Theor Chem Acc (2011) 130:411–417
DOI 10.1007/s00214-011-1008-7

REGULAR ARTICLE

Linear-scaling divide-and-conquer second-order Møller–Plesset perturbation calculation for open-shell systems: implementation and application

Takeshi Yoshikawa · Masato Kobayashi · Hiromi Nakai

Received: 4 April 2011 / Accepted: 16 July 2011 / Published online: 3 August 2011

Abstract We have developed the spin-unrestricted divide-and-conquer (DC)-based linear-scaling self-consistent field method for treating open-shell systems (Kobayashi et al. in Chem Phys Lett 500:172, 2010). Because the method does not require the position of excess spins or charges, it made the treatment of large spin-delocalized systems tractable. The present study extends the DC-based unrestricted open-shell scheme to the correlated second-order Møller–Plesset perturbation (MP2) theory. Numerical applications to polyene cations demonstrate that the present method gives highly accurate results with less computational costs even for spin-delocalized systems.

Keywords Linear-scaling computation · Open-shell system · Unrestricted orbital · Electron correlation · MP2 theory

Dedicated to Professor Shigeru Nagase on the occasion of his 65th birthday and published as part of the Nagase Festschrift Issue.

T. Yoshikawa · M. Kobayashi · H. Nakai (✉)
Department of Chemistry and Biochemistry,
School of Advanced Science and Engineering,
Waseda University, Tokyo 169-8555, Japan
e-mail: nakai@waseda.jp

M. Kobayashi
Department of Theoretical and Computational
Molecular Science, Institute for Molecular Science,
Okazaki 444-8585, Japan

H. Nakai
Research Institute for Science and Engineering,
Waseda University, Tokyo 169-8555, Japan

H. Nakai
CREST, Japan Science and Technology Agency,
Tokyo 102-0075, Japan

1 Introduction

The second-order Møller–Plesset perturbation (MP2) theory [1] has been widely used because it is the most practical (modestly accurate and fast) molecular orbital (MO) method that can deal with electron correlation in nonempirical manner. Therefore, many quantum chemists have practiced the efficient implementations of the MP2 computation to date. Recent trends in developing the efficient MP2 schemes have been to approximate the MP2 computation by using a rapid calculation trick of the standard MP2 formalism (e.g., local correlation method [2–7], Laplace-transformed method [8–15], resolution of the identity (RI) technique [16, 17], and Cholesky decomposed technique [18, 19]) or by fragmenting the system under consideration (e.g., fragment molecular orbital (FMO) method [20–23], molecular tailoring approach [24–26], incremental correlation scheme [27, 28], and divide-and-conquer (DC) method [29–31]). However, an efficient implementation of the straightforward MP2 formalism is indispensable not only to improve the fundamental performance of the approximate treatment but also to evaluate these approximation schemes.

In the past several years, Nagase and coworkers have offered efficient MP2 schemes especially tuned to parallel implementation. They first provided non-approximate MP2 energy calculation scheme [32] and extended it to the nuclear gradient evaluation [33]. This scheme, commonly called IMS-MP2, allowed us to run actual MP2 calculations with $\sim$2,000 basis functions using a moderate-size

PC cluster. In the next place, they presented parallel RI-MP2 implementation in 2009 [34]. This MP2 scheme is further extended to the periodic system calculations with Bloch Gaussian basis functions [35]. These two excellent schemes have been implemented into the GAMESS program package [36] and interfaced to GAMESS-FMO program [37] with an exception of the periodic RI-MP2 method.

We have also implemented fragmentation-based linear-scaling DC-MP2 method into GAMESS package (for review, see Refs. [38, 39]). The DC method was firstly proposed by Yang and coworkers [40, 41] in the framework of the one-body approximation such as the Hartree–Fock (HF) method and density functional theory (DFT). We have investigated its performance for calculations including HF exchange interactions [42–44] and have applied the method to static and dynamic (hyper)polarizability calculations [45, 46]. After our basic assessments of the DC-HF method, He and Merz [47] independently developed the DC-HF code and assessed its effectiveness in calculations of realistic closed-shell proteins. Furthermore, its extension to MP2 and the other electron correlation theories [namely, a series of coupled cluster (CC) methods] was handled by the authors' group in a different fashion [29–31, 48, 49] with the assistance of the energy density analysis (EDA) [50]. Recently, the DC-MP2 module in GAMESS has been interfaced to the IMS-MP2 code [51]. The history of the DC method is well described in a recent review paper [39].

The applications of the DC method were limited to the closed-shell systems until we utilized the unrestricted HF (UHF) or DFT (UDFT) scheme to the DC method [52]. This DC-UHF/UDFT method has an important advantage over the other fragmentation-based linear-scaling open-shell treatments [53–57] that does not require an artificial guess for the position of excess spins or charges. Even in the elongation method [55, 56], where the spin-delocalized π-conjugated systems have been treated with reasonable accuracy, each piecewise calculation is performed for an integer number of electrons, and the number of electrons for the fragment that is frozen in the forthcoming calculation should be specified in integer number. On the other hand, in the DC method, no artificial prediction related to the positions of the spin and/or charge is required because the distribution of electrons in the system under consideration is uniformly settled by the common Fermi level. However, no ab initio electron correlation theories have been practiced in the DC calculations of open-shell systems.

In this paper, we extended the DC-MP2 method to the unrestricted orbital-based open-shell calculations, which we call DC-UMP2. The organization of this article is as follows. Section 2 presents the theoretical aspects of the DC-UMP2 method after a brief summary of the DC-UHF/UDFT method. Numerical applications of the present scheme are given in Sect. 3 in calculations of the charge- and spin-delocalized polyene cation systems. The conclusion follows as Sect. 4.

2 Theory

2.1 DC-UHF method

In the DC method, the system under consideration is spatially divided into disjoint subsystems, which is called the central region. A set of AOs corresponding to the central region α is denoted by $\boldsymbol{S}(\alpha)$. To improve the description of the subsystem, the neighboring region from the central region, called the buffer region, is taken into consideration when expanding subsystem molecular orbitals (MOs) in the DC calculation. A set of AOs corresponding to the buffer region of subsystem α, denoted by $\boldsymbol{B}(\alpha)$, is added to $\boldsymbol{S}(\alpha)$ and one constructs a set of AOs in the localization region of subsystem α, $\boldsymbol{L}(\alpha)$; namely,

$$\boldsymbol{S}(\alpha) \cup \boldsymbol{B}(\alpha) \equiv \boldsymbol{L}(\alpha). \tag{1}$$

In the DC-UHF calculation of a system with $n_\uparrow$ up-spin and $n_\downarrow$ down-spin electrons, the one-electron density matrices for up- and down-spins, $\mathbf{D}^\uparrow$ and $\mathbf{D}^\downarrow$, are given by

$$D^{\uparrow}_{\mu\nu} \approx D^{\uparrow \mathrm{DC}}_{\mu\nu} = \sum_{\alpha} D^{\uparrow\alpha}_{\mu\nu}, \tag{2}$$

$$D^{\downarrow}_{\mu\nu} \approx D^{\downarrow \mathrm{DC}}_{\mu\nu} = \sum_{\alpha} D^{\downarrow\alpha}_{\mu\nu}. \tag{3}$$

$\mathbf{D}^{\sigma\alpha}(\sigma = \uparrow \text{ or } \downarrow)$ represents the σ-spin local density matrix for subsystem α, obtained by using the Fermi level $\varepsilon_{\mathrm{F}}^{\sigma\alpha}$ and Fermi function $f_\beta(x) = [1 + \exp(-\beta x)]^{-1}$ with an inverse temperature parameter β as follows:

$$D^{\sigma\alpha}_{\mu\nu} \approx p^{a}_{\mu\nu} \sum_{q} f_\beta\left(\varepsilon^{\sigma}_{\mathrm{F}} - \varepsilon^{\sigma\alpha}_{q}\right) C^{\sigma\alpha}_{\mu q} C^{\sigma\alpha *}_{\nu q}, \tag{4}$$

where $\mathbf{p}^\alpha$ is the partition matrix with elements of

$$p^{\alpha}_{\mu\nu} = \begin{cases} 1 & [\mu \in \boldsymbol{S}(\alpha) \wedge \nu \in \boldsymbol{S}(\alpha)] \\ 1/2 & [\mu \in \boldsymbol{S}(\alpha) \wedge \nu \in \boldsymbol{B}(\alpha)] \vee [\mu \in \boldsymbol{B}(\alpha) \wedge \nu \in \boldsymbol{S}(\alpha)] \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

$C^{\sigma\alpha}_{\mu q}$ and $\varepsilon^{\sigma\alpha}_{q}$ are the subsystem MO coefficient and orbital energy for σ-spin electrons, which are determined by solving the following Pople–Nesbet equation for subsystem α,

$$\mathbf{F}^{\sigma\alpha}\mathbf{C}^{\sigma\alpha}_{q} = \varepsilon^{\sigma\alpha}_{q}\mathbf{S}^{\alpha}\mathbf{C}^{\sigma\alpha}_{q}. \tag{6}$$

Here, $\mathbf{S}^\alpha$ and $\mathbf{F}^{\sigma\alpha}$ represent local overlap and σ-spin Fock matrices for subsystem α that are the submatrices of the entire overlap and Fock matrices in the basis of $\boldsymbol{L}(\alpha)$. Each Fermi

level $\varepsilon_{\mathrm{F}}^{\sigma}$ can be determined independently and uniquely by the constraint of the total number of σ-spin electrons n_{e}^{σ}:

$$n_{\mathrm{e}}^{\sigma} = \mathrm{Tr}\left[\mathbf{D}^{\sigma\mathrm{DC}}\mathbf{S}\right] = \sum_{\alpha}\sum_{\mu\in\boldsymbol{L}(\alpha)}\left(\mathbf{D}^{\sigma\alpha}\mathbf{S}^{\alpha}\right)_{\mu\mu}. \tag{7}$$

Then, the entire density matrix $\mathbf{D}^{\sigma\mathrm{DC}}$ can be obtained from Eqs. 2 and 3.

2.2 DC-UMP2 method

The electron correlation energy of the MP2 method can be expressed in terms of active occupied orbitals $\{\varphi_i, \varphi_j\}$ and virtual orbitals $\{\varphi_a, \varphi_b\}$ with the two-electron integral notation $\langle ij|ab\rangle = \iint \phi_i^*(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)r_{12}^{-1}\phi_a(\mathbf{r}_1)\phi_b(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2$ as follows:

$$\Delta E_{\mathrm{corr}} = -\sum_{i<j}^{\mathrm{occ}}\sum_{a<b}^{\mathrm{vir}}\frac{|\langle ij \mid ab\rangle - \langle ij \mid ba\rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}, \tag{8}$$

in spin-orbital notation. In the UMP2 theory, the correlation energy can be rewritten with spatial orbitals as sum of up-spin, down-spin, and cross terms as follows:

$$\begin{aligned}\Delta E_{\mathrm{UMP2}} = &\sum_{i^{\uparrow}<j^{\uparrow}}^{\mathrm{occ}}\sum_{a^{\uparrow}<b^{\uparrow}}^{\mathrm{vir}}\left\langle i^{\uparrow}j^{\uparrow}\middle|a^{\uparrow}b^{\uparrow}\right\rangle\left[\tilde{t}_{i^{\uparrow}j^{\uparrow},a^{\uparrow}b^{\uparrow}} - \tilde{t}_{i^{\uparrow}j^{\uparrow},b^{\uparrow}a^{\uparrow}}\right]\\ &+\sum_{i^{\downarrow}<j^{\downarrow}}^{\mathrm{occ}}\sum_{a^{\downarrow}<b^{\downarrow}}^{\mathrm{vir}}\left\langle i^{\downarrow}j^{\downarrow}\middle|a^{\downarrow}b^{\downarrow}\right\rangle\left[\tilde{t}_{i^{\downarrow}j^{\downarrow},a^{\downarrow}b^{\downarrow}} - \tilde{t}_{i^{\downarrow}j^{\downarrow},b^{\downarrow}a^{\downarrow}}\right]\\ &+\sum_{i^{\uparrow}}^{\mathrm{occ}}\sum_{j^{\downarrow}}^{\mathrm{occ}}\sum_{a^{\uparrow}}^{\mathrm{vir}}\sum_{b^{\downarrow}}^{\mathrm{vir}}\left\langle i^{\uparrow}j^{\downarrow}\middle|a^{\uparrow}b^{\downarrow}\right\rangle\tilde{t}_{i^{\uparrow}j^{\downarrow},a^{\uparrow}b^{\downarrow}},\end{aligned} \tag{9}$$

where $\tilde{t}_{i^{\sigma}j^{\sigma'},a^{\sigma}b^{\sigma'}}$ represents an effective two-electron excitation coefficient as follows:

$$\tilde{t}_{i^{\sigma}j^{\sigma'},a^{\sigma}b^{\sigma'}} = -\frac{\left\langle a^{\sigma}b^{\sigma'}\middle|i^{\sigma}j^{\sigma'}\right\rangle}{\varepsilon_a^{\sigma} + \varepsilon_b^{\sigma'} - \varepsilon_i^{\sigma} - \varepsilon_j^{\sigma'}}. \tag{10}$$

In the DC-based correlation theory, the total correlation energy is estimated by summing up correlation energies corresponding to individual subsystems. We have extended this strategy to the UMP2 theory:

$$\Delta E_{\mathrm{DC-UMP2}} = \sum_{\alpha}^{\mathrm{subsystem}}\Delta E_{\mathrm{UMP2}}^{\alpha}. \tag{11}$$

Here, the correlation energy of subsystem α, $\Delta E_{\mathrm{UMP2}}^{\alpha}$, is estimated using subsystem orbitals, which are constructed in the localization region, containing not only the central region but also the buffer region. While the buffer regions overlap in several subsystems, the central ones have no overlap. To avoid double counting, the correlation energies corresponding to the central regions should be estimated. Thus, we adopted the EDA technique [50] applied to the UMP2 correlation energy representation as follows:

$$\begin{aligned}\Delta E_{\mathrm{UMP2}}^{\alpha} = &\sum_{\sigma}^{\uparrow,\downarrow}\sum_{i^{\sigma\alpha}<j^{\sigma\alpha}}^{\mathrm{occ}}\sum_{a^{\sigma\alpha}<b^{\sigma\alpha}}^{\mathrm{vir}}\sum_{\mu\in\boldsymbol{S}(\alpha)}C_{\mu i}^{\sigma\alpha *}\langle\mu j^{\sigma\alpha} \mid a^{\sigma\alpha}b^{\sigma\alpha}\rangle\\ &\times\left[\tilde{t}_{i^{\sigma}j^{\sigma},a^{\sigma}b^{\sigma}}^{\alpha} - \tilde{t}_{i^{\sigma}j^{\sigma},b^{\sigma}a^{\sigma}}^{\alpha}\right]\\ &+\sum_{i^{\uparrow\alpha}}^{\mathrm{occ}}\sum_{j^{\downarrow\alpha}}^{\mathrm{occ}}\sum_{a^{\uparrow\alpha}}^{\mathrm{vir}}\sum_{b^{\downarrow\alpha}}^{\mathrm{vir}}\sum_{\mu\in\boldsymbol{S}(\alpha)}\frac{1}{2}\Big(C_{\mu i}^{\uparrow\alpha *}\langle\mu j^{\downarrow\alpha} \mid a^{\uparrow\alpha}b^{\downarrow\alpha}\rangle\\ &+C_{\mu j}^{\downarrow\alpha *}\langle i^{\uparrow\alpha}\mu \mid a^{\uparrow\alpha}b^{\downarrow\alpha}\rangle\Big)\tilde{t}_{i^{\uparrow}j^{\downarrow},a^{\uparrow}b^{\downarrow}}^{\alpha},\end{aligned} \tag{12}$$

where $\varphi_i^{\sigma\alpha}$ and $\varphi_j^{\sigma\alpha}$ represent occupied subsystem orbitals that have orbital energies $\varepsilon_i^{\sigma\alpha}$ and $\varepsilon_j^{\sigma\alpha}$ less than the Fermi level for σ-spin $\varepsilon_{\mathrm{F}}^{\sigma}$, which is determined by the preceding DC-UHF calculation, while $\varphi_a^{\sigma\alpha}$ and $\varphi_b^{\sigma\alpha}$ represent virtual subsystem orbitals that have orbital energies $\varepsilon_a^{\sigma\alpha}$ and $\varepsilon_b^{\sigma\alpha}$ greater than the Fermi level $\varepsilon_{\mathrm{F}}^{\sigma}$. $\tilde{t}_{i^{\sigma}j^{\sigma'},a^{\sigma}b^{\sigma'}}^{\alpha}$ represents an effective two-electron excitation coefficient for subsystem α as follows:

$$\tilde{t}_{i^{\sigma}j^{\sigma'},a^{\sigma}b^{\sigma'}}^{\alpha} = -\frac{\left\langle a^{\sigma\alpha}b^{\sigma'\alpha} \mid i^{\sigma\alpha}j^{\sigma'\alpha}\right\rangle}{\varepsilon_a^{\sigma\alpha} + \varepsilon_b^{\sigma'\alpha} - \varepsilon_i^{\sigma\alpha} - \varepsilon_j^{\sigma'\alpha}}. \tag{13}$$

In the present calculations, we utilized the dual-buffer DC scheme where the buffer regions used for the DC-based correlation calculation is set to be smaller than those for the DC-HF calculation. This scheme, described in detail in a different paper [31], reduces the computational efforts for the evaluation of the correlation energy with keeping its accuracy. Furthermore, the DC-HF procedure can be substituted with the standard HF by taking the limit of infinite buffer size. Although the computational cost for the conventional HF calculation scales as $O(n^3)$, it is usually significantly less than the cost for the MP2 calculation.

A quantity that specifically appears in unrestricted open-shell calculations is the expected value of the squared spin operator $\hat{S}^2$, which indicates the degree of spin contamination. We should care more about the issue of spin contamination when adopting the UHF or UMP2 method than UDFT. Although the spin-projection methods such as projected UHF and UMP2 [58] are the possible candidates for regaining from the spin contamination, they violate the size consistency, which must be maintained in the DC scheme. $\langle\hat{S}^2\rangle$ is generally given with the reduced two-electron density matrix Γ as [59]

$$\left\langle\hat{S}^2\right\rangle = S_z^2 + \frac{n_{\uparrow}+n_{\downarrow}}{2} + \sum_{p^{\uparrow}q^{\downarrow}r^{\downarrow}s^{\uparrow}}\left\langle p^{\uparrow}\middle|r^{\downarrow}\right\rangle\left\langle q^{\downarrow}\middle|s^{\uparrow}\right\rangle\Gamma_{p^{\uparrow}q^{\downarrow}r^{\downarrow}s^{\uparrow}}, \tag{14}$$

where n_{σ} is the number of σ-spin electrons and $S_z = \left(n_{\uparrow} - n_{\downarrow}\right)/2$. In the DC-UHF method [52], the third term of Eq. 14 can be evaluated with the DC-UHF density

matrices. However, we have not derived the DC-MP2 density matrix so far. We will report the scheme to evaluate $\langle \hat{S}^2 \rangle$ in the DC-UMP2 elsewhere.

3 Illustrative applications

The present DC-UMP2 method was assessed in calculations of polyene cation doublet $C_nH_{n+2}^+$ (see Fig. 1 representing an example for $n = 30$, C_{2h} symmetry). All calculations were performed with the modified version of the GAMESS program package [36]. C_2H_4 (or C_2H_5 for the edges) was adopted as a central region and several adjacent C_2H_4 (or C_2H_5) units were treated as the corresponding buffer region. The size of buffer region is denoted by n_b^{corr} that indicates the number of carbon atoms in each left and right buffer region (see Fig. 1 representing an example for $n_b^{corr} = 8$). The following calculations were performed with the 6-31G** basis set [60] unless otherwise noted. Applying the dual-buffer DC correlation scheme, only the electron correlation was treated with the DC approach after the standard UHF calculations for clearly showing that the errors reported in the present study originate only in the present DC-UMP2 approximation.

Table 1 shows the correlation buffer size $\left(n_b^{corr}\right)$ dependence of the correlation energies obtained by the DC-UMP2 calculations of polyene cation $C_{30}H_{32}^+$. The energies of neutral polyene $C_{30}H_{32}$ obtained by the restricted MP2 calculations are also tabulated for comparison. The conventional MP2 correlation energies are listed on the bottom line, and the differences between DC and standard energies are presented in parentheses in mhartree. As the correlation buffer size n_b^{corr} increases, the energy error becomes small. The energy errors for $n_b^{corr} \geq 6$ are <1.4 mhartree that achieves so-called chemical accuracy (1 kcal/mol). It was also found that the errors by the DC-UMP2 calculations are comparable to those by the closed-shell DC-MP2 calculations except for the case adopting the smallest buffer size of $n_b^{corr} = 4$, where the so-called error cancellation may occur. Note that the correlation energy errors of the neutral polyene system reported in the present study are slightly larger than those in Ref. [48], where the smaller 6-31G basis set was adopted, although they are comparable to those in Ref. [49]

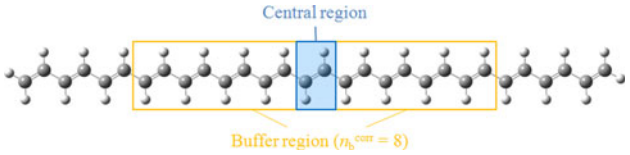

Fig. 1 Structure of the polyene cation $C_{30}H_{32}^+$ and the schematic of the central and buffer regions in the DC calculations with $n_b^{corr} = 8$

Table 1 Correlation buffer-size dependence of DC-MP2 correlation energies (in hartree) of the neutral and cation polyenes, $C_{30}H_{32}$ and $C_{30}H_{32}^+$, at the 6-31G** level

n_b^{corr}	Neutral ΔE_{RMP2} (Diff.)	Cation ΔE_{UMP2} (Diff.)
4	−3.992941 (+2.780)	−3.774790 (−0.452)
6	−3.994383 (+1.339)	−3.772970 (+1.367)
8	−3.995231 (+0.490)	−3.773936 (+0.402)
10	−3.995582 (+0.140)	−3.774274 (+0.064)
Conventional	−3.995721 (–)	−3.774338 (–)

The closed-shell neutral polyene was calculated adopting the restricted orbitals. DC scheme was only applied to the correlation calculation after the standard HF calculation. Energy deviations from conventional MP2 results are shown in parentheses in mhartree

Table 2 System-size dependence of DC and conventional UMP2 total energies (in hartree) of the polyene cation $C_nH_{n+2}^+$ at the 6-31G** level with $n_b^{corr} = 8$

n	E_{UMP2}	$E_{DC-UMP2}$ (Diff.)
10	−386.646383	−386.646383 (−0.000)
20	−772.401788	−772.401678 (+0.105)
26	−1,003.812448	−1,003.812359 (+0.090)
30	−1,158.112941	−1,158.112539 (+0.402)
40	−1,543.817715	−1,543.817507 (+0.209)
60	–	−2,315.223720 (–)

DC scheme was only applied to the correlation calculation after the standard UHF calculation. Energy deviations from conventional UMP2 results are shown in parentheses in mhartree

adopting the same 6-31G** basis set in the CCSD(T) level of theory.

Table 2 compares the total energy of polyene cation systems $C_nH_{n+2}^+$ ($n = 10$, 20, 26, 30, 40, and 60) obtained by the DC and conventional UMP2 method. The correlation buffer size was fixed at $n_b^{corr} = 8$ in the DC calculations, which is larger than in the DC-CC studies [48, 49] because the cost for the MP2 correlation calculation is significantly lower than that for CC calculation. The differences between DC and conventional energies are shown in parentheses in mhartree. The energy error for $n = 10$ becomes zero because all localization regions contain the entire system when using $n_b^{corr} = 8$. It was found that the errors introduced by the DC method keep <0.5 mhartree for $n = 20$–40 by using $n_b^{corr} = 8$.

The efficiency of the DC-UMP2 method was examined by measuring the central processing unit (CPU) time. An Intel Xeon X5470 (3.33 GHz) processor was used on a single core. Table 3 shows the system-size dependence of the CPU times and required memory size for the DC and conventional UMP2 calculations of polyene cation systems $C_nH_{n+2}^+$ ($n = 10$, 20, 26, 30, 40, and 60) with the

Table 3 System-size dependence of DC and conventional UMP2 CPU time (min) and required memory size (Mbyte) of the polyene cation $C_nH_{n+2}^+$ at the 6-31G** level with $n_b^{corr} = 8$

n	CPU time		Required memory size	
	Conventional UMP2	DC-UMP2	Conventional UMP2	DC-UMP2
10	2.0	16.7	76.4	150.2
20	31.3	139.8	558.8	783.0
26	319.6	264.3	1,203.1	751.4
30	697.1	347.0	1,831.8	751.4
40	3,514.8	551.2	4,278.5	751.4
60	–	960.9	–	751.4

An Intel Xeon X5470 (3.33 GHz) processor was used on a single core

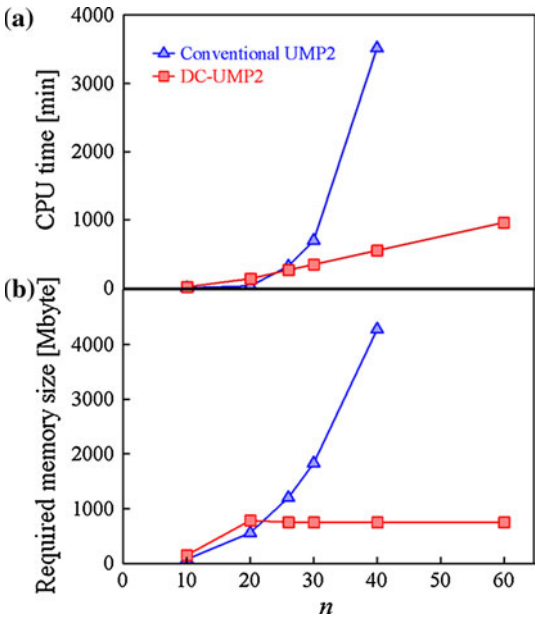

Fig. 2 System size dependence of UMP2 **a** CPU time (min) and **b** required memory size (Mbyte) for the DC and conventional UMP2 calculations of the polyene cation $C_nH_{n+2}^+$ at the 6-31G** level with $n_b^{corr} = 8$. An Intel Xeon X5470 (3.33 GHz) processor was used on a single core

correlation buffer size $n_b^{corr} = 8$. The CPU times and required memory sizes are plotted in Fig. 2a and b, respectively. It should be noted that the time for the HF iterations preceding the MP2 calculation is not included because we utilized the standard full SCF procedure for both DC and conventional calculations. As expected from the closed-shell DC-MP2 presented in the previous paper [30], the DC-UMP2 method drastically reduces the CPU time from the conventional UMP2 time. For $n \leq 20$, the times for DC-UMP2 calculations were larger than those for the conventional calculations, because the median localization region contains up to 18 carbon atoms. However, for $n \geq 26$, the DC method becomes faster than the conventional method. According to the scaling analysis by the double-logarithmic plot, the CPU times scale with $O(n^{5.57})$ and $O(n^{1.53})$ for conventional and DC-UMP2 calculations, respectively, which are slightly larger than the theoretical asymptotic values of $O(n^5)$ and $O(n^1)$. The present implementation is based on the UMP2 code in the present GAMESS program, which requires heavy disk I/O. The disk I/O may reduce the effectiveness of both conventional and DC-MP2 from their theoretical limit. For $n = 10$, where all localization regions contain the entire system, the required memory size for the DC-UMP2 calculation is approximately twice as large as that for the conventional UMP2 calculation because the evaluation of the correlation energy corresponding to the central region requires additional memories to the conventional UMP2 calculation. For $n \geq 26$, however, the memory size required for the DC-UMP2 calculation is constant with respect to the system size n, while that for the conventional calculation increases roughly proportionally to n^3.

Table 4 Basis-set dependence of DC and conventional UMP2 total energies (in hartree) of the polyene cation $C_{40}H_{42}^+$ with $n_b^{corr} = 8$

Basis set	E_{UMP2}	$E_{DC-UMP2}$ (Diff.)
STO-6G	−1,536.939441	−1,536.939354 (+0.087)
6-31G	−1,541.380225	−1,541.380017 (+0.208)
6-31G**	−1,543.817715	−1,543.817507 (+0.209)
6-311G	−1,541.873844	−1,541.873865 (−0.021)
6-311G**	−1,544.368953	−1,544.368583 (+0.370)

DC scheme was only applied to the correlation calculation after the standard UHF calculation. Energy deviations from conventional UMP2 results are shown in parentheses in mhartree

Finally, Table 4 shows the basis-set dependence of the total energies obtained by the DC-UMP2 calculations of polyene cation $C_{40}H_{42}^+$. STO-6G [61], 6-31G [62], 6-311G [63], and 6-311G** [63] basis sets were adopted in addition to the 6-31G** set. The correlation buffer-size was fixed at $n_b^{corr} = 8$. The differences between DC and conventional energies are presented in parentheses in mhartree. The energy errors do not show significant dependence on the basis set adopted and are comparatively small: <0.4 mhartree. Therefore, the use of larger basis set does not deteriorate the effectiveness of the present method, unless diffuse functions are added. The issue on the diffuse functions in the DC method should be resolved elsewhere.

4 Conclusion

In the previous paper [52], we have introduced the unrestricted orbital scheme to the DC SCF method for treating large open-shell systems. In this study, we enabled the

correlated open-shell treatment in the framework of the DC-based correlation method. The present DC-UMP2 method was implemented into the GAMESS program and was assessed in calculations of the spin- and charge-delocalized polyene cation systems $C_nH_{n+2}^+$. Numerical assessments revealed that the DC-UMP2 method has the advantageous features of the closed-shell DC-MP2 method: the correlation energy errors are generally small achieving the chemical accuracy and are controllable with the buffer size, the CPU time scales quasi-linearly with respect to the system size, and the required memory size becomes constant.

In the recent computer architecture, the development of an efficient parallelization scheme is important as well as the acceleration on a single core. Because the correlation energy of a subsystem is evaluated independently of the other subsystems in the DC-MP2 method, straightforward parallelization over subsystems will enhance the parallel efficiency in DC-MP2 calculation. In case of closed-shell systems, the DC-MP2 code applied to the generalized distributed data interface was developed in order to achieve a two-level hierarchical parallelization under the collaboration with Katouda and Nagase [51]. The use of this parallelization scheme as well as the development of the individual MP2 algorithm that is appropriate for the recent computer architecture will extend the applicability of the DC-(U)MP2 method to huge systems, including nanomagnetic materials, carbon materials, metalloenzymes, and heme proteins.

Acknowledgments Some of the present calculations were performed at the Research Center for Computational Science (RCCS), Okazaki Research Facilities, National Institutes of Natural Sciences (NINS). This study was supported in part by Grants-in-Aid for Challenging Exploratory Research "KAKENHI 22655008" and for Young Scientists (B) "KAKENHI 22750016" from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan; by the Nanoscience Program in the Next Generation Super Computing Project of the MEXT; by the Global Center Of Excellence (COE) "Practical Chemical Wisdom" from the MEXT; and by a project research grant for "Practical in-silico chemistry for material design" from the Research Institute for Science and Engineering (RISE), Waseda University.

References

1. Møller C, Plesset MS (1934) Phys Rev 46:618
2. Pulay P, Saebø S (1986) Theor Chim Acta 69:357
3. Saebø S, Pulay P (1993) Annu Rev Phys Chem 44:213
4. Saebø S, Pulay P (2001) J Chem Phys 115:3975
5. Schütz M, Hetzer G, Werner H-J (1999) J Chem Phys 111:5691
6. Maslen PE, Head-Gordon M (1998) Chem Phys Lett 283:102
7. Nielsen IMB, Janssen CL (2007) J Chem Theory Comput 3:71
8. Almlöf J (1991) Chem Phys Lett 181:319
9. Häser M (1993) Theor Chim Acta 87:147
10. Ayala PY, Scuseria GE (1999) J Chem Phys 110:3660
11. Scuseria GE, Ayala PY (1999) J Chem Phys 111:8330
12. Lambrecht DS, Doser B, Ochsenfeld C (2005) J Chem Phys 123:184102
13. Surján PR (2005) Chem Phys Lett 406:318
14. Kobayashi M, Nakai H (2006) Chem Phys Lett 420:250
15. Doser B, Lambrecht DS, Kussmann J, Ochsenfeld C (2009) J Chem Phys 130:064107
16. Feyereisen M, Fitzgerald G, Komornicki A (1993) Chem Phys Lett 208:359
17. Rendell AP, Lee TJ (1994) J Chem Phys 101:400
18. Koch H, Sánchez de Merás A, Pedersen TB (2003) J Chem Phys 118:9481
19. Pedersen TB, Sánchez de Merás AMJ, Koch H (2004) J Chem Phys 120:8887
20. Fedorov DG, Kitaura K (2004) J Chem Phys 121:2483
21. Mochizuki Y, Koikegami S, Nakano T, Amari S, Kitaura K (2004) Chem Phys Lett 396:473
22. Ishikawa T, Kuwata K (2009) Chem Phys Lett 474:195
23. Okiyama Y, Nakano T, Yamashita K, Mochizuki Y, Taguchi N (2010) Chem Phys Lett 490:84
24. Babu K, Gadre SR (2003) J Comput Chem 24:484
25. Ganesh V, Dongare RK, Balanarayan P, Gadre SR (2006) J Chem Phys 125:104109
26. Rahalkar AP, Katouda M, Gadre SR, Nagase S (2010) J Comput Chem 31:2405
27. Friedrich J, Dolg M (2008) J Chem Phys 129:244105
28. Friedrich J, Dolg M (2009) J Chem Theory Comput 5:287
29. Kobayashi M, Akama T, Nakai H (2006) J Chem Phys 125:204106
30. Kobayashi M, Imamura Y, Nakai H (2007) J Chem Phys 127:074103
31. Kobayashi M, Nakai H (2009) Int J Quantum Chem 109:2227
32. Ishimura K, Pulay P, Nagase S (2006) J Comput Chem 27:407
33. Ishimura K, Pulay P, Nagase S (2007) J Comput Chem 28:2034
34. Katouda M, Nagase S (2009) Int J Quantum Chem 109:2121
35. Katouda M, Nagase S (2010) J Chem Phys 133:184103
36. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA Jr (1993) J Comput Chem 14:1347
37. Fedorov DG, Kitaura K (2007) J Phys Chem A 111:6904
38. Kobayashi M, Akama T, Nakai H (2009) J Comput Chem Jpn 8:1
39. Kobayashi M, Nakai H (2011) Divide-and-conquer approaches to quantum chemistry: theory and implementation. In: Papadopoulos MG, Zalesny R, Mezey PG, Leszczynski J (eds) Linear-scaling techniques in computational chemistry and physics. Springer, Berlin, pp 97–127
40. Yang W (1991) Phys Rev Lett 66:1438
41. Yang W, Lee TS (1995) J Chem Phys 103:5674
42. Akama T, Kobayashi M, Nakai H (2007) J Comput Chem 28:2003
43. Akama T, Fujii A, Kobayashi M, Nakai H (2007) Mol Phys 105:2799
44. Kobayashi M, Kunisada T, Akama T, Sakura D, Nakai H (2011) J Chem Phys 134:034105
45. Touma T, Kobayashi M, Nakai H (2010) Chem Phys Lett 485:247
46. Touma T, Kobayashi M, Nakai H (2011) Theor Chem Acc. doi: 10.1007/s00214-011-0964-2
47. He X, Merz KM Jr (2010) J Chem Theory Comput 6:405
48. Kobayashi M, Nakai H (2008) J Chem Phys 129:044103
49. Kobayashi M, Nakai H (2009) J Chem Phys 131:114108
50. Nakai H (2002) Chem Phys Lett 363:73
51. Katouda M, Kobayashi M, Nakai H, Nagase S (2011) J Comput Chem. doi:10.1002/jcc.21855

52. Kobayashi M, Yoshikawa T, Nakai H (2010) Chem Phys Lett 500:172
53. Pruitt SR, Fedorov DG, Kitaura K, Gordon MS (2010) J Chem Theory Comput 6:1
54. Advance/BioStation version 3.2, Advance Soft, Tokyo
55. Korchowiec J, Gu FL, Aoki Y (2005) Int J Quantum Chem 105:875
56. Orimoto Y, Gu FL, Korchowiec J, Imamura A, Aoki Y (2010) Theor Chem Acc 125:493
57. Friedrich J, Hanrath M, Dolg M (2008) J Phys Chem A 112:8762
58. Schlegel HB (1986) J Chem Phys 84:4530
59. Krylov AI (2000) J Chem Phys 113:6052
60. Hariharan PC, Pople JA (1973) Theor Chim Acta 28:213
61. Hehre WJ, Stewart RF, Pople JA (1969) J Chem Phys 51:2657
62. Hehre WJ, Ditchfield R, Pople JA (1972) J Chem Phys 56:2257
63. Krishman R, Binkley JS, Seeger R, Pople JA (1980) J Chem Phys 72:650