Advanced Electronic Structure Theory in Quantum Chemistry

Department of Theoretical and Computational Molecular Science
Division of Theoretical Molecular Science I

YANAI, Takeshi   Associate Professor
YAMADA, Mariko   Secretary

Aimed at predictive computational modelings of molecular electronic structures with *ab initio* quantum chemistry calculations, our scientific exploration is to establish a cutting-edge theoretical methodology that allows one to compute accurately and efficiently the complex electronic structures where the substantial multireference character in the wave functions has to be handled for the qualitative and quantitative descriptions. Our resultant works to be reported here are (1) to develop a new type of the multireference correlation model named Canonical Transformation (CT) theory, which can efficiently describe short-range dynamic correlation on top of the multi-configurational staring wave function, and (2) to construct the extensive complete active space self-consistent field (CASSCF) method combined with *ab initio* density matrix renormalization group (DMRG) method for making unprecedentedly larger active spaces available for the CASSCF calculations. These two pivotal developments are tailored to eventually be incorporated for solving large-scale multi-reference electronic structure problems.

1. Canonical Transformation (CT) Theory from Extended Normal Ordering[1]

We have presented a canonical transformation (CT) theory which is based on an exponential ansatz, is rigorously size extensive, and which may easily be combined with any multi-reference starting wave function such as CASSCF or DMRG wave functions. This study has derived a new formulation of the theory based on the extended normal ordering procedure of Mukherjee and Kutzelnigg.

Assuming that a reference wave function Ψ_0 is available that describes the nondynamic correlation in the problem, we incorporate the remaining dynamic correlation on top of the reference wave function Ψ_0 *via* an exponential operator that generates excitations between the active and external spaces, yielding

$$\Psi = e^A \Psi_0 \qquad \text{(eq. 1)}$$
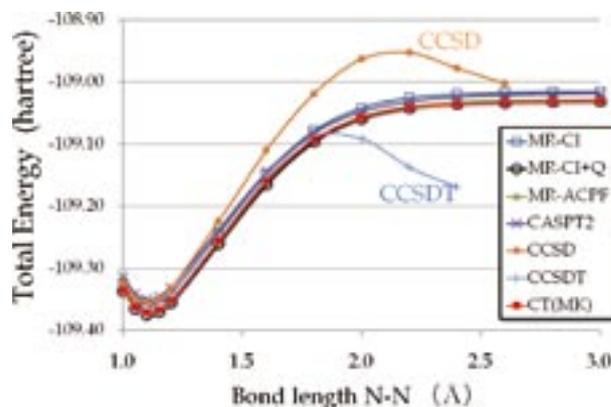
Figure 1. Bond-breaking curve of N_2 molecule with CAS($6e,6o$) and cc-pVTZ Gaussian basis sets.

We will be concerned with a *unitary* formulation, where $A^\dagger = -A$. The excitations are understood to be both of external and semi-internal forms of up to two-particle operators. In a related picture, we can also view e^A as generating an effective *canonically transformed* Hamiltonian H^{CT} that acts only in the active space, but which has dynamic correlation folded in from the external space, where

$$H^{CT} = e^{-A} H e^A \qquad \text{(eq. 2)}$$
$$H^{CT}\Psi = E\Psi \qquad \text{(eq. 3)}$$

A central feature of the canonical transformation theory is the use of an *operator decomposition*, both to close the infinite expansions associated with an exponential ansatz and to reduce the complexity of the energy and amplitude equations that arise when working with a complicated reference function. Starting from the Baker-Campbell-Hausdorff expansion of the exact effective Hamiltonian, we replace each commutator by an approximate *decomposed* commutator to yield an approximate effective Hamiltonian,

$$H_{1,2}^{CT} = H_{1,2} + [H,A]_{1,2} + \tfrac{1}{2}[[H,A]_{1,2},A]_{1,2} + \dots \quad \text{(eq. 4)}$$
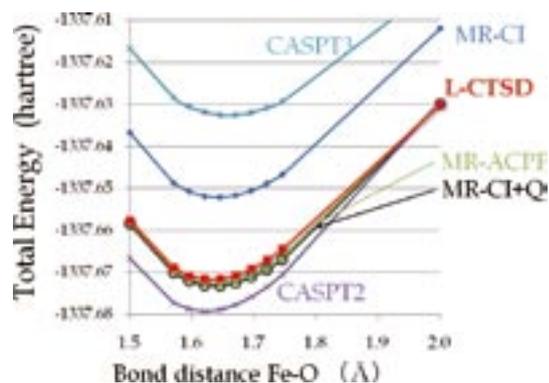
Figure 2. Bond-breaking curve of FeO molecule with CAS(12*e*,12*o*) and ANO-DZP Gaussian basis sets.

Table 1. Timings for different multireference methods for a single point calculation on the FeO curve. The time for the CASSCF calculation is not included. L-CTSD (present work) did not exploit the point-group symmetry of the molecule in the calculation.

	Time (sec)
CASPT2	5,900
CASPT3	17,000
MR-CI+Q	158,000
MR-ACPF	168,000
L-CTSD (present)	4,500

Each subscript denotes a decomposition, and the numbers "1,2" denote the particle ranks of the operators that remain after the decomposition. In this study, we have introduced a new-type operator decomposition, with some formal advantages, that is based on the concept of *extended normal ordering* as introduced by Mukherjee and Kutzelnigg. The study presented and exploited a form of the extended normal-ordered decomposition for three-particle operators. In the form implemented in the work, the computational cost is $O(a^2e^4)$, which is essentially the same as that of the single-reference coupled-cluster single and double (CCSD) model, where *a* is the number of active orbitals and *e* is the number of external orbitals.

A second focus of this work is to investigate in detail the behavior of the canonical transformation theory in a variety of chemical problems. For example, we study, with a range of basis sets, the bond-breaking potential energy curves of water, nitrogen, and iron oxide and compare our results against state-of-the-art multireference configuration interaction and perturbation theories. In addition, we examine numerically the size-extensivity and density-scaling properties of the canonical transformation energies. The results in the present study are much improved, in large part, because of improvements we have made to our numerical algorithms, and we describe in detail the numerical aspects of efficiently implementing and converging the CT equations.

Figure 1 shows the bond-breaking curve of the nitrogen molecule, which is a prototype multireference problem where the complication of breaking a triple bond occurs. For this application, the prevalent single-reference methods CCSD or

CCSDT are known to typically fail to describe the bond breaking. Figure 2 presents the potential curves of FeO computed by various multireference methods. The overall performance of our linearized CT method with single and double substitutes (L-CTSD) for these potential curves was competitive with the best multireference methods such as MR-ACPF. It is found that, for the FeO calculations, the multireference perturbation series CASPT2 and CASPT3 seemed to break down. Timings for the multireference calculations, which are listed in Table 1, reveal that CT method is two to three orders of magnitude faster than the most accurate MR-ACPF method, which though the CT is competitive with in accuracy.

2. Extensive Complete Active Space Self-Consistent Field (CASSCF) with *ab initio* Density Matrix Renormalization Group (DMRG)

To perform the large-scale multireference calculations with CT or other multireference methods, the extensive active-spaced reference wave functions where a large number of active electrons are highly (or fully) correlated within the active orbitals must be found. The CASSCF method provides the most desirable, optimal reference wave functions, in which the static correlations are effectively captured with the relaxed active orbitals that are optimized self-consistently at a high computational cost. We have developed a parallelized CAS optimization method that enables to handle unprecedentedly larger CAS with high-quality basis sets. Figure 3 shows timings for the orbital optimization steps of a single CASSCF iteration for all-trans $C_{12}H_{14}$ chain with a full π active space and cc-pVTZ Gaussian basis sets. The timings were measured on Pentum 4 2.2 GHz PC clusters connected within 100 MB bandwidth network. We have again performed the CASSCF calculation on the $C_{20}H_{22}$ chain with *ab initio* DMRG for an exact diagonalization of CAS(20*e*,20*o*). A single iteration of the extensive CASSCF calculation took just a couple of hours.

CAS(12*e*,12*o*) full π active
- 37 closed orbitals
- 507 external orbitals

	Integral Transformation	Augmented Hessian Optimization	Total
4 CPUs	Out of memory for execution		
8 CPUs	1092 sec	34 sec	1126 sec
12 CPUs	711 sec	20 sec	731 sec
16 CPUs	560 sec	14 sec	574 sec
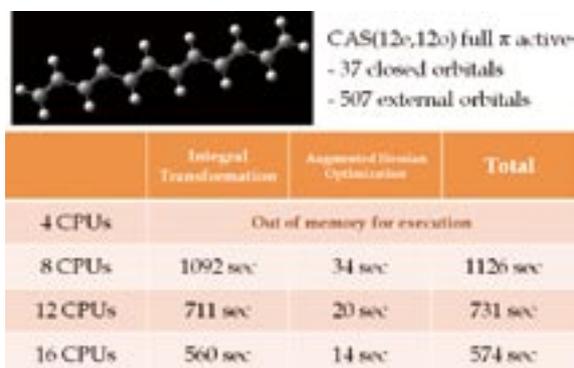
Figure 3. Timings for the parallelized CASSCF method with full π active space CAS(12*e*,12*o*) and cc-pVTZ basis.

References

1) T. Yanai and G. K. -L. Chan, *J. Chem. Phys.* **127**, 104107 (2007).