Regular Article

Superposition of nonorthogonal Slater determinants towards electron correlation problems

Seiichiro Ten-no*

Institute for Molecular Science, Myodaiji, Okazaki, 444 Japan

Received: 7 July 1997 / Accepted: 12 September 1997

Abstract. We propose variational and nonvariational methods based on the superposition of nonorthogonal Slater determinants. Properties of the reference functions are discussed. In the nonorthogonal configuration interaction method, all the excited configurations of multiple determinants are integrated into a variational space. An efficient way to manipulate matrix elements over determinants of distinct vacuums is presented by introducing similarity transformed operator and bracket transformations. The method enables us to map a matrix multiplication in the nonorthogonal problem to an orthogonal one, and thus maintains a fundamental scaling property along with the amount of data processed in the corresponding orthogonal configuration interaction method. Furthermore, we discuss a coupledcluster theory employing a vacuum-dependent wave operator, which is entirely size consistent as well as core extensive. These methods are applied to $H_2O + nHe(n = 0-2)$ and a single-bond dissociation of the HF molecule, compared with conventional methods including full and multireference configuration interaction methods.

Key words: Nonorthogonal Slater determinants $-$ Orbital relaxation $-\overline{\text{Similarity}}$ transformation $-\overline{\text{$ Bracket transformation $-$ Coupled-cluster theory

1 Introduction

Ab initio molecular orbital (MO) methods have been enjoying increasing popularity owing to their wide applicability and reliability. Single reference (SR) methods mostly based on the Hartree-Fock (HF) orbitals have become established for predicting electronic structures, especially around equilibrium bond distances. Descriptions of bond dissociations and of excited states with chemical accuracy require multireference (MR) approaches usually based on the multiconfiguration selfconsistent field (MCSCF) method. The multireference configuration interaction (MRCI) method is conceptually straightforward and has been the most popular tool for such purposes. However, even with this kind of method, it is laborious to generate flexible as well as competent computer codes capable of treating important excitations such as semi-internal ones. Full valence complete active space (CAS) [1] has been used as a conventional model to treat potential energy surfaces of ground and valence excited states. However, the Rydberg series generally exists in the energy range and it is nontrivial to choose model spaces in an optimal way. The size of the active space and accordingly the subsequent MRCI length immediately increase with numbers of active orbitals. Furthermore, besides the scaling property of the MRCI method, it suffers from a lack of proper extensivity with respect to the number of electrons. This makes descriptions of large systems inadequate.

SR many-body perturbation (MBPT) theories and their resemblance to coupled-cluster (CC) methods [2, 3] are the most widely used electron correlation procedures. Some of these methods feature size consistency, viz. a numerically proper description of a fragmentation, or a related terminology size extensivity, viz. an explicit extensivity ensured by connectedness of equations to justify an arbitrary fragmentation. It is nontrivial to extend such methods to MR cases. Mostly on the basis of the Bloch wave operator equation in complete model spaces (CMS), MR-MBPT [4] and MRCC approaches [5-8] have been investigated by many authors for treating quasi-degenerate states. More recently, size-extensive MRCC methods for incomplete model spaces (IMS) have been developed using the Fock [9] and the Hilbert space approaches [10]. Despite these theoretical developments, it was not until quite recently that even the perturbational approaches started to be used routinely in practical applications. One of the main reasons is that

^{*}Present address: Institute of Physical Chemistry,

Academy of Sciences of the Czech Republic, CZ-182 23 Prague, Dolejskova 3, Czech Republic

there has been no systematic way to avoid intruder state problems maintaining size extensivity. In most of the perturbational methods which are now readily available in quantum chemistry codes $[11-14]$, it is difficult to maintain size extensivity. Recently, Heully et al. analyzed the origin of the size-inconsistency error in the second-order state-specific effective Hamiltonian method [15]. However, a more general and rigorous investigation is one of the most important subjects in this area. (For the recent development emphasizing size extensivity, see Ref. [16].)

The above methods all use only one set of spin or spatial orbitals. We recently performed ab initio calculations employing another type of multiconfiguration functions, i.e. superposition of nonorthogonal Slater determinants [17]. In the preliminary results of the ground-state CO molecule, it has been found that the method with a few generating Slater determinants can reproduce the correlation energy to the same extent as the full valence CASSCF one. Furthermore, the method is capable of describing long-range behavior of wave functions better than the MCSCF method, as implied by the calculated one-electron properties. One can expect that the allowance of nonorthogonality leads to a wide variety of methods. The most obvious applications will be in electron correlation problems and in properties among states.

In the present article, we propose new multireference methods based on the superposition of nonorthogonal Slater determinants. The properties of the reference wave functions are examined in the following section. In order to realize a truncated CI expansion, an efficient way to evaluate operator matrix elements over nonorthogonal excited determinants is proposed in Sect. 3. Furthermore, we discuss a nonorthogonal CC method and its aspect as a many electron theory in Sect. 4. We present results of some applications using the nonorthogonal CI and linearized CC methods in Sect. 5. Conclusions are given in Sect. 6.

2 Superposition of nonorthogonal Slater determinants

Let us consider an N-electron system and look at some solutions of the Schrödinger equation,

$$
H\psi_{\mu} = E_{\mu}\psi_{\mu}, \qquad (\mu = 1, \dots, m) , \qquad (1)
$$

which is described by a Hamiltonian involving at most two-body interactions,

$$
H = \sum_{pq} h_{pq} X_q^p + \frac{1}{2} \sum_{pqrs} (pq|rs) (X_q^p X_s^r - \delta_{qr} X_s^p) , \qquad (2)
$$

where the orbital replacement operator is $X_{rs}^{pq...} =$ $X_p^+ X_q^+ \dots X_s X_r$. For reference wave functions, we approximate the exact solutions in terms of the superposition of Slater determinants,

$$
\psi_{\mu}^{(0)} = \sum_{f=1}^{N_s} \phi(f) c_{\mu f}^{(0)} \tag{3}
$$

where $c_{\mu f}^{(0)}$ are CI coefficients of the wave function. Maximizing the flexibility of the approximation, we assume that each of the N-electron determinants, $\{\phi(f)\}\$ $(f = 1, \ldots, N_s \geq m)$, consists of spin orbitals $\{\varphi_n(f)\}\$ which are not necessarily orthogonal to those belonging to the other determinants,

$$
\left\langle \phi_p^{(f)} \middle| \varphi_q^{(g)} \right\rangle = \delta_{pq} \delta_{fg} + u_{pq} (f, g) (1 - \delta_{fg}) \quad , \tag{4}
$$

where $u_{pq}(f,g)$ are overlaps of the nonorthogonal orbitals. When the CI and the orbital coefficients are determined variationally with respect to the ground state energy,

$$
\left\langle \phi(f)|H - E_1^{(0)}|\psi_1^{(0)} \right\rangle = 0, \quad \forall f \quad , \tag{5}
$$

$$
\langle X_i^a(f)\phi(f)|H - E_1^{(0)}|\psi_1^{(0)}\rangle = 0, \quad \forall a, i, f \quad , \tag{6}
$$

the method is called the resonating HF approximation [17], which was recently applied to ab initio MO calculations. Here the spin-orbital indices a,b, \ldots and i, j, \ldots denote virtual and occupied ones in each determinant, respectively, and the approximate energy is $E_{\mu}^{(0)} = \langle \psi_{\mu}^{(0)} | H | \psi_{\mu}^{(0)} \rangle$. Frequently, the variational conditions cannot be satisfied due to near-singularities of the reference space, which will be discussed in Sect. 5. A manifest advantage of allowing the nonorthogonality is that it enables us to describe systems the main configurations of which include large hole-particle differences (case 1). It is well known that the instantaneous repolarization effects [18] are not reproduced by the MCSCF method in which a unique set of MOs is used for descriptions of all the valence bond (VB) configurations. Since the nonorthogonality of determinants is capable of introducing multideterminant orbital relaxation, such a choice of reference functions would accelerate a CI convergence. Note that the individual orbital relaxation is long-ranged and thus is totally nonmultiplyable.

In the CC methods, the orbital relaxation effects are accounted for by inclusions of e^{T_1} . Similarly, it is convenient to relate nonorthogonal Slater determinants via exponential transformations of one-particle operators in intermediate normalization with respect to each vacuum,

$$
|\phi(g)\rangle = \exp(R_g(f))|\phi(f)\rangle S_{fg} \quad , \tag{7}
$$

where

$$
R_g(f) = \sum_{ai} r_{g,ai}(f) X_i^a \quad , \tag{8}
$$

and S denotes the overlap matrix over determinants,

$$
S_{fg} = \langle \phi(f) | \phi(g) \rangle . \tag{9}
$$

According to the Löwdin formula [19], one can easily see that the overlap matrix and the amplitudes in the operator are related to those over spin orbitals belonging to parent determinants as

$$
S_{fg} = \langle \phi(f) | \phi(g) \rangle = \det(\tilde{\mathbf{u}}(f,g)) \quad , \tag{10}
$$

$$
r_{g,ai}(f) = \sum_{j} u_{aj}(f,g) (\tilde{\mathbf{u}}^{-1}(f,g))_{ji} ,
$$
 (11)

Table 1. Main cases of two nonorthogonal HF functions^{a,b}

^a Formally expressed by $\psi^{(0)} = [c_1^{(0)} + S_{12}c_2^{(0)}(1 + R_2(1) + R_2^2)]$

¹ M means modest magnitudes $\frac{1}{c}$ Hole-particle differences between determinants. Case 3 is in the natural orbital representation and it is always possible to reduce the difference to 1

where the tilde denotes the occupied portion of the metric matrix. Note that the nonzero overlap is not a necessary condition to relate nonorthogonal determinants. It can always be satisfied by inserting operators for orbital replacements. In the following section, we will see that this expression plays a central role in evaluating matrix elements over excited configurations of nonorthogonal determinants, which is required in the nonorthogonal CI method.

In order to see the flexibility of the wave function more precisely, let us focus on the minimum model consisting of two nonorthogonal Slater determinants (2NODs). We summarize the main domains of 2NODs in Table 1. In case 1, the resonance among structures, for instance in three-electron bond $A^{+}B$ and AB^{+} , induces different polarization of electrons below the active level [18]. Therefore, successive orbital rotations of instantaneous repolarization would also be important in addition to the single-electron excitations. On the other hand, the orbital relaxation effects among structures are less important in the closed shell molecules. When the determinants are similar to each other, neglecting the nonlinear terms in Eq. (7), the total wave function is approximately expressed by

$$
\psi^{(0)} \cong (c_1^{(0)} + S_{12}c_2^{(0)} + S_{12}c_2^{(0)}R_2(1))\phi(1) , \qquad (12)
$$

(case 2). For pure spin states, the wave function corresponds to the MCSCF wave function at the single-electron excitation (SE) CI level, whose orbitals are generally quite different from the usual SCF or the natural orbitals (NOs). In Fig. 1, we illustrate the convergence behaviors of SR-CI expansions using a few different sets of orbitals; the HF molecule around the equilibrium geometry is calculated as a representative closed shell molecule.

The CI expansion improves the correlation energy at odd-electron excitations when the MCSCF (singles) orbitals are used, whereas the expansions with the usual SCF orbitals or NOs improve it at even-electron excitations. It is apparent that the former set is more advantageous in the CI including up to triple excitations. Since the 2NODs function can reproduce the SECI wave function, we can expect the nonorthogonal CI singles and doubles to be almost as accurate as the SR-CI up to triples with the non-HF orbitals (illustrated by arrows). The method would maintain the scaling property of the SDCI method if the matrix elements over nonorthogonal bases were evaluated as efficiently as orthogonal ones. We will inspect the possibility in the following section.

Fig. 1. Convergence of the HF molecule energies in the SR CI using various spatial orbitals. All electrons are correlated. The natural orbitals (NOs) set is based on the full CI for the space spanned by the $[3s2p/2s]$ basis functions (Ref. [33])

On the other hand, the 2NODs function is also capable of reproducing the generalized valence bond (GVB) wave function which is, in the NO representation,

$$
\psi^{(0)} = (c_g^{(0)} X_g^+ X_g^+ + c_u^{(0)} X_u^+ X_u^+) \phi_{\text{core}} \tag{13}
$$

Only in the (single-bond) dissociation limit (case 3) does the variational solution of 2 NODs become the GVB wave function since there is no mechanism to further optimize orbitals of dissociated fragments. Therefore the theory illustrated in the following sections includes some MR treatment for dynamic correlation effects. When there is no particle-hole difference, the wave function is a single determinant (case 4). The flexibility of the approximation is in the order, $2NODs$ > $MCSCF$ $(singles) > GVB > HF$. When functions consist of more than three determinants, several crosses will appear.

3 Matrix element evaluation

Our purpose in this section is to realize CI expansions based on the wave function in Eq. (3) as a reference, aimed at not only reproducing dynamic correlation effects but also at optimizing orbitals utilizing the

Brillouin-type condition in Eq. (6). The CI wave function based on the superposition of nonorthogonal Slater determinants is

$$
\psi_{\mu} = \sum_{f} \left(1 + C_{1\mu}(f) + \dots + C_{n\mu}(f) \right) \phi(f) c_{\mu f} , \qquad (14)
$$

where the state-dependent CI operators are linear combinations of orbital replacement operators. The conventional intermediate normalization condition does not hold in the CI expansion due to the nonorthogonality of the parent determinants with excited ones, $\langle X_{ii}^{ab...}(f)\phi(f)|\phi(g)\rangle \neq 0$. The CI coefficients are determined in a variational way,

$$
\langle \phi(f)|H - E_{\mu}|\psi_{\mu}\rangle = 0, \quad \forall f \quad , \tag{15}
$$

$$
\left\langle X_{ij...}^{ab...}(f)\phi(f)|H - E_{\mu}|\psi_{\mu}\right\rangle = 0, \quad \forall \ a, b, \dots, i, j, \dots \tag{16}
$$

imposing orthogonalization constraints among the states. The projection space includes all the functions induced by the CI operators with respect to each determinant.

Solving the equations requires evaluating the energy and overlap matrix elements. Usually spin-adapted configuration state functions (CSFs) are chosen as a basis, which are linear combinations of determinants. For diagonal blocks $(f = g)$, ways to evaluate the matrix elements are well known for any base [20, 21]. However, the off-diagonal part is not so simple in structure since it includes different vacuums in the quantity. Much effort has been devoted to developing efficient algorithms for the evaluation, especially in the classical VB methods. They are, however, essentially on the basis of the Löwdin formula and concern how to simplify the calculation of cofactors of the orbital overlap matrix. The computational task is much greater than the orthogonal one even if the cofactors can be evaluated efficiently. Furthermore, the matrices are not sparse even for a zero-particle operator (overlaps). This means we handle much more data than for the orthogonal problems.

Corresponding orbitals have been employed for computing the matrix elements in nonorthogonal CI and transition moment calculations [22]. In the full CI (CAS) case, the model is invariant to unitary transformations of orbitals. Due to the diagonal nature of the corresponding orbital overlaps, two occupations which differ by more than $N + 1$ orbitals do not contribute to an N-particle matrix between CAS spaces. For general CI problems, however, nonzero overlaps including external orbitals significantly increase the computational effort. Since the present purpose is to develop methods including large-scale truncated CI expansion based on the nonorthogonal reference determinants, a more efficient algorithm is required.

For a while, we will focus on a certain block of the matrix elements, whose parents are f 'th and g 'th determinants on the bra and ket sides, respectively. We assume that frozen core orbitals are common in the reference determinants. As we have seen in the previous section, two different determinants are related via Eq. (7) . We can therefore express the desired matrix elements as,

$$
\left\langle X_{ij...}^{ab...}(f)\phi(f)|O|X_{kl...}^{cd...}(g)\phi(g)\right\rangle
$$

=\left\langle X_{ij...}^{ab...}(f)\phi(f)|e^{R_g(f)}\hat{O}|\hat{X}_{kl...}^{cd...}(g)\phi(f)\right\rangle S_{fg} , (17)

where the similarity transformed operators are,

$$
\hat{O} = \exp(-R_g(f))O\exp(R_g(f)), \qquad (18)
$$

$$
\hat{X}_{kl...}^{cd...}(g) = \exp(-R_g(f))X_{kl...}^{cd...}(g) \exp(R_g(f)) . \qquad (19)
$$

Since the exponent includes only one-particle operators, these transformations do not change the particle ranks of the original operators. This kind of transformation was recently employed by Koch et al. [23] to reformulate CCSD equations as effective CCD equations in their direct atomic orbital integral driven approach. Following their notation, the transformed Hamiltonian is

$$
\hat{H} = \sum_{pq} \hat{h}_{pq} X_q^p(f) + \frac{1}{2} \sum_{pqrs} (pq \hat{r} s) (X_q^p(f) X_s^r(f) - \delta_{qr} X_s^p(f)) \tag{20}
$$

The hat integrals are modified one- and two-electron ones corresponding to the transformation,

$$
\hat{h}_{pq} = \sum_{rs} x_{pr} h_{rs} y_{qs} \quad , \tag{21}
$$

$$
(pq\hat{|}rs) = \sum_{tuvw} x_{pt} y_{qu}(tu|vw)x_{rv}y_{sw} , \qquad (22)
$$

where

$$
x_{pq} = \delta_{pq} - \sum_{ai} r_{g,ai}(f) \delta_{pa} \delta_{qi} , \qquad (23)
$$

$$
y_{qp} = \delta_{pq} + \sum_{ai} r_{g,ai}(f)\delta_{pa}\delta_{qi} . \qquad (24)
$$

It is apparent that the transformed Hamiltonian and accordingly the modified integral matrices are non-Hermitian except $R_g(f) = 0$, $exp(R_g(f))$ which is nonunitary. Similarly, the transformed orbital replacement operator is expressed by

$$
\hat{X}_{ij...}^{ab...}(g) = \sum_{pq...rs} X_{rs...}^{pq...}(f) \hat{u}_{pa} \hat{u}_{qb} \dots \hat{u}_{sj} \hat{u}_{ri} ,
$$
\n(25)

where the effective metric is

$$
\hat{u}_{pa} = \sum_{r} x_{pr} u_{ra}(f, g) \quad , \tag{26}
$$

$$
\hat{u}_{qi} = \sum_{r} y_{qr} u_{ri}(f, g) \tag{27}
$$

Additional operations left are to relate the bra and ket parts with the usual CI basis in the corresponding orthogonal problems. The bra parts are straightforwardly expanded as

$$
\langle X_i^a(f)\phi(f)|\exp(R_g(f)) = \langle X_i^a(f)\phi(f)| + r_{g,ai}(f)\langle\phi(f)|, \qquad (28)
$$

and

$$
\langle X_{ij}^{ab}(f)\phi(f)| \exp(R_g(f)) = \langle X_{ij}^{ab}(f)\phi(f)|
$$

+ $(1 + p_{ij}^{ab}) \Big[r_{g,ai}(f) \langle X_j^b(f)\phi(f)| - r_{g,aj}(f) \langle X_i^b(f)\phi(f)| \Big]$
+ $(r_{g,ai}(f)r_{g,bj}(f) - r_{g,aj}(f)r_{g,bi}(f)) \langle \phi(f)|$,
... etc., (29)

where p_{ij}^{ab} means permutation of (*ai*) and (*bj*) indices. For explicit expressions in the ket part, we rewrite the excitation operator in a normal product form and its contractions [24],

$$
\hat{X}_{ij...}^{ab...}(g) = \sum_{pq...rs} (\{X_{rs...}^{pq...}(f)\}_f.
$$
\n
$$
+ \{\underline{X_{rs...}^{pq...}(f)}\}_f) \hat{u}_{pa} \hat{u}_{qb} \dots \hat{u}_{sj} \hat{u}_{ri} ,
$$
\n(30)

where the brace $\{\}f$ means normal ordering with respect to the f' th determinant as a vacuum, and the underline symbolizes a sum of all contractions. Taking account of the fact that each single contraction gives rise to

$$
w_{ck} = \sum_{i} \hat{u}_{ic} \hat{u}_{ik} = -\sum_{a} \hat{u}_{ac} \hat{u}_{ak} , \qquad (31)
$$

we obtain expressions that are similar to the bra portion except for successive one-index transformations with the modified metric,

$$
|\hat{X}_i^a(g)\phi(f)\rangle = \sum_{ck} |\hat{X}_k^c(f)\phi(f)\rangle \hat{u}_{ca}\hat{u}_{ki} + |\phi(f)\rangle w_{ai} ,
$$
\n(32)

and

$$
\begin{aligned}\n\left|\hat{X}_{ij}^{ab}(g)\phi(f)\right\rangle &= \sum_{cdkl} \left|\hat{X}_{kl}^{cd}(f)\phi(f)\right\rangle \hat{u}_{ca}\hat{u}_{db}\hat{u}_{ki}\hat{u}_{ij} \\
&+ (1 + p_{ij}^{ab}) \sum_{ck} \left|\hat{X}_{k}^{c}(f)\phi(f)\right\rangle (\hat{u}_{ca}\hat{u}_{ki}w_{bj} - \hat{u}_{ca}\hat{u}_{kj}w_{bi}) \\
&+ \left|\phi(f)\right\rangle (w_{ai}w_{bj} - w_{aj}w_{bi}), \\
\text{etc.},\n\end{aligned} \tag{33}
$$

The most important conclusion in this paragraph is that any matrix elements over nonorthogonal determinants can be derived from those of similarity transformed operators over orthogonal determinants up to the same excitation level. This is also true for elements over any spin-adapted basis derived from nonorthogonal restricted HF (RHF)-type determinants. When the parent determinants consist of only doubly occupied orbitals, the operator $R_q(f)$ can be written in terms of the unitary group generator, and thus the similarity transformed Hamiltonian is spin-free as long as the original one is so. We will mention a simple spin-adaption procedure in Sect. 5.

In practical computer implementations, operations including the off-diagonal block can commonly be written by

$$
\sigma_{\kappa}(f) = \sum_{\lambda} \langle X_{\kappa}(f)\phi(f)|O|X_{\lambda}(g)\phi(g)\rangle c_{\lambda}(g) , \qquad (34)
$$

where κ and λ denote sets of hole and particle indices of replacements. In the CI case, $\sigma_{\kappa}(f)$ and $c_{\lambda}(g)$ correspond to residuals and CI vectors. The bra and ket transformations can be performed in an indirect way by effectively transforming CI vectors, then multiplying and transforming residuals,

$$
\sigma'_{\kappa}(f) = \sum_{\lambda} \langle X_{\kappa}(f)\phi(f)|\hat{O}|X_{\lambda}(f)\phi(f)\rangle c'_{\lambda}(f) , \qquad (35)
$$

where $\sigma'_k(f)$ and $c'_\lambda(f)$ are transformed vectors and residuals. There is one-to-one correspondence between the original and transformed quantities in each block according to the matrix element transformation illustrated above. This procedure is quite attractive because a complete mapping of nonorthogonal problems to orthogonal ones is attained. The amount of data to be processed is as compact as the usual CI data and thus the scaling property of the orthogonal problem is maintained. Besides the matrix multiplication, it is necessary to generate preconditioning vectors in the Davidson scheme [25]. Generally, the Hamiltonian and overlap matrices are not diagonal dominant in the nonorthogonal basis. Therefore it is preferable to solve linear equations including energy resolvents explicitly or in an approximate way using the diagonal part of the matrices in the corresponding orthogonal basis. These operations can also be performed efficiently using Eq. (35).

4 Nonorthogonal CC method

Truncated CI expansions including the nonorthogonal CI illustrated in the previous section involve sizeinconsistency errors due to a lack of simultaneous descriptions for correlated events. To resolve this defect, we need an explicit consideration of the exponential ansatz on top of approximations. Owing to the present choice of nonorthogonal reference functions, a resultant theory should essentially be an IMS one. In the Hilbert space state-universal approaches as developed by Jeziorski and Monkhorst [7], the vacuum-dependent wave operator is expressed as

$$
\Omega P = \sum_{f} \Omega_f P_f \quad , \tag{36}
$$

$$
\Omega_f = e^{T(f)} \tag{37}
$$

where P_f denotes a projector of the f'th Slater determinant spanning the model space. Actually, the form of the wave operator originates from earlier work by Silverstone and Sinanogolu [26]. Meissner and Bartlett extended the method to a general model-space one [10], and the CC theories based on the wave operator are closely related to the MR-MBPT of Hose and Kaldor [27]. We can follow the ansatz by using a projector in the nonorthogonal model space,

$$
P_f = \sum_{g} |\phi(f)\rangle \left(\mathbf{S}^{-1}\right)_{fg} \langle \phi(g)| \tag{38}
$$

Generally, equations necessary for the state-universal CC methods are obtained, starting with the waveoperator equation [10],

$$
H\Omega P = \Omega P \tilde{H} P \t{39}
$$

186

$$
P = \sum_{f} P_f \tag{40}
$$

which is equivalent to a set of Schrödinger equations in the model space. Equating contributions for each determinant by multiplying P_f to the right and Ω_f^{-1} to the left, one obtains

$$
P\tilde{H}P_f = \overline{H\Omega}_f P_f - (\Omega_f^{-1}\Omega - 1)P\tilde{H}P_f , \qquad (41)
$$

where the bar over $H\Omega_f$ denotes all terms obtained by joining operators. Amplitude equations are obtained by projecting Eq. (41) with excited determinants whose parent determinant matches the vacuum. Obviously, it is not a unique way to determine amplitudes. The state universality of the wave operator assumes that projections with excited determinants of distinct vacuums are automatically satisfied. Unfortunately, to show the connectedness property of the equation is not so straightforward as the usual IMS one, since the reduced resolvent cannot be defined using a single partitioning of the Hamiltonian. This fact prohibits an order-by-order expansion of the nonorthogonal cluster operator. Furthermore, all of the members in the model space are generally not well separated from virtual functions, and thus the procedure will face an intruder state problem, as especially in the CMS case.

In the state-selective approaches, there are not many methods employing the vacuum dependent wave operator, Eq. (36). Based on the state-universal approach, MRCC equations were formulated by collecting contributions in Eq. (41) to the residual of a contracted function [28]. Manifestly, this formulation is not satisfactory to avoid intruders since the state universality is embedded in the original framework. We have no meaningful numerical result in this line except for the use of approximate expressions which can also be derived directly from a single-root Schrödinger equation,

$$
\sum_{f} (H - E)\Omega_f |\phi_f\rangle \tilde{c}_f = 0 \tag{42}
$$

Here the \tilde{c}_f are contraction coefficients in the model space, and the root index is ignored. Recently, Meller et al. [29] proposed an approximate dressing to the MRSDCI method based on the ansatz [29]. They showed that the method satisfies a size consistency by inspecting the separability conditions. More recently, Mahapatra et al. [30] derived a sufficiency condition of the Schrödinger equation, which satisfies a connectedness in the CMS. However, it is difficult to show if the sufficiency condition works in a state-selective manner avoiding intruder state problems. We shall consider a more straightforward application of the single-root Schrödinger equation in this paper. The P_q projection of the equation is

$$
|\phi_g\rangle E\tilde{c}_g = \sum_f P_g \left[\overline{H\Omega}_f | \phi_f \rangle + \chi_f (\overline{H\Omega}_f - E) | \phi_f \rangle \right] \tilde{c}_f ,
$$
\n(43)

where $\chi_f = \Omega_f - 1$. By multiplying $q_f \Omega_f^{-1}$ to the Schrödinger equation and substituting terms in the above equation such that all the explicit energy expression disappears, one yields

$$
q_f \overline{H\Omega}_f |\phi_f\rangle + q_f \sum_{g \neq f} \Omega_f^{-1} \Omega_g \overline{H\Omega}_g |\phi_g\rangle \tilde{c}_g / \tilde{c}_f - q_f
$$

$$
\times \sum_{g \neq f} \sum_h \Omega_f^{-1} \Omega_g P_g \overline{[H\Omega}_h + \chi_h (\overline{H\Omega}_h - \ldots)] \phi_h \rangle \tilde{c}_h / \tilde{c}_f = 0 ,
$$
 (44)

where $q_f = 1 - |\phi_f\rangle\langle\phi_f|$. In order to simplify the problem, let us consider the CMS limit and assume that purely internal excitations are not included in the cluster operator,

$$
q_f \overline{H\Omega}_f |\phi_f\rangle + q_f \sum_{g \neq f} \left[\overline{H\Omega}_g + (\Omega_f^{-1} \Omega_g - 1) q_g \overline{H\Omega}_g \right] \phi_g \rangle \tilde{c}_g / \tilde{c}_f
$$

$$
- q_f \sum_{g \neq f} \sum_{h \neq g} (\Omega_f^{-1} \Omega_g - 1) P_g \overline{H\Omega}_h |\phi_h \rangle \tilde{c}_h / \tilde{c}_f = 0 \quad . \quad (45)
$$

We used the fact that the portions including $P_g\chi_f$ or q_fP_g vanish in the limit. The first terms are trivially connected. For the other terms, the CI coefficients in the CMS are related in terms of cluster expansions [30], $\tilde{c}_q/\tilde{c}_f=\langle \phi_a | \exp(\tau_f) | \phi_f \rangle$, which is connected via indices including only valence labels distinguishing determinants. The τ_f includes all excitations needed to span the CAS space. The connectedness of τ_f is most evident in the fixed coefficients case. Therefore, the requirement for the entire connectedness of the terms is that all of the q_f projections are connected with respect to each intermediate vacuum as well as including at least one valence index distinguishing determinants in the CI coefficients. The second term, $q_f H\Omega_g |\phi_a\rangle$, satisfies the requirement and is connected since the terms having no such index result in null. The third and fourth terms involve the operator, $\Omega_f^{-1} \Omega_g - 1$, which also appeared in the state-universal approaches [7, 10]. Using the explicit expression for the wave operator and the Baker-Campbell-Hausdorff formula,

$$
\ln(e^{-T(f)}e^{T(g)}) = T(g) - T(f) + [T(g), T(f)]/2 + \cdots,
$$
\n(46)

one can see that the operator can be expressed as a polynomial in $T(g) - T(f)$ and the commutators. In the Hilbert space MRCC methods, it has been a common technique to prove connectivity that the operator, $e^{-T(f)}e^{T(g)} - 1$, leads to terms which include at least one valence index distinguishing determinant, making use of the property of the commutators and functional forms of amplitudes. They are also required in the present formulation. The third term is, however, disconnected even if such conditions are satisfied. For $T(q)$, types of cluster operators whose annihilation index distinguish determinants are not members of $T(f)$. Thus they survive and are necessarily connected with the portion of CI coefficients. On the other hand, $H\Omega$ _a involves arbitrary types of operators such that nonzero q_f projections are generated without joining $T(g)$. They do not cancel with the other diagrams in the third and the fourth terms, and thus an entire connectivity is not obtained a priori from the single-root Schrödinger equation itself even in the CMS case.

On the positive side, Eq. (44) has the following properties: (1) it is more straightforward to devise a formula for complicated model spaces like the present nonorthogonal one, (2) core extensivity is always satis fied, and (3) resulting energies are size consistent as long as the references are size consistent. Property 2 is not carried directly from the connectivity, as seen in the previous paragraph, but it is apparent that any separation of core electrons leads to N_s -folded SRCC states. The system equations become connected and thus core extensive. Note that the connectivity of the core portion can be shown only after the fragmentation. It is easy to show property 3. We can follow an investigation similar to that in the intermediate Hamiltonian methods [15, 29]. Consider a super-system, $A + B$, with its Hamiltonian, $H^{AB} = H^A + H^B$. Spin orbitals can be chosen such that each of them is localized in one of the fragments. We assume there are N_s^A and N_s^B nonorthogonal determinants to describe *A* and *B*. Accordingly $N_s^A \times N_s^B$ determinants arise to describe the entire system that are anti-symmetrized products of subsystem determinants. For reference functions, one can see that the secular equation (5) for the entire system is satisfied if the CI coefficients are written as products of those in the subsystem equations, $c_{fg}^{AB} = c_f^A c_g^B$. This requirement leads to a couple of secular equations. Similarly, the Brillouinlike condition in Eq. (6) is also separable into another orbital equation and a secular equation. In the CC part, the contraction coefficients in Eq. (42) can be obtained by the model space projections,

$$
\sum_{g} \langle \phi_f | (H - E) \Omega_g | \phi_g \rangle \tilde{c}_g = 0, \quad \forall f \quad , \tag{47}
$$

whose separability can be shown assuming that the wave operator has a form, $\Omega_{fg}^{AB} = \Omega_f^A \Omega_g^B$, and using the fact that the operators with different fragment labels commute. As long as the CC equations are solved without any approximation, it is sufficient to check the q_f projection of the Schrödinger equation to test the separability of Eq. (44). The multiplication of Ω_f^{-1} just induces lower classes of equations. The q_f projection separates into a couple of amplitude equations. Since the subsystem equations include all derivative projections in the entire equation, any truncation like CCD breaks size consistency. This also implies the size-inextensive nature of the single-root Schrödinger equation with the vacuum-dependent wave operator.

5 Results

5.1 Computer implementation

We now present some numerical results using the methods presented in the previous sections. Reference functions are optimized essentially according to Eqs. (5) and (6). An explicit second-order procedure includes terms much more complicated than the usual quadratically convergent HF method [31]. Fortunately, a superCI method generally gives smooth optimizations of the parameters, starting with distinct determinants. In the nonorthogonal CI, we solved linear equations without introducing any diagonal denominators for correction vectors. Especially around the equilibrium distances, the conditions in Eqs. (5) and (6) often lead to numerical problems. Determinants become similar to the others due to the unimportance of nonlinear $R_f(g)$ products, and the model space approaches near-singularity. There are several possibilities to avoid the problem. In this particular work, we introduce constraints such that absolute values of the CI coefficients do not exceed a threshold, $|c_j^{(0)}| \leq \varsigma$. The CI coefficients are subtracted
according to ς and the maximum absolute coefficient according to ς and the maximum absolute coefficient. Although a larger ς gives a lower energy, numerical uncertainty increases due to a large cancellation of contributions from determinants with opposite signs. In subsequent CI expansions, all the singles and doubles excitations are included.

The nonorthogonal CC method, Eqs. (44) and (47), requires evaluations of rather complicated diagrams which include truncation problems. In order to maximize the similarity with the CI method, we perform some approximations to the CC equations. The CI coefficients in the reference functions are reoptimized and used as fixed contraction coefficients. This can be justified since the inclusion of all hole-particle excitations is expected to induce principal internal excitations recovering the relaxation of coefficients. Linearizing the amplitudes equation, one obtains a simple expression,

$$
q_f \sum_{g=1}^{N_s} [\mathbf{H} - T(f)\mathbf{H} + (1 - P_0)\mathbf{H}T(g)] \phi_g \rangle c_g^{(0)} = 0 \quad , \quad (48)
$$

which maintains the core extensivity, where $P_0 =$ $|\Psi^{(0)}\rangle\langle\Psi^{(0)}|$ and $H = H - E^{(0)}$. We also used the fact that only the projector P_0 survives in the amplitude equation after fixing the coefficients and the linearization. The term including the projector originates from the nonorthogonality of the model functions to the external ones. Manifestly, this expression reduces to the LCC equation in the SR case $(N_s = 1)$. All the external determinants induced by orbital replacements are multiplied according to vacuum indices of q_f . Although the approximation does not ensure the entire size consistency explicitly, essential contributions are included in the energy-independent expression. Equation (48) is not a unique approximation to the CC method. If the premultiplication of Ω_f^{-1} is ignored in deriving Eq. (44), one obtains a slightly different equation,

$$
q_f \sum_{g=1}^{N_s} [\mathbf{H} + (1 - P_0)\mathbf{H}T(g)] \phi_g \rangle c_g^{(0)} = 0 \quad , \tag{49}
$$

which is similar to that in the MR linearized CC (MR-LCC) method [28, 32]. We also use the equation for comparisons. If the resonating HF conditions, Eqs. (5) and (6), are satisfied in the reference function, both approximations become identical in the single and doubles (LCCSD). This is similar to the disappearance

of some linked-disconnected diagrams in the SR-CC with HF orbitals. In principle, we assume that all the external determinants are linearly independent. However, as the magnitude of near-degeneracy correlation effects increases, particular excitations including a few electrons become important. This fact may cause redundancies in the cluster operator. Since the approximate equations, (48) and (49), do not include nonlinear terms of the amplitudes, we also use a linearized energy expression,

$$
\sum_{f=1}^{N_s} \langle \Psi^{(0)} | (H - E) T(f) | \phi_f \rangle c_f^{(0)} = \Delta E \quad , \tag{50}
$$

such that the resulting energy is invariant to the redundancy, where $\Delta E = E - E^{(0)}$. The entire equations in the LCC and the CI methods are represented in Goldstone diagrams. For a diagram with L loops, we replace the spin orbital indices by spatial orbital ones and multiply a factor 2^L to obtain spin free equations [24]. Extra matrix elements required in the LCC method can be manipulated with a minor modification of the corresponding CI expansion.

5.2 Illustrative applications

We apply the above methods to super-systems consisting of isolated molecules, $H_2O + nHe$ $(n = 0-2)$, using splitvalence basis sets [33, 34] to numerically investigate the size consistency, which is important to extract pure interactions of subsystems. Henceforward, RHF-type 2NODs functions with $\varsigma = 1.5$ are used as references in

Table 2. Individual energy components of H₂O + nHe $(n = 0-2)^{a}$

System	Method	Energy/ E_h
He	RHF LCCSD FCI	-2.835680 -2.850625 -2.850 577
H ₂ O	RHF CISD $CISD+O$ CISDTO 2NODs-RHF 2NODs-RHF(relaxed) 2NODs-CISD $2NODs-LCCSD(A)$ $2NODs-LCCSD(B)$ FCI	-76.009 145 -76.134914 -76.140310 -76.141651 -76.072 878 -76.072951 -76.140984 -76.141093 -76.141789 -76.141 855
$H2O + He$	CISD $CISD+Q$ CISDTO 2NODs-CISD	-78.968 133 -78.988537 -78.991378 -78.990991
$H2O + 2He$	CISD $CISD + O$ CISDTO 2NODs-CISD	-81.816875 -81.838889 -81.841802 -81.840914

^a The equilibrium bond distance and the angle of the water molecule are assumed to be 0.95781 Å and 104.4776° , respectively

the nonorthogonal methods. In Table 2, energy components of the molecules are listed. The full CI of the water molecule gives rise to 109,857 CSFs in the C_{2v} framework. The numbers of CSFs are 282, 533 and 937 at the CISD and 11,078, 38,623 and 115,350 at the CISDTQ for $n = 0, 1$ and 2, respectively. The 2NODs-CISD method includes explicitly twice as many configurations as the SR-CISD ones. In the RHF and full CI (FCI), energies of the super-systems are sums of fragment energies. Although we approximate the secular equation in the 2NODs-RHF methods, the relaxation of the CI coefficients scarcely affects the total energies. The total wave functions are antisymmetrized products of the 2NODs-RHF function of the water and the RHF functions of the He atoms owing to the Brillouin-like condition in Eq. (6). The reference energies are size consistent explicitly in both cases, the coefficients fixed and relaxed. Similarly, in the 2NODs-LCCSD methods, the total energies are explicitly sums of the corresponding ones of the water and the LCCSD ones of the He atoms, the methods being core extensive. "A" and "B" mean that Eqs. (48) and (49) are used respectively. The other truncated CI methods including the 2NODs-CISD are size inconsistent. In Fig. 2, we compare the performances of the methods.

The size-inconsistency error in the CISD method is prominent. The Davidson correction [35], referred to as $CISD+Q$, adjusts a large amount of the error. Although the CISDTQ and the 2NODs-CISD methods are more accurate, the errors becomes nontrivial as n increases. The curves of the 2NODs-LCCSD methods are satisfactorily flat. Between methods A and B, there is a difference as large as 0.44 kcal/mol. The difference is due partly to neglecting nonlinear terms.

Fig. 2. Size-inconsistency error in calculations of the super-system, $H_2O + nHe (n = 0-2)$.

Table 3. Comparison of energies in the single-bond dissocia-
tion of the HF molecule^{a,b}

Method	$R_{\rm e}$	$2R_e$	$3R_e$
RHF CISD CISDT CISDTO	-100.021 895 (0.00) -100.140 831 (95.48) (96.33) -100.141 887 -100.146252 (99.84)	$-99.815.395$ (0.00) -99.989956 (90.69) -99.994799 (93.20) -100.007273 (99.68)	-99.687648 (0.00) -99.932832 (86.16) -99.949135 (91.89) -99.970946 (99.55)
CCSD CCSD(T)	-100.144 859 (98.71) -100.146 139 (99.75)	-100.001 917 (96.90) -100.007805 (99.96)	(95.99) -99.960793 -99.994445 (107.81)
CASSCF MRCI	-100.047 717 (20.73) -100.142 822 (97.08)	-99.920565 (54.64) $-100,004,568$ (98.28)	-99.893933 (72.49) -99.969651 (99.10)
$2NODs-RHF$ 2NODs-RHF(relaxed) 2NODs-CISD	-100.097271 (60.51) -100.097352 (60.58) -100.146079 (99.70)	(58.92) -99.928 804 (58.94) -99.928859 -100.006361 (99.21)	-99.894218 (72.59) -99.894219 (72.59) -99.969707 (99.12)
$2NODs-LCCSD(A)$ $2NODs-LCCSD(B)$	-100.146 167 (99.77) -100.146 180 (99.78)	-100.007 773 (99.94) -100.007 989 (100.05)	-99.971383 (99.71) -99.971844 (99.87)
FCI	-100.146 457 (100.00)	$-100.007886(100.00)$	-99.972218 (100.00)

^aThe equilibrium bond distance is assumed to be 1.733 a.u.
^b Energies are in E_h . Numbers in parentheses are the percentage of the correlation energies

Fig. 3. Deviations from the FCI energies for the single-bond dissociation of the HF molecule

We also calculate a single-bond dissociation of the HF molecule using the same basis set. For MRCI, fullvalence CASSCF calculations are performed. The 1s orbitals in fluoride are added to the reference space in the MRCI [36] for fair comparisons. In Table 3, we compare calculated energies based on the SR, MR and nonorthogonal methods. All of the SR methods are based on the RHF references. The variational 2NODs-CISD is more accurate than the CCSD and MRCI methods. This agrees with the evaluation in Sect. 2. To see the performances more closely, we show deviations from the FCI energies in Fig. 3.

The deviation of the CCSD increases as the nondynamic correlation effects become important at larger internuclear distances. Although inclusion of perturbational triples [37] adjusts the error, the modification is too large at $3R_e$. The MRCI curve is more or less flat, including the same amount of deviation. This is due mostly to the size-inconsistency error according to the present choice of the active space. In the 2NODsCISD method, the size-inconsistency error is also nonnegligible except at R_e . This indicates an alternation of the reference function from case 2 to case 3, as discussed in Sect. 2. On the other hand, the LCC methods reproduce the entire correlation energies accurately. Both of the 2NODs-LCCSD methods are always accurate within 0.3% of the correlation energies.

6 Conclusions

We proposed the CI and the CC methods based on the superposition of nonorthogonal Slater determinants. It has been shown that the matrix elements over excited determinants can be efficiently generated and manipulated by introducing similarity transformed operators and bracket transformations. The method maintains the fundamental scaling property of the corresponding orthogonal problem. In the CC methods, we employ a multi-exponential wave operator including vacuum dependencies. The resulting CC method is size consistent as well as core extensive. In the calculation of the singlebond dissociation of the HF molecule, the variational 2NODs-CISD method is entirely more accurate than the SR-CCSD method. Moreover, approximate versions of the CC, the 2NODs-LCCSD methods, are capable of reproducing accurate potential energies in comparison with the conventional SR and MR methods.

Acknowledgements. The author is greatly indebted to D. Mukherjee. He is also grateful to S. Pal for helpful discussions. The present work is partly supported by the Grant-in Aids for Scientific Research (A) (No. 09304057) from the Ministry of Education, Science, Sports and Culture, Japan.

References

- 1. Roos BO (1987) In: Lawley KP (ed) Ab initio methods in quantum chemistry II. Advances in chemical physics 69. Wiley, New York, pp 399
- 2. Cizek J (1966) J Chem Phys 45:4151; (1969) Adv Chem Phys 14:35
- 3. Barlett RJ (1995) In: Yarkony DR (ed) Modern electronic structure theory, vol 2. World Scientific, Singapore, pp 1047 and references therein
- 4. Brandow B (1967) Rev Mod Phys 39:771; (1977) Adv Quantum Chem 10:187
- 5. Mukherjee D, Moitra RK, Mukhopadhyay A (1975) Mol Phys 30:1961; (1977) Mol Phys 33:955; Haque M, Mukherjee D (1984) J Chem Phys 80:5058
- 6. Lindgren I (1978) Int J Quantum Chem Symp 12:33
- Jeziorski B, Monkhorst HJ (1981) Phys Rev A24:1668
- 8. Mukherjee D, Pal S (1989) Adv Quantum Chem 20:291
- 9. Mukherjee D (1986) Chem Phys Lett 125:207; (1986) Int J Quantum Chem Symp 20:409
- 10. Meissner L, Kucharski SA, Bartlett RJ (1989) J Chem Phys 91:6187; Meissner L, Bartlett RJ (1990) J Chem Phys 92:561
- 11. Andersson K, Malmqvist P-A, Roos BO, Sadlej AJ, Wolinski K (1990) J Phys Chem 94:5483; Andersson K, Malmqvist P-Å, Roos BO (1992) J Chem Phys 96:1218
- 12. Hirao K (1992) Chem Phys Lett 190:374; (1992) Chem Phys Lett 196:397
- 13. Nakano H (1993) J Chem Phys 99:7983; (1993) Chem Phys Lett 207:372
- 14. Kozlowski PM, Davisdon ER (1994) Chem Phys Lett 222:615
- 15. Heully JL, Malrieu J-P, Zaitsevskii A (1996) J Chem Phys 105:6887
- 16. Mukhopadhyay D, Datta B, Mukherjee D (1992) Chem Phys Lett 197:236; Datta B, Mukherjee D (1995) Chem Phys Lett 235:31
- 17. Fukutome H (1988) Prog Theor Phys 80:611; Tomita N, Ten-no S, Tanimura Y (1996) Chem Phys Lett 263:687
- 18. Clark T (1988) J Am Chem Soc 110:1672; Ghailane R, Lepetit M-B, Malrieu J-P (1993) J Phys Chem 97:94; Hiberty PC, Humbel S, Archirel P (1994) J Phys Chem 98:11697
- 19. Löwdin PO (1955) Phys Rev 97:1474
- 20. Pauncz R (1995) The symmetric group in quantum chemistry. CRC Press, Boca Raton
- 21. Hinze J (ed) (1981) The unitary group for the evaluation of electronic energy matrix elements. Springer, Berlin Heidelberg New York
- 22. King HF, Stanton RE, Kim J, Wyatt RE, Parr RG (1967) J Chem Phys 47:1936; Lengsfield BH III, Jafri JA, Phillips DH, Bauschlicher CW Jr (1981) J Chem Phys 74:6849
- 23. Koch H, Christiansen O, Kobayashi R, Jørgensen P, Helgaker T (1994) Chem Phys Lett 30:233; Koch H, Sánchez de Merás A, Helgaker T, Christiansen O (1996) J Chem Phys 104:4157
- 24. Paldus J, Cizek J (1975) Adv Quantum Chem 9:105; Lindgren I, Morrison J (1986) Atomic many-body theory, 2nd edn. Springer, Berlin Heidelberg New York
- 25. Davidson ER (1975) J Comp Phys 17:87
- 26. Silverstone H, Sinanogolu O (1966) J Chem Phys 44:1899, 3608
- 27. Hose G, Kaldor U (1979) J Phys B12:3827
- 28. Laidig WD, Bartlett RJ (1984) Chem Phys Lett 104:424
- 29. Meller J, Malrieu J-P, Caballol R (1996) J Chem Phys 104:4068
- 30. Mahapatra US, Datta B, Bandyopadhyay B, Mukherjee D, Adv Quantum Chem (in press)
- 31. Bacskay GB (1981) Chem Phys 61:385
- 32. Laidig WD, Saxe P, Bartlett RJ (1987) J Chem Phys 86:887
- 33. Dunning TH Jr, Hay PJ (1977) In: Schaefer HF III (ed) Methods of electronic structure theory. Plenum, New York, pp
- 34. Huzinaga S, Andzelm J, Klobukowski M, Radzio-Andzelm E, Sakai Y, Tatewaki H (1984) Gaussian basis sets for molecular calculations. Elsevier, Amsterdam. This set is used for the isolated He atoms in this work
- 35. Langhoff SR, Davidson ER (1974) Int J Quantum Chem 8:61
- 36. Werner HJ, Knowles PJ (1988) J Chem Phys 89:5803
- 37. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M (1989) Chem Phys Lett 157:479; Raghavachari K, Pople JA, Head-Gordon M (1989) In: Kaldor U (ed) Many-body methods in quantum chemistry. Springer, Berlin Heidelberg New York, pp 215