 *Annu. Rev. Phys. Chem.* **62** 465
[11] Kurashige Y 2014 *Mol. Phys.* **112** 1485
[12] Baiardi A and Reiher M 2020 *J. Chem. Phys.* **152** 040903
[13] Cheng Y, Xie Z and Ma H 2022 *J. Phys. Chem. Lett.* **13** 904
[14] Schollwöck U 2011 *Ann. Phys., NY* **326** 96
[15] Cazalilla M A and Marston J B 2002 *Phys. Rev. Lett.* **88** 256403
[16] Luo H, Xiang T and Wang X 2003 *Phys. Rev. Lett.* **91** 049701
[17] White S R and Feiguin A E 2004 *Phys. Rev. Lett.* **93** 076401
[18] Vidal G 2004 *Phys. Rev. Lett.* **93** 040502
[19] Daley A J, Kollath C, Schollwöck U and Vidal G 2004 *J. Stat. Mech.* P04005
[20] Ma H, Luo Z and Yao Y 2018 *Mol. Phys.* **116** 854
[21] Paeckel S, Köhler T, Swoboda A, Manmana S R, Schollwöck U and Hubig C 2019 *Ann. Phys., NY* **411** 167998
[22] Borrelli R and Gelin M F 2021 *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **11** e1539
[23] Li W, Ren J and Shuai Z 2021 *Chem. J. Chinese Universities* **42** 2085
[24] Haegeman J, Lubich C, Oseledets I, Vandereycken B and Verstraete F 2016 *Phys. Rev.* B **94** 165116
[25] Dirac P A M 1930 *Math. Proc. Camb. Phil. Soc.* **26** 376–85
[26] Haegeman J, Cirac J I, Osborne T J, Pižorn I, Verschelde H and Verstraete F 2011 *Phys. Rev. Lett.* **107** 070601
[27] Baiardi A and Reiher M 2019 *J. Chem. Theory Comput.* **15** 3481–98
[28] Xie X, Liu Y, Yao Y, Schollwöck U, Liu C and Ma H 2019 *J. Chem. Phys.* **151** 224101
[29] Li W, Ren J and Shuai Z 2020 *J. Phys. Chem. Lett.* **11** 4930
[30] Li W, Ren J and Shuai Z 2021 *Nat. Commun.* **12** 4260
[31] Li W, Ren J and Shuai Z 2020 *J. Chem. Phys.* **152** 024127
[32] Secular P, Gourianov N, Lubasch M, Dolgov S, Clark S R and Jaksch D 2020 *Phys. Rev.* B **101** 235123
[33] Zhai H and Chan G K-L 2021 *J. Chem. Phys.* **154** 224116
[34] Brabec J, Brandejs J, Kowalski K, Xantheas S, Legeza Ö and Veis L 2021 *J. Comput. Chem.* **42** 534
[35] Krumnow C, Veis L, Legeza Ö and Eisert J 2016 *Phys. Rev. Lett.* **117** 210402
[36] Rams M M and Zwolak M 2020 *Phys. Rev. Lett.* **124** 137701
[37] Xiang T 1996 *Phys. Rev.* B **53** R10445
[38] Xiang T, Lou J and Su Z 2001 *Phys. Rev.* B **64** 104414
[39] Nakatani N and Chan G K-L 2013 *J. Chem. Phys.* **138** 134113
[40] Gunst K, Verstraete F, Wouters S, Legeza Ö and Van Neck D 2018 *J. Chem. Theory Comput.* **14** 2026
[41] Larsson H R 2019 *J. Chem. Phys.* **151** 204102

[42] Chepiga N and White S R 2019 *Phys. Rev.* B **99** 235426
[43] Chan G K-L and Head-Gordon M 2002 *J. Chem. Phys.* **116** 4462
[44] Legeza Ö and Sólyom J 2003 *Phys. Rev.* B **68** 195116
[45] Moritz G, Hess B A and Reiher M 2005 *J. Chem. Phys.* **122** 024107
[46] Rissler J, Noack R M and White S R 2006 *Chem. Phys.* **323** 519
[47] Barcza G, Legeza O, Marti K H and Reiher M 2011 *Phys. Rev.* A **83** 012508
[48] Olivares-Amaya R, Hu W, Nakatani N, Sharma S, Yang J and Chan G K-L 2015 *J. Chem. Phys.* **142** 034102
[49] Dupuy M-S and Friesecke G 2021 *SIAM J. Sci. Comput.* **43** B108
[50] Roemelt M, Krewald V and Pantazis D A 2018 *J. Chem. Theory Comput.* **14** 166
[51] Feldt M, Martín-Fernández C and Harvey J N 2020 *Phys. Chem. Chem. Phys.* **22** 23908
[52] Chan G K-L, Keselman A, Nakatani N, Li Z and White S R 2016 *J. Chem. Phys.* **145** 014102
[53] Jiang T, Li W, Ren J and Shuai Z 2020 *J. Phys. Chem. Lett.* **11** 3761
[54] Jiang T, Ren J and Shuai Z 2021 *J. Phys. Chem. Lett.* **12** 9344
[55] Schröder F A Y N, Turban D H P, Musser A J, Hine N D M and Chin A W 2019 *Nat. Commun.* **10** 1062
[56] Kloss B and Bar Lev Y 2019 *Phys. Rev.* A **99** 032114
[57] Verstraete F and Cirac J I 2006 *Phys. Rev.* B **73** 094423
[58] Baiardi A, Stein C J, Barone V and Reiher M 2017 *J. Chem. Theory Comput.* **13** 3764
[59] Ren J, Li W, Jiang T and Shuai Z 2020 *J. Chem. Phys.* **153** 084118
[60] Hopcroft J E and Karp R M 1973 *SIAM J. Comput.* **2** 225
[61] Ren J, Li W, Jiang T, Wang Y and Shuai Z 2021 The Renormalizer package https://github.com/shuaigroup/renormalizer
[62] Chan G K-L 2004 *J. Chem. Phys.* **120** 3172
[63] Cuthill E and McKee J 1969 *Proc. 1969 24th National Conf., ACM '69* (New York: Association for Computing Machinery) pp 157–72
[64] Fiedler M 1973 *Czech. Math. J.* **23** 298
[65] Wouters S, Poelmans W, De Baerdemacker S, Ayers P W and Van Neck D 2015 *Comput. Phys. Commun.* **191** 235
[66] Dunning T H 1989 *J. Chem. Phys.* **90** 1007
[67] Sun Q *et al* 2020 *J. Chem. Phys.* **153** 024109
[68] Larsen H, Olsen J, Jørgensen P and Christiansen O 2000 *J. Chem. Phys.* **113** 6677
[69] Chan G K-L, Kállay M and Gauss J 2004 *J. Chem. Phys.* **121** 6110
[70] Fortin F-A, De Rainville F-M, Gardner M-A, Parizeau M and Gagné C 2012 *J. Mach. Learn. Res.* **13** 2171–5
[71] Meyer H-D, Manthe U and Cederbaum L S 1990 *Chem. Phys. Lett.* **165** 73
[72] Beck M H, Jäckle A, Worth G A and Meyer H-D 2000 *Phys. Rep.* **324** 1–105
[73] Worth G A, Meyer H D and Cederbaum L S 1996 *J. Chem. Phys.* **105** 4412
[74] Raab A, Worth G A, Meyer H-D and Cederbaum L S 1999 *J. Chem. Phys.* **110** 936
[75] Greene S M and Batista V S 2017 *J. Chem. Theory Comput.* **13** 4034
[76] Baiardi A and Reiher M 2019 *J. Chem. Theory Comput.* **15** 3481
[77] Vendrell O and Meyer H-D 2011 *J. Chem. Phys.* **134** 044135
[78] Orús R 2019 *Nat. Rev. Phys.* **1** 538