THEORETICA CHIMICA ACTA

Theoret. Chim. Acta (Berl.) 66, 311-332 (1984)


# Development of a size-consistent energy functional for open shell states

**Sourav Pal*, M. Durga Prasad and Debashis Mukherjee****

Theory Group, Department of Physical Chemistry, Indian Association for the Cultivation of Science, Calcutta-700 032, India

In this paper we develop a size-consistent energy functional for a manifold of open-shell states using a specific form of the cluster expansion representation of the wave-function. Starting with a set of determinants $\{\Phi_A\}$ spanning a "full valence" manifold of a model space the cluster operator generates a set of correlated exact functions. The expectation value for energy for each such function consists of a numerator and a denominator. By using two different kinds of cluster operators, one set, $T_v$, including excitations out of the model space and the other, $X$, including transitions within the model space, the cancellation of the denominator leading to a size-consistent energy functional is proved. It is also shown that if the cluster operator exp $(\tilde{S}^{(n)})$ for an $n$-valence problem is built up hierarchically from the zero-valence core problem, where $\tilde{S}^{(n)}$ consists of $m<n$ cluster amplitudes frozen at their $m$-valence problem values and an additional set $S^{(n)}$ introduced at the $n$-valence level, the shift in energy upon addition of extra valence holes/particles can be calculated in a transparent manner. This build-up procedure, to be called "Subsystem Embedding Condition" (SEC) by us, allows calculation of difference energies as well. For calculating the cluster-amplitudes, and the combining coefficients for $\{\Phi_A\}$ we invoke Euler's variational principle for the energy-functional. First, the size-consistent expressions obtained from $\bar{H}^{(n)}_{AA}=\langle\chi^{(n)}_A|H|\chi^{(n)}_A\rangle|\langle\chi^{(n)}_A|\chi^{(n)}_A\rangle$ where $|\chi^{(n)}_A\rangle=\exp\tilde{S}^{(n)}|\Phi^{(n)}_A\rangle$ for the $n$-valence problem, are each varied with respect to the cluster-amplitudes $S^{(n)}$. The amplitudes are thus optimal for the functions $\{\chi_A\}$. Next, the combining coefficients for $\Phi^{(n)}_A$ are determined by varying the size-consistent energy

* *Present address*: Theoretical Chemistry Group, Physical Chemistry Division, National Chemical Laboratory, Pune 411008, India.

** Author to whom correspondence should be addressed

expression for each function $\Psi_K^{(n)} = \sum_A C_{AK}\chi_A^{(n)}$ with respect to the combining coefficients. The resulting equation is analogous to a secular equation of a CI problem with a hermitian effective Hamiltonian $\bar{H}$ replacing $H$ and having the dimension of the model space. The eigenvalues $E_K^{(n)}$ are the correlated energies for the functions. As the operators $T_v$ and $X$ are not all independent - being related by the conditions guaranteeing cancellation of the denominator - the variation of the quantities $\bar{H}_{AA}$ with respect to the cluster amplitudes have to be done using Lagrange's method of undetermined multipliers. The working of the theory has been illustrated explicitly deriving the equations for such difference energies as IP or EA for a closed shell system.

**Key words:** Open shell many body theory—size consistency—coupled cluster theory

## 1. Introduction

The development of theories for electron correlation and accurate difference energies of chemical interest including correlation has always been a major goal of quantum chemists. Several highly successful formulations of the correlation energy for the closed-shell problem have emerged over the past two decades. Sinanoglu's many theory (MET) [1] and the consequent development leading to coupled-cluster method (CCM) [2–4] remain a fundamental step in this direction. Subsequently, there have been attempts to generalize the cluster expansion techniques to open-shell systems. Silverstone and Sinanoglu generalized the closed-shell MET to open shell systems [7] and the open shell analogues of the coupled-cluster method were developed by Mukherjee et al. [8, 9], Offerman, Ey and Kummel [10], Lindgren [11], Jeziorski and Monkhorst [12] and others [13–18].

The important structural difference of this development vis-a-vis the linear variation method of CI stems from the cluster representation of the wave function which ensures the separability criterion for many electron systems [5] and the concomitant property of size-consistency [6].

Recently we introduced a variation-based size-consistent theory for the closed-shells [19], where the closed-shells wave-function was taken to be of coupled-cluster form. The expectation value for the energy becomes size-consistent when the norm of the function appearing in the denominator of the expectation value is cancelled with the disconnected terms factoring out of the numerator. The cluster components were determined by invoking an Euler Variational principle for the full (or truncated) size-consistent energy functional - similar in spirit to the "Varied Portion Approach" (VPA) of Sinanoglu [1]. In the present paper we want to pursue the same idea in an attempt to incorporate the treatment of the open-shell states. We shall thus start with expectation-value like quantities for the energies of a manifold of open-shell states using a cluster expansion of the wave-function akin to the Silverstone–Sinanoglu approach [7], and prove that it is possible to cancel the norm in the denominator with the disconnected terms factored out from the numerator exactly as in the closed-shell situation.

The development of a size-consistent energy functional and a consequent Euler variation method for the determination of a manifold of energy levels has not been systematically pursued so far. Only Reitz and Kutzelnigg [13] used a common unitary ansatz description for both a parent closed and the derived open shell states and developed a size-consistent theory for the respective energies. The unitarity of the cluster operator rendered a multicommutator structure of the numerator through Hausdorff formula guaranteeing size-consistency [4, 38], and the denominator was identically taken to be unity. This has since been generalized by Kutzelnigg and coworkers [36]. As we shall discuss in Sect. 5, the unitary cluster operator tends to proliferate the number of terms involved in the energy calculation (see also, Ref. [39]). A nonunitary cluster operator, suitably chosen, would eliminate most of them. The norm cancellation is then not automatic, however, and suitable manipulations are called for. Basically, what we shall do is to introduce two sets of cluster operators in the wave-operator $W$ for the open shell states. If the starting functions are said to span a model space, the one set of cluster amplitude excites to "virtual states" orthogonal to those in the model space. The other, auxiliary set "scatters" only between the model space states. The role of the auxiliary set is to force $W$ to behave as a norm-conserving operator in the model space: $[W^+W]_{MS} = 1_{MS}$, where $[\ ]_{MS}$ stands for the projection on to the model space. This is the principal point of our departure from the closed shell development.

## 2. The theory

### *2.1. Preliminary considerations*

In our variational development for the open-shell states, we shall follow closely the general idea of quasidegenerate MBPT [29, 36, 38]: viz, an effective Hamiltonian will be constructed first by means of a wave-operator that applies to all the matrix-elements of this effective Hamiltonian. The energies are obtained by diagonalizing this effective Hamiltonian over suitable model space functions at the end (see e.g. Ref. [36] for an extensive discussion on this point). Clearly, such an approach demands that we treat the various $N$-electron sectors of the Hilbert space on the same footing. This is implicit and automatic in the quasidegenerate MBPT [29], but requires a careful and explicit consideration for nonperturbative methods, as has been done, for example, in the nonvariational works of Mukherjee et al. [9], Ey et al. [10] and Kutzelnigg [36]. For the variational treatment to be envisaged by us, a similar development is warranted.

We adopt a hole-particle representation of the functions entering our theory. A closed-shell "core" function defines the "hole" and "particle" orbitals of the problem. The open-shell reference functions spanning the "model space" are built up by the action of products of appropriate hole and/or particle operators on the core function. The orbital labels on these operators are taken as valence (hole valence or particle valence). The number $n$ of these valence occupancy signifies that we are dealing with an $n$-valence problem. Given an $n$, and the total number of electrons $N$, the type of $n$-valence determinants spanning the

model space is unique and may be obtained by assigning unit occupancies to a set of $n$-valence orbitals (valence holes and/or valence particles) chosen from among all the valence orbital labels in all possible manner. The model space for a specific $n$-valence problem thus spans a "full valence" space. In the hole-particle formulation, existence of a unique effective Hamiltonian furnishing energies for the various $N$-electron sectors implies that we have a unique transformation procedure for the Hamiltonian such that the various $n$-valence sectors are treated on the same footing. If we confine ourselves to a particular $n$-valence sector, then the various cluster amplitudes become linearly dependent, as has been noted by several workers before [9, 12, 20, 36]. This *redundancy problem* is a feature quite special to the open-shell situation.

To avoid the redundancy problem, we clearly need to postulate the existence of a unique wave-operator $W$ for all the $n_h$ hole-$n_p$ particle sectors of the Hilbert space. There are two possible ways to achieve this, and both have been tried. One way is to postulate that same cluster amplitudes entering the cluster expansion of $W$ apply to all the $n$-valence sectors of the Hilbert space. This *couples* the various $n$-valence sectors of the Hilbert space and one has to solve for the cluster amplitudes simultaneously for the coupled problems [9, 20, 36]. The other is to postulate an ansatz for the wave-operator as an exponential operator written in *normal order* with respect to the core-function as the vacuum [11]. In that case there is a *decoupling* of the various $n$-valence sectors of the Hilbert space [11, 20, 35, 40][1]. We shall follow the second strategy in our development. We shall, however, construct the wave-operator in a hierarchical manner, taking advantage of the fact that only the $n$-valence projection of the wave-operator is active in a particular $n$-valence problem.

Let us assume that $W$ can be written in the form

$$W = N[\exp(S)] \tag{1}$$

where $N[\ ]$ signifies the normal ordering with respect to the core-function as the vacuum. The cluster amplitudes appearing in $S$ can all be classified depending on the number of valence destruction operators they contain. Thus, we may denote operators in $S$ having $n$ valence-destruction operators as $S^{(n)}$. The simplifying role of the normal ordering in Eq. (1) is now evident: for an $n$-valence problem, only those components $S^{(m)}$ of $S$ can give non-vanishing contribution upon acting on the $n$-valence model space functions that satisfy $m \leq n$; the $m > nS^{(m)}$ operators will give identically zero by acting on the model space functions. Defining operators $\tilde{S}^{(n)}$ and $W^{(n)}$ as

$$\tilde{S}^{(n)} = \sum_{m=0}^{n} S^{(m)} \tag{2a}$$

$$W^{(n)} = N[\exp(\tilde{S}^{(n)})] \tag{2b}$$

[1] In case there are valence holes in the problem, the effective Hamiltonian *can* couple various $n$-valence sectors because it conserves number of electrons and *not* holes or particles separately. For an $n_v$ valence problem with $n_h$ valence holes, all the lower valence sectors differing by $h$-$p$, $2h$-$2p$... etc are coupled. But the basic physics we are going to describe is not altered by this complication

we have

$$WP^{(n)} = W^{(n)}P^{(n)} \tag{3}$$

where $P^{(n)}$ is the projector onto the $n$-valence model space. $W^{(n)}$ is thus the active component of $W$ for the $n$-valence problem. The norm-preservation condition within the model space also becomes

$$P^{(n)}W^{\dagger}WP^{(n)} = P^{(n)}W^{\dagger(n)}W^{(n)}P^{(n)}. \tag{4}$$

We now describe the hierarchical build-up of $W$ successively through various $W^{(m)}$, $m = 0 \ldots n$, following our earlier developments of nonperturbative open-shell coupled cluster theory [9, 20, 35, 40]. This build-up has been termed "The Subsystem Embedding Condition" (SEC) by us. SEC gives a practical and unique procedure for generating higher rank matrix-elements of the effective Hamiltonian as one goes to states having increasing number of valence occupancies. As we shall also see, SEC is very useful for ensuring easy and transparent cancellation of common terms for difference energy calculations.

We start with the zero-valence core problem first, and build up hierarchically, finally stopping at the desired $n$-valence level, say. At the zero-valence core level, the cluster operators $\tilde{S}^{(0)} = S^{(0)}$ induce hole to particle excitations only, as in the closed-shell coupled cluster theory [2] and may be denoted by $T_c$. The associated cluster amplitudes can be obtained by variationally solving the closed-shell variational equations derived by us [19]. The one-valence problem is solved next. The operator $\tilde{S}^{(1)}$ may be written as

$$\begin{aligned} \tilde{S}^{(1)} &= \tilde{S}^{(0)} + S^{(1)} \\ &= T_c + S^{(1)} \end{aligned} \tag{5a}$$

where $S^{(1)}$ contain two distinct kinds of operators. Calling the operators inducing excitations from the one-valence model space to virtual space as $T_v^{(1)}$, and those inducing model space to model space "scatterings" as $X^{(1)}$, we may write

$$S^{(1)} = T_v^{(1)} + X^{(1)}. \tag{5b}$$

The major dynamical role of $T_v^{(1)}$ is to introduce relaxation and correlation effects through (core-valence) $\rightarrow$ (particle, valence) excitations and it must involve apart from core excitation, one valence destruction operator. Suitable size-consistent energy functional will be constructed from $W^{(1)}$, defined through (2) and (3), and the amplitudes for $T_v^{(1)}$ and $X^{(1)}$ will be obtained from variational equations corresponding to the stationarity of the energy functional and the norm preservation condition (Eq. (4)) for $n = 1$. The amplitudes for $T_c$ will be frozen at their core values. One proceeds to the two-valence problem next, and introduce operators $S^{(2)} = T_v^{(2)} + X^{(2)}$ with two-valence destruction operators. In the determination of $S^{(2)}$ amplitude $S^{(1)}$ amplitude are kept frozen. Thus, generally, we may write

$$\tilde{S}^{(n)} = \tilde{S}^{(n-1)} + S^{(n)} \tag{6a}$$

with

$$S^{(n)} = T_v^{(n)} + X^{(n)} \tag{6b}$$

where $S^{(n)}$ is the new cluster operator at the $n$-valence level. SEC thus generates a unique $W$ needed for constructing a unique effective Hamiltonian but generates it hierarchically by essentially a decoupling procedure for the various $n$-valence sectors.

We have still not indicated how the factorization leading to a connected series is obtained, nor have we discussed how the variational equations determining the $T_v$ amplitudes are obtained. These aspects we discuss in Sects. 2.2 and 2.3.

### *2.2. The SEC and the norm-cancellation*

For an $n$-valence problem, we assume that there are starting (reference) functions $\Psi_K^{0(n)}$ which get converted into exact functions $\Psi_K^{(n)}$ by the action of $W$. Using Eq. (3), we have

$$\begin{aligned} &W\Psi_k^{0(n)} = W^{(n)}\Psi_K^{0(n)} = \Psi_K^{(n)} \\ &K = 1, \ldots, N_D^{(n)} \end{aligned} \tag{6}$$

where $N_D^{(n)}$ is the dimension of the $n$-valence model space. Clearly, the functions $\Psi_K^{0(n)}$ can be expressed as linear combination of the $n$-valence model space determinants $\Phi_A^{(n)}$:

$$\Psi_K^{0(n)} = \sum_A C_{AK}^{(n)} \Phi_A^{(n)} \tag{7}$$

with $C_{AK}^{(n)}$ as expansion coefficients. We leave the specification of $C_{AK}^{(n)}$'s for the present. Each $\Phi_A^{(n)}$ is obtained by the action of a product of $n$ appropriate "valence creation" operators on the core function $\Phi_c$:

$$\Phi_A^{(n)} = \Omega_A^{(n)\dagger} \Phi_c \tag{8}$$

with $\Omega_A^{(n)\dagger}$ signifying the product of $n$-valence creation operators.

The energy functional $E_K^{(n)}$ may be written as

$$E_K^{(n)} = \frac{\langle \Psi_K^{(n)} | H | \Psi_K^{(n)} \rangle}{\langle \Psi_K^{(n)} | \Psi_K^{(n)} \rangle}, \quad K = 1, \ldots, N_D^{(n)}. \tag{9}$$

Substituting Eq. (6) in Eq. (9), and using Eqs. (2) and (7), we have

$$E_K^{(n)} = \frac{\sum_{A,B} \langle \Phi_A^{(n)} | N[\exp(\tilde{S}^{(n)\dagger})] H N[\exp(\tilde{S}^{(n)})] | \Phi_B^{(n)} \rangle C_{AK}^{(n)} C_{BK}^{(n)}}{\sum_{A,B} \langle \Phi_A^{(n)} | N[\exp(\tilde{S}^{(n)\dagger})] N[\exp(\tilde{S}^{(n)})] | \Phi_B^{(n)} \rangle C_{AK}^{(n)} C_{BK}^{(n)}}. \tag{10}$$

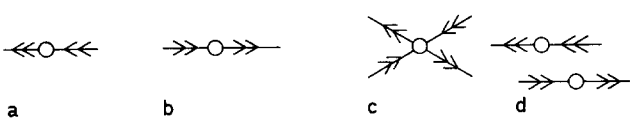
Our motivation is to get rid of the denominator in Eq. (10) leading to a connected series.

Using the core-function $\Phi_c$ as the vacuum, we can analyse the structure of the denominator in Eq. (10), as a sum over various contracted operators using

Generalized Wick's Theorem (GWT). Clearly, only those terms in the expansion will survive which have only valence operators left uncontracted, leading to model space to model space transitions.

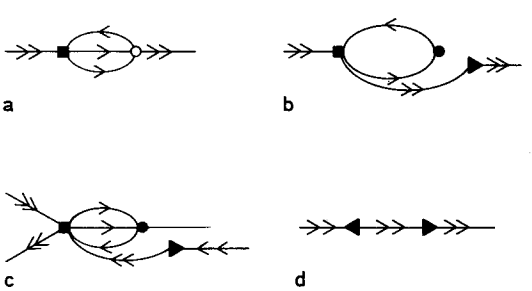
All such terms may be depicted as open diagrams having open lines labelled by valence indices only. A typical term will contain products of several composite operators in normal order. Each composite operator will involve contractions among various cluster operators. Figs. 1 depict typical box structures of the denominator for one and two valence situations, where we have followed the usual double-arrow representation for the valence lines. Also we shall use the single-arrow representation for the internal lines in the diagrams to follow. Each box contains connected terms obtained by contracting $T_c^+$, $T_v^+$ and $X^+$ operators with $T_c$, $T_v$ and $X$ operators which have shapes like valence to valence "scattering". The normal ordering ansatz, Eq. (2b) ensures that the contractions within the $\tilde{S}^{(n)\dagger}$ operators or the $\tilde{S}^{(n)}$ operators are absent. It follows that in a box having $n$-valence lines, each operator $T_v$ or $X$ can have at most $n$ valence destruction operators and similarly each operator $T_v^+$ or $X^+$ can have at most $n$ valence creation operators. Figs. 2 illustrate this diagrammatically, taking typical examples leading to one and two valence boxes. We have followed the Hugenholtz convention of representing the antisymmetrized matrix-elements by vertices.

We shall now see that SEC allows us to cancel the denominator in a unique manner for any $n$-valence problem. We start from the zero valence problem. The boxes would then be all closed diagrams (corresponding to $n = 0$ valence problem) and because of the normal ordering in the wave-operator, would contain only $T_c^+$ and $T_c$ operators; any $T_v$, $T_v^+$, $X$ and $X^+$ operator would have left at least one valence operator uncontracted. The set of all such boxes clearly constitutes




**Figs. 1** Typical diagrammatic structures of the denominator of Eq. (10) for one and two valence problems. Each circle indicates a series of connected terms having the orientation of valence lines as shown. (a) and (b) depict one-particle and one-hole situations. (c) and (d) are the corresponding sets for two-valence (hole-particle) situation. Fig. (d) is disconnected

**Figs. 2** Some typical diagrams contributing to the circles of Figs 1. Figs. (a), (b) and (d) contribute to Fig. 1(b), while (c) contributes to Fig. 1(c). The vertices with filled circles denote $T_v$ or $T_c$ operators, as the cases may be, and those with filled squares depict the $T_v^\dagger$ or $T_c^\dagger$ operators, ► corresponds to an $X$ vertex, and ◄ corresponds to an $X^\dagger$ vertex



the norm of the core function $x_c = \exp(T_c)\,\Phi_c$ quite analogous to the closed-shell theory [19, 21], because the norm consists of all the closed boxes obtained by contracting $T_c^+$ and $T_c$ operators in all possible manner. Let us write the norm $\langle\Phi_c|\exp(T_c^+)\exp(T_c)|\Phi_c\rangle$ as $N_{cl}$.

Now, for a one-valence problem, the nonvanishing diagrams will have either closed boxes or at most one box having one valence line with the rest having closed boxes. The overall contribution $D_1$ from all the diagrams to the denominator of a one-valence problem then factorizes as follows:

$$D_1 = N_{op}^{(1)} N_{cl} \tag{11}$$

where $N_{op}^{(1)}$ contains the total contribution from all the boxes having open valence lines only. The lines on the right and left of all the boxes will be labelled by valence indices. Thus, for example, the lines on Fig. 1(a) will have particle-valence labels; those on Fig. 1(b) will have hole-valence labels, and so on. Noting that there can be at most one $X^{(1)}$ and/or $X^{+(1)}$ operator in the box, we may invoke a condition that each $(a, b)$ element of all the terms coming from all the one valence boxes cancel among each other. Thus, if $N_{OP}^{(1)}$ be written as

$$N_{op}^{(1)} = \sum_{b,a} (M_{op}^{(1)})_{ba} N[a_b^+ a_a] \tag{12}$$

we demand that

$$(M_{op}^{(1)})_{ba} = 0 \text{ for all valence labels } b, a. \tag{13}$$

The total number of such equations equals the number of $X$ or $X^+$ operators. Eqs. (11) and (12) imply that for a one-valence problem, the denominator in Eq. (10) can be written as

$$\langle\Psi_K^{0(1)}|D_1|\Psi_K^{0(1)}\rangle = \sum_{A,B} C_{AK}^{(1)} C_{BK}^{(1)} \langle\Phi_A^{(1)}|\Phi_B^{(1)}\rangle \cdot N_{cl} = \sum_A C_{AK}^{(1)} C_{AK}^{(1)} N_{cl} \tag{14}$$

for the one valence problem. Choosing the coefficient matrix $C^{(1)}$ as orthonormal, we have

$$\langle\Psi_K^{0(1)}|D_1|\Psi_K^{0(1)}\rangle = N_{cl}. \tag{15}$$

For the two-valence problem, likewise, the overall nonvanishing operator $D_2$ for the denominator may be written as

$$D_2 = N_{op}^{(2)} N_{cl} \tag{16}$$

where $N_{op}^{(2)}$ consists of one box with two valence lines and two boxes having one valence line each. A typical situation is depicted in Figs. 1(c) and 1(d). 1(c) is a two-valence box, and 1(d) is a product of two boxes. The one-valence boxes have the matrix-elements $M_{op}^{(1)}$, since no $T_v^{(2)}$, $T_v^{+(2)}$, $X^{(2)}$ or $X^{+(2)}$ can contribute to a one-valence box, and contribution of all such one-valence boxes is zero from Eq. (13). The contribution of the matrix-element of the form $(M_{op}^{(2)})$ in two-valence box will come from diagrams having at most one $T_v^{(2)}/X^{(2)}$ and/or one $T_v^{+(2)}/X^{+(2)}$ operator. Again we choose the matrix-elements of the operators $X^{(2)}$ or $X^{+(2)}$ in

such a way that

$$(M_{op}^{(2)})_{ab,cd} = 0 \tag{17}$$

for all choices of the outgoing and incoming valence labels $(a, b)$ and $(c, d)$. The denominator of Eq. (10) for the two-valence problem is then

$$\langle \Psi_K^{0(2)} | D_2 | \Psi_K^{0(2)} \rangle = \sum_A C_{AK}^{(2)} C_{AK}^{(2)} N_{cl} = N_{cl} \tag{18}$$

for orthonormal choice of $C^{(2)}$. Proceeding hierarchically in the manner, we thus find quite generally that for any $n$-valence problem, we have

$$(M_{op}^{(n)}) a_1 \ldots a_n, b_1 \ldots b_n = 0 \tag{19}$$

and

$$\langle \Psi_K^{0(n)} | D_n | \Psi_K^{0(n)} \rangle = N_{cl}. \tag{20}$$

Let us note that the constraints like (19) imposed on matrix-elements of the operators $X^{(n)}$ ensure that the denominator for *any* state $K$ of *any* $n$-valence manifold has the value $N_{cl}$. This choice thus reduces the denominator simply to the norm $N_{cl}$ for each function $\Psi_K^{0(n)}$ for all $K$ and $(n)$.

Let us now consider the numerator of the energy functional, Eq. (10). Using GWT, we may again expand the product operators $N[\exp(\tilde{S}^{(n)\dagger})]HN[\exp(\tilde{S}^{(n)})]$ in normal order containing contractions among various operators, excluding again contractions within operators in each group of $S^{(n)\dagger}$ and $S^{(n)}$ operators. Clearly, a typical term in one such contracted component obtained from the expansion using GWT will have a set of operators contracted together containing the Hamiltonian and the rest having contractions containing $S^{(n)\dagger}$ and $S^{(n)}$ operators. A typical term may then be depicted schematically as in Fig. 3 having one square-box with open valence lines and containing the Hamiltonian vertex and a sequence of circled boxes just as generated in the denominators. Because of the constraints (19) the overall contribution of the nonvanishing terms of the operators contained in the numerator of Eq. (10) will come from all the square boxes containing at most $n$ open lines for an $n$-valence problem apart from a factor $N_{cl}$. Some typical diagrams contributing to the square box are depicted in

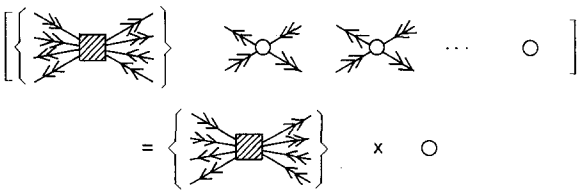
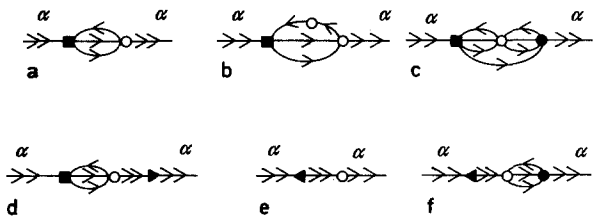


**Fig. 3** A typical diagrammatic structure of the numerator of Eq. (10). The square box depicts the set of connected diagrams containing the Hamiltonian vertex. The rest are circled boxes having the same structure as in Figs 1. The circle with no external line depicts $N_{cl}$




**Figs. 4** Some typical diagrams contributing to the square box of Fig. 3 for one valence hole situation. The unfilled circles with two/four lines denote the one/two body operator of $H$

Fig. 4. The numerator $N'$ may thus be written as

$$N' = \langle \Psi_K^{0(n)} | N[\exp(\tilde{S}^{(n)\dagger}) H \exp(\tilde{S}^{(n)})] | \Psi_K^{0(n)} \rangle_{\text{conn}}, \; N_{cl} \tag{21}$$

which clearly shows that $N'$ factors out into a size-consistent connected term containing the Hamiltonian times the denominator $N_{cl}$. The energy functional for $E_K^{(n)}$ finally takes the form

$$E_K^{(n)} = \langle \Psi_K^{0(n)} | N[\exp(\tilde{S}^{(n)\dagger}) H \exp(\tilde{S}^{(n)})] | \Psi_0^{(n)} \rangle_{\text{conn}}. \tag{22}$$

Eq. (22) contains linked terms only and is thus size consistent. Calling the operator in $N[\;]_{\text{conn}}$ in Eq. (22) as an operator $\bar{H}$, we may write Eq. (22) after expanding in terms of $\Phi_A^{(n)}$'s using Eq. (8) as,

$$\begin{aligned} E_K^{(n)} &= \sum_{A,B} C_{AK}^{(n)} C_{BK}^{(n)} \langle \Phi_A^{(n)} | \bar{H} | \Phi_B^{(n)} \rangle \\ &= \sum_{AB} C_{AK}^{(n)} C_{BK}^{(n)} \bar{H}_{AB} \end{aligned} \tag{23}$$

$\bar{H}$ may thus be viewed as an *effective Hamiltonian* acting on the model space functions only but generating the exact energies $E_K^{(n)}$. From Eq. (22) it follows in a straightforward manner, that $\tilde{S}^{(n)\dagger}$ operators appear symmetrically with $\tilde{S}^{(n)}$ in $\bar{H}$, and thus $\bar{H}$ is by construction automatically hermitian. For any truncated expansion also, $\bar{H}$ remains hermitian if all the terms containing $\tilde{S}^{(n)\dagger}$ and $\tilde{S}^{(n)}$ are kept to the same power in the expansion. Clearly a variation of $E_K^{(n)}$'s with respect to the $C_{AK}^{(n)}$'s would lead to an eigenvalue equation with $C_K^{(n)}$ as an eigen-vector. It is therefore consistent to choose $C^{(n)}$'s as orthonormal, as was indeed assumed by us.

We now observe that, as a consequence of the SEC (and, as a special case thereof, the core-valence separation) appropriate difference energies emerge as size-consistent expressions. If by $E_c$ we denote the correlated core energy, corresponding to the function $\chi_c = e^{T_c}\Phi_c$, then the difference energies $\Delta E_K^{(n)}$ defined as $(E_K^{(n)} - E_c)$, can be obtained simply by dropping the completely closed diagrams obtained from $N[\;]_{\text{conn}}$ or equivalently from the operator $\bar{H}$, from the expression for $E_K^{(n)}$. This is because, $E_c$ consists of all the closed diagrams obtained by connecting powers of $T_c^+$ and $T_c$ with an $H$ vertex from left and right of $H$ respectively. In $\bar{H}$, the closed diagrams cannot come from terms containing even

a single $T_v^{+(n)}/T_v^{(n)}$ or $X^{(n)}/X^{+(n)}$ operator. Hence, if

$$\begin{aligned}\bar{H} &= \bar{H}_{cl} + \bar{H}_{op} \\ &= E_c + \bar{H}_{op},\end{aligned} \tag{24}$$

where $\bar{H}_{op}$ consists of open diagrams only, we have

$$\Delta E_k^{(n)} = \langle \Psi_K^{0(n)} | \bar{H}_{op} | \Psi_K^{0(n)} \rangle. \tag{25}$$

For a one-valence problem, $\Delta E_K^{(1)}$ for $(N \mp 1)$ electron problem, with $\Phi_c$ as the ground $N$-electron Hartree-Fock function, will yield IP's and EA's. This case will be further illustrated in Sect. 3 as a concrete application.

### *2.3. Euler variational equations from the energy functional*

If we now postulate an Euler variational principle for the energy functional $E_K^{(n)}$, then the equations obtained by varying $E_K^{(n)}$ with respect to the $\tilde{S}^{(n)}$ cluster amplitudes will contain the coefficients $C_{AK}^{(n)}$ in them. *We now want to be able to decouple the variational equations for $\tilde{S}^{(n)}$ from those for the coefficients $C^{(n)}$.* In that case there will be a single set of variational equations for $\tilde{S}^{(n)}$ independent of $K$. The variational equations for $C_{AK}^{(n)}$'s will provide us next with the $E_K^{(n)}$'s and all the $N_D^{(n)}$ energies $E_K^{(n)}$'s will be generated from a single diagonalization. This will also enable us to guarantee both the orthogonality and the noninteracting nature of the functions $\Psi_K^{0(n)}$ even in a truncated calculation:

$$\langle \Psi_K^{0(n)} | \Psi_L^{0(n)} \rangle = \delta_{KL}. \tag{26}$$

$$\langle \Psi_K^{0(n)} | \bar{H} | \Psi_L^{0(n)} \rangle = E_K^{(n)} \delta_{KL}. \tag{27}$$

In the complete cluster expansion of the functions $\Psi_K^{(n)}$ we have two different sets of parameters: (a) a set of linear parameters $C_K$ and (b) a set of non-linear parameters - the cluster amplitudes of $\tilde{S}^{(n)}$. If we look upon the functions $\chi_A^{(n)} = N[\exp(\tilde{S}^{(n)})]\, \Phi_A^{(n)}$'s as correlated basis-functions whose linear combinations furnish the exact function, then a sensible way to optimize the parameters suitable for generating energies of a manifold is a two-step procedure. We minimize the expectation values

$$\langle \chi_A^{(n)} | H | \chi_A^{(n)} \rangle / \langle \chi_A^{(n)} | \chi_A^{(n)} \rangle$$

with respect to the cluster amplitudes. Keeping the cluster amplitudes fixed, we next optimize the energy expressions with respect to the linear expansion coefficients to get the respective energies for the manifold. This recipe, for example, has recently been advocated in a different context by Hendecovic [37].

For the variation of the nonlinear parameters, we again invoke SEC. We start from the zero-valence core problem, and get the cluster-amplitudes $t_c$ by solving the Euler equation for the closed-shell problem:

$$\frac{\partial \bar{H}_{cl}}{\partial t_{c_i}^{+}} = 0 \text{ for all } i \tag{28}$$

as described in our closed-shell development [19]. For one-valence problems the $t_c$-amplitudes are kept fixed and the amplitudes $s_i^{(1)}$ of the operator $\tilde{S}^{(1)}$ are to be treated as variables. We take each $\Phi_A^{(1)}$ in turn, and calculate the size-consistent expression $\bar{H}_{AA}^{(1)}$:

$$\bar{H}_{AA}^{(1)} = \frac{\langle \chi_A^{(1)}|H|\chi_A^{(1)}\rangle}{\langle \chi_A^{(1)}|\chi_A^{(1)}\rangle} = \langle \Phi_A^{(1)}|\bar{H}|\Phi_A^{(1)}\rangle. \tag{29}$$

One may guess that the cluster-amplitudes $s_i^{(1)}$ may be found by directly solving the equations

$$\frac{\partial \bar{H}_{AA}^{(1)}}{\partial s_i^{(1)\dagger}} = 0 \text{ for all } A \text{ and } i. \tag{30}$$

This would have been true provided all $s_i^*$'s are independent, but they are not. The $X^{(1)}$ and $T_v^{(1)}$ operators are related through the constraints, Eq. (13). However, as we show later by using Lagrange's method of undermined multipliers, we can bypass this problem.

The two-valence problem comes next, and we build up successively in a similar fashion reaching finally the $n$-valence level. The general system of equations for the $s_i^{(m)}$ amplitudes would be the constrained variation equations of all the $\bar{H}_{AA}^{(m)}$'s with respect to the $m$-valence cluster amplitudes.

We now derive the Euler–Lagrange constrained variational equations for determining the $s_i^{(n)}$ amplitudes. We designate the $T_v^{(n)}$ cluster amplitudes as $t_v^{(n)}$ and the $X^{(n)}$ as $x_I^{(n)}$. The index $I$ for an $n$-valence problem runs over the pair of labels $(A, B)$ with each $A$ and $B$ built as an ordered set of $n$-valence labels specifying $\Phi_A^{(n)}$ and $\Phi_B^{(n)}$. The constraint equations (19) may thus be compactly written as

$$(M_{op}^{(n)})_{AB} = 0 \text{ for all } A, B. \tag{31}$$

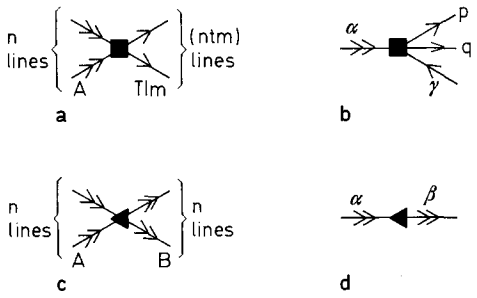
The variational equations for the amplitudes may thus be written as

$$\frac{\partial \bar{H}_{AA}^{(n)}}{\partial t_{v_i}^{(n)\dagger}} - \sum_{C,D} \lambda_{CD}^{A} \frac{\partial (M_{op}^{(n)})_{CD}}{\partial t_{v_i}^{(n)\dagger}} = 0, \tag{32}$$

$$\frac{\partial \bar{H}_{AA}^{(n)}}{\partial x_I^{(n)\dagger}} - \sum_{C,D} \lambda_{CD}^{A} \frac{\partial (M_{op}^{(n)})_{CD}}{\partial x_I^{(n)\dagger}} = 0. \tag{33}$$

Variation with respect to $t_{v_i}$ and $x_I$ amplitudes furnishes the conjugate Eqs. of (32) and (33).

For real amplitudes, the conjugate equations may be dropped, the $t_{v_i}^{(n)\dagger}$ and $x_I^{(n)\dagger}$ amplitudes equated to the corresponding $t_{v_i}^{(n)}$ and $x_I^{(n)}$ amplitudes, and Eqs. (31) to (33) determine $t_{v_i}^{(n)}$, $x_I^{(n)}$ and the Lagranges' multipliers completely. Owing to the normal ordering in the cluster operator, there can at most be one $t_v^{(n)}/x^{(n)}$ and/or $t_v^{(n)\dagger}/x^{(n)\dagger}$ amplitude in any $\bar{H}_{AA}^{(n)}$ or $(M_{op}^{(n)})_{AB}$. There can thus be at the most a quadratic term containing $s_i^{(n)}$ amplitudes in the functional to be differenti-




**Figs. 5** A typical shape of a $T_v^\dagger$ vertex with $n$ valence holes on the left and outgoing $(n+m)$ $p-h$ lines on the right is shown in Fig. 5(a). Fig. 5(b) shows a concrete case for a one-valence-hole situation. A typical shape of $X^{(n)\dagger}$ vertex is shown in Fig. 5(c). A concrete case for a one valence hole situation is depicted in Fig. 5(d)

ated. There are further simplifications in the Eqs. (32) and (33) which can best be described using a diagrammatic analysis.

Let us assume that a typical $T_{v_i}^{(n)\dagger}$ operator has the shape as shown in Fig. 5(a). Its valence lines on the left are labelled by an ordered set of $n$-valence indices, which we designate as $A$. Its lines on the right are naturally not all valence, which are collectively labelled as $\pi_m$. Clearly, all choices of the pair $(\pi_m, A)$ generate all the $t_v^{(n)\dagger}$ matrix-elements. As a concrete example consider the one valence-hole problem. The typical vertex is shown in Fig. 5(b). There is only one valence index in the $n$-ordered valence set (as $n$ is one), and $A$ thus can take on the values $\alpha$ – the hole labels. The set $\pi_m$ will consist of all the ordered triples $(p, \gamma, \delta)$ that can be formed by taking two hole labels and one particle label from among the hole and particle indices respectively. All such possible choices of $\pi_m$ and $A$ will exhaust all distinct $t_v^{(1)\dagger}$ matrix-elements. In a similar vein, a typical $X_I^{(n)\dagger}$ is shown in Fig. 5(c), where the pair $(A, B)$ defines the label $I$. A concrete example for the one valence-hole case is depicted in Fig. 5(d). The variational equations are obtained if we delete the $t_{v_i}^\dagger$ or $X_I^\dagger$ vertices from the square boxes. Thus, typically with the diagrams as shown in Figs. 4(a) to 4(d) as components of a box of Fig. 1(b), the variational expression $\partial \bar{H}_{\alpha\alpha}/\partial t_{v_i}^{(1)\dagger}$ in Eq. (32) for $t_{v_i}^\dagger$ as $\langle \gamma\delta | t_v^{(1)\dagger} | \alpha p \rangle_a$ (for the choice $\pi_m = (p\gamma\delta)$, $A = \alpha$) may be diagrammatically shown as in Fig. 6(a). The variational expression for $\partial \bar{H}_{\alpha\alpha}/\partial x_I^{(1)\dagger}$ in Eq. (33) for $x_I^{(n)} = \langle \gamma | x_I^{(1)\dagger} | \alpha \rangle$ for diagrams 4(e) and 4(f) contributing to the box Fig. 1(b) will likewise look like Fig. 4(b). In the general case, for a variational expression $\partial \bar{H}_{AA}/\partial t_{v_i}^{(n)\dagger}$ the diagrams generated will be labelled by the set $\pi_m$ for the open

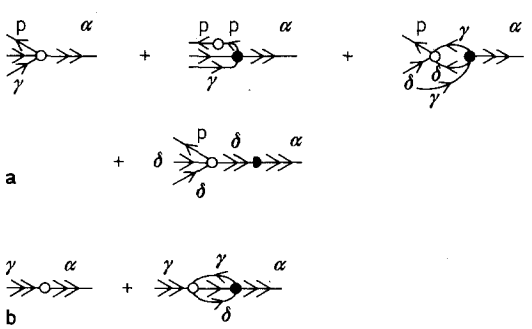



**Fig. 6** (a) The diagrams generated by differentiating the terms representing the diagrams of Figs. 4(a) to 4(d) with respect to $\langle \gamma\delta | t_v^{(1)\dagger} | \alpha p \rangle_a$ amplitude. (b): The diagrams generated by differentiating the terms representing the diagrams of Figs. 4(e) and 4(f) with respect to $\langle \gamma | x^{(1)\dagger} | \alpha \rangle$ amplitude

lines on the left and by the set $A$ for the open lines on the right for the choice $(\pi_m, A)$ of the labels of $t_{v_i}^{(n)}$ matrix-elements. Similarly, for the expression $\partial \bar{H}_{AA}/\partial x_I^{(n)\dagger}$ with $I$ as $(A, C)$ the open lines on the left of the diagrams will be labelled by $C$ and those on the right by $A$.

Let us now note carefully, that in the expressions

$$\frac{\partial (M_{op}^{(n)})_{CD}}{\partial t_{v_i}^{(n)\dagger}} \quad \text{or} \quad \frac{\partial (M_{op}^{(n)})_{CD}}{\partial x_I^{(n)\dagger}}$$

the label $C$ must match label $A$ of the set $i = (\pi_m, A)$ and also of the set $I = (A, B)$ as otherwise $(M_{op}^{(n)})_{CD}$ will not contain the specific $t_{v_i}^{(n)\dagger}$ or $X_I^{(n)\dagger}$ at all, and the derivative will be trivially zero. In other words, the double sum over $C$ and $D$ in Eqs. (32) and (33) really reduces to a single sum over $D$ only. The Lagrange's multipliers $\lambda_{CD}^{A}$ become simply $\lambda_{AD}^{A} = \lambda_{AD}$. We again illustrate this feature more concretely by taking the one valence-hole problem. The expression $\partial (M_{op}^{(1)})_{\delta\mu}/\partial t_{v_i}^{(1)\dagger}$ is non zero only when $\delta = \alpha$ for $i = (p, \gamma, \delta, \alpha)$ and $\partial (M_{op}^{(1)})_{\delta\mu}/\partial x_I^{(1)\dagger}$ is non zero only when $\delta = \alpha$ and $I = (\alpha, \delta)$. Thus the associated Lagrange's multiplier may be denoted by $\lambda_{\alpha\mu}$.

Let us also note that, since $\bar{H}_{AA}$ and $(M_{op}^{(n)})_{CD}$ etc are at most quadratic in $\tilde{s}_i^{(n)}$ amplitudes, the variational equations are linear in $\tilde{s}_i^{(n)}$.

Having found the amplitudes $t_{v_i}^{(n)}$ and $x_I^{(n)}$, we may substitute their values in $\bar{H}_{AB}$ to generate the expression for $E_K^{(n)}$. A variation of $E_K^{(n)}$ with respect to the $C_{AK}^{(n)}$'s then leads to

$$\sum_B \bar{H}_{AB}^{(n)} C_{BK}^{(n)} = E_K^{(n)} C_{AK}^{(n)}, \tag{34}$$

which is the model space eigenvalue equation furnishing the $E_K^{(n)}$'s. Due to the decoupling of the Eqs. (32) and (33) from the $C_{AK}^{(n)}$'s, the matrix-element $\bar{H}_{AB}$ is independent of the states $K$, so that all the $E_K^{(n)}$'s are obtained from a single diagonalization procedure. This also guarantees the orthogonality and noninteracting nature of $\Psi_K^{(n)}$'s (Eqs. (26) and (27)).

We conclude this section by summarizing the various steps of the solution of Eqs. (31) to (33) for an $n$-valence problem: The matrix-elements of $T_c$ are solved first, corresponding to the zero valence problem. With the $t_c$ amplitudes frozen, Eqs. (31), (32) and (33) are solved for the one valence problem next, to get $t_{v_i}^{(1)}$ and $x_I^{(1)}$ amplitudes. These and $t_c$ amplitudes are frozen at the two-valence problem and only the new variables in $\tilde{S}^{(2)}$, i.e. $t_{v_i}^{(2)}$ and $x_I^{(2)}$ are varied next to get these amplitudes. This process should ideally continue hierarchically up to the $n$-valence level. If $n$ is too large, we may stop at an appropriate $m < n$ valence level, and assume that $\tilde{s}_i^{(p)}$'s with $p > m$ are all zero. This truncation may be termed as an "$m$ valence subsystem approximation" [40].

## 3. Calculation of IP and EA: An illustrative example of one-valence problems

In this section we shall work out the actual expressions for $\bar{H}_{AB}$ etc by taking IP and EA as typical one-valence problems. $\Phi_C$ is chosen as the $N$-electron

Hartree-Fock function $\Phi_0$ and the one valence model space functions are the $(N+1)$ electron determinants $a_\alpha \Phi_0$ and $a_p^+ \Phi_0$ for all holes $\alpha$ and particles $p$. IP and EA are obtained by dropping the closed diagrams of $\bar{H}$, i.e. working with $\bar{H}_{op}^{(1)}$ only.

The zero-valence ground state $T_c$-values are calculated as in the variational closed-shell coupled cluster theory [19] and kept frozen at the appropriately calculated values. Only the two-body $T_c$ operator is retained for our purpose. We confine the rank of $T_v^{(1)}$ operators to two-body only. Specifically, for IP calculation, $T_v^{(1)}$ is given by

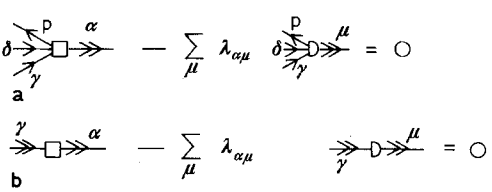
$$T_v^{(1)} = \frac{1}{2!^2} \sum_{\alpha p \gamma \delta} \langle \alpha p | t_v^{(1)} | \gamma \delta \rangle_a N[a_\alpha^\dagger a_p^\dagger a_\delta a_\nu]. \tag{35}$$

When $\gamma = \delta$, *this is a pseudo one-body excitation with only one change in occupancy* viz. $\delta \to p$. Thus effectively this approximation involves one and two body cluster terms. This choice, Eq. (35) also ensures that we have taken all the distinct amplitudes that we can have to generate all the single shakeup states of the manifold free of unnecessary constraints [9, 40]. $\alpha$ in Eq. (35) is, by construction, a valence hole. We elaborate only on the IP's, because the EA diagrams may be generated from the IP diagrams by hole-particle reversal. $T_v^{(1)\dagger}$and $X^{(1)\dagger}$ operators in this case have already been depicted in Figs. 5(b) and 5(d) respectively. For generating the variational equations, the labels $A$, $B$ may be given the hole indices $\alpha, \beta, \ldots$. etc. and $\pi_m$, indices $(p, \gamma, \delta)$ for IP. The variational equations may then be diagrammatically represented as in Fig. 7.

To arrive at the final working equations, we have used reduced Hugenholtz matrix elements of $T_v^{(1)}$ and spin adapted the expressions in a manner analogous to our closed shell work [19]. The graphical methods of spin algebra [22] are used for this purpose. The reduced Hugenholtz matrix elements of $T_v$ are given by

$$\{\alpha p | t_v^{(1)} | \gamma \delta\}_\sigma = \langle \alpha p | t_v^{(1)} | \gamma \delta \rangle + (-1)^\sigma \langle \alpha p | t_v^{(1)} | \delta \gamma \rangle \quad \text{with} \quad \sigma = 0, 1. \tag{36}$$

$\langle \alpha p | t_v^{(1)} | \gamma \delta \rangle$ are Goldstone matrix elements, and $\alpha$, $p$, $\gamma$, $\delta$ etc. are now orbitals. The functions $\Phi_A^{(n)}$ etc for the spin-adapted theory should also be taken as spin-adapted combination of determinants. In the Appendix we have collected the expressions for $\bar{H}_{\alpha\beta}$. From these expressions, the variational equations for IP may be set up in a straightforward manner. Hole-particle reversal generates the corresponding EA equations. For excitation energy (EE) calculations, which is a two-valence problem, the model space may be taken as hole-particle excited determinants $a_p^\dagger a_\alpha \Phi_0$. However, the one-valence level has to be solved first. EE




**Figs. 7** The diagrammatic representation of the variational Eqs. (32) and (33) for the IP problem, corresponding to the one valence hole situation. 7(a) corresponds to Eq. (32) and 7(b) corresponds to Eq. (33). The boxes □ and D indicate connected diagrams from $\partial \bar{H}_{AB}^{(n)} / \partial s_i^{(n)\dagger}$ and $\partial (M_{op}^{(n)})_{AB} / \partial s_i^{(n)\dagger}$ respectively

calculations would thus automatically generate as a by-product the IP and EA values as well.

## 4. Critical discussions

First, we shall try to indicate a relation between the formalism developed above with the relevant open-shell perturbation theories [23–33]. We shall illustrate our points by taking as an example the one-valence problem.

For a perturbative analysis of the resulting system of equations, we have to rewrite the linear equations in $S^{(1)}$ amplitudes depicted in Fig. 7 in a form suitable for iterative solution. If we write the equations in the matrix form

$$(\boldsymbol{A}-\boldsymbol{\lambda})\boldsymbol{t}_v+\boldsymbol{B}\boldsymbol{X}+\boldsymbol{Y}=0, \tag{37a}$$

$$\boldsymbol{C}\boldsymbol{t}_v+(\boldsymbol{D}-\boldsymbol{\lambda})\boldsymbol{X}+\boldsymbol{Z}=0 \tag{37b}$$

corresponding to Figs. 7(a) and 7(b) respectively, then a perturbative series for $t_v$ (and $x$) may be obtained by an iterative procedure where one rewrites Eqns. (37) as

$$(\boldsymbol{A}^d-\boldsymbol{\lambda}^d)\boldsymbol{t}_v^{j+1}=-\boldsymbol{Y}-\boldsymbol{B}\boldsymbol{x}^j-(\boldsymbol{A}^{od}-\boldsymbol{\lambda}^{od})\boldsymbol{t}_v^j \tag{38a}$$

$$(\boldsymbol{D}^d-\boldsymbol{\lambda}^d)\boldsymbol{X}^{j+1}=-\boldsymbol{Z}-\boldsymbol{C}\boldsymbol{t}_v^j-(\boldsymbol{D}^{od}-\boldsymbol{\lambda}^{od})\boldsymbol{X}^j, \tag{38b}$$

where $\boldsymbol{C}^d$ and $\boldsymbol{D}^d$ contains only the one body matrix-elements of $H$ (the Hartree–Fock matrix elements $\varepsilon_p\delta_{pq}$ or $\varepsilon_\alpha\delta_{\alpha\beta}$ etc.) and $\boldsymbol{C}^{od}$ and $\boldsymbol{D}^{od}$ are the remaining components of $\boldsymbol{C}$ and $\boldsymbol{D}$. $\boldsymbol{\lambda}$ is the Lagrange's multiplier matrix, whose diagonal and nondiagonal components are also likewise partitioned. $\boldsymbol{t}_v$ and $\boldsymbol{X}$ are the matrix-elements of $t_{v_i}^{(1)}$ and $x_i^{(1)}$ stored in a column. The superscripts $(j+1)$ etc indicate the matrix-elements at the $(j+1)$ step of iteration. We also have the norm preservation conditions

$$(M_{op}^{(1)})_{\alpha\beta}=0. \tag{39}$$

From Eq. (39), it follows that to the lowest order, we have $t_v=x=0$. From Eq. (39b) we find

$$(\varepsilon_\alpha-\lambda_{\alpha\alpha})(1-x_{\alpha\alpha}^0)=(\varepsilon_a-\lambda_{\alpha\alpha})=0, \tag{40}$$

where the diagram coming from $\boldsymbol{D}^d$ and $\boldsymbol{\lambda}^d$ are merely orbital energies and $\lambda_{\alpha\alpha}$ respectively.

Substituting the value of $\lambda_{\alpha\alpha}$ from Eq. (40) to Eq. (38a) we find that at the lowest order

$$\{\alpha p|t_v^{(1)}|\gamma\delta\}_\sigma=\frac{\{\alpha p|v|\gamma\delta\}_\sigma}{\varepsilon_\gamma+\varepsilon_\delta-\varepsilon_p-\varepsilon_\alpha}. \tag{41}$$

The pertinent point of this analysis is that in the expression for $\{\alpha p|t_v^{(1)}|\gamma\delta\}_\sigma$, the orbital energy of the valence label $\alpha$ appears in the denominator, and this structure is preserved at the higher order iterations. Thus in the expression of $\bar{H}_{\beta\alpha}$, a typical

term $[VT_v^{(1)}]_{\text{conn}}$ will be of the form

$$\sum_{\gamma\delta p\sigma} \frac{(2\sigma+1)\{\gamma\delta|\nu|\beta p\}_\sigma\{\alpha p|\nu|\gamma\delta\}_\sigma}{(\varepsilon_\gamma+\varepsilon_\delta-\varepsilon_p-\varepsilon_\alpha)}$$

*without any reference to the orbital energy $\varepsilon_\beta$ in the denominator.* Thus the expression is unsymmetrical in the labels $\alpha$ and $\beta$. However there will be a counterterm $[T_v^{(1)\dagger}V]_{\text{conn}}$ which will be hermitian adjoint of the term considered above, and the two together will give an expression which is manifestly symmetric in the labels $\alpha$ and $\beta$. The lack of symmetry between $\alpha$ and $\beta$ in *individual* term of $\bar{H}_{\beta\alpha}$ and retention of symmetry in the *complete* expression of the $\bar{H}_{\beta\alpha}$ is reminiscent of the hermitization of the nonhermitian formulation of the open shell perturbation theory of Brandow [29], where the effective Hamiltonian is explicitly hermitized after the wave-function is generated by adding counterterms. Brandow calls this hermitized procedure a theta-expansion [29].

An interesting sidelight to the method developed here is the following: If we adopt an alternative approach to the variation problem by assuming $X$ as a hermitian operator, then we can equate the $X_I^+$ and $X_I$ amplitudes. During the differentiation in Eq. (33) both $X_I^+$ and $X_I$ amplitudes are then differentiated. In that case the variational equations that are generated are different from those considered until now, implying consequent changes in the actual values of $\bar{H}_{AB}$ matrix elements as well. In this alternative procedure, we may eliminate the Lagrange's multipliers in favor of $T_v/T_v^+$ only, and differentiate the $\bar{H}_{AB}^{(n)}$'s with respect to $t_{v_i}^{(n)\dagger}$ amplitudes. As an example, for the IP problem, we have, up to quadratic terms:

$$\begin{aligned}(M_{op}^{(1)})_{\alpha\beta} &= x_{\alpha\beta}^{(1)\dagger}+x_{\alpha\beta}^{(1)}+[\overbrace{T_v^{(1)\dagger}T_v^{(1)}}]_{\alpha\beta}+[\overbrace{T_c^\dagger T_c}]_{\alpha\beta}\\ &= 2x_{\alpha\beta}^{(1)}+[\overbrace{T_v^{(1)\dagger}T_v^{(1)}}]_{\alpha\beta}+[\overbrace{T_c^\dagger T_c}]_{\alpha\beta}\end{aligned} \tag{42}$$

where $[\overbrace{T_v^{(1)\dagger}T_v^{(1)}}]$ etc. are the contractions between $T_v^\dagger$ and $T_v$ operators having open lines labelled $(\alpha, \beta)$. From Eq. (31) we have

$$x_{\alpha\beta}^{(1)} = -\tfrac{1}{2}\{[\overbrace{T_v^{(1)\dagger}T_v^{(1)}}]_{\alpha\beta}+[\overbrace{T_c^\dagger T_c}]_{\alpha\beta}\} \tag{43}$$

to the pertinent order of approximations. Substituting this expression of $x_{\alpha\beta}^{(1)}$ into $\bar{H}_{\alpha\beta}$, and truncating the expression of $\bar{H}_{\alpha\beta}$ up to a total of quadratic power $T_v^{(1)}$ we obtain

$$\begin{aligned}\bar{H}_{\alpha\beta} &= [H_c]_{\alpha\beta}+[\overbrace{T_v^{(1)\dagger}H_c}]_{\alpha\beta}+[\overbrace{H_c T_v^{(1)}}]_{\alpha\beta}\\ &\quad+[\overbrace{T_v^{(1)\dagger}H_c T_v^{(1)}}]-\tfrac{1}{2}[\overbrace{T_v^{(1)\dagger}T_v^{(1)}}H_c]_{\alpha\beta}\\ &\quad-\tfrac{1}{2}[\overbrace{H_c T_v^{(1)\dagger}}T_v^{(1)}]_{\alpha\beta}-\tfrac{1}{2}[\overbrace{T_c^\dagger T_c}H_c]_{\alpha\beta}\\ &\quad-\tfrac{1}{2}[\overbrace{H_c T_c^\dagger}T_c]_{\alpha\beta}.\end{aligned} \tag{44}$$

where

$$H_c = [\mathrm{e}^{T_c^\dagger}H\,\mathrm{e}^{T_c}]_{\text{conn}}. \tag{45}$$

The variational equation with respect to $t_{v_i}^{(1)\dagger}$ amplitudes would generate a set of simultaneous linear equations in $t_v$ amplitudes. To the first order of $V$, this will give

$$\{\alpha p|t_v^{(1)}|\gamma\delta\}_\sigma = \frac{\{\alpha p|t_v^{(1)}|\gamma\delta\}_\sigma}{\varepsilon_\gamma + \varepsilon_\delta - \varepsilon_p - \frac{1}{2}(\varepsilon_\alpha + \varepsilon_\beta)}. \tag{46}$$

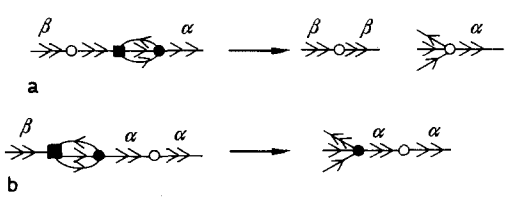
This expression contains the labels $(\alpha, \beta)$ *symmetrically* in the denominator and is reminiscent of the intrinsically hermitian open-shell perturbation theory of Banerjee et al. [28] following an earlier suggestion by Johnson and Baranger [27]. Let us note, however, that this prescription leads to a rather awkward situation in which differentiation with respect to a $t_{v_i}^{(1)\dagger}$ amplitude leads to a disconnected diagram as depicted in Fig. 8(a) coming from $[\overbrace{H_c \overbrace{T_v^{+(1)} T_v^{(1)}}}]_{\alpha\beta}$ and the other connected as in Fig. 8(b) coming from $[\overbrace{T_v^{+(1)} \overbrace{T_v^{(1)} H_c}}]_{\alpha\beta}$. The first diagram, the disconnected type, indicates that due to the presence of label $\beta$, we need an "extra" label on the $t_v$ amplitude to show that it depends explicitly on $\beta$. This is the price we pay for the intrinsically hermitian theory we generate thereby.

We now make a brief comparative study with the pertinent open-shell cluster expansion theories which also invoke variational principle.

One such theory is due to Paldus et al. [17] for difference energy calculations, where the open shell wavefunction $\Psi_K^{(n)}$ is given by

$$\Psi_K^{(n)} = W_K^{(n)} \exp(T_c)\Phi_0. \tag{47}$$

$e^{T_c}\Phi_0$ is the exact ground state as in our work and $W_K^{(n)}$ creates the ionized/excited state $\Psi_K^{(n)}$ by using a linear combination of $\Omega_A^{(n)\dagger}$ of Eq. (9). Direct difference energy, in this work [17] is not a single size-consistent term but instead is a ratio of two connected terms. This is the crucial difference of Paldus formalism from the work presented here. Besides this, the lack of a cluster property of the operator $W_K^{(n)}$ makes the form of direct difference energy progressively unsuitable for higher valence situations.




**Figs. 8** The diagrams generated from the variational expressions of an intrinsically hermitian formulation. Fig. 8(a) is obtained by differentiating a diagrammatic expression of the form $[\overbrace{H_c \overbrace{T_v^{(1)\dagger} T_v^{(1)}}}]_{\alpha\beta}$ with respect to the $t_v^\dagger$ amplitude. This really appears when Eq. (43) is used to replace the $x_i$ amplitudes in $[\overbrace{H_c X^{(1)}}]_{\alpha\beta}$ by $\overbrace{T^\dagger T}$ terms. The variational expression of Fig. 8(a) is clearly disconnected. Fig. 8(b), on the other hand, comes from the variation of $[\overbrace{X^{(1)\dagger} H_c}]_{\alpha\beta}$ with the $t_v^\dagger$ amplitude where again Eq. (43) is used. This expression is connected

The theory due to Nakatsuji [18] uses symmetry-adapted cluster expansion using CI type of reference function (SAC-CI). This ansatz has been applied both variationally and nonvariationally. In the variational type, the functional used by Nakatsuji retains both the numerator and the denominator and hence lacks size-consistency. Nakatsuji uses symmetry-adapted linked cluster operators while an appropriate projector is used to project the appropriate symmetry components from the unlinked terms. Our work differs from the work in [18] additionally in the treatment of spin-variables.

Reitz and Kutzelnigg [13] used a common unitary ansatz for the cluster operator both for a parent closed shell and the appropriate open-shell states and calculated the difference energies directly via a variational principle. The unitary ansatz ensures that the denominator is identically equal to unity. The choice of the cluster operator $W$ is according to the suggestion of Primas [38]:

$$W = \exp(\sigma) \tag{48}$$

with

$$\sigma^{\dagger} = -\sigma \tag{49a}$$

and

$$P\sigma P = 0. \tag{49b}$$

Eqs. (48) and (49) ensure that the expectation value of energy can be expressed as a Hausdorff expansion, and as Primas has shown earlier [38] this guarantees the size-consistency of energy. In later publications, Kutzelnigg [36] analysed the size-consistency problem for such a cluster ansatz in great detail and pointed out the need to consider all the $n$-valence sectors of the Hilbert space to guarantee size-consistency. Our work shares with Kutzelnigg's work [36] this aspect of the problem. It, however, differs essentially from [36] in choosing the form for the wave-operator. Essentially to guarantee hermiticity of the effective Hamiltonian, we need the property of norm preservation only within the model space. This also serves to cancel the denominator. This is the constraint we have adopted. The unitarity on the other hand automatically demands the norm preservation outside the model space as well. This forces the unitary operator to have the exponential form of Eq. (48) which is not in normal order. This form thus has a less compact structure, bringing in contractions among the various operators in the multicommutator expansion. Our choice of the nonunitary cluster ansatz with auxiliary operators forcing the norm preservation within the model space only is less restrictive and it is also possible to retain the more compact normal ordered structure for the ansatz as well. The proof of size-consistency is nontrivial, however, due to lack of a multicommutator expansion. This thus had to be proved, as was done in section 2. We should mention here that Kvasnicka [15], in an analysis of the open shell perturbation theory noted that a brute-force hermitization of the effective Hamiltonian introduces errors only in fourth order, and these can be rectified by introducing suitable hermitian operators which he could

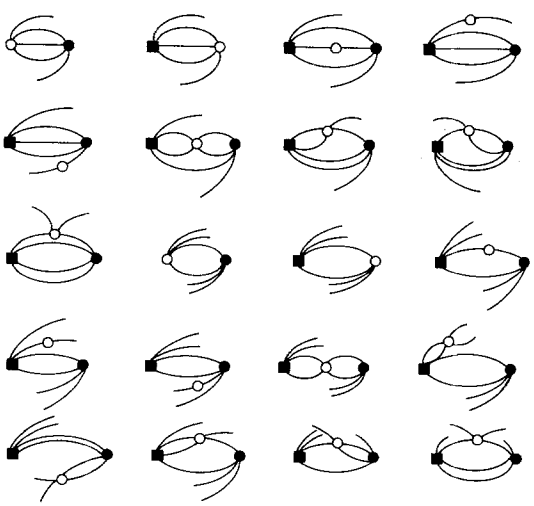
determine self-consistently. The ansatz chosen by us is similar in spirit to Kvasnicka's work [15], but we neither need any explicit hermitization of the effective Hamiltonian nor do we require Bloch Eq. [34] or recursive self-consistent solution for our development. Let us also mention that Soliverez [15] also used analogous ideas in his development of perturbation theory.

*Acknowledgments.* The authors wish to thank the referee for making constructive comments and giving useful suggestions which have led to considerable improvement in the presentation of the material covered in the paper. D.M. acknowledges financial support from INSA and DST (New Delhi).

## Appendix

The expression for $\bar{H}_{\alpha\beta}$ for one-valence IP problems is given below. Greek letters correspond to hole orbitals. The expression is written in terms of spin reduced Hugenholtz matrix elements given by Eq. (36)

$$\begin{aligned}
\bar{H}_{\alpha\beta} = {} & \bar{f}_{\alpha\beta} + \sum_{\gamma} (\bar{f}_{\alpha\gamma} x^{(1)}_{\gamma\beta} + \bar{f}_{\gamma\beta} x^{(1)}_{\alpha\gamma}) + \sum_{\gamma\delta} x^{(1)}_{\gamma\beta} \bar{f}_{\delta\gamma} x^{(1)}_{\alpha\delta} \\
& - \tfrac{1}{2} \sum_{\sigma pq\gamma\delta} \{\beta q | t^{(1)}_v | \gamma\delta\}_\sigma \{\alpha p | t^{(1)}_v | \gamma\delta\}_\sigma \bar{f}_{qp} (2\sigma+1) \\
& + \sum_{\sigma p\gamma\delta\theta} \{\beta p | t^{(1)}_v | \gamma\theta\}_\sigma \{\alpha p | t^{(1)}_v | \delta\theta\}_\sigma \bar{f}_{\delta\gamma} (2\sigma+1) \\
& + \tfrac{1}{2} \sum_{\sigma p\gamma\delta\theta} \{\beta p | t^{(1)}_v | \gamma\delta\}_\sigma \{\theta p | \bar{v} | \gamma\delta\}_\sigma (x^{(1)}_{\alpha\theta} - \delta_{\alpha\theta}) (2\sigma+1) \\
& + \tfrac{1}{2} \sum_{\sigma p\gamma\delta\theta} \{\alpha p | t^{(1)}_v | \gamma\delta\}_\sigma \{\gamma\delta | \bar{v} | \theta p\}_\sigma (x^{(1)}_{\beta\theta} - \delta_{\beta\theta}) (2\sigma+1) \\
& - \tfrac{1}{4} \sum_{\sigma p\gamma\delta\theta\eta} \{\beta p | t^{(1)}_v | \gamma\delta\}_\sigma \{\theta\eta | \bar{v} | \gamma\delta\}_\sigma \{\alpha p | t^{(1)}_v | \theta\eta\}_\sigma (2\sigma+1) \\
& - \sum_{\substack{\sigma_1\sigma_2\sigma_3 \\ pq\gamma\delta\theta}} \{\beta q | t^{(1)}_v | \theta\gamma\}_{\sigma_1} \{\delta q | \bar{v} | p\gamma\}_{\sigma_2} \{\alpha p | t^{(1)}_v | \theta\delta\}_{\sigma_3} \cdot \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & \sigma_1 \\ \sigma_2 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \sigma_3 & \frac{1}{2} \end{Bmatrix} \\
& \times (2\sigma_1+1)(2\sigma_2+1)(2\sigma_3+1)
\end{aligned}$$



**Figs. 9** Some representative skeleton diagrams contributing to $\bar{f}$ and $\bar{v}$ appearing in Eq. (A-1.). The actual diagrams may be obtained from the skeletons by assigning suitable arrow directions to the lines

where $\sigma_1$, $\sigma_2$, $\sigma_3$ are coupled spins and

$$\begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & \sigma_1 \\ \sigma_2 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \sigma_3 & \frac{1}{2} \end{Bmatrix} \tag{A-1}$$

is a 9-$j$ symbol. $\bar{f}$ and $\bar{v}$ are one and two-body operators of the composite $N[e^{T_c^\dagger} H e^{T_c}]_{\text{conn}}$. Some representative diagrams of $\bar{f}$ and $\bar{v}$ are represented skeletally in Fig. 9.

## References


1. Sinanoğlu, O. J.: Chem. Phys. **36**, 706 (1962); Sinanoğlu, O.: Adv. Chem. Phys. 315 (1964); Sinanoğlu, O.: in Three approach to electron correlation in atoms. New Haven and London; Yale University Press 1970
2. Cizek, J. J.: Chem. Phys. **45**, 4256 (1966); Adv. Chem. Phys. **14**, 35 (1969)
3. Cizek, J., Paldus, J., Shavitt, I.: Phys. Rev. **A5**, 50 (1972); Pople, J. A., Krishnan, R., Schlegel, H. B., Binkley, J. S.: Int. J. Quantum Chem. **14**, 545 (1978); Bartlett R. J., Purvis, G. D.: Int. J. Quantum Chem. **13**, 561 (1976)
4. Kutzelnigg, W.: in Methods of electronic structure theory, p. 129, H. F. Schaefer III, ed., New York: Plenum 1977
5. Primas, H.: in Modern quantum chemistry, Vol. 2, p. 45, Sinanoğlu, O., ed. New York: Academic Press (1965)
6. Pople, J. A., Binkley, J. S., Seeger, R.: Int. J. Quantum Chem. **S10**, 1 (1976)
7. Silverstone, H. J., Sinanoğlu, O.: J. Chem. Phys. **44**, 1899; 3608 (1966); J. Chem. Phys. **46**, 854 (1967)
8. Mukherjee, D., Moitra, R. K., Mukhopadhyay, A.: Pramana, **4**, 247 (1975); Mol. Phys. **30**, 1861 (1975)
9. Mukherjee, D., Moitra, R. K., Mukhopadhyay, A.: Mol. Phys. **33**, 955 (1977); Ind. J. Pure Appl. Phys. **15**, 613 (1977); Mukhopadhyay, A., Moitra, R. K., Mukherjee, D.: J. Phys. **B12**, 1 (1979)
10. Offermann, R., Ey, W., Kummel, H.: Nucl. Phys. **A273**, 349 (1976); Offermann R. Nucl. Phys. **A273**, 368 (1976)
11. Lindgren, I.: Int. J. Quantum Chem. **S12**, 33 (1978)
12. Jeziorski, B., Monkhorst, H.: Phys. Rev. **A24**, 1668 (1981)
13. Reitz, H., Kutzelnigg, W.: Chem. Phys. Lett. **66**, 111 (1979)
14. Ey, W.: Nucl. Phys. **A296**, 189 (1978)
15. Soliverez, E.: Phys. Rev. **A24**, 4 (1981); Kvasnicka, V.: Chem. Phys. Lett. **79**, 89 (1981)
16. Mukherjee, D., Mukherjee, P. K.: Chem. Phys. **39**, 325 (1979); Adnan, S. S. Z., Bhattacharyya, S., Mukherjee, D.: Mol. Phys. **39**, 519 (1980); Ghosh, S., Mukherjee, D., Bhattacharyya, S.: Mol. Phys. 43, 173 (1981); Ghosh, S., Mukherjee, D., Bhattacharyya, S.: Chem. Phys. **72**, 161 (1982)
17. Paldus, J., Cizek, J., Saute, M., Laforgue, A.: Phys. Rev. **A17**, 805 (1978)
18. Nakatsuji, H.: Chem. Phys. Lett. **59**, 362 (1978)
19. Pal, S., Durga Prasad, M., Mukherjee, D.: Theoret. Chim. Acta (Berl.) **62**, 523 (1983)
20. Mukherjee, D.: Pramana, **12**, 203 (1979)
21. Pal, S., Durga Prasad, M., Mukherjee, D.: Pramana, **18**, 261 (1982)
22. Elbaz, E., Castel, B.: Graphical methods of spin algebras. New York: Marcel Dekker 1972
23. Bloch, C., Horowitz, J.: Nucl. Phys. **8**, 91 (1958)
24. Morita, T.: Prog. Theoret Phys. **29**, 351 (1963)
25. Oberlechner, G., Owono N-Guema, F., Richert, J.: Nuovo Cimento **B68**, 23 (1970)
26. Kuo, T. T. S., Lee, S. Y., Ratcliff, K. F.: Nucl. Phys. **A126**, 65 (1971)
27. Johnson, M. B., Baranger, M.: Ann. Phys. (N.Y.) **62**, 172 (1971)
28. Banerjee, A., Mukherjee, D., Simons, J.: J. Chem. Phys. **76**, 1972; 1979; 1995 (1982)
29. Brandow, B. H.: Rev. Mod. Phys. **39**, 771 (1967)
30. Sanders, P. G. H.: Adv. Chem. Phys. **14**, 365 (1969)
31. Lindgren, I.: J. Phys. **B7**, 2441 (1974)

32. Kvasnicka, V.: Phys. Rev. **A12**, 1159 (1975)
33. Hose, G., Kaldor, U.: J. Phys. **B12**, 3827 (1979)
34. Bloch, C.: Nucl. Phys. **7**, 451 (1958)
35. Haque, A.: Ph.D. Thesis, Calcutta University 1984.
36. Kutzelnigg, W.: J. Chem. Phys. **77**, 3081 (1982); Kutzelnigg, W., Koch, S.: J. Chem. Phys. **79**, 4315 (1983); Kutzelnigg, W.: J. Chem. Phys. **80**, 822 (1984)
37. Hendekovic, J.: Chem. Phys. Lett. **90**, 198 (1982)
38. Primas, H.: Helv. Phys. Acta. **34**, 331 (1961)
39. Brandow, B.: Int. J. Quantum Chem. **15**, 207 (1979)
40. Haque, M. A., Mukherjee, D.: J. Chem. Phys. **80**, 5058 (1984)

Received September 20, 1983/June 18, 1984