String-based methods for state-selective multi-reference coupled-cluster

Mihály Kállay

Department of Physical Chemistry and Materials Science

Budapest University of Technology and Economics

Outline

- **1** Implementation of single-reference CC theories
- 2 State-specific multi-reference CC theories
- **3 Implementation of the SR-MRCC ansatz**
- 4 Implementation of Mukherjee's SS-MRCC ansatz
- **5** Assessment of state-specific MRCC methods

6 Conclusions

1

Implementation of single-reference CC theories

Coupled-cluster theory

n

Coupled-cluster wave function:

$$|\Psi\rangle = e^{\hat{T}}|0\rangle$$
 $\hat{T} = \sum_{k=1}^{n} \hat{T}_{k}$

$$\hat{T}_k = \sum_{\substack{a_1 < a_2 \dots < a_k \\ i_1 < i_2 \dots < i_k}} t_{i_1 i_2 \dots i_k}^{a_1 a_2 \dots a_k} a_1^+ i_1^- a_2^+ i_2^- \dots a_k^+ i_k^-$$

 $i, j, k \dots (a, b, c \dots)$: occupied (virtual) orbitals $p, q, r \dots$: generic indices $t_{i_1 i_2 \dots i_k}^{a_1 a_2 \dots a_k}$: cluster amplitudes

Coupled-cluster equations:

$$\langle 0|\mathrm{e}^{-\hat{T}}\hat{H}\mathrm{e}^{\hat{T}}|0\rangle = E$$

$$\langle \Psi_{i_1 i_2 \dots i_k}^{a_1 a_2 \dots a_k} | e^{-\hat{T}} \hat{H} e^{\hat{T}} | 0 \rangle = 0 \quad (k = 1, \dots n)$$

$$e^{-\hat{T}}\hat{H}e^{\hat{T}} = 1 + \left[\hat{H},\hat{T}\right] + \frac{1}{2!}\left[\left[\hat{H},\hat{T}\right],\hat{T}\right] + \dots = \\ \{\hat{H}\left(1 + \hat{T}_{1} + \hat{T}_{2} + \hat{T}_{3} + \dots + \frac{1}{2!}\hat{T}_{1}^{2} + \hat{T}_{1}\hat{T}_{2} + \frac{1}{3!}\hat{T}_{1}^{3} + \dots\right)\}_{C}$$

Consequence: only connected terms in operator $e^{-\hat{T}}\hat{H}e^{\hat{T}}$ Two-body Hamiltonian \Rightarrow at most the fourth power of \hat{T}

Solution of CC equations

Our goal is to solve the CC equations up to infinite order minimizing the costs of the calculation.

- 1. Derivation of CC equations in terms of antisymmetrized (Brandow-type) diagrams
- 2. Factorization of these diagrams and the definition of intermediates
- 3. The numerical calculation itself which corresponds to successive evaluation of contractions of cluster amplitudes with intermediates.

Automated diagram generation

Antisymmetrized diagrams are represented by 13-membered integer sequences:

 $\mu_{1,1} \ \mu_{1,2} \ \mu_{1,3} \ \ \mu_{2,1} \ \mu_{2,2} \ \mu_{2,3} \ \ \mu_{3,1} \ \mu_{3,2} \ \mu_{3,3} \ \ \mu_{4,1} \ \mu_{4,2} \ \mu_{4,3} \ \ \mu_{5}$

a triplet of numbers $(\{\mu_{i,j}\}_{j=1}^3)$ for each amplitude vertex (or zeros if there are less than four \hat{T} vertices) and the serial number (μ_5) of the integral list.

 $\mu_{i,1}$: excitation level of the amplitude vertex

- $\mu_{i,2}$: number of internal lines attached to the \hat{T} vertex
- $\mu_{i,3}$: number of internal particle lines

Automated diagram generation: Hamiltonian



Automated diagram generation: examples



Intermediates can also be symbolized by such sequences of integers.

In general: any expression can be represented by such sequences (a multiplet of integers for each factor) provided that there are only excitation operators or only one of them is a de-excitation operator.

Diagram factorization

What does factorization mean?

- diagrams in CC equations are usually non-linear in cluster amplitudes
- however, contractions of Fock-matrix elements or twoelectron integrals with cluster amplitudes can be carried out by linear operations
- \bullet we can sum over internal lines connected to a \hat{T} vertex independently
- a non-linear term can be calculated as successive contractions between appropriately defined intermediates (partial sums) and cluster amplitudes.

 $\left[\sim n_b^8\right]$

Example:

$$W_{ij}^{ab} = -\hat{P}(i|j) \sum_{c < d,k,l} \langle kl || cd \rangle \ t_{ki}^{ab} \ t_{lj}^{cd}$$

It can be factorized in two different ways:

$$\begin{split} W_{l,i}^{cd,ab} &= \sum_{k} \langle kl || cd \rangle \ t_{ki}^{ab} \qquad \left[\sim n_{b}^{7} \right] \\ W_{ij}^{ab} &= -\hat{P}(i|j) \sum_{c < d,l} W_{l,i}^{cd,ab} \ t_{lj}^{cd} \qquad \left[\sim n_{b}^{7} \right] \end{split}$$

or

$$W_{k,j} = \sum_{c < d,l} \langle kl || cd \rangle \ t_{lj}^{cd} \qquad [\sim n_b^5]$$
$$W_{ij}^{ab} = -\hat{P}(i|j) \sum_k W_{k,j} \ t_{ki}^{ab} \qquad [\sim n_b^5]$$

Identification of intermediates

- further intermediates can be defined by adding those ones together that have identical labels and will be contracted with the same amplitude vertices
- rearrange integer triples within each 13-membered sequence: last contracted amplitude vertex to the left then the nextto-the-last etc.
- the 13-membered integer sequences can be regarded as 13digit integers and can be ordered into ascending order
- consider consecutive diagrams: supposing that the mth integer triple is identical in the sequences, but the (m+1)th one is different, intermediates $\mu_{m+1,1}$ $\mu_{m+1,2}$... $\mu_{4,3}$ μ_5 can be added together

Example: above four diagrams (CCSD, \hat{T}_2 equations)



This scheme with the conventional diagrams:



String-based algorithm

What does term "string" mean?

A string is an ordered set of (spin-)orbital indices (p):

$$\mathcal{P} = p_1 \ p_2 \ p_3 \dots \qquad (p_1 < p_2 < p_3 < \dots)$$

E.g., in the case of four orbitals, the three-membered strings are:

The concept of strings was introduced into quantum chemistry by Handy four decades ago for addressing the determinants in a full CI calculation.

String-based algorithm

Why do we need strings?

- in conventional many-body codes, quantities are stored in terms of orbital indices, loops run over orbitals
- a general procedure is desired which is independent of the number of indices.
- solution: string-based algorithm quantities in terms of strings, loops over strings

Cluster amplitudes and intermediates with strings:

 $\begin{array}{lll} t^{a_1 a_2 \dots a_n}_{i_1 i_2 \dots i_n} & \to & t^{\mathcal{A}}_{\mathcal{I}} \\ W^{c_1 c_2, a_1 a_2 \dots a_n}_{k_1 k_2, i_1 i_2 \dots i_n} & \to & W^{\mathcal{C} \mathcal{A}}_{\mathcal{K} \mathcal{I}} \\ \mathcal{I}, \mathcal{J}, \mathcal{K} \dots (\mathcal{A}, \mathcal{B}, \mathcal{C} \dots) : \text{ strings of occupied (virtual) indices} \end{array}$

Concept of string multiplication

Product of two strings of the same type:

 $\mathcal{R} = sgn_{\mathcal{PQ}} \mathcal{P} \cdot \mathcal{Q}$

 ${\cal R}$ is a string that consists of the elements of ${\cal P}$ and ${\cal Q}$

 $sgn_{\mathcal{PQ}} = 0$ if \mathcal{P} and \mathcal{Q} contain any common index

 $sgn_{\mathcal{PQ}} = (-1)^p$ otherwise

 \boldsymbol{p} is the parity of the permutation which arranges elements into ascending order

All operations with cluster amplitudes and intermediates can be expressed in terms of string multiplications.

Evaluation of contractions

An example:

In terms of spin-orbital indices:

$$V^{a_3 a_4 \dots a_n}_{i_3 i_4 \dots i_n} = \sum_{a_1 < a_2 \ i_1 < i_2} \langle a_1 a_2 || i_1 i_2 \rangle \ t^{a_1 a_2 \dots a_n}_{i_1 i_2 \dots i_n}$$

In terms of strings:

$$V_{\mathcal{I}}^{\mathcal{A}} = \sum_{\mathcal{B}\mathcal{J}} sgn_{\mathcal{B}\mathcal{A}} sgn_{\mathcal{J}\mathcal{I}} W_{\mathcal{J}}^{\mathcal{B}} t_{\mathcal{J}\cdot\mathcal{I}}^{\mathcal{B}\cdot\mathcal{A}}$$

$$\mathcal{A} = a_3 a_4 \dots a_n \qquad \mathcal{I} = i_3 i_4 \dots i_n \qquad \mathcal{B} = a_1 a_2 \qquad \mathcal{J} = i_1 i_2$$

Algorithm for contractions

```
Loop over \mathcal{A}
              Loop over \mathcal{I}
                            Initialize an array F to zero
                            Loop over \mathcal{J} (\mathcal{J} \cdot \mathcal{I} \neq 0)
                                          \mathcal{K} = \mathcal{J} \cdot \mathcal{I}
                                           Loop over \mathcal{B} (\mathcal{B} \cdot \mathcal{A} \neq 0)
                                                         \mathcal{C} = \mathcal{B} \cdot \mathcal{A}
                                                        F_{\mathcal{T}}^{\mathcal{B}} = sgn_{\mathcal{B}\mathcal{A}} \ sgn_{\mathcal{J}\mathcal{I}} \ t_{\mathcal{K}}^{\mathcal{C}}
                                           End of loop over \mathcal{B}
                             End of loop over \mathcal{J}
                            V_{\mathcal{I}}^{\mathcal{A}} = \sum_{\mathcal{B}\mathcal{J}} W_{\mathcal{J}}^{\mathcal{B}} F_{\mathcal{J}}^{\mathcal{B}}
              End of loop over \mathcal{I}
End of loop over \mathcal{A}
```

2

State-specific multi-reference CC theories

Why are multi-reference coupled-cluster models required?

- the majority of the quantum chemical methods only work at the equilibrium
 - E.g., the dissociation curve of the fluorine molecule



• configuration interaction (CI) methods, which can describe multi-reference problems, are not size-extensive, thus not applicable to large molecules

The Jeziorski–Monkhorst ansatz

Jeziorski and Monkhorst, *Phys. Rev. A* **24**, 1668 (1981). Silverstone and Sinanoglu, *J. Chem. Phys.* **44**, 1899 (1966).

The wave function ansatz:

$$\psi = \sum_{\mu} \exp(T^{\mu}) \phi_{\mu} c_{\mu}$$

 ϕ_{μ} - reference function c_{μ} - combining coefficients T^{μ} - cluster operator for ϕ_{μ} : $T^{\mu} = \sum_{l} t^{\mu}_{l\mu} a^{+} \cdots i^{-} \cdots$ where the summation runs through all single and double excitations out of ϕ_{μ} generating a virtual determinant $|\chi_{l}\rangle$

The effective Hamiltonian:

$$\bar{H}_{\mu} = \exp(-T^{\mu})H\exp(T^{\mu})$$

The equations for the c_{μ} coefficients: diagonalize the effective Hamiltonian in the model space

$$\sum_{\nu} \tilde{H}_{\mu\nu} c_{\nu} = E c_{\mu}$$

where $\tilde{H}_{\mu\nu} = \langle \phi_{\mu} | \bar{H}_{\nu} | \phi_{\nu} \rangle$.

The equations for the cluster amplitudes: derived from the Bloch equation, no Schrödinger equation

$$\langle \chi_l | \bar{H}_\mu | \phi_\mu \rangle - \sum_\nu \langle \chi_l | \exp(-T^\mu) \exp(T^\nu) | \phi_\nu \rangle \tilde{H}_{\nu\mu} = 0$$

Problems with the ansatz:

- one has to consider several eigenstates simultaneously whose majority is irrelevant to the problem—it is not state-specific
- high number of cluster amplitudes
- appearance of intruder states—convergence problems

$$t^{\mu}_{l\mu} = \frac{H_{l\mu} + \sum_{m \neq l} H_{lm} t^{\mu}_{m\mu} + \dots + \sum_{\nu} (t^{\nu}_{l\nu} - t^{\mu}_{l\nu} + \dots) \tilde{H}_{\nu\mu}}{\tilde{H}_{\mu\mu} - H_{ll}}$$

where $H_{lm} = \langle \chi_l | H | \chi_m \rangle$

State-specific MRCC theories based on the Jeziorski–Monkhorst ansatz

• Mukherjee's state-specific MRCC theory (SS-MRCC)

Mahapatra, Datta, and Mukherjee, J. Chem. Phys. **110**, 6171 (1999).

- Multi-reference Brillouin-Wigner CC theory (MRBWCC)
 Mášik, Hubač, and Mach, J. Chem. Phys. 108, 6571 (1998).
- Hanrath's MRexpT model

Hanrath, J. Chem. Phys. 123, 084102 (2005).

Mukherjee's state-specific MRCC theory

The equations for the cluster amplitudes: derived from the Schrödinger equation invoking sufficiency conditions

$$\langle \chi_l | \bar{H}_\mu | \phi_\mu \rangle c_\mu + \sum_{\nu} \langle \chi_l | \exp(-T^\mu) \exp(T^\nu) | \phi_\mu \rangle \tilde{H}_{\mu\nu} c_\nu = 0$$

- state-specific
- size-extensive and size-consistent with localized orbitals
- not invariant to transformation among the active orbitals
- no intruder states

$$t^{\mu}_{l\mu} = \frac{H_{l\mu} + \sum_{m \neq l} H_{lm} t^{\mu}_{m\mu} + \dots + \sum_{\nu} (t^{\nu}_{l\mu} + \dots) \tilde{H}_{\mu\nu} c_{\nu} / c_{\mu}}{E - H_{ll}}$$

Multi-reference Brillouin-Wigner CC theory

The equations for the cluster amplitudes: derived from the Schrödinger equation invoking sufficiency conditions

 $\langle \chi_l | H \exp(T^{\mu}) | \phi_{\mu} \rangle = E \langle \chi_l | \exp(T^{\mu}) | \phi_{\mu} \rangle$

- state-specific
- not size-extensive and size-consistent
- no intruder states
- size-extensivity corrections are possible but result in intruder state problems

Hanrath's MRexpT model

Modified wave function ansatz: cluster amplitudes generating $|\chi_l\rangle$ from different reference determinants are constrained to be equal

The equations for the cluster amplitudes: derived from the Schrödinger equation

$$\sum_{\mu} \langle \chi_l | H \exp(T^{\mu}) | \phi_{\mu} \rangle c_{\mu} = E \sum_{\mu} \langle \chi_l | \exp(T^{\mu}) | \phi_{\mu} \rangle c_{\mu}$$

- state-specific
- not size-extensive
- no intruder states

An alternative formulation: the "single-reference" multi-reference CC ansatz

Oliphant and Adamowicz, J. Chem. Phys. 94, 1229 (1991).

Piecuch, Oliphant, and Adamowicz, J. Chem. Phys. 99, 1875 (1993).

MK, Szalay, and Surján, J. Chem. Phys. 117, 980 (2002).

- assuming a complete active space, one of the reference determinants can be chosen as a formal Fermi-vacuum
- all single, double (or higher) excitations out of the other reference functions can be regarded as higher excitations out of this single determinant



- the single-reference formalism can be applied, however, only selected excitations are included
- it retains the simplicity and the size-extensivity of the singlereference CC methods, state-specific, no intruder states
- disturbing feature: the method is not invariant with respect to the choice of the Fermi-vacuum
- however, the error resulting from the lack of invariance is usually smaller than the absolute error of the method
- it can be regarded as a selection scheme for single reference wave functions

3

Implementation of the SR-MRCC ansatz

Automated diagram generation

Multi-reference case: amplitudes are classified according to the number of inactive labels

Two additional integers per \hat{T} vertex \rightarrow 21-membered integer strings

 $\mu_{1,1} \ \mu_{1,2} \ \mu_{1,3} \ \mu_{1,4} \ \mu_{1,5} \quad \dots \quad \mu_{4,1} \ \mu_{4,2} \ \mu_{4,3} \ \mu_{4,4} \ \mu_{4,5} \ \mu_{5}$

 $\mu_{i,1}$: excitation level of the amplitude vertex

- $\mu_{i,2}$: number of internal lines attached to the \hat{T} vertex
- $\mu_{i,3}$: number of internal particle lines
- $\mu_{i,4}$: number of inactive fixed particle lines
- $\mu_{i,5}$: number of inactive fixed hole lines

Automated diagram generation: examples



Inactive labels: thick lines

The projecting amplitude has one particle and one hole inactive index.

With these modifications, the single-reference machinery can be employed.

4

Implementation of Mukherjee's SS-MRCC ansatz

Previous implementations

• Mukherjee and co-workers: the active space is limited to two closed-shell determinants

Mahapatra, Datta, and Mukherjee, J. Chem. Phys. **110**, 6171 (1999).

• Evangelista and co-workers: the active space is limited to two electrons

Evangelista, Allen, and Schaefer, *J. Chem. Phys.* **125**, 154113 (2006); **127**, 024102 (2007).

Pittner and co-workers: general active spaces; single, double, and triple excitations
 Bhaskaran-Nair, Demel, and Pittner, J. Chem. Phys. 132, 154105 (2010).

Our implementation

Features:

- string-based techniques
- arbitrary complete active spaces
- no restriction on the number of electrons, number of active orbitals, spin multiplicity
- arbitrarily high excitations in the cluster operators
- other implemented methods: SS-MRCEPA, SS-MRPT, MRBWCC

The working equations:

$$\langle \chi_l | \bar{H}_\mu | \phi_\mu \rangle c_\mu + \sum_\nu \langle \chi_l | \exp(-T^\mu) \exp(T^\nu) | \phi_\mu \rangle \tilde{H}_{\mu\nu} c_\nu = 0$$

$$\sum_{\nu} \tilde{H}_{\mu\nu} c_{\nu} = E c_{\mu} \quad \text{where} \quad \tilde{H}_{\mu\nu} = \langle \phi_{\mu} | \bar{H}_{\nu} | \phi_{\nu} \rangle$$

Implementation of $\langle \chi_l | \bar{H}_\mu | \phi_\mu \rangle$ and $\tilde{H}_{\mu\nu}$:

- similar to the matrix elements for the SRMRCC approach
- the all-active excitations are dropped from T^{μ}
- the projecting space is limited to excitations corresponding to T^{μ} as well as to the reference determinants

Implementation of $\langle \chi_l | \exp(-T^{\mu}) \exp(T^{\nu}) | \phi_{\mu} \rangle$:

- new term
- the non-vanishing component of $T^{\nu},\,\overline{T^{\nu}}(\mu),$ has exactly the same excitation structure as T^{μ}
- $\overline{T^{\nu}}(\mu)$ and T^{μ} commute \Rightarrow one can rewrite the matrix element as $\langle \chi_l | \exp[-T^{\mu} + \overline{T^{\nu}}(\mu)] | \phi_{\mu} \rangle$
- operator $-T^{\mu} + \overline{T^{\nu}}(\mu)$ is constructed by calculating the addresses of excitation operators in $\overline{T^{\nu}}(\mu)$ wrt ϕ_{μ} using graphical techniques
- the matrix element is calculated by means of the power series expansion of the exponential
- the powers of operator $-T^{\mu}+\overline{T^{\nu}}(\mu)$ are evaluated by a string-based algorithm

Convergence problems

Iteration formula for the cluster amplitudes:

$$t^{\mu}_{l\mu} = \frac{H_{l\mu} + \sum_{m \neq l} H_{lm} t^{\mu}_{m\mu} + \dots + \sum_{\nu} (t^{\nu}_{l\mu} + \dots) \tilde{H}_{\mu\nu} c_{\nu} / c_{\mu}}{E - H_{ll}}$$

- numerical instability when c_{μ} is much smaller than c_{ν}
- this effect is different from the inherent intruder problem of the Jeziorski–Monkhorst ansatz and is less severe
- it can be resolved using various techniques, however, our experience shows that no universal solution exists and each system requires special attention

Possible solutions:

- (i) Ignore problematic reference determinants: drop the cluster amplitudes for the determinants with combining coefficient less than 10^{-6} , results in negligible change in the energy
- (ii) Tikhonov regularization: replace c_{ν}/c_{μ} in the iteration formula by $c_{\nu}c_{\mu}/(c_{\mu}^2 + \omega^2)$, no change in the energy if used only in the first iteration cycles, small change in the energy if used throughout, $\omega = 0.01$ is sufficient in most cases
- (iii) Restarting cluster amplitudes: restart from the converged SS-MRCEPA(0) cluster amplitudes, no change in the energy

5

Assessment of state-specific MRCC methods

The lithium hydride molecule

- single bond dissociation
- active space: 2×2 CAS including the 2σ and 3σ orbitals
- 4 reference determinants
- basis set: cc-pVTZ
- Tikhonov regularization with ω =0.01 for the first ten iteration cycles, dropping cluster amplitudes for determinants with c_{μ} 's less than 10^{-6}



The hydrogen fluoride molecule

- single bond dissociation
- active space: 2×2 CAS including the $3a_1$ and $4a_1$ orbitals
- 4 reference determinants
- basis set: valence double-zeta (Dunning-Hay)
- no regularization, but cluster amplitudes for determinants with c_{μ} 's less than 10^{-6} are dropped



The water molecule

- simultaneous dissociation of two single bonds
- active space: 4×4 CAS including the $3a_1$, $4a_1$, $1b_2$, and $2b_2$ orbitals
- 20 reference determinants
- basis set: cc-pVDZ
- Tikhonov regularization with ω =0.01 for the first ten iteration cycles, dropping cluster amplitudes for determinants with c_{μ} 's less than 10^{-6}





The nitrogen molecule

- triple bond dissociation
- active space: 6×6 CAS including the $3\sigma_g^+$, $1\pi_u$, $1\pi_g$, and $3\sigma_u^+$ orbitals
- 36 reference determinants
- basis set: cc-pVDZ
- Tikhonov regularization with ω =0.01 for all the iteration cycles, dropping cluster amplitudes for determinants with c_{μ} 's less than 10^{-6}



Performance in multi-reference regions

- the performance of SS-MRCC usually falls behind that of extensivity-corrected MRCI variants, SRMRCC, and MRexpT
- fundamental difference between SS-MRCC and the other ansätze: different coupling terms for virtual determinants
- in the SS-MRCC equations the virtual determinants derived from different reference functions are only indirectly coupled
- in the case of the other methods these determinants are coupled directly through the Hamiltonian
- localized orbitals improve the performance

6 Conclusions

- the string-based approach enables the general implementation of single- and multi-reference CC methods
- the full implementation of the SR-MRCC anstatz and Mukherjee's SS-MRCC theory has been presented
- the performance of various state-specific MRCC methods has been assessed
- the iterative solution of the SS-MRCC equations is prone to convergence problems
- the performance of the SS-MRCC method is often not satisfactory in the multi-reference regions of potential energy surfaces especially for multiple bond cleavage

Acknowledgments

Sanghamitra Das, József Csóka, Jürgen Gauss, László Gyevi-Nagy, Bence Hégely, Réka Horváth, Bence Ladóczki, Dávid Mester, Debashis Mukherjee, Péter Nagy, Zoltán Rolik, Gyula Samu







Program availability http://www.mrcc.hu