Symmetric Tensor Decomposition Description of Fermionic Many-Body Wave Functions

Wataru Uemura and Osamu Sugino

The Institute for Solid State Physics, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba, 277-8581, Japan

(Received 7 August 2012; published 21 December 2012)

The configuration interaction (CI) is a versatile wave function theory for interacting fermions, but it involves an extremely long CI series. Using a symmetric tensor decomposition method, we convert the CI series into a compact and numerically tractable form. The converted series encompasses the Hartree-Fock state in the first term and rapidly converges to the full-CI state, as numerically tested by using small molecules. Provided that the length of the symmetric tensor decomposition CI series grows only moderately with the increasing complexity of the system, the new method will serve as one of the alternative variational methods to achieve full CI with enhanced practicability.

DOI: 10.1103/PhysRevLett.109.253001 PACS numbers: 31.15.V−, 31.10.+z

An accurate description of the ground-state wave function of an interacting fermion system is one of the central goals of modern science. The most straightforward and versatile approach to describing this wave function is the configuration interaction (CI), but its numerical application is greatly limited by the fact that the full-CI series consists of ${}_{M}C_{N}$ Slater determinants (SDs) when describing an N-electron system by using M basis functions. To truncate this extremely long CI series without compromising on chemical accuracy, many methods have been developed, such as the multireference CI, which uses a part of the SDs derived from a few of the most important ones, or the complete active space CI, which uses all SDs generated from a selected set of orbitals [1]. Even so, the application has been hampered by the slow convergency of the CI series.

In this context, the many-body perturbation approaches to treat all SDs have attracted attention; these approaches include the coupled cluster theory [2,3], which is used to represent the wave function in terms of an SD (or a few SDs) applied with the exponential of an excitation operator. The coupled cluster theory has proven accurate for a number of molecules, although it occasionally provides qualitatively incorrect potential surfaces [4]. The density matrix renormalization group method [5] has also attracted attention as a variational method within the space of the matrix product state [6]. It has been extensively applied to correlated electron systems [7,8]; however, this method was originally formulated only for one-dimensional systems, and its extension to three-dimensional systems is not very straightforward.

Recent tensor analyses have shown that, despite the large number, the CI coefficients may be described by a tractable number of variational parameters. For example, the full-CI results of some molecules were accurately reproduced by the complete-graph tensor network state containing $\sim M^2$ variational parameters [9,10]. Tensor decomposition (TD) [11] methods such as the Tucker decomposition [12] and the canonical decomposition combined with the parallel factor decomposition, abbreviated as CP [13,14], have also been applied to molecules. These methods were used to analyze the double excitation tensor $\mathcal{T}_2$ originating from the electron-electron interaction [15,16]. The results showed that the $\mathcal{T}_2$ tensor of rank 4, consisting of $\sim M^4$ terms, can be described by $\sim MK$ parameters, where K denotes the length of the tensor decomposition [16]. The TD method was also suggested as being effective in greatly reducing the variational parameters required for full CI [15].

In this context, we formulate a practical scheme to perform full-CI level calculation using a TD method. In this study, we describe the CI coefficients as a product of a symmetric tensor and the permutation tensor, and following the CP procedure we expand the former into K symmetric Kronecker product states, which are composed of vectors of dimension M. Subsequently, we calculate the second-order density matrix consisting of $\sim K^2M^4$ elements using Viète's formula [17], thereby performing $\sim M^2$ operations for each element. This allows us to perform the total energy calculation variationally by using $\sim K^2M^6$ operations. Our test calculations for the potential surface of simple diatomic molecules and for a Hubbard cluster model with different parameters show that, with increasing K, the total energy rapidly converges to the full-CI result. This shows that our symmetric tensor decomposition (STD) CI scheme will greatly extend the applicability of the full-CI level calculation, provided that K increases only moderately with N, M, or the complexity of the electron correlation.

In treating the interacting indistinguishable particles, the (anti)symmetric nature of the wave function plays a key role. In the STD CI scheme, the antisymmetric nature of electrons is directly treated by explicitly introducing the permutation tensor, reducing thereby the variational degree of freedom to a power of N, although it is $N!$ in the conventional CI. Moreover, the application of this scheme is not restricted to one-dimensional systems, contrary to the density matrix renormalization group that is based on the assumption of the matrix product state. In the rest of this Letter, we provide the details of the STD CI method.

We begin by describing the CI-series representation of the many-body wave function

$$\Psi(x_1 \cdots x_N) = \sum_{i_1,\ldots,i_N=1}^{M} A_{i_1,\ldots,i_N} \psi_{i_1}(x_1) \cdots \psi_{i_N}(x_N), \quad (1)$$

where $x_1 \cdots x_N$ represent the space and spin coordinates of the electrons and ψ_{i_k}'s are the orthonormal orbitals which are represented as a linear combination of orthonormalized basis functions as

$$\psi_i(x) = \sum_j U_{ij} \phi_j(x).$$

The antisymmetric tensor $A_{i_1,\ldots,i_N}$ can be described as the product of a symmetric tensor ($S_{i_1,\ldots,i_N}$) of rank N and dimension M and a product of $N(N-1)/2$ permutation tensors (ϵ_{ij}'s) of rank 2 as

$$A_{i_1,\ldots,i_N} = S_{i_1,\ldots,i_N} \epsilon_{i_1 i_2} \epsilon_{i_1 i_3} \cdots \epsilon_{i_{N-1} i_N}. \quad (2)$$

Next, $S_{i_1,\ldots,i_N}$ is decomposed into a minimal linear combination of symmetric Kronecker product states using vectors of dimension M, $c^1_{i_k}, \cdots, c^K_{i_k}$, as

$$S_{i_1,\ldots,i_N} = \sum_{j=1}^{K} \lambda_j c^j_{i_1} \cdots c^j_{i_N}. \quad (3)$$

This STD is a symmetric version of CP, which is also a special case of the symmetric Tucker decomposition

$$S_{i_1,\ldots,i_N} = \sum_{j_1,\ldots,j_N} s_{j_1,\ldots,j_N} u^{j_1}_{i_1} \cdots u^{j_N}_{i_N}, \quad (4)$$

in that the transformed tensor $s_{j_1 j_2,\ldots,j_N}$ is the superdiagonal λ_j in CP. The total energy is optimized by varying the vectors c^j_i and the unitary matrix U_{ij}, so that no approximation is made in our STD CI method apart from the truncation of the series at K. It is noteworthy that each term in the STD series contains all the SDs generated from the orbitals ψ_i, although the degrees of freedom are only $MK + M(M+1)/2$ as a whole, thereby indicating that we are treating the entangled states and that the degree of entanglement is reduced with increasing K. It can be shown that the Hartree-Fock (HF) approximation corresponds to taking $K = 1$ and $c^1_{N+1} = \cdots = c^1_M = 0$; therefore, the approximation with $K = 1$ is already a natural extension of the HF approximation. When treating a weakly correlated system, an HF-like solution is obtained, and, on the other hand, when treating a strongly correlated system, an orbital-ordered solution is obtained, provided that a sufficiently large value of K is considered. In this manner, we can bridge the HF solution with the fully correlated state by increasing the value of K. The STD CI method is exact when $K = {}_M C_N$.

Our numerical procedure begins by constructing the second-order density matrix, which has the form

$$\gamma_2(x_1 x_2, x_3 x_4) = \sum_{i_1 i_2 i_3 i_4=1}^{M} \Gamma_{i_1 i_2 i_3 i_4} \psi^*_{i_1}(x_1) \psi^*_{i_2}(x_2) \times \psi_{i_3}(x_3) \psi_{i_4}(x_4).$$

By using (2) and (3), the density matrix coefficient can be rewritten as

$$\Gamma_{i_1 i_2 i_3 i_4} = \sum_{ij=1}^{K} \lambda_i \lambda_j c^{i*}_{i_1} c^{i*}_{i_2} c^j_{i_3} c^j_{i_4} \epsilon_{i_1 i_2} \epsilon_{i_3 i_4} I^{ij}_{i_1 i_2 i_3 i_4},$$

where I for each set of indices $\{ij i_1 i_2 i_3 i_4\}$ is expressed, by using $a_{k_l} \equiv c^{i*}_{k_l} c^j_{k_l} \epsilon_{i_1 k_l} \epsilon_{i_2 k_l} \epsilon_{i_3 k_l} \epsilon_{i_4 k_l}$, as

$$I = \sum_{k_3,\ldots,k_N=1}^{M} a_{k_3} \cdots a_{k_N} (\epsilon_{k_3,\ldots,k_N})^2. \quad (5)$$

Based on the fact that the permutation tensor squared is equal to 1 when all the indices are different and 0 otherwise, it can be shown by using Viète's formula that the value of I is equal to the $M - (N-2)$th order coefficients of the polynomial $(N-2)! f_M(t)$ with $f_M(t) \equiv (t + a_1) \cdots (t + a_M)$ [18]. The coefficient can be easily obtained by using a list manipulation, where the coefficients of $f_p(t)$ with $0 \leq p \leq M$ are described by a row vector of dimension p as $\boldsymbol{f}_p = (f_{p,0}, f_{p,1}, \cdots, f_{p,p-1})$ and are applied with the iterative equation $\boldsymbol{f}_p = a_p(\boldsymbol{f}_{p-1}, 0) + (0, \boldsymbol{f}_{p-1})$, considering $\boldsymbol{f}_0 = 1$. $\propto M^2$ operations are required to obtain the coefficient of $f_M(t)$. Therefore, the total number of operations needed to obtain all I's is $\propto K^2 M^6$. In practical coding, one may use the fact that $a_{i_k} = 0$ when i_k is equal to one of the four indices $\{i_1 i_2 i_3 i_4\}$ to achieve further efficiency.

Subsequently the parameters $c^j_{i_k}$, λ^j, and U_{ij} are varied to minimize the total energy $E_{\text{tot}} = \sum h_{i_1 i_2 i_3 i_4} \Gamma_{i_3 i_4 i_1 i_2} / \sum \Gamma_{i_1 i_2 i_1 i_2}$ with

$$h_{i_1 i_2 i_3 i_4} = \int dx_1 dx_2 \psi^*_{i_1}(x_1) \psi^*_{i_2}(x_2) \left[N\left(-\frac{1}{2} \boldsymbol{\nabla}^2_1 + v_{\text{ext}}(\boldsymbol{r}_1) \right) + \frac{N(N-1)}{2} \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \right] \psi_{i_3}(x_1) \psi_{i_4}(x_2), \quad (6)$$

where v_{ext} denotes the external potential. In the variation, we require derivatives of $I^{ij}_{i_1 i_2 i_3 i_4}$ with respect to a_{i_k} for those i_k not in $\{i_1 i_2 i_3 i_4\}$. To obtain the derivatives, we need to differentiate $f_M(t)$ by a_{i_k} and obtain its $[M - (N-1)]$th coefficient. When this is done simply by using the list manipulation, $\propto K^2 M^7$ operations are required for each i_k; however, the number of operations can be reduced when using $f_M(t)/(t + a_{i_k})$ for the differentiation. When the series $(t + a_{i_k})^{-1} = \sum_{m=0}^{\infty} a^{-m-1}_{i_k} (-t)^m$ is multiplied with $f_M(t)$, the $[M - (N-1)]$th coefficient can be obtained as

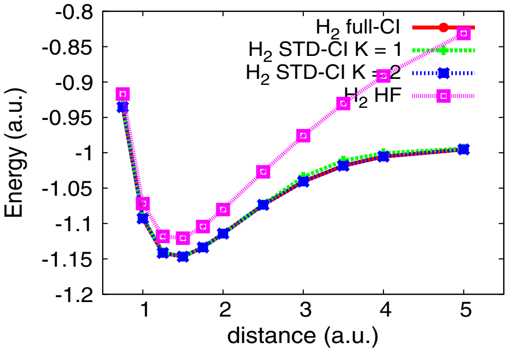

FIG. 1 (color online). Potential curve of H_2 molecules for full CI (solid line with circle), STD CI with $K = 1$ (broken line with cross) and $K = 2$ (broken line with asterisk), and HF (dotted line with rectangle).

$$- \sum_{s=0}^{M-(N-1)} (-a_{i_k})^{s-1-M+(N-1)} f_{M,s},$$

thereby requiring $\propto M$ operations for each k. Therefore, $\sim K^2 M^6$ operations are required to obtain all the derivatives of $I^{ij}_{i_1 i_2 i_3 i_4}$. By applying the same technique to the expression $f(t)/(t + a_i)(t + a_j)$, the second derivatives are similarly obtained with $\propto K^2 M^6$ operations. It should be noted that the calculation of the derivatives is the rate-determining step in our calculation.

In a manner similar to HF [19,20], STD CI can be applied to a crystalline solid by taking a linear combination of the atomic orbitals χ_i as

$$\phi_{ik}(r) = \sum_{\tau} e^{ik\cdot R_\tau} \chi_i(r - R_\tau),$$

where k and R_τ denote the reciprocal vector in the Brillouin zone and the nuclear coordinate, respectively. Thus, M should be read as the number of k values multiplied by the number of basis functions.

To assess the efficiency of the STD CI scheme, we investigate how many terms in Eq. (3) are required to achieve convergence in E_{tot}. This investigation is carried out for simple diatomic molecules (H_2, He_2, and LiH) and a four-site Hubbard model. Relativistic effects are neglected, and only the spin unpolarized state is calculated by using the same number of orbitals with an α and β spin. In testing the convergence, the calculated results are compared with the full-CI calculation performed by using the same basis functions. In our calculations, the Newton-Raphson method is used to variationally determine the parameters.

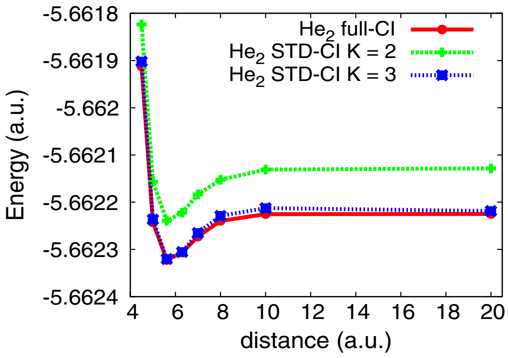

FIG. 2 (color online). Potential curve of He_2 molecules for full CI (solid line with circle) and STD CI with $K = 2$ (broken line with cross) and $K = 3$ (broken line with asterisk).

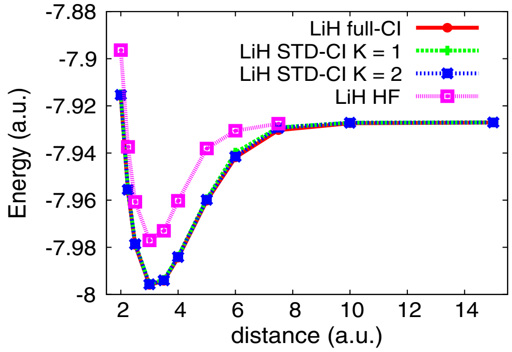

FIG. 3 (color online). Potential curve of LiH molecules for full CI (solid line with circle), STD CI with $K = 1$ (broken line with cross) and $K = 2$ (broken line with asterisk), and HF (dotted line with rectangle).

H_2 is the simplest molecule where the molecular orbital picture, valid near the equilibrium bond length, is switched to the Heitler-London picture, as the interatomic distance increases to infinity. The calculation with the Slater-type orbital 3G basis set shows that $K = 1$ reproduces the full-CI potential curve within an error of 0.01 hartree, while the error is less than 0.01 mhartree when $K = 2$ (Fig. 1). The molecule He_2 is weakly bound to the dispersion forces, and the test is more stringent in this case. The calculation with the 6–311G basis set shows that $K = 3$ is sufficient to reproduce the full-CI result within an error of 0.01 mhartree, while $K = 2$ is already sufficient to obtain the binding

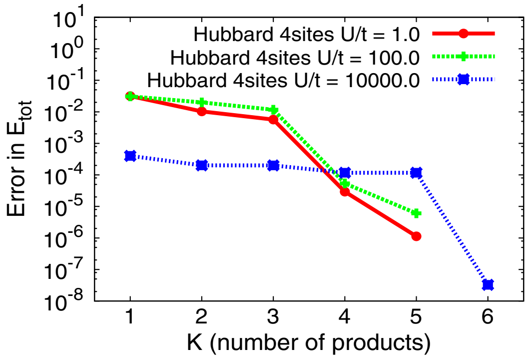

FIG. 4 (color online). Error in E_{tot} for the four-site Hubbard model for various parameters $U/t = 1$ (solid line with circle), $U/t = 100$ (dotted line with cross), and $U/t = 10\,000$ (dotted line with asterisk).

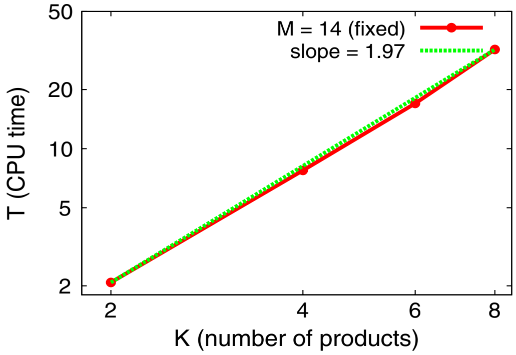

FIG. 5 (color online). Total CPU time (T) versus K. The solid line with circle indicates the result after measurement for one iteration step, and the broken line indicates that after fitting to $T = aK^b$. b is obtained as 1.97.

energy within the same accuracy, although the absolute value of E_{tot} is always larger by 0.1 mhartree (Fig. 2). The binding energy is about 3 times larger than the accurate quantum chemical calculation [21] and the experimental results [22], which is presumably due to the insufficient number of basis functions; obtaining an accurate value of the binding energy is beyond the scope of our comparative study, and this must be the consideration of future studies. LiH is a typical heteronuclear diatomic molecule. The 4–31G calculation for this case shows that $K = 1$ nearly sufficiently reproduces the full-CI result while the HF calculation significantly underestimates the binding energy (Fig. 3). The final test is the application of our idea to the four-site Hubbard model in the tetrahedron structure under the half filled condition. As the Hubbard U over the transfer t increases, larger K values are required; however, $K = 6$ is found sufficient even in the large U/t limit (Fig. 4).

The computational time theoretically scales as K^2M^6. We tested the time scaling with our numerical code to find that CPU time indeed scales as $K^{1.97}M^{5.97}$ on average (Figs. 5 and 6). Because the operations involved in the calculation can be performed independently, this method is suitable for massively parallel computers.

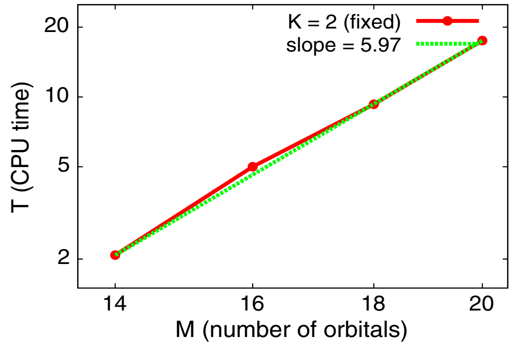

FIG. 6 (color online). Total CPU time (T) versus M. The solid line with circle indicates the result after measurement for one iteration step, and the broken line indicates that after fitting to $T = aM^b$. b is 5.97 after the fitting.

We formulated a practical scheme to perform the full-CI level calculation using the STD method. In the STD CI method, we expand the CI coefficients as the product of a symmetric tensor and the permutation tensor, and we further expand the symmetric tensor into Kronecker product states composed of vectors of dimension M. By varying the vectors and the unitary transformation matrix U_{ij}, the total energy is minimized by using the second-order density matrix technique. The STD CI method, which involves taking the length of the series K as the only input parameter, allows us to perform a full-CI level calculation rigorously by using $KM + M(M + 1)/2$ variational parameters and $\sim K^2M^6$ operations. By applying the scheme to the potential curve of small diatomic molecules such as H_2, He_2, and LiH and the four-site Hubbard model for various parameters, we found that a very small K value is required to reproduce the full-CI results within millihartree accuracy. If K increases moderately with N, M, or the degree of correlation, the scheme will greatly extend the applicability of the full-CI level calculation. We believe that application of the scheme to a crystalline solid and the use of the scheme as a building block of the fragment molecular orbital scheme can be of high significance [23].

The authors thank Professor Y. Mochizuki (Rikkyo University) and M. Nakata (RIKEN) for their valuable discussions.

[1] I. Shavitt, Mol. Phys. **94**, 3 (1998).
[2] J. Cízek, J. Chem. Phys. **45**, 4256 (1966); Adv. Chem. Phys. **143**, 35 (1969).
[3] J. Cízek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).
[4] M. Nakata, M. Ehara, and H. Nakatsuji, J. Chem. Phys. **116**, 5432 (2002).
[5] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[6] S. Rommer and S. Ostlund, Phys. Rev. B **55**, 2164 (1997).
[7] S. Daul, I. Ciofini, C. Daul, and S. R. White, Int. J. Quantum Chem. **79**, 331 (2000).
[8] S. Sharma and G. K.-L. Chan, J. Chem. Phys. **136**, 124121 (2012).
[9] K. H. Marti, B. Bauer, M. Reiher, M. Troyer, and F. Verstraete, New J. Phys. **12**, 103008 (2010).
[10] K. H. Marti and M. Reiher, Phys. Chem. Chem. Phys. **13**, 6750 (2011).
[11] F. Hitchcock, J. Math. Phys. (Cambridge, Mass.) **6**, 164 (1927); **7**, 9 (1927).
[12] L. R. Tucker, Psychometrika **31**, 279 (1966).
[13] J. D. Carroll and J. J. Chang, Psychometrika **35**, 283 (1970).
[14] R. A. Harshman, UCLA Working Papers in Phonetics 16, 1 (1970).
[15] F. Bell, D. S. Lambrecht, and M. Head-Gordon, Mol. Phys. **108**, 2759 (2010).
[16] U. Benedikt, A. A. Auer, M. Espig, and W. Hackbusch, J. Chem. Phys. **134**, 054118 (2011).

[17] F. Viète, *Opera Mathematica*, edited by F. van Schooten (1579; reprinted Leiden, Netherlands, 1646), p. 162.

[18] M. Hasse, SIAM J. Appl. Math. **40**, 523 (1960).

[19] N. E. Brener and J. L. Fry, Phys. Rev. B **17**, 506 (1978).

[20] J. L. Fry, N. E. Brener, and R. K. Bruyere, Phys. Rev. B **16**, 5225 (1977).

[21] T. van Mourik, A. K. Wilson, and T. H. Dunning, Jr., Mol. Phys. **96**, 529 (1999).

[22] U. Kleinekathofer, K. T. Tang, J. P. Toennies, and C. L. Yiu, Chem. Phys. Lett. **249**, 257 (1996).

[23] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999).