The extension of the fragment molecular orbital method with the many-particle Green's function

Koji Yasuda[a] and Daisuke Yamaki
Graduate School of Information Science, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan

By using the many-particle Green's function (GF) the extension of the fragment molecular orbital (FMO) method by Kitaura *et al.* [Chem. Phys. Lett. **313**, 701 (1999)] is proposed. It is shown that the partial summation of the cluster expansion of GF reproduces the same extrapolation formula as that of FMO. Therefore we can determine the excitation energy, the transition moment, and the linear response of a molecule from GF approximated with the FMO procedure. It is also shown that no wave function exists which is consistent to the FMO results. The perturbation expansion in which the self-consistent charge approximation defines the unperturbed state is reported. By using it the three-body effects missing in the pair approximation of FMO are analyzed and the corrections to the energy and the reduced density matrices are proposed. In contrast to the previous works these new corrections are not expressed as the addition or the subtraction of the energies of fragments. They are size extensive and require only the quantities available by the FMO calculation. The accuracy of these corrections is validated with the extended Hubbard model and the several test molecules. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2358978]

I. INTRODUCTION

In this paper we report the extension of the fragment molecular orbital method (FMO) proposed by Kitaura and co-workers[1–5] by using the many-particle Green's function (GF). The extension to the excited states and the perturbative correction are derived by reformulating FMO as the cluster expansion of GF. FMO is one of the promising *ab initio* methods for large systems, like biomolecules. It approximates the total energy of a molecule from the energies of the molecular fragments and the pairs of the fragments. These fragments are generated by cutting the localized σ bonds at the sp^3 carbon atoms. The implementation of FMO is rather easy and the computational cost scales formally the square of the system size. This square scaling can be reduced to linear by using the cutoff techniques.[2] This method was applied to various systems including proteins.[2–4] The Hartree-Fock total energies and the dipole moments of polypeptides were estimated very accurately. The electron correlation effects within the pairs of the fragments were also taken into account.[2,6]

The accuracy FMO achieved seems surprising. The bond fragmentation is a drastic procedure. Moreover the truncation of the localized molecular orbitals (LMOs) and the basis overlap between fragments may introduce the error. FMO was derived with the energy decomposition analysis of Kitaura and Morokuma.[7] However, we think that it is difficult to ascribe the quantitative success of FMO to this analysis. In addition there remain the fundamental questions; the existence of the wave function that is consistent to the FMO results and the extension to the less localized systems or the excited states. These questions will be resolved naturally by reformulating FMO as the perturbation expansion of the intersystem interactions.

In this paper we first show that the extrapolation formulas of FMO are reproduced from the summation procedure of GF. In other words we can also approximate GFs of various orders with the FMO procedure. This result is significant because GFs determine the linear response and the excitation energies. We also present the perturbation expansion in which the self-consistent charge (SCC) approximation defines the unperturbed state. This expansion is not trivial because no unperturbed total wave function exists that is consistent to the SCC results. This rigorous reformulation of FMO is very useful for the development of the theory. As an example we analyze the three-body effects missing in the pair approximation of FMO in detail. It is known that in several systems, for example, in the hydrogen-bonded systems, the pair approximation does not give good results.[4] The only way to overcome this deficiency known today is to adopt the more expensive "trimer approximation," in which the energies of the trimers of fragments are also calculated to estimate the total energy.[4,8] By using the technique developed in this paper we derive the three-body corrections to the energy and the reduced density matrices (RDMs) perturbatively. These corrections are inherently size extensive and the results of the extended Hubbard model and molecules are very promising.

II. METHOD

A. Fragment molecular orbital method

We briefly summarize the procedure of FMO.[1] A molecule is first divided into N fragments by using the truncated LMOs. Electrons are distributed to each fragment. The energy E_I and the molecular orbitals (MOs) of the Ith fragment are calculated self-consistently under the electrostatic potential generated by other fragments (SCC). The auxiliary po-

[a]Electronic mail: yasudak@is.nagoya-u.ac.jp

     125, 154101-1

tential is added to localize MOs within each fragment if necessary. When the self-consistent charge is obtained, the energy E_{IJ} of the pair of the fragments I and J is calculated under the electrostatic potential generated by other fragments. FMO approximates the total energy as

$$E = \sum_{I>J} E_{IJ} - (N-2)\sum_{I} E_{I}. \qquad (1)$$

The energies E_I and E_{IJ} can be the correlated ones.[2,6] We assume that the expectation value of an operator $\hat{O}$, for example, the electron density ρ, the dipole moment,[2] or RDM, is given by the same formula as the energy.

$$O = \sum_{I>J} O_{IJ} - (N-2)\sum_{I} O_{I}. \qquad (2)$$

We point out the subtle aspects of FMO. Basis functions on the different fragments are not orthogonal, but the monomer SCC calculation assumes the orthogonality. Thus the converged MOs on a fragment are not orthogonal to those on other fragments. The SCC calculation needs the electrostatic potential generated by the total electron density. This density is calculated with the occupied MOs of all the fragments. However, since these MOs are not orthogonal, the total electron density contains the unphysical term which violates the exclusion principle. The eigenvalues of the total 1-RDM of the SCC calculation reveal this violation. This total 1-RDM is given as the direct sum of the 1-RDMs of the fragments. The eigenvalues of 1-RDM derived from the antisymmetric wave function should be in the range between 0 and 1.[9] The eigenvalues of the Hartree-Fock 1-RDM should be either 0 or 1. However, we found that the largest eigenvalue becomes about 1.3 in a certain system (see the Numerical Results section). Thus no antisymmetric wave function exists that reproduces 1-RDMs of the fragments. The violation of the exclusion principle in the 1-RDM of pair approximation is much smaller, yet it exists, implying the nonexistence of the wave function.

B. Cluster expansion of Green's function

The FMO is derived naturally from the cluster expansion of the energy and Green's functions. This cluster expansion differs from the usual coupled cluster expansion of the wave function.[10] For simplicity we focus on the trimer composed of the fragments A, B, and C. The total Hamiltonian is

$$H = H_0 + H_1,$$
$$H_0 = H_A + H_B + H_C, \qquad (3)$$
$$H_1 = V_{AB} + V_{BC} + V_{CA} + V_{ABC},$$

where $H_I(I=A,B,C)$ describes the isolated monomer fragment under the electrostatic potential of other fragments; V_{IJ} is the difference between the dimer Hamiltonian H_{IJ} and the sum of the monomer Hamiltonians H_I+H_J. We neglect the term V_{ABC} and treat H_1 as the perturbation. We adopt as the one-particle basis the localized orthonormal spin-orbital and express the creation and the annihilation operators as $x_i^\dagger$ and x_i, respectively. The operators in the Heisenberg representa-

tion are denoted as $x_i^\dagger(t)$ and $x_i(t)$. Later we will extend the results to the nonorthogonal basis.

To derive the pair approximation [Eq. (2)] the perturbation expansion of the one-particle Green's function g (Ref. 11) is the most convenient,

$$g(jt,is) = (-i)\langle \Psi | T[x_j(t)x_i^\dagger(s)] | \Psi \rangle, \qquad (4)$$

where T and Ψ are the time ordering and the exact ground state, respectively. The limit of the zero time interval ($s \to 0^+, t \to 0^-$) of g gives 1-RDM,

$$\gamma(j,i) = \langle \Psi | x_i^\dagger x_j | \Psi \rangle. \qquad (5)$$

The cluster expansion divides the total g,

$$g = \underset{(a)}{g_A + g_B + g_C} + \underset{(b)}{\delta g_{AB} + \delta g_{BC} + \delta g_{CA}} + \underset{(c)}{\delta g_{ABC}}, \qquad (6)$$

into the contribution from the isolated fragments (a), the two-body correction (b), and the three-body correction (c). The direct sum of $g_I(I=A,B,C)$ gives the unperturbed Green's function. The term δg_{AB} is defined as

$$\delta g_{AB} = g_{AB} - (g_A + g_B), \qquad (7)$$

where g_{AB} is GF of the pair of the fragments. Equation (7), which is the sum of the diagrams that contain V_{AB} alone, describes the multiple scattering process of electrons by V_{AB}. By substituting Eq. (7) into Eq. (6) and by setting $\delta g_{ABC}=0$, we can reproduce the pair FMO formula,

$$g = g_{AB} + g_{BC} + g_{CA} - (g_A + g_B + g_C). \qquad (8)$$

Since each term in these equations represents the sum of the distinct Feynman diagrams there is no overcounting of the terms. The FMO formula [Eq. (2)] holds for the every quantity O which we can represent as the sum of the distinct diagrams.

When the basis functions on a fragment are not orthogonal to those on other fragments, the difference of the overlap matrices also contributes to the perturbation.

$$S = S_0 + S_1,$$
$$S_0 = S_A + S_B + S_C, \qquad (9)$$
$$S_1 = \delta S_{AB} + \delta S_{BC} + \delta S_{CA}.$$

$S_I(I=A,B,C)$ is the atomic orbital (AO) overlap matrix of the fragment, while $\delta S_{IJ}=S_{IJ}-(S_I+S_J)$ is the off-diagonal block of the overlap matrix of the pair of the fragments $I+J$. For simplicity we focus on the FMO calculation under the Hartree-Fock approximation. Extension to the general case is straightforward by using the double perturbation. We assume that the Hamiltonians in Eq. (3) represent the AO Fock matrices and we ignore the three-body term δV_{ABC}.

FMO is reproduced as the cluster expansion of the one-particle Green's function in AO basis,

$$g(\omega)_{kl} = \sum_{i} \frac{C_k^{(i)} C_l^{(i)}}{\omega - \epsilon_i + i\eta\,\mathrm{sgn}(\epsilon_i - \mu)}, \qquad (10)$$

where ϵ_i and $C_k^{(i)}$ are the orbital energy and the MO coefficient, respectively. As usual η is the positive infinitesimals, $\mathrm{sgn}(x)$ is the sign of the argument x, and μ is the chemical

potential, which we choose an arbitrary energy between highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO). This Green's function is the Fourier transform of Eq. (4) with respect to the time interval $t-s$. The 1-RDM elements in AO basis are given by the complex integral along the closed contour C in the upper half-plane,

$$\gamma_{kl} = \frac{1}{2\pi i} \int_C g(\omega)_{kl}\, d\omega. \qquad (11)$$

The exact and the unperturbed Green's functions satisfy the following matrix equations:

$$(\omega S - H)g(\omega) = \mathbf{1}, \qquad (12a)$$

$$(\omega S_0 - H_0)g_0(\omega) = \mathbf{1}, \qquad (12b)$$

where $\mathbf{1}$ denotes the unit matrix. The unperturbed Green's function is given by substituting ϵ_i and $C_k^{(i)}$ of Eq. (10) with the orbital energies and the MO coefficients of the monomer SCC calculation, respectively. Hence the 1-RDM derived from this g_0 equals to the direct sum of the 1-RDMs of the fragments.

By using Eq. (12) one can verify that the same Dyson equation holds as the orthogonal basis functions,

$$g = g_0 + g_0(H_1 - \omega S_1)g, \qquad (13)$$

except that the perturbation contains the extra term $-\omega S_1$. Thus we conclude that the FMO formula [Eq. (2)] holds also for Green's function in AO basis.

C. Related approximations

By using the energies of a fragment and a pair of fragments the original FMO energy E_{FMO} is given by Eq. (1). The energies E_I and E_{IJ} are calculated under the self-consistent Coulomb potential of other fragments. Here we discuss other three approximation methods. In the first method we require the self-consistency under the Coulomb and exchange potentials of other fragments [self-consistent Coulomb and exchange (SCCE)]. Equation (1) gives the slightly different energy E_{SCCE}. In the second method by using Eq. (2) we extrapolate the total 1-RDM γ and the 2-RDM,

$$\Gamma^{(2)}(j_1,j_2;i_1,i_2) = \tfrac{1}{2}\langle \Psi | x_{i1}^\dagger x_{i2}^\dagger x_{j2} x_{j1} | \Psi \rangle, \qquad (14)$$

from those of a fragment and a pair of fragments [density matrix (DM) expansion].

$$\gamma = \sum_{I>J} \gamma_{IJ} - (N-2)\sum_I \gamma_I, \qquad (15a)$$

$$\Gamma_{DM}^{(2)} = \sum_{I>J} \Gamma_{IJ}^{(2)} - (N-2)\sum_I \Gamma_I^{(2)}. \qquad (15b)$$

The total energy E_{DM} is calculated as the expectation value of the total Hamiltonian; $E_{DM} = \mathrm{Tr}(v\,\gamma + w\Gamma_{DM}^{(2)})$, where v and w are the one- and two-electron integrals, respectively. In the third method instead of 2-RDM we extrapolate the cumulant of 2-RDM,

$$\Delta = \sum_{I>J} \Delta_{IJ} - (N-2)\sum_I \Delta_I \qquad (15c)$$

(cumulant expansion). Cumulant is defined as the difference between 2-RDM and the antisymmetric product of 1-RDM.

$$\Delta_I = \Gamma_I^{(2)} - \gamma_I \wedge \gamma_I,$$

$$\gamma \wedge \gamma = \tfrac{1}{2}\{ \gamma(j_1,i_1)\gamma(j_2,i_2) - \gamma(j_1,i_2)\gamma(j_2,i_1)\}. \qquad (16)$$

The 1-RDM and the cumulant determine the 2-RDM,

$$\Gamma_{cum}^{(2)} = \Delta + \gamma \wedge \gamma, \qquad (17a)$$

and the energy $E_{cum} = \mathrm{Tr}(v\,\gamma + w\Gamma_{cum}^{(2)})$.

Generally these energies are different to each other even if we use the same 1- and 2-RDMs of fragments and pairs of fragments. The energy difference is attributed to the difference among the 2-RDMs. After short calculation we have

$$\Gamma_{DM}^{(2)} = \Delta + \gamma_0 \wedge \gamma_0$$
$$+ \sum_{I>J} \{(\gamma_I + \gamma_J) \wedge \delta\gamma_{IJ} + \delta\gamma_{IJ} \wedge (\gamma_I + \gamma_J) + \delta\gamma_{IJ} \wedge \delta\gamma_{IJ}\}, \qquad (17b)$$

where $\gamma_0 = \sum_I \gamma_I$ and $\delta\gamma_{IJ} = \gamma_{IJ} - (\gamma_I + \gamma_J)$. Similarly we can show that the 2-RDMs,

$$\Gamma_{FMO}^{(2)} = \Gamma_{DM}^{(2)} + \tfrac{1}{2}\sum_{I>J}\sum_{K\neq I,J} \{\delta\gamma_{IJ}(j_1,i_1)\gamma_K(j_2,i_2)$$
$$+ \gamma_K(j_1,i_1)\delta\gamma_{IJ}(j_2,i_2)\}, \qquad (17c)$$

$$\Gamma_{SCCE}^{(2)} = \Gamma_{DM}^{(2)} + \sum_{I>J}\sum_{K\neq I,J} (\delta\gamma_{IJ} \wedge \gamma_K + \gamma_K \wedge \delta\gamma_{IJ})$$
$$= \Gamma_{cum}^{(2)} - (\gamma - \gamma_0) \wedge (\gamma - \gamma_0), \qquad (17d)$$

give the energies $E_{FMO} = \mathrm{Tr}(v\gamma + \omega\Gamma_{FMO}^{(2)})$ and $E_{SCCE} = \mathrm{Tr}(v\gamma + w\Gamma_{SCCE}^{(2)})$, respectively. By comparing them the most and the least accurate energies are expected to be E_{cum} and E_{DM}, respectively. The difference between E_{SCCE} and E_{cum} is the second order in $\delta\gamma_{IJ}$.

Inadomi *et al.* compared the accuracy of E_{cum} with E_{FMO} under the Hartree-Fock approximation for polypeptides.[12] They found that E_{cum} is worse than E_{FMO}. They attributed this error to the superfluous terms such as Coulomb and exchange terms between two fragments. However, E_{cum} does not have the superfluous terms, like the overcounted diagrams, and it contains more diagrams than that of FMO theoretically. The results by Inadomi *et al.* would be attributed to the error cancellation.

FMO and the related approximations discussed above inevitably suffer from the nonexistence of the wave function. Each of them does not have wave functions because it ignores the basis overlap between fragments. Moreover even if the basis functions are orthogonal there exists no wave function. This is the direct consequence of Marcinkiewicz's theorem.[13] It states that no wave function reproduces the truncated cluster expansion of RDMs [Eq. (2)]. A consequence of this nonexistence is the lack of the consistency between RDMs approximated by Eq. (2). The 2-RDMs $\Gamma_{cum}^{(2)}$, $\Gamma_{SCCE}^{(2)}$, and $\Gamma_{FMO}^{(2)}$ do not reduce to the total 1-RDM γ by the

contraction. Except for the cumulant expansion the total RDMs and cumulant [Eq. (15)] do not satisfy the general definition of the cumulant [Eq. (16)]. Under the Hartree-Fock approximation all the cumulants Δ_I and Δ_{IJ} vanish. Then the breakdown of Eq. (16) indicates that the 2-RDMs are not equal to the antisymmetric product of 1-RDM γ, indicating the nonexistence of the wave function.

D. Excitation energy

The FMO formulas [Eqs. (1) and (2)] hold for the total energy and GF, because they are given as the sums of the distinct terms (the Feynman[11] or the Goldstone diagrams.[14]). On the other hand since the excitation energy is not given as the sum of diagrams, we cannot justify applying Eq. (1) for the excitation energy. It is the self-energy and the polarization propagator which FMO can yield valid approximation of. For clarity we focus on the ionization energy. The same story holds for the excitation energy. The self-energy Σ of a molecule determines the ionization energy ϵ and the transition amplitude $\psi(r)$ as

$$\{F + \Sigma(\epsilon)\}\psi = \epsilon\psi,$$

$$\psi(r) = \langle \Psi_i | \phi_\sigma(r) | \Psi_g \rangle,$$

$$(18)$$

where F is the total Fock operator of the ground state, $\phi_\sigma(r)$ is the field annihilation operator, and Ψ_g and Ψ_i are the ground and the ionized states, respectively. The self-energy takes account of the effects beyond Koopman's theorem. Several methods were reported[15] to calculate the self-energy of a fragment Σ_I or that of a pair of fragments Σ_{IJ}. We can approximate the total self-energy with the FMO formula,

$$\Sigma = \sum_{I>J} \Sigma_{IJ} - (N-2)\sum_I \Sigma_I. \qquad (19)$$

Because of the energy dependence Eq. (18) should be solved iteratively. The size of the matrices also causes problem for large molecules. We can split this equation into the small equations as follows.

The matrix $F' = F + \Sigma(\epsilon)$ and the vector ψ are divided into small pieces denoted as F'_{IJ} and ψ_I. They satisfy the equation

$$\{F'_{AA} + F'_{AB}(\epsilon - F'_{BB})^{-1}F'_{BA}\}\psi_A = \epsilon\psi_A, \qquad (20)$$

where A is a fragment of a molecule and B is the rest. Hence instead of a large equation we have many small equations. Since F'_{BB} is close to the block-diagonal matrix we can approximate the second term of Eq. (20) as

$$F'_{AB}(\epsilon - F'_{BB})^{-1}F'_{BA} \approx \sum_{I \neq A} F'_{AI}(\epsilon - F'_{II})^{-1}F'_{IA}, \qquad (21)$$

where the sum I runs over the molecular fragments. The series expansion of the inverse matrix in terms of the off-diagonal elements gives the systematic correction to it.

E. Three-body correction to the density matrices

The reformulation of FMO in terms of GF enables us to derive the explicit formula of the three-body correction to the pair FMO approximation. The leading term of it is the sec-

ond order of the perturbation, which contains V_{AB} and V_{BC}, for example. This term describes the following one-electron process: V_{AB} scatters an electron from the fragment A to B, and then V_{BC} scatters it from B to C. Since the correction δg_{AB} describes the scattering amplitude by V_{AB}, such sequential process is expressed as

$$i\sum_b \delta g_{AB}(as,b0)\{\theta(s)\theta(-t) - \theta(t)\theta(-s)\}\delta g_{BC}(b0,ct),$$

$$(22)$$

where θ is the step function and the sum runs over the spin-orbitals on the fragment B. The first term with $\theta(s)\theta(-t)$ describes the following process: an electron on C at time $t < 0$ is scattered to B, then it is scattered to reach A at time $s > 0$.[16] The second term describes the process of a hole. Equation (22) also contains some higher-order terms of the perturbation, for example, first order of V_{AB} and second order of V_{BC}. Clearly there is no overcounting of the diagram and all of them are missing in the pair approximation. Because of the linked-cluster theorem the correction is size extensive. Thus the sum of the sequential process provides the most important correction to the pair approximation.

By taking the proper time limit the following matrix product gives the correction to 1-RDM:

$$\delta\gamma_{3\,\text{body}} = \sum_{I>J}\sum_{K\neq I,J} \delta\gamma_{IK}P\delta\gamma_{KJ} = \delta\gamma P\delta\gamma - \sum_{I,J}\delta\gamma_{IJ}P\delta\gamma_{JI},$$

$$(23a)$$

$$\delta\gamma = \sum_{I\neq J}\delta\gamma_{IJ}, \qquad (23b)$$

where P is the projector. It is given as

$$P = \theta(2\gamma_0 - \mathbf{1}), \qquad (24a)$$

$$P = 2S_0\gamma_0 S_0 - S_0, \qquad (24b)$$

in orthonormal spin-orbital and in AO basis, respectively. Note that all the quantities needed are available from the pair FMO calculation. The block structure of P and $\delta\gamma_{IJ}$ may be used to reduce the computational cost.

By using the cumulant wave operator Kladko and Fulde derived the three-body correction[8] to the incremental method.[5] The difference between FMO and the incremental method was discussed in Ref. 2. They made a valuable observation that Eqs. (1) and (2) were derived as the approximation of the Faddeev three-body equation of the scattering amplitude.[8] Their equation might produce a correction similar to Eq. (23). However, the present Green's function method yields the feasible approximation much easily. It establishes the clear connection between the scattering amplitude and RDMs. In addition the diagrammatic rule provides the convenient way to sum up distinct terms.

F. Electron correlations

When we have the correlated wave functions of a fragment and a pair of fragments we can also approximate the three-body effects to the electron correlations. We derive the corrections to the connected two-particle Green's function

and the connected 2-RDM, in which information on the correlations is encoded. Our result is different to the previous ones, in which electron correlations within the pairs of the fragments were taken into account.[2] First define the vertices of GF and RDM as

$$
\begin{aligned}
G(j_1t_1, j_2t_2; i_1s_1, i_2s_2) &- G(j_1t_1, j_2t_2; i_2s_2, i_1s_1) = \\
&- \langle \Psi | T[x_{j_1}(t_1) x_{j_2}(t_2) x_{i_2}^\dagger(s_2) x_{i_1}^\dagger(s_1)] | \Psi \rangle \\
&- g(j_1t_1, i_1s_1) g(j_2t_2, i_2s_2) + g(j_1t_1, i_2s_2) g(j_2t_2, i_1s_1),
\end{aligned}
$$
$$(25a)$$

$$
\begin{aligned}
2\Gamma^{(2)}(j_1, j_2; i_1, i_2) &= \gamma(j_1, i_1)\gamma(j_2, i_2) - \gamma(j_1, i_2)\gamma(j_2, i_1) \\
&+ \Gamma(j_1, j_2; i_1, i_2) - \Gamma(j_1, j_2; i_2, i_1). \quad (25b)
\end{aligned}
$$

The vertex G represents the scattering amplitude of the two electrons. The proper time limit of G gives Γ, in the same way as the one-particle Green's function. The same cluster expansion as Eq. (6) recovers Eq. (2) for the vertex. We can also approximate the missing three-body terms as follows.

Since the perturbation H_1 contains the one- and two-electron operators, there are two kinds of processes. The first one is bilinear in δg_{IK} and δG_{KJ} (or in $\delta\gamma_{IK}$ and $\delta\Gamma_{KJ}$). It describes the process that two electrons on the fragment J collide and one of them is scattered to K, then it is scattered to I. We express this amplitude by connecting one of four external legs of δG_{KJ} with that of δg_{IK} or equivalently $\delta\Gamma_{KJ}$ with $\delta\gamma_{IK}$. The following matrix product connects the first leg of $\delta\Gamma_{KJ}$ with the second one of $\delta\gamma_{IK}$:

$$
\sum_{k,k'} \delta\gamma_{IK}(i_1, k') P(k', k) \delta\Gamma_{KJ}(k, j_2; j_3, j_4). \quad (26)
$$

Other three operations that connect one of three other external legs of $\delta\Gamma$ with $\delta\gamma$ are defined similarly. We express the sum of these four terms symbolically as $\delta\gamma_{IK}P\delta\Gamma_{KJ}$. The sum of such terms over the pairs of the fragments gives the correction bilinear in $\delta\gamma$ and $\delta\Gamma$,

$$
\delta\Gamma_{3body-1} = \delta\gamma P \delta\Gamma - \sum_{I,J} \delta\gamma_{IJ} P \delta\Gamma_{JI}, \quad (27a)
$$

$$
\delta\Gamma = \sum_{I \neq J} \delta\Gamma_{IJ}. \quad (27b)
$$

Next we consider the correction quadratic in $\delta\Gamma$. It represents the successive two-electron scatterings on the pairs of the fragments, $I+K$ and $J+K$. The total amplitude is given by connecting the two amplitudes δG_{IK} and δG_{KJ} (or $\delta\Gamma_{IK}$ and $\delta\Gamma_{KJ}$). As shown in Fig. 1, there are five methods, labeled as s, u, c, r, and l, to connect two vertices.[17] For example, the s operation gives the following sequential amplitude:

$$
\begin{aligned}
\delta\Gamma_{IK} s \delta\Gamma_{KJ} = \sum_{k_i} \delta\Gamma_{IK}(i_1, i_2; k_1, k_2) P(k_1, k_3) \\
\times P(k_2, k_4) \delta\Gamma_{KJ}(k_3, k_4; j_1, j_2). \quad (28)
\end{aligned}
$$

Other connecting operations are defined similarly. Note that the c operation has the spin-sum factor of -2.[17] The sum of such terms over the pairs of the fragments gives the correction quadratic in $\delta\Gamma$,
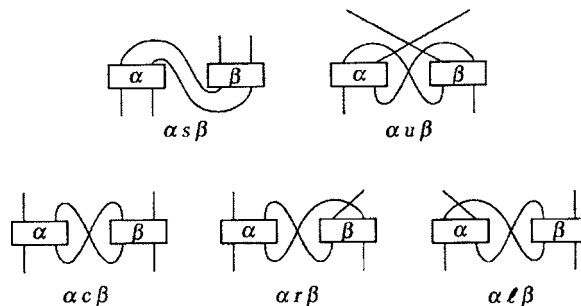

FIG. 1. Five connecting operations of s, u, c, r, and l channels in Ref. 17. α and β are the generic vertex diagrams.

$$
\delta\Gamma_{3body-2} = \sum_{x \in s,u,c,r,l} \left\{ \delta\Gamma x \delta\Gamma - \sum_{I,J} \delta\Gamma_{IJ} x \delta\Gamma_{JI} \right\}. \quad (29)
$$

The sum $\delta\Gamma_{3body-1} + \delta\Gamma_{3body-2}$ gives the three-body correction to the vertex Γ. Finally, 2-RDM is given by Eq. (25b) and the energy is calculated as usual,

$$
E = \mathrm{Tr}(v\gamma + w\Gamma^{(2)}). \quad (30)
$$

G. Cluster expansion of the energy

When one is interested in the accurate energy alone, one can calculate the energy correction by applying Goldstone's diagrammatic technique.[14] To derive the perturbation series of the ground-state energy we take the Fock operator of the SCC calculation as the unperturbed Hamiltonian. Two complicated issues should be solved. (i) Because of the violation of the exclusion principle no wave function exists that is consistent to the SCC results. (ii) The difference S_1 of the overlap matrices between the unperturbed and the exact Hamiltonians should be also treated as the perturbation.

We can resolve the first issue by introducing the fictitious unperturbed system; it has the same AO Hamiltonian and overlap matrices, but the basis functions are replaced with the new fictitious ones χ_k'. In matrix notation they are given as

$$
\chi' = S_0^{1/2}(S_0 + S_1)^{-1/2}\chi. \quad (31)
$$

Note that the unperturbed energy does not change because matrix elements completely determine it. The overlap matrix of the basis functions χ_k is $S_0 + S_1$, while that of χ_k' is S_0. The second issue can be resolved by changing the basis functions χ_k of the exact Hamiltonian to the fictitious ones χ_k'. To keep the exact energy invariant this basis transformation induces the changes of the one- and two-electron integrals of the exact Hamiltonian. New one-electron integrals are given as

$$
v' = S_0^{1/2}(S_0 + S_1)^{-1/2} v (S_0 + S_1)^{-1/2} S_0^{1/2}. \quad (32)
$$

New two-electron integrals are defined similarly.

We take as the perturbation the difference between these two fictitious Hamiltonians to derive the perturbation series of the exact ground-state energy. The two-electron part of the perturbation is the transformed two-electron integrals w', while the one-electron part is the transformed one-electron integrals v' minus the Fock operator of the SCC calculation $f_{monomer}$. The zeroth-order energy E_0 is just the twice of the sum of the occupied orbital energies of the SCC calculation.

The first-order energy E_1 is the expectation value of the perturbation with respect to the unperturbed state.

$$E_0 + E_1 = 2\,\mathrm{Tr}(v'\,\gamma_0 + w'\,\gamma_0 \wedge \gamma_0)$$
$$= 2\,\mathrm{Tr}(v\,\gamma_0' + w\,\gamma_0' \wedge \gamma_0'),$$

$$\gamma_0 \wedge \gamma_0 = \gamma_0(j_1|i_1)\,\gamma_0(j_2|i_2) - \tfrac{1}{2}\,\gamma_0(j_1|i_2)\,\gamma_0(j_2|i_1). \tag{33}$$

We used the following transformed 1-RDM:

$$\gamma_0' = (S_0 + S_1)^{-1/2} S_0^{1/2} \gamma_0 S_0^{1/2} (S_0 + S_1)^{-1/2}. \tag{34}$$

Now we show that the partial summation of this perturbation expansion reproduces the same FMO energy as in Ref. 1. If we sum up all the diagrams except for those representing the electron correlations, we have the total RHF energy. We substitute the perturbation V in these diagrams with $\Sigma_{I>J} V_{IJ}$. The sum of the terms that contain only V_{IJ} reproduces the energy $E_{IJ} - (E_I + E_J)$, because it represents the pair of fragments under the electrostatic potential of other fragments. Thus it is clear that by selecting the terms we can reproduce the FMO energy.

The three-body correction to the FMO energy under the RHF approximation is derived as follows. The two-electron part of the perturbation is absorbed in the one-electron part, and hence the AO matrix elements of the perturbation are given as

$$V(i|j) = v'(i|j) + \{2w'(ij|kl) - w'(ik|jl)\}\gamma_0(k|l) - f_{\mathrm{monomer}}. \tag{35}$$

The second- and third-order perturbation energies are expressed concisely in the MO basis,

$$E_2 = 2\sum_{ia} \frac{V(i|a)^2}{\epsilon_i - \epsilon_a}, \tag{36}$$

$$E_3 = -2\sum_{ija} \frac{V(a|i)V(i|j)V(j|a)}{(\epsilon_i - \epsilon_a)(\epsilon_j - \epsilon_a)} + 2\sum_{iab} \frac{V(a|i)V(i|b)V(b|a)}{(\epsilon_i - \epsilon_b)(\epsilon_i - \epsilon_a)}, \tag{37}$$

where as usual $i, j, \ldots$ denote the occupied orbitals and $a, b, \ldots$ the virtual orbitals. Because of the linked-cluster theorem the energy corrections are size extensive. To calculate the correction of a large molecule the square inverse of $S_0 + S_1$ can be obtained with the series expansion or the iterative formula, and AO Laplace technique [18] can be used to avoid the integral transform.

Although the perturbation expansion converges rather rapidly, it is desirable to sum up the selected infinite series. Since the pair FMO approximation takes the two-body scatterings into account, it is worth from Eqs. (33), (36), and (37) to extract the three-body correction missing in it. A simple method is to calculate the energy $E_0 + E_1 + E_2 + E_3$ for the fragments, pairs of the fragments, and the target molecule. By using these energies the pair energy is calculated with Eq. (1). The difference of the total energy and this pair energy provides the leading three-body correction to the FMO energy.

III. NUMERICAL RESULTS

A. Extended Hubbard model

In this section we examine the numerical accuracy of the three-body corrections derived in the previous section. The three-body effects missing in the pair approximation of FMO originate from (i) the three-body term of the Hamiltonian V_{ABC}, (ii) the three-body scattering analyzed in the previous section, and (iii) others, for example, the basis set superposition.

To make the transparent analysis we first apply the present method to the six-site periodic one-dimensional extended Hubbard model, a simple model of the interacting three hydrogen molecules.

$$H = H_1 + H_2,$$

$$H_1 = -\sum_{i=1,\sigma}^{6} \{1 - \delta(-1)^i\}(a_{i\sigma}^{\dagger} a_{i+1\sigma} + \text{c.c.}), \tag{38}$$

$$H_2 = U\sum_{i=1}^{6} n_{i\alpha} n_{i\beta} + V\sum_{i=1,\sigma}^{6} \{1 - \delta(-1)^i\} n_i n_{i+1},$$

We adopt the convention, $n_{i\sigma} = a_{i\sigma}^{\dagger} a_{i\sigma}$, $n_i = n_{i\alpha} + n_{i\beta}$, $a_{7\sigma} = a_{1\sigma}$, and $n_7 = n_1$. The purpose of the present study is not to compare the relative importance of the three origins but to distinguish the second from the rest and examine the approximation of it. This Hubbard model is suitable because it is free from the first and the third origins.

Because of the symmetry the occupied and the empty orbitals of the isolated ith fragment are $(a_{2i-1}^{\dagger} + a_{2i}^{\dagger})/\sqrt{2}$ and $(a_{2i-1}^{\dagger} - a_{2i}^{\dagger})/\sqrt{2}$, respectively. These six MOs are orthonormal and were used to define the three fragments. We calculated the exact energy, 1- and 2-RDMs, and the vertex Γ of Eq. (25b) with the full configuration interaction (CI) method [19] for the fragments and the pairs of the fragments. The energy and 1- and 2-RDMs of the pair approximation (labeled as "pair") were calculated with Eqs. (1) and (2). The vertex function of 2-RDM Γ_{pair} was also calculated with Eq. (2). This and 1-RDM γ_{pair} define another 2-RDM $\Gamma_{\mathrm{cum}}^{(2)}$ and the energy E_{cum} through Eq. (30). The three-body correction [Eq. (23)] was added to γ_{cum} to get a new 1-RDM γ_{3body}. Similarly Eqs. (27a) and (29) were added to Γ_{cum} to get a new vertex Γ_{3body}. They define 2-RDM $\Gamma_{\mathrm{3body}}^{(2)}$ and the energy E_{3body} through Eq. (30).

We examined the errors of the energy and 1- and 2-RDMs for two systems: Hückel model with bond alternation ($U = V = 0$, no electron correlations) and the extended Hubbard model with $U = 1$ and $V = 1/2$. Table I compares these errors as a function of bond alternation δ; $\delta = 1$ corresponds to the isolated two-site model, while $\delta = 0$ the fully delocalized one. As shown our "3body" results were always the most accurate for both the Hückel and the Hubbard models. Energy errors were about half of the pair approximation for the strong intersystem interaction ($\delta = 0$) and two to three orders of magnitude smaller for the weak interaction ($\delta = 0.95$). The energy of the cumulant expansion E_{cum} is worse than those of pair FMO E_{pair} for the Hubbard model.

TABLE I. Errors of the energy, spin-reduced 1- and 2-RDMs of the Hückel model, and the extended Hubbard model as a function of the bond alternation δ. The error of RDM is measured by the square norm of the difference between the approximate and the exact ones. Numbers in the square brackets indicate powers of 10.

	Energy			1-RDM		2-RDM		
δ	Pair	Cumulant	3body	Pair	3body	Pair	Cumulant	3body
	Hückel model ($U=V=0$)							
0	5.84[−1]	5.84[−1]	2.00[−1]	6.76[−1]	5.29[−1]	3.33	1.58	1.24
0.5	4.32[−2]	4.32[−2]	7.72[−3]	1.09[−1]	9.74[−2]	1.06	2.55[−1]	2.28[−1]
0.9	2.10[−4]	2.10[−4]	6.79[−6]	2.71[−3]	2.43[−3]	1.48[−1]	6.35[−3]	5.69[−3]
0.95	2.48[−5]	2.48[−5]	3.94[−7]	6.40[−4]	5.73[−4]	7.11[−2]	1.50[−3]	1.34[−3]
	Extended Hubbard model ($U=1$, $V=1/2$)							
0	4.36[−1]	7.88[−1]	2.82[−1]	7.73[−1]	6.75[−1]	3.40	1.85	1.57
0.5	−1.91[−2]	8.75[−2]	6.81[−3]	9.09[−2]	7.83[−2]	9.58[−1]	2.26[−1]	1.85[−1]
0.9	−1.78[−3]	2.48[−3]	6.09[−6]	2.05[−3]	1.60[−3]	1.38[−1]	5.21[−3]	3.82[−3]
0.95	−4.58[−4]	6.02[−4]	4.46[−7]	4.82[−4]	3.72[−4]	6.67[−2]	1.23[−3]	8.88[−4]

In the Hückel model the pair approximation and the cumulant expansion give the same energy because only the 1-RDM enters the energy formula. Interestingly they give different 2-RDM. $\Gamma^{(2)}_{\text{pair}}$ shows larger error than $\Gamma^{(2)}_{\text{cum}}$. The error was attributed to the 2-RDM elements $\Gamma^{(2)}(I,J,I,K)$ over the three fragments I, J, and K. Although these 2-RDM elements are not zero, in general, they are zero by definition in the pair approximation. On the other hand in the cumulant expansion the antisymmetric product of 1-RDM gives the nonzero contribution to it. This implies that the connected quantities should be approximated, like the energy, 1-RDM, and the cumulant, not 2-RDM. On the other hand a good aspect of the pair approximation of 2-RDM is that the 1- and 2-RDMs satisfy the correct contraction relations: partial integration of 2-RDM gives the 1-RDM. This does not hold for the cumulant expansion and 3body approximations, although the errors become smaller as we apply higher-order corrections.

B. Cyclohexane and homologues

Next we apply the corrections to the cyclohexane C_6H_{12} and the homologues $C_5BH_{12}^-$ and $C_5NH_{12}^+$. They were chosen to highlight the three-body effects. It was reported that the water clusters and the polypeptides show the measurable three-body effects.[2] However, they are not convenient for benchmark because of their sizes. Cyclohexane is small enough, yet it shows the measurable three-body effects.

STO-3G basis set[20] and the experimental geometry of cyclohexane were used. A molecule was divided in three fragments, three ethylene units. We used the same procedure as in Ref. 1 to define the molecular fragments. AOs on C_1, C_3, and C_5 were transformed to the natural hybridized orbitals (HOs) of methane which were taken from the ABINIT-MP program.[21] These HOs and other AOs were divided in three to define the fragments. The first fragment consists of (i) all the HOs and AOs on the C_2H_4 units except for a HO on C_1 toward C_6 and (ii) a HO on C_3 towards C_2. The rest of the fragments were defined similarly. The Hamiltonian and the overlap matrices in terms of them define the splitting of the total Hamiltonian as Eq. (3). The molecular fragments of

$C_5BH_{12}^-$ and $C_5NH_{12}^+$ are defined similarly. FMO calculation was done essentially the same way as in Ref. 1. FMO energy and 1-RDM were calculated with Eqs. (1) and (2). The three-body correction of 1-RDM was calculated with Eqs. (23) and (24b). The third-order perturbation energies, Eqs. (33), (36), and (37), were calculated for fragments, the pairs of the fragments, and the target molecule. From these energies we extracted the three-body correction to the FMO energy, as described in the previous section.

Table II compares the energies and the 1-RDMs of various approximations with those of the exact Hartree-Fock. As shown the monomer SCC energy is always lower than the exact one by the amount of 0.14−0.19 a.u. This is due to the violation of the exclusion principle. The pair FMO approximation corrects most of this error. The ABINIT-MP program[21] yielded the energies (denoted as "pair penalty") whose errors

TABLE II. Errors of the energy and the spin-reduced 1-RDMs of the cyclohexane and the homologues. STO-3G basis and the experimental structure of cyclohexane were used. The error of RDM is measured by the square norm of the difference between the approximate and the exact ones. Numbers in the square brackets indicate powers of 10.

	C_6H_{12}	$C_5NH_{12}^+$	$C_5BH_{12}^-$
Exact energy	−231.474 78	−247.642 93	−218.307 82
	Energy error (10^{-3} a.u.)		
Monomer SCC	−177.62	−140.42	−186.29
Pair penalty	24.60	20.41	25.76
Pair	8.77	6.08	10.82
Cumulant	6.16	4.05	8.22
Pair+3body	1.92	0.26	2.49
Without V_{ABC}	0.18	0.14	0.28
	1-RDM error (square norm)		
Monomer SCC	1.233	1.220	1.340
Pair	0.160	0.163	0.177
Pair+3body	0.107	0.106	0.122
	Maximum error of 1-RDM eigenvalue		
Monomer	0.31	0.28	0.33
Pair	3.4[−2]	2.7[−2]	4.0[−2]
Pair+3body	7.7[−3]	6.6[−3]	1.2[−2]

were $2 \times 10^{-2} - 2.5 \times 10^{-2}$ a.u. This program localizes MOs within each fragment by adding a penalty projection operator. On the other hand if the hybridized orbitals are used to define the fragments, these errors are reduced to half (denoted as "pair"). The use of the penalty function seems to introduce another sizable three-body effects.

The importance of the three-body term in the Hamiltonian, V_{ABC} in Eq. (3), was quantified as follows. From the monomer Fock operators H_A, H_B, and H_C and the dimer Fock operators H_{AB}, H_{BC}, and H_{CA}, we calculated the perturbations $V_{IJ} = H_{IJ} - (H_I + H_J)$. Then we diagonalize the Fock operator, $F = H_A + H_B + H_C + V_{AB} + V_{BC} + V_{CA}$, to calculate 1-RDM. The expectation value of the true Hamiltonian with respect to this 1-RDM differs from the exact energy. This difference (denoted as without V_{ABC}) shows the importance of the three-body terms of the Hamiltonian. We found it rather small (about 2×10^{-4} a.u.), so that we can safely ignore them in this analysis. Thus the three-body energy originated from the three-body scattering dominates the errors of the pair approximation. Note that our perturbative correction treats the basis set superposition effect simultaneously.

The results in Table II indicate that our correction reproduced this three-body energy accurately: the errors were reduced to $1/4 - 1/20$. It is interesting that some portion (25%–35%) of the three-body energy can be recovered by just calculating the expectation value of the Hamiltonian in terms of the pair 1-RDM (denoted as "cumulant"). We also examined the errors of the 1-RDM. The error was measured with the square norm of the difference between the approximate and the exact 1-RDMs. The error of the idempotency of 1-RDM was measured by the maximum deviation of the eigenvalues from zero or one. The results are summarized in Table II. The errors of the monomer SCC are rather large: the maximum eigenvalue of 1-RDM reaches about 1.3, which indicates that 1.3 electrons occupy this spin-orbital. The pair approximation reduces the error of 1-RDM by an order of magnitude. The three-body correction [Eq. (23)] further reduces them, giving the best results.

IV. CONCLUSIONS

In this paper we showed that the partial summation of the cluster expansion of Green's function (GF) yields the same extrapolation formula as that of the fragment molecular orbital method (FMO).[1] We can determine the excitation energies from GFs approximated by the FMO procedure. Since the cluster expansion is truncated at finite order there exists no wave function that is consistent to the FMO results. The perturbation expansion in which the self-consistent charge approximation defines the unperturbed state was reported. We proposed the three-body corrections to the energy and RDMs that are missing in the pair FMO approximation. They are size extensive and require only the quantities available by the pair approximation. We examined the accuracy of these corrections with the extended Hubbard model and the test molecules including cyclohexane. It was found that the

present three-body correction reduced the errors of pair approximation systematically. Green's function technique developed in this paper provides valuable tool for the extension of FMO to the delocalized systems and the excitation energies. Because of the larger intersystem interaction in the delocalized system the summation of the important higher-order terms is necessary. We can do it concisely by using the standard diagrammatic technique. The self-energy and the polarization propagator provide the transparent way to calculate the excitation energy.

[1] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999); K. Kitaura, T. Sawai, T. Asada, T. Nakano, and M. Uebayasi, *ibid.* **312**, 319 (1999).

[2] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **121**, 2483 (2004); **122**, 054108 (2005); **123**, 134103 (2005).

[3] T. Nakano, T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayasi, and K. Kitaura, Chem. Phys. Lett. **318**, 614 (2000).

[4] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **120**, 6832 (2004).

[5] H. Stoll, J. Chem. Phys. **97**, 8449 (1992); Phys. Rev. B **46**, 6700 (1992); H. Stoll, B. Paulus, and P. Fulde, J. Chem. Phys. **123**, 144108 (2005).

[6] B. Paulus, P. Fulde, and H. Stoll, Phys. Rev. B **54**, 2556 (1996); P. Fulde, H. Stoll, and K. Kladko, Chem. Phys. Lett. **299**, 481 (1999).

[7] K. Kitaura and K. Morokuma, Int. J. Quantum Chem. **10**, 325 (1976).

[8] B. Paulus, P. Fulde, and H. Stoll, Phys. Rev. B **51**, 10572 (1995); K. Kladko and P. Fulde, Int. J. Quantum Chem. **66**, 377 (1998).

[9] A. J. Coleman, Rev. Mod. Phys. **35**, 668 (1963); E. R. Davidson, *ibid.* **44**, 451 (1972); A. J. Coleman and V. I. Yukalov, *Reduced Density Matrices, Lecture Notes in Chemistry* (Springer, Berlin, 2000).

[10] N. Flocke and R. J. Bartlett, J. Chem. Phys. **121**, 10935 (2004); M. Schütz and H. Werner, *ibid.* **114**, 661 (2001); S. Li, J. Ma, and Y. Jiang, J. Comput. Chem. **23**, 237 (2002).

[11] A. A. Abrikosov, L. P. Gor'kov, and E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood, Cliffs, NJ, 1963); J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, New York, 1988); R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[12] Y. Inadomi, T. Nakano, K. Kitaura, and U. Nagashima, Chem. Phys. Lett. **364**, 139 (2002).

[13] J. Marcinkiewicz, Math. Z. **44**, 612 (1939); A. K. Rajagopal and E. C. G. Sudarshan, Phys. Rev. A **10**, 1852 (1974).

[14] J. Goldstone, Proc. R. Soc. London, Ser. A **239**, 267 (1957).

[15] L. S. Cederbaum, Theor. Chim. Acta **31**, 239 (1973); J. Schirmer, L. S. Cederbaum, and O. Walter, Phys. Rev. A **28**, 1237 (1983).

[16] H. Nakatsuji and K. Yasuda, Phys. Rev. Lett. **76**, 1039 (1996); K. Yasuda and H. Nakatsuji, Phys. Rev. A **56**, 2648 (1997).

[17] A. Lande and R. A. Smith, Phys. Rev. A **45**, 913 (1992); K. Yasuda, Phys. Rev. A **59**, 4133 (1999).

[18] J. Almlof, Chem. Phys. Lett. **181**, 319 (1991).

[19] P. J. Knowles and N. C. Handy, Chem. Phys. Lett. **111**, 315 (1984).

[20] W. J. Hehre, R. F. Stewart, and J. A. Pople, J. Chem. Phys. **51**, 2657 (1969); J. Hehre, R. Ditchfield, R. F. Stewart, and J. A. Pople, *ibid.* **52**, 2769 (1970).

[21] T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, and K. Kitaura, Chem. Phys. Lett. **351**, 475 (2002); O. Kikuchi, T. Nakano, and K. Morihashi, ABINIT, University of Tsukuba, 1992.