

Configuration interaction with antisymmetrized geminal powers

Wataru Uemura, Shusuke Kasamatsu, and Osamu Sugino

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

(Received 24 April 2015; published 16 June 2015)

To avoid the combinatorial computational cost of configuration interaction (CI), we previously introduced the symmetric tensor decomposition CI (STD-CI) method, which takes advantage of the antisymmetric nature of the electronic wave function and expresses the CI coefficients compactly as a series of Kronecker product states (STD series) [W. Uemura and O. Sugino, *Phys. Rev. Lett.* **109**, 253001 (2012)]. Here we extend the variational degrees of freedom by using different molecular orbitals for different terms in the STD series. This scheme is equivalent to the linear combination of the Hartree-Fock-Bogoliubov state or the antisymmetrized geminal powers (AGPs). The total energy converges very rapidly within 0.72 μ hartree taking only 10 terms for the water molecule, and the convergence is likewise fast for Hubbard tetramers. The computational cost scales as the fifth power of the number of electrons and the square of the number of terms in the STD series, indicating the promise of this AGP-based scheme for highly accurate and efficient computation of quantum systems.

DOI: 10.1103/PhysRevA.91.062504

PACS number(s): 31.10.+z, 31.15.V–

I. INTRODUCTION

Determining the ground state of a many-body system is the most basic problem in many fields of science, such as condensed matter physics, nuclear physics, and quantum chemistry. Many numerical approaches have been developed so far to obtain the wave function. In the configuration interaction (CI) [1], multiple quasiparticle states are generated from a mean-field wave function and the coefficients for their linear combination are determined by solving an eigenvalue problem, while in the variational Monte Carlo (VMC) [2], the quasiparticles are generated stochastically to obtain the expectation value of the total energy to be variationally determined. In the tensor network (TN) framework [3], the total energy is variationally determined using specific quasiparticle states generated according to an assumed TN. In those popular schemes, accessible degrees of freedom are usually not very large. This is particularly the case for CI albeit being the most versatile in that CI is free from the well-known negative sign problem of QMC and the dimensional restriction of the tensor network [3]. In this view, extending the degrees of freedom accessible by CI is a very important problem.

In the field of quantum chemistry, the CI wave function of an N -electron system is expanded as

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

in terms of the molecular orbitals (MOs) $\psi_i(x)$, which are represented as a linear combination of the atomic orbitals (AOs) ϕ_a as

$$\psi_i(x) = \sum_{a=1}^M U_{ia} \phi_a(x), \quad (2)$$

using elements of an orthogonal matrix U . The CI coefficients $A_{i_1 \cdots i_N}$, which are elements of an antisymmetric tensor of order N [4] and dimension M , are varied to minimize the total energy and thus the full-CI calculation is, in practice, hampered by the degrees of freedom that grow combinatorially with N . Usually, the CI series is truncated to make the computation tractable; however, when starting from the Hartree-Fock (HF)

wave function by taking the permutation tensor $\epsilon_{i_1 \cdots i_N}$ in place of $A_{i_1 \cdots i_N}$ and the HF orbital ψ_i^{HF} in place of ψ_i , the CI series usually converges very slowly [1]. The convergence does not, in general, speed up drastically even when starting from multiple Slater determinants. Instead, use of localized MOs in place of the canonical HF orbitals is known to be effective for this purpose [5], and, recently, significant speedup was achieved by using nonorthonormal Slater determinants comprised of nonorthonormal MOs that augment the HF orbitals [6]. It is therefore expected that CI can be made more practicable by using optimal MOs for this purpose.

Apart from CI, the many-body perturbation theory such as the coupled cluster (CC) [7] has been developed to collect infinite correction terms, thereby enabling the correction of the total energy of the water molecule, for example, within the error of 0.53 mhartree by taking up to triple excitations (CCSDT). However, the computational cost scales as $O(N^8)$, which is still prohibitively high for application to many systems of interest. Thus, accelerating the convergence of the CI series should open up the possibility to treat, with unprecedented accuracy, a whole new class of problems in electronic structure theory of strongly correlated systems.

The key to the development that we present here is in a compact representation of the antisymmetric tensor $A_{i_1 \cdots i_N}$. With this in mind, we proposed in our previous work [8] to represent $A_{i_1 \cdots i_N}$ as a product of $\epsilon_{i_1 \cdots i_N}$ and a symmetric tensor $S_{i_1 \cdots i_N}$ and then to decompose the latter into a series of Kronecker products of a set of vectors such that

$$S_{i_1 \cdots i_N} = \sum_{r=1}^K c_{i_1}^r \cdots c_{i_N}^r. \quad (3)$$

Such decomposition of an order- N tensor into a linear combination of rank-1 tensors is known in the literature as the canonical decomposition [9] or the parallel factor decomposition [10]. In our approach, which we call the symmetric tensor decomposition CI (STD-CI), variational parameters are the set of vectors c and the orthogonal matrix U and thus the degrees of freedom are $KM + M(M-1)/2$. The number of order-1 tensors (i.e., the rank of the tensor decomposition K [4]), was found to be relatively few for small molecules

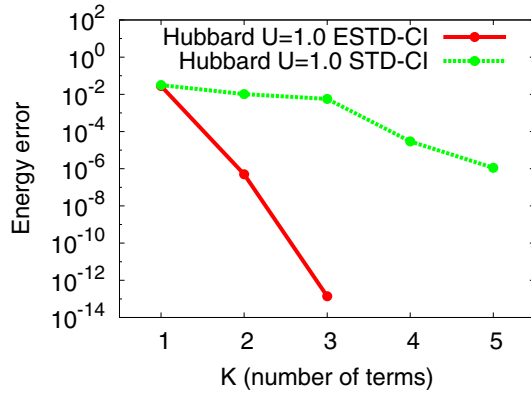


FIG. 1. (Color online) Energy error (in units of t) of the four-site $U = 1.0$ Hubbard model for ESTD and STD-CI.

(H_2 , He_2 , and LiH) and the Hubbard tetramer (Figs. 1–4). This suggests effectiveness of the tensor decomposition in compactly describing the wave function with a rapidly converging series (STD series) when the molecular orbitals are optimized together. Note that each term in the STD series contains all possible Slater determinants, which means that all the terms of the CI series are regrouped differently to form different terms of the STD series. The computational cost thereby required is proportional to $K^2 M^6$ when using the algorithm to handle the permutation tensor developed in Ref. [8], so that the effectiveness of STD-CI depends on the the rank of the decomposition K required to accurately express the wave function. When applied to larger systems, however, we have found that the rank K needed for sufficient convergence is increased in general. In addition, the computation suffers from loss of significance in the floating-point arithmetic as will be shown below. Therefore, an improvement in the tensor decomposition is clearly desirable.

In this context, one may use a more elaborate tensor decomposition, such as the tensor train network (TTN) suitable for one-dimensional systems [11,12] or the projected entangled pair states (PEPS) extended to two-dimensional systems [13]. Indeed, the former approach was applied to several molecular systems [14]. However, here we take a different approach:

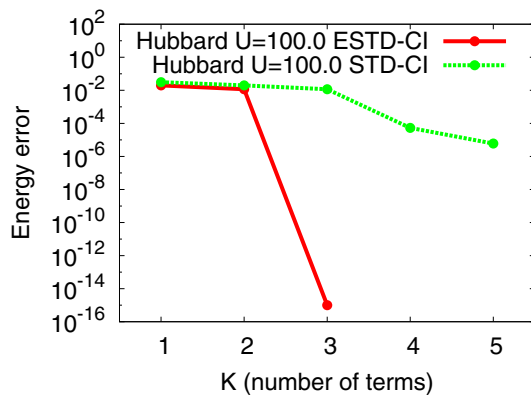


FIG. 2. (Color online) Same as Fig. 1, but for $U = 100.0$. For $K \leq 2$, the error is within 0.01 to 0.1 for both methods, which is within 1 to 10 in units of U .

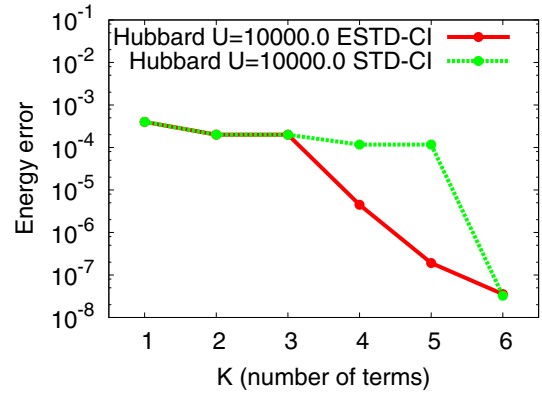


FIG. 3. (Color online) Same as Fig. 1, but for $U = 10000.0$. For $K \leq 3$, the error is within 0.0001 to 0.001 for both methods, which is within 1 to 10 in units of U .

within the canonical decomposition technique, the degrees of freedom are extended by using different nonorthonormal MO coefficients for different terms in the STD series. That is, we use a general complex matrix U_{ia}^r for the MO coefficient although it was described as $c_i^r U_{ia}$ in the original STD-CI scheme, extending thereby the degrees of freedom to $K M^2$. We will call our new scheme the extended STD (ESTD). As will be shown below, each term in the STD series corresponds to a different Hartree-Fock-Bogoliubov (HFB) wave function [15]. Since the HFB wave function is called alternatively the antisymmetrized geminal power (AGP) [16,17], ESTD may also be referred to as a linear combination of AGP.

The AGP wave function has been used for variational calculations in a different context. In quantum chemistry, a single AGP (i.e., $K = 1$) has been used as the trial wave function to obtain the potential energy surface of small molecules [18–21]. In VMC calculations, single AGP [22–25] or multi-AGP [26,27] is multiplied by a correlation factor of the Jastrow type to form a Jastrow-AGP (JAGP) trial wave function; the optimization is done only for the Jastrow parameters with the MO coefficients kept unchanged from the initial HF orbitals or the initial natural orbitals constructed by performing a small preparatory CI calculation. In nuclear physics, multi-quasiparticle states are generated from a single AGP and are linearly combined to describe the wave function.

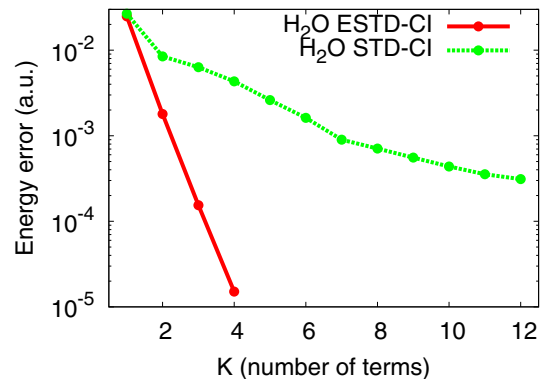


FIG. 4. (Color online) Residual error of the total energy of the H_2O molecule obtained by ESTD and STD-CI.

This computational scheme is called the generator coordinate method (GCM) [28]. The total-energy scheme of GCM was formulated by Onishi and Yoshida [29], and the mathematical structure was studied recently by other groups [30,31]. Contrary to those earlier works, we fully optimize the MO coefficients in our ESTD scheme without introducing the correlation factor. In the present scheme, we use an algorithm to decompose the permutation tensor into products of the second-order permutation tensors thereby allowing us to reduce the computational scaling to K^2M^5 . Contrary to STD-CI, this scheme does not require orthonormalization of AOs so that one can take advantage of the AO basis that is spatially localized and chemically comprehensible. We will show below that the STD series can be drastically shortened for the H₂O molecule and Hubbard tetramers. This may open up the possibility of interpreting strongly correlated systems in terms of MOs. In the next section, we introduce the outline of our formalism. We restrict our study to systems with even numbers of electrons throughout this paper.

II. FORMALISM

In the original STD-CI, we applied the canonical decomposition to the symmetric part of the CI coefficient [Eq. (3)], so that the wave function Eq. (1) is rewritten using Eq. (2) as

$$\Psi(x_1 \cdots x_N) = \sum_{a_1 \cdots a_N=1}^M A_{a_1 \cdots a_N} \phi_{a_1}(x_1) \cdots \phi_{a_N}(x_N), \quad (4)$$

with the CI coefficient becoming

$$A_{a_1 \cdots a_N} = \sum_{i_1 \cdots i_N=1}^M \sum_{r=1}^K c_{i_1}^r \cdots c_{i_N}^r \epsilon_{i_1 \cdots i_N} U_{i_1 a_1} \cdots U_{i_N a_N}, \quad (5)$$

where K is the rank of the symmetric tensor decomposition and c_i^r is a set of complex vectors. We will take the convention for the subscript such that i, j, \dots correspond to MOs and a, b, \dots to AOs. In the extended STD-CI (ESTD), we expand the degrees of freedom by replacing the product $c_i^r U_{ia}$ with a general complex matrix $U_{i\alpha}^r$ such that

$$A_{a_1 \cdots a_N} = \sum_{r=1}^K \sum_{i_1 \cdots i_N=1}^M \epsilon_{i_1 \cdots i_N} U_{i_1 a_1}^r \cdots U_{i_N a_N}^r \equiv \sum_{r=1}^K A_{a_1 \cdots a_N}^r. \quad (6)$$

Using the fact that the permutation tensor of order N for even N can be decomposed into a product of the second-order permutation tensors as

$$\epsilon_{i_1 \cdots i_N} = \frac{1}{(N/2)!2^{N/2}} \sum_{\sigma \in S_N} \text{sgn}(\sigma) \epsilon_{i_{\sigma(1)} i_{\sigma(2)}} \cdots \epsilon_{i_{\sigma(N-1)} i_{\sigma(N)}} \quad (7)$$

$$= \frac{1}{(N/2)!2^{N/2}} \hat{A}(\epsilon_{i_1 i_2} \epsilon_{i_3 i_4} \cdots \epsilon_{i_{N-1} i_N}), \quad (8)$$

where S_N is the symmetric group of degree N and \hat{A} is the antisymmetrizer, Eq. (6) can be represented as

$$A_{a_1 \cdots a_N}^r = \frac{1}{(N/2)!2^{N/2}} \hat{A}(\gamma_{a_1 a_2}^r \gamma_{a_3 a_4}^r \cdots \gamma_{a_{N-1} a_N}^r) \quad (9)$$

using the antisymmetrized geminal

$$\gamma_{ab}^r \equiv \sum_{ij} \epsilon_{ij} U_{ia}^r U_{jb}^r. \quad (10)$$

This shows that our ESTD is an extension of the antisymmetrized geminal power [16,17] in that AGPs are linearly combined. In the second-quantization form, the state vector is given by

$$|\Psi\rangle = \sum_{r=1}^K |\gamma^r\rangle \quad (11)$$

with

$$|\gamma^r\rangle = \frac{1}{(N/2)!2^{N/2}} \left[\sum_{ab=1}^M \gamma_{ab}^r c_a^\dagger c_b^\dagger \right]^{N/2} |0\rangle \quad (12)$$

$$= \exp \left[\frac{t}{2} \sum_{ab} \gamma_{ab}^r c_a^\dagger c_b^\dagger \right] |0\rangle \Big|_{t^{N/2}}, \quad (13)$$

where the subscript $t^{N/2}$ means taking a coefficient of degree $N/2$ from the Hartree-Fock-Bogoliubov state [15]. This indicates a formal similarity of the ESTD wave function with that of the generator coordinate method [28], and thus one can take advantage of the formulas developed for the GCM. For example, we can use the matrix elements derived by Onishi and Yoshida [29] such as

$$\langle \gamma^{r_2} | \gamma^{r_1} \rangle = \exp \left[\frac{1}{2} \text{tr} \log(1 + \gamma^{r_2 \dagger} \gamma^{r_1} t) \right] \Big|_{t^{N/2}} \quad (14)$$

$$\equiv \text{Pf}(1 + \gamma^{r_2 \dagger} \gamma^{r_1} t) \Big|_{t^{N/2}} \quad (15)$$

$$\langle \gamma^{r_2} | c_a^\dagger c_b | \gamma^{r_1} \rangle = \left(\frac{\gamma^{r_1} \gamma^{r_2 \dagger} t}{1 + \gamma^{r_1} \gamma^{r_2 \dagger} t} \right)_{ba} \text{Pf}(1 + \gamma^{r_2 \dagger} \gamma^{r_1} t) \Big|_{t^{N/2}} \quad (16)$$

$$\begin{aligned} \langle \gamma^{r_2} | c_a^\dagger c_b^\dagger c_d c_c | \gamma^{r_1} \rangle &= \left(2 \left[\frac{\gamma^{r_1} \gamma^{r_2 \dagger} t}{1 + \gamma^{r_1} \gamma^{r_2 \dagger} t} \right]_{ca} \left[\frac{\gamma^{r_1} \gamma^{r_2 \dagger} t}{1 + \gamma^{r_1} \gamma^{r_2 \dagger} t} \right]_{db} \right. \\ &\quad \left. + t \left[\frac{1}{1 + \gamma^{r_1} \gamma^{r_2 \dagger} t} \right]_{cd} \left[\frac{1}{1 + \gamma^{r_1} \gamma^{r_2 \dagger} t} \right]_{ba} \right) \\ &\quad \times \text{Pf}(1 + \gamma^{r_2 \dagger} \gamma^{r_1} t) \Big|_{t^{N/2}}, \end{aligned} \quad (17)$$

where γ^\dagger corresponds to the Hermitian conjugate of γ and Pf denotes the Fredholm Pfaffian which is the square root of the Fredholm determinant

$$\text{Pf}(1 + \gamma^{r_2 \dagger} \gamma^{r_1} t) = \sqrt{\det(1 + \gamma^{r_2 \dagger} \gamma^{r_1} t)}. \quad (18)$$

It was shown in Ref. [32] that $\text{Pf}(1 + \gamma^{r_2 \dagger} \gamma^{r_1} t)$ is a polynomial of degree $M/2$ since roots of the characteristic polynomial of a product of two antisymmetric matrices are pairwise degenerate.

The total energy can be obtained from the matrix element of the Hamiltonian (\mathcal{H})

$$\langle \gamma^{r_2} | \mathcal{H} | \gamma^{r_1} \rangle = \sum_{ab} h_{ab} \langle \gamma^{r_2} | c_b^\dagger c_a | \gamma^{r_1} \rangle + \frac{1}{2} \sum_{abcd} V_{abcd} \langle \gamma^{r_2} | c_a^\dagger c_b^\dagger c_d c_c | \gamma^{r_1} \rangle, \quad (19)$$

where h corresponds to the sum of the kinetic energy and the external potential and V is the Coulomb repulsion. Contrary to standard GCM schemes, we explicitly obtain the coefficient of $t^{N/2}$ in Eqs. (14)–(17) instead of using the particle projection method [29,32]. A similar projection method [29,32] can be applied to the HFB state to adapt the wave function to the spatial and spin symmetries, which is necessary to obtain accurate ground-state wave functions, but we do not use it in the present study to compare ESTD and full-CI without the symmetry adaptation. We also explicitly obtained the derivatives of the energy with respect to γ , with which to optimize the parameters U_{ia}^r using the Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS) [33–37]. We emphasize that ESTD allows the handling of nonorthogonal AOs by replacing γ_{ab}^r by $(s^{1/2} \gamma^r s^{1/2})_{ab}$ where s is the overlap matrix of AOs, although orthogonality is assumed for simplicity in the present formulation.

When using general complex matrices for U_{ia}^r in practical numerical works, the computation is sometimes impossible to carry on because of the loss of significance in the floating-point arithmetic. To avoid this problem, we make the geminal complex unitary as well as antisymmetric so that all the eigenvalues are equal to unity in the absolute values although the convergence speed is slightly affected. Note that the geminal can be made unitary by the following redefinition of U_{ia}^r : noticing that the second-order permutation tensor ϵ is real antisymmetric, we can apply the canonical transformation [38] by which

$$\epsilon = \bar{U}^T \begin{pmatrix} 0 & \Lambda \\ -\Lambda & 0 \end{pmatrix} \bar{U}, \quad \bar{U} = \bar{U}^T \begin{pmatrix} 0 & \Lambda^{1/2} \\ \Lambda^{1/2} & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} 0 & \Lambda^{1/2} \\ \Lambda^{1/2} & 0 \end{pmatrix} \bar{U}, \quad (20)$$

where Λ and I are, respectively, the real diagonal and identity matrices, \bar{U} is a real orthogonal matrix, and the superscript T means taking the matrix transpose. The geminal [Eq. (10)] then becomes

$$\gamma^r = (U^r)^T \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} U^r, \quad (21)$$

where U^r has been redefined as

$$U^r \rightarrow \begin{pmatrix} 0 & \Lambda^{1/2} \\ \Lambda^{1/2} & 0 \end{pmatrix} \bar{U} U^r. \quad (22)$$

Since U^r was introduced in Eq. (6) as a general matrix to be determined variationally, this redefinition does not affect the resulting formulas. In this way the geminal becomes complex unitary when restricting U_{ia}^r to be complex unitary. We call such restricted ESTD the antisymmetrized unitary geminal powers (AUGPs) hereafter. The variation of the unitary matrix

can be done using the Cayley transform

$$U = (1 + X)^{-1} (1 - X) \quad (23)$$

with a skew-Hermitian matrix X .

III. RESULTS

We have applied ESTD and AUGP to the water molecule and half-filled Hubbard tetramers and compared the total energy with the STD-CI and the full-CI calculations. In the water molecule, we set the geometry condition to

$$O = (0, 0, 0), H = (-1.809, 0, 0), (0.453549, 1.751221, 0) \quad (24)$$

in atomic units. The calculation was done using the basis set STO-3G. The minimal basis set is sufficient for comparison between the different schemes, although larger basis sets will be required for highly quantitative calculations. The Hubbard tetramer has a tetrahedral geometry, and all the sites are connected by the transfer integral of unity ($t = 1$). The Hubbard U was taken as 1, 10^2 , and 10^4 .

Figures 1–3 show the total energy referred to the full-CI calculation, or the residual error vs the tensor rank of the decomposition. The convergence is much faster for ESTD than for STD-CI. For larger U cases, the residual error is initially almost insensitive to the rank and the value amounts to 1–10 when rescaled in units of U . Then the residual error drops suddenly when K is larger than 2 for $U = 10^2$ and 3 for $U = 10^4$. This might indicate that the antiferromagnetic (AF) ground state cannot be described even at a qualitative level when the number of parameters is unreasonably small. For $U = 10^4$, the total energy is almost the same for ESTD, STD-CI, and full-CI when the rank is 6. We postulate that the parameter set for STD-CI happens to become suitable for describing the AF limit only at $K \geq 6$.

Table I shows detailed comparison of the total energy calculated by AUGP and full-CI for the Hubbard tetramer with $U = 10^2$. The calculation was done using double precision with $K = 3$. The residual error is 2.2×10^{-11} while the error was around 10^{-15} for ESTD [Fig. (2)]; in either scheme the error is close to the double precision limit.

Figure 4 shows the residual error of the total energy of the H_2O molecule. ESTD shows strikingly faster convergence than STD-CI. In the ESTD and STD-CI calculations, we used quadruple precision since serious loss of significance occurred with double precision. Because of this problem, we could not easily obtain accurate results using K larger than 4 for ESTD, and thus we could not reduce the residual error below 10^{-5} hartree. Table II shows the result of the same system obtained by using AUGP with double precision. In obtaining the residual

TABLE I. Total energy (in units of t) of the Hubbard tetramer with $U = 100$ obtained by the AUGP with $K = 3$. The result obtained by the full-CI is also shown.

Method	Total energy
AUGP	-0.119880248924950
Exact	-0.119880248946222

TABLE II. Total energy (in hartrees) of H₂O obtained by AUGP. For comparison, the full-CI calculation was done using our own code and the CASSCF calculation using the GAUSSIAN09 package.

Method	Total energy
AUGP, $K = 1$	-61.508355
AUGP, $K = 2$	-72.985876
AUGP, $K = 3$	-73.457483
AUGP, $K = 4$	-75.006008
AUGP, $K = 6$	-75.011870
AUGP, $K = 8$	-75.012385975
AUGP, $K = 10$	-75.012425100
Exact (ours)	-75.012425818
Exact (Gaussian)	-75.012425839

error, we used the total energy from the full-CI calculation as the reference, but the result is almost the same when using the complete active space self-consistent field (CASSCF) calculation performed using GAUSSIAN09 package [39]. The table shows that the error is 6.4×10^{-3} hartree with $K = 4$ and is reduced to 7.2×10^{-7} hartree with $K = 10$. Figure 5 shows a comparison of the residual error of the energy of H₂O for ESTD and AUGP. The figure shows that the residual energy for AUGP is reduced almost linearly in log-scale with respect to K , indicating an exponential convergence. The overall slope of two methods is almost the same. Although AUGP yields slower convergence than ESTD, the AUGP is very stable. It is not difficult to obtain the optimal parameters U'_{ia} even when the calculation is started from random initial values.

IV. CONCLUSION

To efficiently represent the CI wave function, we have applied the canonical decomposition algorithm to the symmetric part of the CI coefficient. Therein the molecular orbitals are fully optimized, without imposing the orthonormality condition, differently for different terms in the decomposition series. The computational scheme, which we call the extended symmetric tensor decomposition (ESTD), is equivalent to the linear combination of the Hartree-Fock-Bogoliubov states or the linear combination of the antisymmetrized geminal power (AGP). By this, we can rearrange the full-CI series into the canonical decomposition series (STD series). By applying ESTD to the water molecule and Hubbard tetramers, we found

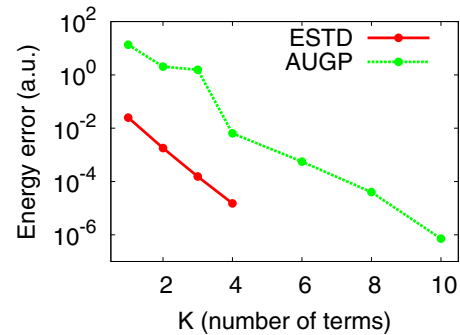


FIG. 5. (Color online) Residual error of the total energy of H₂O obtained by ESTD and AUGP.

that the total energy rapidly converges well within ten terms ($K = 10$) for the STD series. The ESTD calculation was found to be numerically unstable when the number of electrons is increased to as large as 10 because of the loss of significance in the floating-point arithmetic. This problem could be avoided by restricting the MO coefficients to be complex unitary although the convergence speed with respect to tensor rank K was slightly affected. The computational cost of ESTD scales as $K^2 M^5$ where M is the number of MOs. The result suggests that the AGP-based scheme is a promising computational tool for quantum systems. In this context, it will be an important target of future study to clarify how K depends on the complexity and the size of the system. Our calculation was done without parallelization, but an acceleration by a factor of K^2 can be expected because of the almost independent nature of the computation. Further acceleration is expected by applying the tensor decomposition scheme to the two-electron integrals as done in Ref. [40]; such technique may possibly reduce the scaling to $K^2 M^4$.

ACKNOWLEDGMENTS

We acknowledge the Strategic Programs for Innovative Research by the Ministry of Education, Culture, Sports, Science and Technology of Japan and the Computational Materials Science Initiative for the financial support during our research. Some of the calculations were performed in the supercomputer facility at the Institute for Solid State Physics, the University of Tokyo. S.K. is supported by Grant-in-Aid for Research Activity Start-up (Grant No. 25887017) by the Japan Society for the Promotion of Science.

- [1] I. Shavitt, *Mol. Phys.* **94**, 3 (1998).
- [2] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [3] D. C. Cabra, A. Honecker, and P. Pujol, *Modern Theories of Many-Particle Systems in Condensed Matter Physics* (Springer-Verlag, Berlin, 2012).
- [4] In this paper, we use “order” for the number of indices required to represent a tensor such that a matrix is a tensor of order 2 and use “rank” for the number of terms required for the tensor decomposition.
- [5] T. Neuhauser, M. von Armin, and S. D. Peyermhoff, *Theor. Chim. Acta.* **83**, 123 (1992).
- [6] H. Goto, M. Kojo, A. Sasaki, and K. Hirose, *Nanoscale Res. Lett.* **8**, 200 (2013).
- [7] R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- [8] W. Uemura and O. Sugino, *Phys. Rev. Lett.* **109**, 253001 (2012).
- [9] R. A. Harshman and M. E. Lundy, *Comput. Stat. Data Anal.* **18**, 39 (1994).
- [10] J. D. Carroll and J. J. Chang, *Psychometrika* **35**, 283 (1970).
- [11] S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).

- [12] S. R. White, *Phys. Rev. B* **48**, 10345 (1993).
- [13] F. Verstraete and J. Cirac, [arXiv:cond-mat/0407066](https://arxiv.org/abs/cond-mat/0407066); V. Murg, F. Verstraete, and J. I. Cirac, *Phys. Rev. A* **75**, 033605 (2007).
- [14] G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002).
- [15] N. N. Bogoliubov, *Sov. Phys. Usp.* **2**, 236 (1959); M. Baranger, *1962 Cargese Lectures in Theoretical Physics* (Benjamin, New York, 1963).
- [16] A. J. Coleman, *Rev. Mod. Phys.* **35**, 668 (1963).
- [17] A. J. Coleman, *J. Math. Phys.* **6**, 1425 (1965).
- [18] J. V. Ortiz, B. Weiner, and Y. Ohn, *Int. J. Quantum Chem.* **S15**, 113 (1981).
- [19] V. N. Starovecrov and G. E. Scuseria, *J. Chem. Phys.* **117**, 11107 (2002).
- [20] G. E. Scuseria, C. A. Jimenez-Hoyos, T. M. Henderson, K. Samanta, and J. K. Ellis, *J. Chem. Phys.* **135**, 124108 (2011).
- [21] M. Kobayashi, *J. Chem. Phys.* **140**, 084115 (2014).
- [22] B. Weiner and J. V. Ortiz, *J. Chem. Phys.* **117**, 5135 (2002).
- [23] E. Neuscamman, *Phys. Rev. Lett.* **109**, 203001 (2012).
- [24] A. Zen, E. Coccia, Y. Luo, S. Sorella, and L. Guidoni, *J. Chem. Theory Comput.* **10**, 1048 (2014).
- [25] D. Taraha and M. Imada, *J. Phys. Soc. Jpn.* **77**, 114701 (2008).
- [26] M. Bajdich, L. Mitas, G. Drobny, L. K. Wagner, and K. E. Schmidt, *Phys. Rev. Lett.* **96**, 130201 (2006).
- [27] M. Bajdich, L. Mitas, L. K. Wagner, and K. E. Schmidt, *Phys. Rev. B* **77**, 115112 (2008).
- [28] J. J. Griffin and J. A. Wheeler, *Phys. Rev.* **108**, 311 (1957).
- [29] N. Onishi and S. Yoshida, *Nucl. Phys.* **80**, 367 (1966).
- [30] J. Dobaczewski, *Phys. Rev. C* **62**, 017301 (2000).
- [31] T. Mizusaki and M. Oi, *Phys. Lett. B* **715**, 219 (2012).
- [32] K. Neergård and W. Wüst, *Nucl. Phys. A* **402**, 311 (1983).
- [33] C. G. Broyden, *J. Inst. Math. Appl.* **6**, 76 (1970).
- [34] R. Fletcher, *Comp. J.* **13**, 317 (1970).
- [35] D. Goldfarb, *Math. Comp.* **24**, 23 (1970).
- [36] D. F. Shanno, *Math Comp.* **24**, 647 (1970).
- [37] R. Fletcher, in *Practical Methods of Optimization*, Vol. 1 (Wiley, New York, 1980).
- [38] C. Bloch and A. Messiah, *Nucl. Phys.* **39**, 95 (1962).
- [39] M. J. Frisch *et al.*, GAUSSIAN09 (Gaussian, Inc., Wallingford, CT, 2009).
- [40] U. Benedikt, A. A. Auer, M. Espig, and W. Hackbusch, *J. Chem. Phys.* **134**, 054118 (2011).