Doctoral Thesis

Development of Black-Box Static Electron Correlation and Extremely Large-Scale Electronic Structure Methods

（ブラックボックスな静的電子相関計算法と
超大規模電子状態計算手法の開発）

Masatsugu Nishida
Doctor of Philosophy

Graduate School of Chemical Sciences and Engineering
Hokkaido University
March 2025

Contents

Chapter 1

General Introduction

P.A.M Dirac said, *"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."*[15]

Quantum chemistry is the study of seeking methods to approximate the exact Schrödinger equation into solvable forms. Constraints are accuracy and cost. In other words, the calculation must be accurate enough to capture the essence of the phenomenon and finished within a realistic time. Quantum chemists have been challenging to elucidate chemical phenomena fighting against these constraints.

Even in the most basic quantum chemical calculation method; Hartree-Fock (HF) method, its cost scales $O(N^{3-4})$, where N is the number of basis functions. Specifically, the computation of the two-electron integral scales $O(N^4)$, and the diagonalization of the Fock matrix scales $O(N^3)$. Although the HF method can give

more than 99 % of the exact energy, the remaining 1 % is critical in the discussion of chemical reactions. Löwdin defined the difference between the HF solution and the exact one as the electron correlation. Sinanoğlu further classified electron correlations into two types[104]. One is dynamical correlation, the instantaneous repulsion between electrons. The other is static correlation. It is multi-configurational effects caused by quasi-degeneracy around the frontier orbitals. For dynamical electron correlation, Møller and Plesset developed a perturbation correction method. Møller-Plesset (MP) method can incorporate dynamic electron correlation by perturbation with the HF state as a zeroth-order wave function[63]. Especially, the MP2 method, which incorporates perturbations up to the second order, is often used. In MP2 calculation, the most time-consuming step is the integral transformation, which scales $O(N^5)$. In addition, a coupled-cluster (CC) method was developed that approximates higher-order electronic excitations by a product of lower-order ones using an exponential operator[12, 66]. A typical method for static correlation is the CASSCF method. In this method, all possible electronic excitation configurations for a given set of orbitals are considered in the calculation. This set of orbitals and the belonging electrons are called active orbitals and active electrons. These combined active spaces must be set appropriately. However, as the active space is expanded, the computational cost increases exponentially. In addition to high costs, the difficulty of definition for the active space is a serious problem. The fact that

chemical intuition is required to know which electron configurations are important for the description of the system. In addition, the limitation of expanding the active space originating from high costs makes this configuration selection even more difficult. In particular, when important electron configurations change during chemical reactions, the choice of configurations must be made carefully. There is a need for black-box methods that can automatically capture the important electron configurations.

Broken symmetry methods are candidates for incorporating static correlation in a black-box manner. In the unrestricted Hartree-Fock (UHF), static electronic correlations can be partially incorporated in exchange for the loss of spin symmetry. Spin contamination causes problems in physical properties and excited state calculations. The Hartree-Fock-Bogoliubov (HFB) method can automatically incorporate multiple electron configurations by using Bogoliubov quasiparticles, which are expressed as a linear combination of electron production and annihilation operators. Although the HFB wave function preserves the spin symmetry, it violates number symmetry, i.e. different electron number states are contaminated. The projection method is valid for restoring symmetry. The spin-projected unrestricted Hartree-Fock (PUHF) can include static correlation in a black-box way. However, it is still difficult to incorporate static electron correlations in a black-box manner in excited-state calculations. The high scaling computational time required for electron correlation calculations was also an obstacle.

However, since the 1990s, linear scaling methods have been developed to reduce computational scaling to a linear with molecule size. These are based on the "nearsightedness" principle proposed by Kohn[52]. For example, the fragment molecular orbital (FMO) method was proposed by Kitaura and Fedrov[40], the divide-and-conquer (DC) method was proposed by Yang[127]. These methods utilize the sparsity of the density and Hamiltonian matrices to divide a large molecule into smaller fragments, perform calculations, and approximate the energy and properties of the whole molecule from the fragment calculation results. Furthermore, since each fragment is treated separately, further speed-up can be achieved by parallel computation. With the advent of massively parallel computers, such as supercomputers, these methods can handle even larger molecules. These methods have been extended to electron correlation calculations. Dynamic electron correlation is particularly compatible with these large-scale computational methods because it is a local effect. However, static electron correlation is a long-range effect in some cases. In addition, explicitly multiconfigurational wave functions using the active space make division difficult. Therefore, large-scale calculations of static electron correlations have been difficult so far.

Even with linear scaling methods, the computational cost can be high. It is because of the integral calculation. In particular, the computation of two-electron integrals is a bottleneck for linear scaling methods. In recent years, enormously large-scale calculations

that exceed tens of thousands of atoms have been possible because of combination linear scaling methods with semi-empirical methods. In semi-empirical methods, the integral values are replaced by parameters. Nishimura and Nakai developed the DC-DFTB method, which applied the DFTB method to the DC method. The DC-DFTB method can treat very large systems from tens of thousands to hundreds of million atoms. However, the DFTB parameters are complex and the number of available element species is limited. Thus, a very large-scale calculation method, which has the versatility, is necessary.

This thesis consists of six chapters. In Chapter 2, I review two directions of electronic structure calculations: accuracy and larger scale. As for accuracy, I mainly discuss static correlation calculation methods in a black-box way. Regarding larger-scale, linear-scaling/semi-empirical methods are described as large-scale calculation methods. In Chapter 3, I describe an excited-state calculation method that incorporates static correlations in a black-box manner. The TDHFB method, which applies the HFB method that can treat static electron correlations in a black-box manner to excited-state calculations, is described. In Chapter 4, I describe the DC-PUHF method, which combines the PUHF method, which recovers spin symmetry by projection, and the DC method, which is a large-scale electronic structure calculation method. In Chapter 5, I describe the development of the DC-xTB method, which combines the DC method with the xTB method, a general-purpose semi-empirical

calculation method. In Chapter 6, I summarized this study and described future prospects.

Chapter 2

Theoretical Background

2.1 Hartree-Fock method

The Hartree-Fock (HF) method[110] approximates N-electron wave function as a single Slater determinant,

$$\Phi_{\mathrm{HF}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \ldots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \ldots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \ldots & \psi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.1}$$

where ψ_i, $\mathbf{x}_i$ are spin-orbital and spin coordinate. The HF Slater determinant is composed of the occupied spin-orbital set. The goal of the HF method is the optimization of occupied spin orbitals set by the following HF equation,

$$f(i)\psi(\mathbf{x}_i) = \varepsilon\psi(\mathbf{x}_i) \tag{2.2}$$

where $f(i)$ and ε are the effective one-electron operator called the Fock operator and the orbital energy.

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{N_{\mathrm{nuc}}} \frac{Z_A}{r_{iA}} + v^{\mathrm{HF}}(i) \tag{2.3}$$

where v^{HF} is the average potential created by the other electrons, called the mean-field. In addition, the spin-orbital is decomposed into space and spin components.

$$\psi_i(\mathbf{x}) = \begin{cases} \phi_j^{\alpha}(\mathbf{r})\alpha(\omega) \\ \phi_j^{\beta}(\mathbf{r})\beta(\omega) \end{cases} \tag{2.4}$$

In practice, only spatial orbital are considered and the basis function set is introduced. The molecular orbitals are approximated by the linear combination of the atomic orbital (LCAO approximation).

$$\phi_i(\mathbf{r}) = \sum_{\mu}^{N_{\mathrm{basis}}} C_{\mu i}\chi_{\mu}(\mathbf{r}) \tag{2.5}$$

where χ_μ is the atomic orbital. The HF equation is expanded by the basis function.

$$f(i)\sum_{\mu}^{N_{\mathrm{basis}}} C_{\mu i}\chi_{\mu} = \varepsilon_i \sum_{\mu}^{N_{\mathrm{basis}}} C_{\mu i}\chi_{\mu} \tag{2.6}$$

Finally, the Hartree-Fock-Roothaan equation[93, 30] is followed as,

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{2.7}$$

The Fock matrix expanded by the basis function is followed as,

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} D_{\lambda\sigma} \langle \mu\sigma || \nu\lambda \rangle \tag{2.8}$$

where $h_{\mu\nu}$ is core Hamiltonian as follows,

$$h_{\mu\nu}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{N_{\mathrm{nuc}}} \frac{Z_A}{r_{iA}} \tag{2.9}$$

$D_{\mu\nu}$ is the element of the density matrix is represented by the MO coefficients as follows,

$$D_{\mu\nu} = \sum_{i}^{N_{\mathrm{occ}}} C_{\mu i} C_{\nu i}^{*} \tag{2.10}$$

$\langle \mu\sigma || \nu\lambda \rangle$ is the anti-symmetrized two-electron integral as follows,

$$\int\int dr_1 dr_2 \chi_{\mu}^{*}(r_1)\chi_{\sigma}^{*}(r_2) r_{12}^{-1}(1 - \mathcal{P}_{12})\chi_{\nu}(r_1)\chi_{\lambda}(r_2) \tag{2.11}$$

where r_{12} is the distance between electron 1 and electron 2, and $\mathcal{P}_{12}$ is operators that exchange the coordinates of electron 1 and electron 2.

2.2 Density functional theory

The density functional theory (DFT) is based on the Hohenberg-Kohn theorem that the energy in the ground state is determined by only the electron density ρ. The advantage of describing the system in terms of electron density is that the electron density depends on only three spatial coordinates, so the number of variables remains three even when the system grows. On the other hand, the wave function theories (WFT), such as the HF method, have $4N$ variables. Thus, the wave function becomes more complicated as the number of electrons increases. In most cases, the Kohn-Sham orbital[52] is introduced to better describe the kinetic energy in the DFT. The remaining issue is the description of exchange and correlation energies. Various exchange-correlation functionalities have been developed. However, no exact exchange-correlation functional

has been found in 2024.

The exchange/correlation functional described by only the electron density $\rho(\mathbf{r})$ is called the local density approximation (LDA) functional, and the functional corrected by the electron density gradient $\nabla\rho(r)$ is called the Generalized Gradient Approximation (GGA) functional. Furthermore, functional mixed a certain amount of HF exchange energy based on the adiabatic connections is called hybrid functional.

This paper describes the Perdew-Burke-Ernzerhof (PBE) functional, a GGA-type functional that restricts the fundamental physical conditions to be satisfied and the fundamental physical constants to be used to two[76]. In addition, the PBE0 functional was developed by mixing the PBE exchange functional with the HF exchange integral by 25%[1].

$$E_{\mathrm{XC}}^{\mathrm{PBE0}} = E_{\mathrm{XC}}^{\mathrm{PBE}} + \frac{1}{4}(E_{\mathrm{X}}^{\mathrm{HF}} - E_{\mathrm{X}}^{\mathrm{PBE}}) \tag{2.12}$$

The PBEh-3c method has also been proposed to correct for dispersion interaction and basis set superposition error, which are problems with DFT[29].

$$E_{\mathrm{tot}}^{\mathrm{PBEh-3c}} = E_{\mathrm{tot}}^{\mathrm{PBEh}} + E_{\mathrm{disp}} + E_{\mathrm{BSSE}} \tag{2.13}$$

2.3 Electron correlation method

The electron correlation energy is defined as the difference between the exact energy and the HF one[60]. The electron correlation can be classified into two types[104]. The first one is the dynamic cor-

relation originating from the electron repulsion in instantaneous proximity. There are several methods for the dynamic correlation. For example, the Møller-Plesset (MP) method, which is based on the Rayleigh-Schrödinger perturbation theory, can account for the dynamic correlation[63]. The MP method adopts the HF wave function as the zeroth-order wave function and the sum of the Fock operators as the 0th Hamiltonian. The simplest MP method is the second-order perturbation version, MP2. The bottleneck of the MP2 method is the AO-MO transformation whose cost is $O(N^5)$. The coupled cluster (CC) method can also capture the dynamic correlations. In the CC method, the higher excitation is approximated by the products of the lower excitation using the exponential type operator.

On the other hand, another type of electron correlation is the quasi-degenerate effect of the orbital energies around the Frontier Orbital, called the static correlation. In these degenerate systems, the contribution of excited configurations becomes large. In this case, the single configurational wave function such as the HF/DFT gives qualitatively wrong results. Although the multi-configurational wave function can describe the static correlation accurately, there are problems; the complicated selection of the active space and its high computational cost. The black-box and low computational methods for the static correlation are necessary.

2.4 Excited state calculation method

The excited state is important in the specific chemical reaction; the photochemical reaction. The information of the excited states is obtained by solving the time-dependent Schrödinger equation.

$$i\hbar\frac{\partial}{\partial t}\Psi(t) = \hat{H}(t)\Psi(t) \tag{2.14}$$

It is impossible to obtain the exact solution of the Schrödinger equation. Thus, various approximations have been developed. The typical approximation method is the Time-Dependent HF/DFT (TDHF/DFT) method. Especially, the TDHF/DFT equation in the frequency domain, called the Casida equation, is the most popular[9]. In this study, the real-time (RT) method for the solution of the time-dependent equation is adopted. This method has the advantage of tracking the time variation of the electronic states. In the RT method, the TDHF wave function is the time-dependent Slater determinant composed of the time-dependent MOs as follows,

$$i\frac{\partial}{\partial t}\psi_m(\mathbf{r},t) = \hat{F}(\mathbf{r},t) + V^{\mathrm{ext}(\mathbf{r},t)}\psi_m(\mathbf{r},t) \tag{2.15}$$

where $\psi_m(\mathbf{r},t)$, $\hat{F}$ are the m-th time-dependent MO and the Fock operator.

Hereafter, we use the density matrix form of the TD equation, called quantum Liouville's equation, as follows,

$$i\hbar\frac{\partial}{\partial t}\mathbf{D}(t) = [\mathbf{F}(t) + \mathbf{V}^{\mathrm{ext}}(t), \mathbf{D}(t)] \tag{2.16}$$

where $\mathbf{V}^{\mathrm{ext}}(t)$, $\mathbf{F}(t)$ are the time-dependent external fields and Fock

matrices and its element follows as,

$$F_{ij} = \langle \phi_i | \hat{F} | \phi_j \rangle = h_{ij} + \sum_{kl} \langle ik||jl \rangle D_{lk}(t) \tag{2.17}$$

where h_{ij} is one-electron integral and $\langle ik||jl \rangle$ is the two-electron integral. Note that i, j, k, l represents the molecular orbital index.

$$V_{ij}^{\text{ext}} = \langle \phi_i | \hat{V}^{\text{ext}} | \phi_j \rangle \tag{2.18}$$

The time-dependent density matrix $\mathbf{D}(t)$ and $\mathbf{F}(t)$ are represented by the ground states' ones $\mathbf{D}^{(0)}$, $\mathbf{F}^{(0)}$ without external field and induced density $\delta\mathbf{D}(t)$and Fock matrices $\delta\mathbf{F}(t)$ as follows,

$$\mathbf{D}(t) = \mathbf{D}^{(0)} + \delta\mathbf{D}(t) = \mathbf{D}^{(0)} + \sum_{n=1} \delta\mathbf{D}^{(n)}(t) \tag{2.19}$$

$$\mathbf{F}(t) = \mathbf{F}^{(0)} + \delta\mathbf{F}(t) = \mathbf{F}^{(0)} + \sum_{n=1} \delta\mathbf{F}^{(n)}(t) \tag{2.20}$$

where $\delta\mathbf{D}^{(n)}(t)$, $\delta\mathbf{F}^{(n)}(t)$ are n-th order induced density and Fock matrices, respectively.

The equation of motion for the density matrix is as follows,

$$\begin{aligned} & i\hbar \frac{\partial}{\partial t} \left(\mathbf{D}^{(0)} + \sum_{n=1} \delta\mathbf{D}^{(n)}(t) \right) \\ & = \left[\mathbf{F}^{(0)} + \sum_{n=1} \delta\mathbf{F}^{(n)}(t) + \mathbf{V}^{\text{ext}}(t), \mathbf{D}^{(0)} + \sum_{n=1} \delta\mathbf{D}^{(n)}(t) \right] \end{aligned} \tag{2.21}$$

Considering up to the first order, the above equation reduces to the linear-scaling excited state calculation. The equation of motion for

first order induced density matrix $\delta\mathbf{D}^{(1)}$ follows as,

$$i\hbar\frac{\partial}{\partial t}\left(\mathbf{D}^{(1)}(t)\right) = [\mathbf{F}^{(0)}, \delta\mathbf{D}^{(1)}(t)] + [\delta\mathbf{F}^{(1)}, \mathbf{D}^{(0)}] + [\mathbf{V}^{\mathrm{ext}}(t), \mathbf{D}^{(0)}] \tag{2.22}$$

This equation corresponds to Casida's equation. Although only linear response calculations are considered in this study, a non-linear response approach is possible in the RT method. In this study, the dipole approximation for the external fields is adopted,

$$V_{ij}^{\mathrm{ext}}(t) = \varepsilon(t)e^{\mathrm{ext}}\langle\phi_i|\mathbf{r}|\phi_j\rangle \tag{2.23}$$

where e^{ext} and $\langle\phi_i|\mathbf{r}|\phi_j\rangle$, $\varepsilon(t)$ are the direction of the external fields, the dipole integral and the time-dependent part of the external fields. In this study, the external fields of the δ function type are adopted.

$$\varepsilon(t) = \delta(t) \tag{2.24}$$

The time variation of the polarized vector is obtained from the induced density matrices at each time. The polarized vector is represented as follows,

$$\begin{aligned}\mathbf{P}(t) =& -\sum_{ij} D_{ij}(t)\langle\phi_i|\mathbf{r}|\phi_j\rangle \\ =& -\sum_{ij}\left(D_{ij}^{(0)} + \sum_{n=1}\delta D_{ij}^{(n)}(t)\langle\phi_i|\mathbf{r}|\phi_j\rangle\right) \\ =& \mathbf{P}^{(0)} + \sum_{n=1}\delta\mathbf{P}^{(n)}(t)\end{aligned} \tag{2.25}$$

where $\mathbf{P}^{(0)}$, $\delta\mathbf{P}^{(n)}$ static polarized vector and n-th order induced polarized vector. The time variation of the induced polarized vector

gives the absorption spectrum through the Fourier transformation.

$$P(\omega) = \int_{-\infty}^{\infty} dt P(t) e^{-i\omega t} \tag{2.26}$$

The imaginary part of $\mathbf{P}(\omega)$ corresponds to the absorption spectrum and the real part corresponds to polarization.

2.5 Black-Box and Low cost static correlation

2.5.1 Unrestricted Hartree-Fock method

For considering the static correlation, the Unrestricted Hartree-Fock (UHF) method[80] removes the restriction of spin-symmetry. It allows the different spin-orbital to feel the different mean-field.

$$\phi_i^{\alpha}(\mathbf{r}) \neq \phi_i^{\beta}(\mathbf{r}) \tag{2.27}$$

Furthermore, the UHF equation using the Fock operators for α, β orbitals as follows,

$$f^{\alpha}\phi_i^{\alpha} = \varepsilon_i^{\alpha}\phi_i^{\alpha} \tag{2.28}$$

$$f^{\beta}\phi_i^{\beta} = \varepsilon_i^{\beta}\phi_i^{\beta} \tag{2.29}$$

In addition, the following Pople-Nesbet equation are obtained by the basis set expansion,

$$\mathbf{F}^{\alpha}\mathbf{C}^{\alpha} = \mathbf{S}\mathbf{C}^{\alpha}\varepsilon^{\alpha} \tag{2.30}$$

$$\mathbf{F}^{\beta}\mathbf{C}^{\beta} = \mathbf{S}\mathbf{C}^{\beta}\varepsilon^{\beta} \tag{2.31}$$

In terms of the energetic property, the UHF method can describe qualitatively correct for the bond-dissociation. However, the wave function is not correct because the spin symmetry is broken. For example, the property associated with the spin density and the excited state calculation based on the UHF wave function is not qualitatively correct.

2.5.2 Hartree-Fock-Bogoliubov method

The HFB method is one of the black-box static correlation calculation methods. The HFB method has the feature that the wave function is constructed not by electrons but by Bogoliubov quasiparticles, which are described by a linear combination of electron creation/annihilation operators as follows,

$$\hat{b}_i^\dagger = \sum_j U_{ji}\hat{a}_j^\dagger + V_{ji}\hat{a}_j \tag{2.32}$$

$$\hat{b}_i = \sum_j U_{ji}\hat{a}_j + V_{ji}\hat{a}_j^\dagger \tag{2.33}$$

where $\hat{a}^\dagger$, $\hat{a}$ are the electron creation/annihilation operators respectively and $\mathbf{U}$, $\mathbf{V}$ are Bogoliubov transformation matrices. The HFB wave function is constructed as the total product of quasiparticle annihilation operators,

$$|\Psi_{\mathrm{HFB}}\rangle = \prod_i^M \hat{b}_i\,|\rangle \tag{2.34}$$

where $|\rangle$, M and $\hat{b}_i$ are the vacuum state, the number of the quasiparticle. Note that the Bogoliubov transformation is performed

between all orbitals, including unoccupied orbits, not just occupied orbitals. Thus, the HFB wave function becomes multi-configurational.

For simplicity, the Bogboliubov quasiparticle operator in the canonical form follows as,

$$\hat{b}_i = u_i \hat{a}_i + v_i \hat{a}_{\bar{i}}^\dagger \tag{2.35}$$

$$\hat{b}_{\bar{i}} = u_i \hat{a}_{\bar{i}} + v_i \hat{a}_i^\dagger \tag{2.36}$$

Then, the HFB wave function is

$$|\Psi_{\mathrm{HFB}}\rangle = \prod_k^M (u_k + v_k \hat{a}_i^\dagger \hat{a}_{\bar{i}}^\dagger)\,|\rangle\,. \tag{2.37}$$

Thus, the HFB wave function is represented by the electron-pair creation operator. It indicates that the HFB method is connected with the two-electron wave function theories, called geminal[107]. Note that the HFB wave function contaminates the different electron number states.

$$|\Psi_{\mathrm{HFB}}\rangle = \sum_N |N\rangle \tag{2.38}$$

where $|N\rangle$ is the N-electron wave function. The HFB energy is expressed as a spin-orbital representation as follows[48]:

$$E_{\mathrm{HFB}} = \langle \Psi_{\mathrm{HFB}} | \hat{H} | \Psi_{\mathrm{HFB}} \rangle \tag{2.39}$$

where the second quantization form of the Hamiltonian follows as,

$$\hat{H} = \sum_{pq} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \frac{1}{4} \sum_{pqrs} \langle pq||rs\rangle\, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s \tag{2.40}$$

Finally, the HFB energy is represented as follows,

$$E_{\mathrm{HFB}} = \sum_{ij} h_{ij} D_{ji} + \frac{1}{2} \sum_{ijkl} \langle ik||jl\rangle D_{ji} D_{lk} - \frac{1}{4} \zeta \sum_{ijkl} \langle ij||kl\rangle K^*_{ij} K_{kl} \tag{2.41}$$

where h_{ij} is the ijth element of the one-electron integral matrix and $\langle ij||kl\rangle$ are the anti-symmetrized two-electron integrals. D_{ij} and K_{ij} are the ijth elements of the electron density matrix $\mathbf{D}$ and electron pairing matrix $\mathbf{K}$, respectively. The elements of these matrices are expressed with the HFB wavefunction and electron creation and annihilation operators $\hat{a}_i^\dagger$ and $\hat{a}_i$, as

$$D_{ij} = \langle \Psi_{\mathrm{HFB}} | \hat{a}_j^\dagger \hat{a}_i | \Psi_{\mathrm{HFB}} \rangle = D^*_{ji} \tag{2.42}$$

$$K_{ij} = \langle \Psi_{\mathrm{HFB}} | \hat{a}_j \hat{a}_i | \Psi_{\mathrm{HFB}} \rangle = -K_{ji} \tag{2.43}$$

and the third term in Eq. 2.41 represents the pair energy, which effectively considers static correlation. However, the HFB method overestimates the static correlation, so previous study introduced tunable parameter ζ that describes the magnitude of the static correlation[105]. In this study, we focused only on the spin-restricted closed-shell case (i.e., RHFB[41]). In the RHFB method, density and pairing matrices satisfy the following relationship:

$$\bar{\mathbf{D}} \equiv \mathbf{D}^{\alpha\alpha} = \mathbf{D}^{\beta\beta} \tag{2.44}$$

$$\mathbf{D}^{\alpha\beta} = \mathbf{D}^{\beta\alpha} = \mathbf{0} \tag{2.45}$$

$$\bar{\mathbf{D}}^{\dagger} = \bar{\mathbf{D}} \tag{2.46}$$

$$\bar{\mathbf{K}} \equiv \mathbf{K}^{\alpha\beta} = -\mathbf{K}^{\beta\alpha} \tag{2.47}$$

$$\mathbf{D}^{\alpha\beta} = \mathbf{D}^{\beta\alpha} = \mathbf{0} \tag{2.48}$$

$$\mathbf{K}^{\alpha\alpha} = \mathbf{K}^{\beta\beta} = \mathbf{0} \tag{2.49}$$

$$\bar{\mathbf{K}}^{\mathrm{T}} = \bar{\mathbf{K}} \tag{2.50}$$

Then, the RHFB energy is written with spatial orthogonal orbitals:

$$\begin{aligned} E_{\mathrm{RHFB}} = 2 \sum_{ij}^{N_{\mathrm{basis}}} h_{ij} \bar{D}_{ji} + \sum_{ijkl}^{N_{\mathrm{basis}}} \left(2 \langle ik|jl\rangle - \langle ik|lj\rangle\right) \bar{D}_{ji} \bar{D}_{lk} \\ -\zeta \sum_{ijkl}^{N_{\mathrm{basis}}} \langle ij||kl\rangle \bar{K}^{*}_{ij} \bar{K}_{kl} + 2\Lambda \left(N_{\mathrm{elec}} - \sum_{i}^{N_{\mathrm{basis}}} \bar{D}_{ii} \right) \end{aligned} \tag{2.51}$$

where Λ is the chemical potential that conserves the average of the electron number. In the RHFB formalism, the generalized density matrix R can be used, which is expressed in the case of an orthogonal basis set as

$$\bar{\mathbf{R}} = \begin{pmatrix} \bar{\mathbf{D}} & \bar{\mathbf{K}} \\ \bar{\mathbf{K}} & \mathbf{1} - \bar{\mathbf{D}} \end{pmatrix} \tag{2.52}$$

Note that the following equation can be derived from the idempotency of $\bar{\mathbf{R}}$:

$$(\bar{\mathbf{K}})^2 = \bar{\mathbf{D}}(\mathbf{1} - \bar{\mathbf{D}}) \tag{2.53}$$

The energy variation in Eq. 2.51 with respect to $\bar{\mathbf{D}}$ and $\bar{\mathbf{K}}$ yields the following mean-field Hamiltonian matrix:

$$\mathbf{H}_{\mathrm{HFB}} = \begin{pmatrix} \tilde{\mathbf{F}} & \boldsymbol{\Delta} \\ \boldsymbol{\Delta} & -\tilde{\mathbf{F}} \end{pmatrix} \tag{2.54}$$

where $\tilde{\mathbf{F}}$ is the shifted Fock matrix that represents the electron mean-field

$$\tilde{F}_{ij} = \frac{\delta E_{\mathrm{HFB}}}{\delta \bar{D}_{ji}} = h_{ij} - \Lambda \delta_{ij} + \sum_{kl}^{N_{\mathrm{basis}}} \left(2\langle ij|jl\rangle - \langle ik|lj\rangle\right) \bar{D}_{lk} \tag{2.55}$$

and Δ is the pairing-field matrix representing the quasiparticle mean field

$$\Delta_{ij} = \frac{\delta E_{\mathrm{HFB}}}{\delta \bar{K}_{ij}} = -\zeta \sum_{kl}^{N_{\mathrm{basis}}} \bar{K}_{kl} \tag{2.56}$$

From Eq. 2.41, the pairing energy, which is the energy of static electron correlation, can also be expressed using the pairing matrix and the pairing field matrix as follows:

$$E_{\mathrm{pair}} = \mathrm{Tr}(\bar{\mathbf{K}}^{*}\boldsymbol{\Delta}) = \mathrm{Tr}(\bar{\mathbf{K}}\boldsymbol{\Delta}^{*}) \tag{2.57}$$

However, when using a real orthogonal basis, it is simply a trace of the product of the pairing matrix and the pairing field matrix. The quasiparticle orbital energies and their coefficients were obtained by solving the following HFB equation:

Only half of the solutions with negative eigenvalues are used for the Bogoliubov transformation, leading to the following $\bar{\mathbf{D}}$ and $\bar{\mathbf{K}}$:

$$\bar{D}_{ij} = \sum_{k}^{M} V_{ik} V_{jk} \tag{2.58}$$

$$\bar{K}_{ij} = \sum_{k}^{M} V_{ik} U_{jk} \tag{2.59}$$

Note that the HFB wave function conserves the spatial and spin symmetries, but violates the particle-number symmetry. That is, the HFB wave function is not the eigenfunction of the number operator, but rather a superposition of different particle number determinants. Therefore, the chemical potential in Eq. 2.51 was introduced to adjust the average number of electrons in the wave function to the correct value N.

2.6 Projection scheme for broken symmetry wave function

In the quantum chemistry community, the unrestricted methods, which allow alpha and beta orbitals to have different spatial orbitals, have been widely used for describing static correlation. However, this wave function is not an eigenfunction of spin operator $\hat{S}^2$. This problem is called "spin-contamination". Löwdin developed the extended HF (EHF) method by the projection scheme to remove contaminated spin states. In the EHF method, a contaminated spin state is removed by the following spin annihilation operator.

$$\hat{A}^l = \frac{\hat{S}^2 - l(l+1)}{s(s+1) - l(l+1)} \tag{2.60}$$

where l, s are contaminated and target spin states. If we want to remove all contaminated spin states, it is necessary to use the

production of all spin annihilation operators.

$$\hat{P}^s = \prod_{l \neq s} \hat{A}^l \tag{2.61}$$

However, the exact projection is high cost. Sometimes, only the most contaminated state is removed. In the bond dissociation limit, the EHF can be described correctly. However, in the middle region, called the HF instability region, the EHF method gives unphysical results[50]. it is because the EHF wave function doesn't optimize variational. In this way, only once projection after the SCF calculation is called the Projection-After-Variation (PAV) method. The perturbation correction against EHF energy can describe full PES for bond-dissociation[96]. Mayer has developed the variation scheme for EHF wavefunction including projection operator[61]. Such projection schemes are called Variation-After-Projection (VAP) method. However, Mayer's method is complicated, and the computational cost is high.

In nuclear physics, the HFB method is popular for the strong correlation between nuclei[88]. The other particle number states are contaminated in the HFB wave function. To remove contaminated number states, the Generator Coordinate Method (GCM) was developed in this field[103, 102]. This method adopts the following projection operator

$$\hat{P}^\lambda = \frac{1}{L} \int_L d\phi e^{i\phi(\hat{\Lambda} - \lambda)} \tag{2.62}$$

where $\hat{\Lambda}$ and λ, L are the operator associated with the specific symmetry, the target quantum number, and the volume of integral

space. This method has the advantage that all contaminated states are removed at one time. In addition, the GCM scheme is simply extended to the VAP scheme. So, the GCM-VAP method satisfies the variational principle and considers the orbital relaxation effect.

In the quantum chemistry, Scuseria et al. developed the GCM-projection scheme for broken symmetries wavefunction for the first time. They developed the number projection scheme for the HFB wave function as the number projected HFB (PHFB) method[99]. As a result, the parameter ζ in the HFB method is not necessary. In addition, the PHFB wavefunction is equivalent to the anti-symmetrized Geminal Power (AGP) wavefunction, which is one of the geminal theories. Scuseria et al. also extended the spin projection. They developed the Spin-Projected Unrestricted Hartree-Fock (PUHF) method[37]. In this thesis, I focus on the PUHF method. Scuseria *et al.* developed the Projected Hartree-Fock using the GCM-type projection operator[37]. They confirmed that the Jastrow type operator, the local projection, recovers the size-consistency[67, 34].

2.6.1 Spin-projected unrestricted Hartree-Fock method

In the GCM, the following projection operator is performed,

$$\hat{P}^{\lambda} = \frac{1}{L}\int_{L} d\phi e^{i\phi(\hat{\Lambda}-\lambda)} \tag{2.63}$$

where $\hat{\Lambda}$ is operator connected with symmetry that restore, and λ is desire quantum number. In the spin projection case, $\hat{\Lambda} = \hat{S^2}$, and

λ is the spin quantum number function $s(s+1)$. In the PUHF case, Φ_{UHF} is an eigenfunction of $\hat{S}_{\mathrm{z}}$. Therefore, the projection operator is represented as follows,

$$\hat{P}^{s}_{mm} = \frac{2s+1}{2}\int_{0}^{\pi} d\beta \sin\beta d^{s}_{mm}(\beta) e^{i\beta\hat{S}_{\mathrm{y}}} \tag{2.64}$$

where s, m are eigenvalues for $\hat{S}$ and $\hat{S}_{\mathrm{z}}$, respectively. $d^{s}_{mm}(\beta) = \langle j, m|\exp(\frac{-i\hat{S}_{\mathrm{y}}\beta}{\hbar})|j, m\rangle$ is wigner's small d function. The rotation matrix in ortho-AO basis is represented as follows,

$$\mathbf{R}(\beta, \hat{S}_{\mathrm{y}}) = \begin{pmatrix} \cos(\beta/2) & & O & & -\sin(\beta/2) & & O \\ & \ddots & & & & \ddots & \\ O & & \cos(\beta/2) & & O & & -\sin(\beta/2) \\ \sin(\beta/2) & & O & & \cos(\beta/2) & & O \\ & \ddots & & & & \ddots & \\ O & & \sin(\beta/2) & & O & & \cos(\beta/2) \end{pmatrix} \tag{2.65}$$

In the PUHF method, projecting the deformed states obtained by rotating the UHF wavefunction removes all the spin states that are contaminating. The PUHF energy is represented as follows,

$$E_{\mathrm{PUHF}} = \frac{\langle\Phi|\hat{P}^{\dagger}\hat{H}\hat{P}|\Phi\rangle}{\langle\Phi|\hat{P}^{\dagger}\hat{P}|\Phi\rangle} = \frac{\langle\Phi|\hat{H}\hat{P}|\Phi\rangle}{\langle\Phi|\hat{P}|\Phi\rangle} \tag{2.66}$$

where I made use of the property of the projection operator, the Hermicity ($\hat{P}^{\dagger} = \hat{P}$), and commutative with Hamiltonian ($[\hat{H}, \hat{P}] = 0$), and the idempotency ($\hat{P}^{2} = \hat{P}$) In practice, exponential term($e^{i\phi\hat{\Lambda}}$) corresponds to rotation, and $e^{i\phi\lambda}$ are weight factors.

$$E_{\mathrm{PUHF}} = \frac{\int d\Omega w(\Omega)\langle\Phi|\hat{H}\hat{R}(\Omega)|\Phi\rangle}{\int d\Omega w(\Omega)\langle\Phi|\hat{R}(\Omega)|\Phi\rangle} \tag{2.67}$$

The denominator of the energy is the determinant of overlap between normal UHF wavefunction and rotated UHF wavefunction $\mathbf{M}(\Omega)$.

$$\mathbf{M}(\Omega) = \langle \Phi | \hat{R}(\Omega) | \Phi \rangle \tag{2.68}$$

In practice, $\mathbf{M}(\Omega)$ are calculated from molecular orbital coefficients $\mathbf{C}$, and rotation matrix $\mathbf{R}(\Omega)$.

$$\mathbf{M}(\Omega) = \mathbf{C}^{\dagger}\mathbf{R}(\Omega)\mathbf{C} \tag{2.69}$$

The numerator of the energy is projected Hamiltonian matrices constructed from the transition density matrix.

$$\langle \Phi | HR(\Omega) | \Phi \rangle = \sum_{ik} h_{ik} (\mathbf{D}_{\Omega})_{ki} + \frac{1}{2} \sum_{ijkl} \langle ij||kl \rangle (\mathbf{D}_{\Omega})_{ki} (\mathbf{D}_{\Omega})_{lj} \tag{2.70}$$

where transition density matrices are represented as follows,

$$(\mathbf{D}_{\Omega})_{kl} = \frac{\langle \Phi | a_l^{\dagger} a_k \hat{R}(\Omega) | \Phi \rangle}{\langle \Phi | \hat{R}(\Omega) | \Phi \rangle}. \tag{2.71}$$

In practice, the transition density matrices are calculated from MO coefficients, rotation matrices, and overlap matrices $\mathbf{M}(\Omega)$

$$\mathbf{D}_{\Omega} = \mathbf{R}_{\Omega}\mathbf{C}\mathbf{M}_{\Omega}^{-1}\mathbf{C}^{\dagger} \tag{2.72}$$

I implemented the PUHF method using the natural orbital basis for simplicity. First, overlap matrices $\mathbf{M}(\Omega)$ are divided into occu-

pied and virtual blocks and rewritten density matrix forms.

$$\begin{aligned}\mathbf{M}(\Omega) &= \begin{pmatrix}\mathbf{C}_{\text{occ}}\\ \mathbf{C}_{\text{vir}}\end{pmatrix}^{\dagger} \mathbf{R}(\Omega) \begin{pmatrix}\mathbf{C}_{\text{occ}}\\ \mathbf{C}_{\text{vir}}\end{pmatrix} \\ &= (\mathbf{C}_{\text{occ}})^{-1}\mathbf{C}_{\text{occ}} \begin{pmatrix}\mathbf{C}_{\text{occ}}\\ \mathbf{C}_{\text{vir}}\end{pmatrix}^{\dagger} \mathbf{R}(\Omega) \begin{pmatrix}\mathbf{C}_{\text{occ}}\\ \mathbf{C}_{\text{vir}}\end{pmatrix} \mathbf{C}_{\text{occ}}^{\dagger}(\mathbf{C}_{\text{occ}}^{\dagger})^{-1} \\ &= (\mathbf{C}_{\text{occ}})^{-1}\mathbf{N}_{\Omega}^{-1}(\mathbf{C}_{\text{occ}}^{\dagger})^{-1}\end{aligned} \tag{2.73}$$

In implementation, I adopted $\mathbf{N}_{\Omega}$ as follows,

$$\mathbf{N}_{\Omega} = \left[\begin{pmatrix}\mathbf{D}_{oo} & \mathbf{D}_{ov}\end{pmatrix} \mathbf{R}_{\Omega} \begin{pmatrix}\mathbf{D}_{oo}\\ \mathbf{D}_{vo}\end{pmatrix}\right]^{-1}. \tag{2.74}$$

After all, the overlap between normal and rotated state is following,

$$\langle\Phi|\hat{R}(\Omega)|\Phi\rangle = \frac{1}{\det(\mathbf{N}_{\Omega}\mathbf{D}_{oo})}. \tag{2.75}$$

In addition, the VAP procedure is necessary effective Fock matrix, the derivative of energy to the density matrix.

$$\mathcal{F}_{kl} = \frac{\partial}{\partial D_{lk}} E[\mathbf{D}] \tag{2.76}$$

New MOs are obtained by diagonalization of $\mathcal{F}$.

$$\mathcal{F}\mathbf{C} = \mathbf{C}\epsilon \tag{2.77}$$

And a new effective Fock matrix is constructed. Note that $\mathcal{F}$ commutes with density matrix $\mathbf{D}$

$$[\mathcal{F}, \mathbf{D}] = 0. \tag{2.78}$$

So we can use direct inversion in the iterative subspace (DIIS).

The effective Fock matrix is represented as follows,

$$\mathcal{F}_\Omega = \frac{1}{2}\mathbf{Y}_\Omega \mathrm{Tr}[(\mathbf{h}+\mathbf{F}_\Omega)\mathbf{D}_\Omega] + \tilde{\mathcal{F}}_\Omega \tag{2.79}$$

where Y_Ω is calculated as follows,

$$\mathbf{Y}_\Omega = \mathcal{Y}_\Omega - \int d\Omega' y(\Omega')\mathcal{Y}_{\Omega'} \tag{2.80}$$

Furthermore, $\mathcal{Y}_\Omega$ is represented using the NO basis.

$$\mathcal{Y}_\Omega = \begin{pmatrix} \mathbf{2} & \mathbf{R}_{oo}^{-1}\mathbf{R}_{ov} \\ \mathbf{R}_{vo}\mathbf{R}_{oo}^{-1} & \mathbf{0} \end{pmatrix} \tag{2.81}$$

Dividing rotation matrix no NO basis into occupied and virtual blocks.

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{oo} & \mathbf{R}_{ov} \\ \mathbf{R}_{vo} & \mathbf{R}_{vv} \end{pmatrix} \tag{2.82}$$

Next, $\tilde{F}_\Omega$ is also divided into occupied and virtual blocks as follows,

$$\left(\tilde{\mathcal{F}}_\Omega\right)_{oo} = \left(\tilde{\mathcal{F}}_\Omega\right)_{vv} = \mathbf{0} \tag{2.83}$$

$$\left(\tilde{\mathcal{F}}_\Omega\right)_{ov} = \mathbf{R}_{oo}^{-1}\mathbf{F}_{ov}\left[\mathbf{R}_{vv} - \mathbf{R}_{vo}\mathbf{R}_{oo}^{-1}\mathbf{R}_{ov}\right] \tag{2.84}$$

$$\left(\tilde{\mathcal{F}}_\Omega\right)_{vo} = \mathbf{F}_{vo} - \mathbf{R}_{vo}\mathbf{R}_{oo}^{-1}\mathbf{F}_{oo} + \left[\mathbf{F}_{vv} - \mathbf{R}_{vo}\mathbf{R}_{oo}^{-1}\mathbf{F}_{ov}\right]\mathbf{R}_{vo}\mathbf{R}_{oo}^{-1} \tag{2.85}$$

Here, Brillouin condition Eq. 2.78 is satisfied only with occupied-virtual and virtual-occupied blocks. So occupied-occupied and virtual-virtual blocks are not defined. In the previous study [37], these blocks of normal Fock are defined as the effective Fock blocks.

$$\mathcal{F}_{pp} = \mathbf{F}_{pp} \tag{2.86}$$

$$\mathcal{F}_{qq} = \mathbf{F}_{qq} \tag{2.87}$$

Finally, the effective Fock is represented as follows,

$$\mathcal{F} \rightarrow \mathcal{F} + \rho\mathbf{F}\rho + (\mathbf{1} - \rho)\mathbf{F}(\mathbf{1} - \rho). \tag{2.88}$$

Figure 2.1 shows the PUHF scheme. The matrices at each rotation angle can be separately calculated, so it can make use of the parallel computation.

The PUHF is not size-consistent. For molecules A and B separate infinity, the energy is inconsistent with the total system calculation and the sum of the individual calculation ($E^{A+B}_{\mathrm{PUHF}} \neq E^{A}_{\mathrm{PUHF}} + E^{B}_{\mathrm{PUHF}}$). This originates from the fact that the electron correlation obtained from the projection is different. The previous study[67, 34] shows that the Jastraw type operator can reduce size-consistency error.

2.6.2 Reduced density matrix for PUHF wave function

The projected UHF wave function has multi-configurational character. The reduced density matrix (RDM) is useful for analyzing the PUHF wave function[90]. The one-particle RDM (1-RDM) is represented as follows,

$$\gamma_{ij} = \frac{\langle\Phi|\hat{P}^{\dagger}\hat{a}^{\dagger}_{j}\hat{a}_{i}\hat{P}|\Phi\rangle}{\langle\Phi|\hat{P}^{\dagger}\hat{P}|\Phi\rangle} \tag{2.89}$$

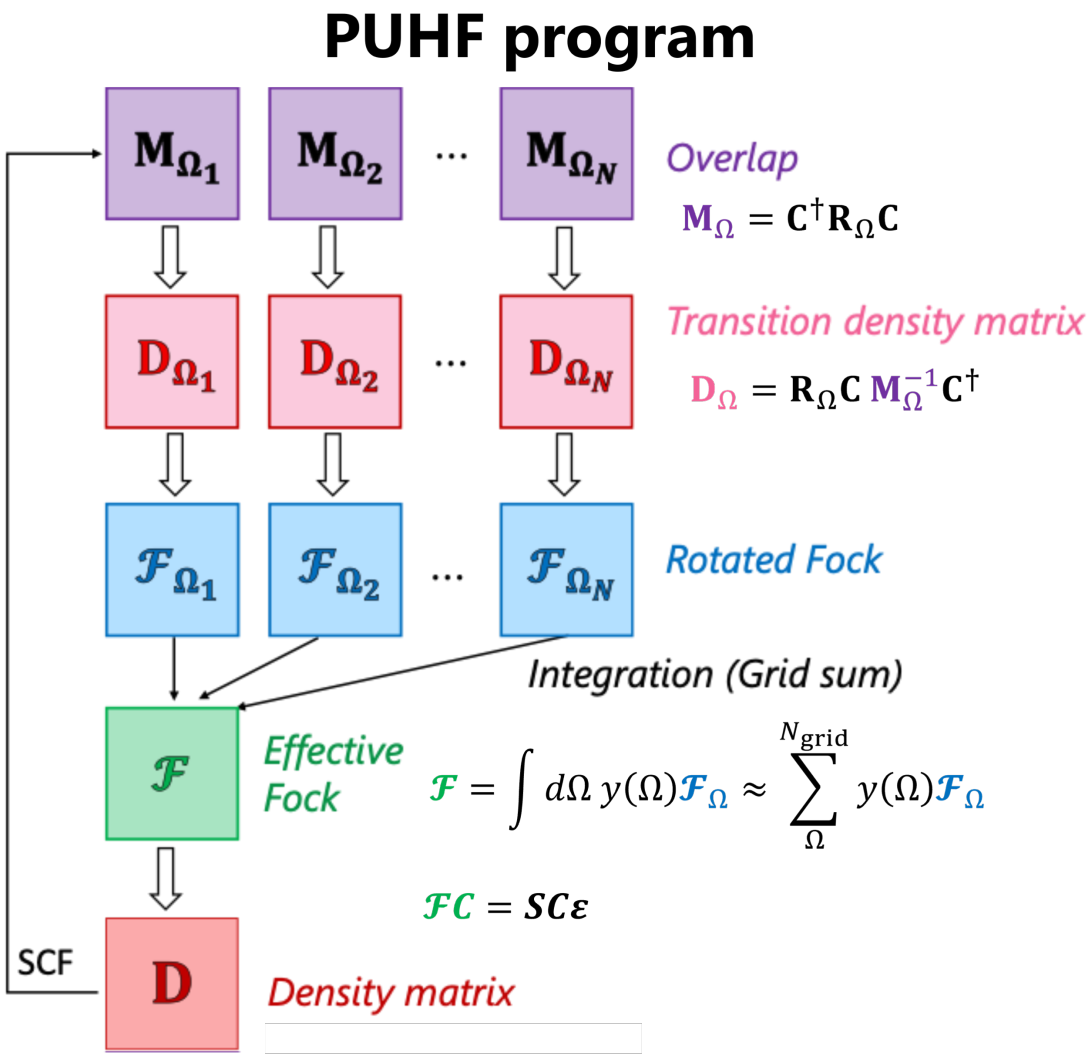

Figure 2.1: The PUHF scheme

If T_q^k is rank-k tensor, it makes simple form,

$$\begin{aligned}&\hat{P}^s_{mm} T^k_q \hat{P}^s_{mm} = \delta_{q,0} \langle smk0|sm\rangle \\ &\sum_\mu (-1)^\mu \langle s, m, k, -\mu | s, m-\mu\rangle T^k_\mu \hat{P}^s_{m-\mu,m}\end{aligned} \tag{2.90}$$

where $\langle j_1, m_1, j_2, m_2 | j, m\rangle$ are Clebsh-Gordan coefficients. Natural orbitals (NOs) are obtained by diagonalizing 1-RDM. PUHF's NOs are consistent with UHF ones. However, the occupation number is fractional[32]. In singlet case,

$$\gamma^\alpha = \gamma^\beta = \int d\Omega y(\Omega) \frac{1}{2} \left(\mathbf{D}^{\alpha\alpha}_\Omega + \mathbf{D}^{\beta\beta}_\Omega \right) \tag{2.91}$$

If the operator $\hat{O}$ commutes with projection operator $\hat{P}$, the expectation value is simply calculated as follows,

$$\langle \hat{O} \rangle = \frac{\langle \Phi | \hat{P}\hat{O}\hat{P} | \Phi \rangle}{\langle \Phi | \hat{P} | \Phi \rangle} = \frac{\langle \Phi | \hat{O}\hat{P} | \Phi \rangle}{\langle \Phi | \hat{P} | \Phi \rangle}. \tag{2.92}$$

For example, the evaluation of $\langle \hat{S}^2 \rangle$ is represented as follows[99],

$$\langle \hat{S}^2 \rangle = \int d\Omega y(\Omega) \sum_i \left([\mathrm{Tr}\mathbf{M}_i(\Omega)]^2 + \frac{1}{2}\mathrm{Tr}[\mathbf{M}_i^2(\Omega)] \right) + \frac{3}{2}\mathrm{Tr}[\mathbf{P}(\Omega) - \mathbf{P}^2(\Omega)] \tag{2.93}$$

where $\mathbf{P}(\Omega)$, $\mathbf{M}_x(\Omega)$, $\mathbf{M}_y(\Omega)$, and $\mathbf{M}_z(\Omega)$ are transition density matrix as follows,

$$\mathbf{P}(\Omega) = \frac{1}{2} \left(\mathbf{D}^{\alpha\alpha}_\Omega + \mathbf{D}^{\beta\beta}_\Omega \right) \tag{2.94}$$

$$\mathbf{M}_z(\Omega) = \frac{1}{2} \left(\mathbf{D}^{\alpha\alpha}_\Omega - \mathbf{D}^{\beta\beta}_\Omega \right) \tag{2.95}$$

$$\mathbf{M}_x(\Omega) = \frac{1}{2} \left(\mathbf{D}^{\alpha\beta}_\Omega + \mathbf{D}^{\beta\alpha}_\Omega \right) \tag{2.96}$$

$$\mathbf{M}_y(\Omega) = \frac{1}{2}\left(\mathbf{D}_{\Omega}^{\beta\alpha} - \mathbf{D}_{\Omega}^{\alpha\beta}\right) \tag{2.97}$$

2.6.3 The extension to the projection for post-HF method

The PUHF can describe the static correlation. However, for a quantitative description, capturing dynamic correlation is vital. Thus, the GCM-projection scheme is applied to various post-HF methods, for example, the configuration interaction[118], the coupled cluster[83], the perturbation theory[119], the excited states[120], and quantum computing[115].

2.7 Linear-scaling technique for large-scale calculation

2.7.1 Nearsightedness in chemistry

In quantum chemistry, the HF method and DFT, the fundamental methods, scale $O(N^3)$. For reducing scaling costs, linear scaling schemes have been developed. Linear scaling schemes based on the "Nearsightedness" principle proposed by Kohn[51, 82]. Consider the static properties $F(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e})$ of N_e electron system in motion under a certain potential $v(\mathbf{r})$. Suppose that all electrons exist within a certain distance λ and that their center of gravity is as follows,

$$\bar{\mathbf{r}} \equiv \frac{1}{N_e}\sum_{n=1}^{N_e} \mathbf{r}_n. \tag{2.98}$$

Under a constant chemical potential, no matter how large the change in the external potential $\delta v(\mathbf{r}')$ is, if the change occurs

farther away than a certain distance $|\mathbf{r}' - \bar{\mathbf{r}}| \gg \lambda$, the effect on F is negligible. This means that changes in the external potential that occur very far away are 'invisible'. This is the 'nearsightedness' principle. It states that changes in potential at a distance greater than a certain distance have negligible effect. This principle is a general that applies not only to insulators but also to metals. This corresponds to the fact that the real-space density matrix decays with distance and is negligible over long distances. The behavior of the decay differs for metals and insulators. In the case of insulators, the attenuation is exponential,

$$\rho(\mathbf{r}, \mathbf{r}') \approx \exp(-K|\mathbf{r} - \mathbf{r}'|) \tag{2.99}$$

On the other hand, for metals, at a certain temperature T

$$\rho(\mathbf{r}, \mathbf{r}') \approx k_F \frac{\cos(k_F|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|^2} \exp(-c\frac{k_B T}{k_F}|\mathbf{r} - \mathbf{r}'|) \tag{2.100}$$

where k_F is Fermi wavenumber, and k_B is the Boltzmann constant. At finite temperatures, it decays exponentially.

2.7.2 Fragment molecular orbital method

The Fragment Molecular Orbital (FMO) method is one of the fragment-based large-scale calculation methods developed by Kitaura and Fedrov[40]. In the FMO method, the whole molecule is divided into fragments, and each fragment's energy estimates the energy of the whole molecule. The energy of a fragment is calculated under the electrostatic fields produced by other fragments.

If the sum of the fragment monomer approximates the total energy, it is called the FMO1 method. Assigning electron number and spin multiplicity to each fragment in the FMO method is necessary. Thus, it is impossible to treat charge transfer between fragments. To solve this problem, the dimer of fragments is considered in the FMO2 method.

$$E_{\mathrm{FMO2}} = \sum_{I} E_I + \sum_{I<J} (E_{IJ} - E_I - E_J) \tag{2.101}$$

Furthermore, if trimer and tetramer are considered, these methods are called FMO3 and FMO4, respectively. The feature of the FMO method is the analysis of interaction energy between fragments. This analysis, called Pair Interaction Energy (PIE) can be decomposed into the following energy components; electrostatic (ES), exchange (EX), charge-transfer(CT), and dispersion (DI). This Pair Interaction Energy Decomposition Analysis (PIEDA) is a very powerful tool for bio-molecule computation. The FMO method and the PIEDA analysis can reveal the interaction between amino acid residues or protein-ligands.

$$\Delta\tilde{E}_{IJ} = \Delta\tilde{E}_{IJ}^{ES} + \Delta\tilde{E}_{IJ}^{EX} + \Delta\tilde{E}_{IJ}^{CT+mix} + \Delta\tilde{E}_{IJ}^{DI} \tag{2.102}$$

2.7.3 Density matrix minimization method

In addition to fragment division type linear scaling methods, other methods have been developed based on minimizing the density matrix and avoiding direct diagonalization of the full Hamiltonian mat-

rix[55, 58, 72]. In this method, the ground potential Ω is minimized

$$\Omega = 2\mathrm{Tr}[\mathbf{D}(\mathbf{F} - \mu)] \tag{2.103}$$

where μ is the chemical potential. The following $\tilde{\mathbf{D}}$ using the McWeeny transformation are adopted to satisfy the power equality of the density matrix.

$$\tilde{\mathbf{D}} = 3\mathbf{DSD} - 2\mathbf{DSDSD} \tag{2.104}$$

After all, $\tilde{\Omega}$ is minimized.

$$\tilde{\Omega} = 2\mathrm{Tr}[3\mathbf{DSD} - 2\mathbf{DSDSD}(\mathbf{F} - \mu)] \tag{2.105}$$

2.7.4 Divide-and-conquer method

The divide-and-conquer (DC) method is the one of the fragment type linear-scaling method. Yang et al. developed the DC method and applied for pure DFT calculation[127]. The DC method is also applied for wave-function theory. For example, HF[2], MP2[44], CC[47]. In addition, the open-shell calculation[49], the gradient calculation[131, 45, 46], the excited state calculation[113, 128], In the DC scheme, the basis sets belonging to the subsystem, represented as $\mathcal{S}(\alpha)$, are exclusive with other basis sets belonging to other subsystems. i.e. $\mathcal{S}(\alpha) \cap \mathcal{S}(\beta) = 0$. The buffer region, having overlaps with other center regions, is set to improve the accuracy. The basis set belonging to the buffer region is denoted as $\mathcal{B}(\alpha)$. Combining $\mathcal{S}(\alpha)$ with $\mathcal{B}(\alpha)$ constructs the localization region as

follows,

$$\mathcal{L}(\alpha) = \mathcal{S}(\alpha) \cup \mathcal{B}(\alpha). \tag{2.106}$$

In the HF case, the local Fock matrices are constructed as following,

$$F^{\alpha}_{\mu\nu} = h_{\mu\nu} + \frac{1}{2}\sum_{\lambda\sigma} \langle \mu\sigma || \nu\lambda \rangle D^{\mathrm{DC}}_{\lambda\sigma} \qquad (\mu, \nu \in \mathcal{L}(\alpha)), \tag{2.107}$$

where $h_{\mu\nu}$ and $D^{\mathrm{DC}}_{\mu\nu}$ is one-electron integral and the approximated total density matrix by the DC scheme. The local eigenvalue equations are solved as follows,

$$\mathbf{F}^{\alpha}\mathbf{C}^{\alpha}_{i} = \varepsilon^{\alpha}_{i}\mathbf{S}^{\alpha}\mathbf{C}^{\alpha}_{i}. \tag{2.108}$$

where $\mathbf{C}_i$, and ε_i are local molecular orbital coefficients, and orbital energy, respectively. The local density matrix is calculated from the local MO coefficients,

$$D^{\alpha}_{\mu\nu} = 2p^{\alpha}_{\mu\nu}\sum_{i=1} f_{\beta}(\varepsilon_F - \varepsilon^{\alpha}_{i}) C^{\alpha}_{\mu\nu} C^{\alpha *}_{\mu\nu}. \tag{2.109}$$

ε_F, f_β are fermi level and fermi function. $p^{\alpha}_{\mu\nu}$ is a partition matrix to avoid overlap between localization regions as follows,

$$p^{\alpha}_{\mu\nu} = \begin{cases} 1 & \text{if } \mu \in \mathcal{S}(\alpha) \text{ and } \nu \in \mathcal{S}(\alpha), \\ 1/2 & \text{if } (\mu \in \mathcal{S}(\alpha) \text{ and } \nu \in \mathcal{B}(\alpha)) \text{ or vice versa}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.110}$$

The sum of the partition matrix is one.

$$\sum_{\alpha} p^{\alpha}_{\mu\nu} = 1 \tag{2.111}$$

The total density matrix is constructed from the local density matrices and partition matrix,

$$D_{\mu\nu}^{\mathrm{DC}} = \sum_{\alpha} p_{\mu\nu}^{\alpha} D_{\mu\nu}^{\alpha}. \tag{2.112}$$

The Fermi level is determined so that the number of electrons is conserved in the entire system,

$$N_{\mathrm{e}} = \sum_{\alpha} \sum_{\mu} (\mathbf{D}^{\alpha}\mathbf{S}^{\alpha})_{\mu\mu}. \tag{2.113}$$

In the DC method, spin multiplicity and the number of electrons to each subsystem are determined automatically because the fermi level in each subsystem is unified. It indicates that the DC scheme is suited for molecular dynamics simulation, which accompanies the structural changes. As for its accuracy, the DC method can control its error originating from faragmentation by setting appropriate buffer regions. Although appropriate buffer regions depend on the target system, we have developed the auto-determination for appropriate buffer regions[42, 23, 43].

2.8 Semi-empirical quantum chemical calculation

Although the *ab initio* method is accurate, its cost is high and its application has been limited to small molecules. Thus, semi-empirical molecular orbital methods have been developed and used for large molecules. In exchange for its low accuracy, its computational speed is hundreds of times faster than *ab initio* one. Many semi-empirical methods have the following features.

1. Deal only with some electrons, not all electrons. For example, valence electrons.

2. Avoid calculating the electron-repulsion integrals explicitly.

3. Restricting the number of integrals in the Fock matrix. For example, neglect the overlap between different basis functions at the point $\mathbf{r}_1$ (zero differential overlap: ZDO).

$$\chi_\mu^*(\mathbf{r_1})\chi_\nu(\mathbf{r_1}) = 0 \qquad (\mu \neq \nu) \tag{2.114}$$

2.8.1 Hückel method

The Hückel method is the first semi-empirical method in quantum chemistry[36]. In Hückel method, only π electrons are treated, and the off-diagonal elements of the Hamiltonian matrix are neglected for those not directly bonded. In addition, the overlap matrix is ignored and the orthogonality of the basis is assumed. Despite this bold approximation, the Hüchel method succeeded in qualitatively describing the π-conjugated system.

In 1963, Hoffmann developed the extended Hückel theory (EHT)[35]. In the EHT, the overlap matrix is calculated explicitly. In addition, Fock matrix is approximated by the ionization potential and overlap matrix.

$$F_{\mu\mu} = -I_\mu \tag{2.115}$$

$$F_{\mu\nu} = -\frac{1}{2}K(I_\mu + I_\nu)S_{\mu\nu} \tag{2.116}$$

where I_μ is ionization potential of atomic orbital μ, $S_{\mu\nu}$ is the element of overlap matrix. K is a parameter. Adopting ionization potential is based on Koopmans' theorem. The parameter K is determined to reproduce the experimental value. For considering neighboring environments, Mulliken charges are determined in self-consistently manner,

$$\Delta Q_A = Z_A - \rho_A \tag{2.117}$$

$$F_{\mu\mu} = -I_\mu + \omega Q_A \tag{2.118}$$

where Q_A is effective charge on atom A and Z_A is original atomic nuclear charge, and ρ_A is Mulliken charge on atom A, ω is a parameter.

2.8.2 PPP method

Pariser-Parr-Pople (PPP) method was proposed in 1953[74, 79]. The PPP method is the first semi-empirical method to consider electron repulsion explicitly. In the PPP method, only π electrons are considered and the ZDO approximation is assumed. The electron repulsion is approximated as follows,

$$(\mu\nu|\lambda\sigma) = \delta_{\mu\nu}\delta_{\lambda\sigma}(\mu\mu|\lambda\lambda) = \delta_{\mu\nu}\delta_{\lambda\sigma}\gamma_{\mu\lambda} \tag{2.119}$$

$$\gamma_{\mu\lambda} = (\mu\mu|\lambda\lambda) \tag{2.120}$$

The diagonal term of γ is represented using ionization potential and electron affinity as follows,

$$\gamma_{\mu\mu} = (\mu\mu|\mu\mu) = I_\mu - A_\mu \tag{2.121}$$

where I_μ and A_μ are the ionization potential and electron affinity of atomic orbital μ. These values are adopted from experimental values. Several approximation schemes exist in the off-diagonal term of γ. I will show the following two schemes,

$$\gamma_{\mu\nu} = \frac{1}{R_{\mu\nu} + a} \qquad (Nishimoto - Magata) \tag{2.122}$$

$$\gamma_{\mu\nu} = \frac{1}{\sqrt{R_{\mu\nu}^2 + a^2}} \qquad (Ohno - Klopmans) \tag{2.123}$$

where

$$a = \frac{2}{\gamma_{\mu\mu} + \gamma_{\nu\nu}} \tag{2.124}$$

1-electron Hamiltonian is approximated as follows,

$$H_{\mu\mu} = -I_\mu - \sum_{\nu \neq \mu} n_\nu \gamma_{\mu\nu} \tag{2.125}$$

$$H_{\mu\nu} = b \exp(-cR_{\mu\nu}) \tag{2.126}$$

where n_ν is a occupation number, and b, c is parameter fitted for experimental value.

2.8.3 Density-functional tight-binding method

The density-functional tight-binding (DFTB) method[81, 101] is a semi-empirical method based on DFT. In the DFTB, a minimal

basis function is adopted, neglecting the three- and four-center integrals. The zeroth-order energy is the repulsion between shells and is equivalent to the repulsion between bare nuclei.

$$E_{\mathrm{DFTB0}} = \frac{1}{2} \sum_{A \neq B} E_{\mathrm{core}}^{AB} \tag{2.127}$$

The repulsion between shells is parameterized as the spline function and fitted to reproduce the reference DFT results. The 1-electron and 2-electron term is parametrized as follows,

$$F_{\mu\nu} = \begin{cases} \varepsilon_\mu & \mu = \nu \\ \langle \mu_A | -\frac{1}{2}\nabla^2 + \hat{V}_A + \hat{V}_B | \nu_B \rangle & A \neq B \\ 0 & otherwise \end{cases} \tag{2.128}$$

where ε_μ corresponds to the orbital energy of the free atom. The DFTB1 energy is represented as follows,

$$E_{\mathrm{DFTB1}} = \sum_{i=1}^{N} n_i \varepsilon_i + E_{\mathrm{DFTB0}} \tag{2.129}$$

where n_i is the occupation number. However, the DFTB1 energy cannot take into account the fluctuation of the electron distribution. Thus, the DFTB2/3 methods consider the fluctuation of the electron distribution for reference charge.

$$F_{\mu\nu} = F_{\mu\nu}^0 + \frac{1}{2} S_{\mu\nu} \sum_C Q_C (\gamma_{\mathrm{AC}} + \gamma_{\mathrm{BC}}) \qquad \mu \in A, \nu \in B \tag{2.130}$$

where Q_C is the Mulliken charge and γfunction is represented as

follows,

$$\gamma_{\mathrm{AB}} = \left(R_{AB}^2 + \frac{1}{4}(\mathcal{U}_A^{-1} + \mathcal{U}_B^{-1}) \right)^{-\frac{1}{2}} \tag{2.131}$$

γ function decays as $1/r$ in the long-range limit. In the short range, it approaches chemical hardness $\mathcal{U} = 2\eta = IP - EA$ asymptotically.

The DFTB2 energy is represented as follows,

$$E_{\mathrm{DFTB2}} = E_{\mathrm{DFTB1}} + \frac{1}{2} \sum_{A \neq B} \Delta Q_A \Delta Q_B \gamma_{AB} \tag{2.132}$$

where ΔQ is the charge fluctuation from the reference charge. Tne DFTB3 energy is represented as follows,

$$E_{\mathrm{DFTB3}} = E_{\mathrm{DFTB2}} + \frac{1}{3} \sum_{A,B} Q_A^2 Q_B \left. \frac{\partial \gamma_{AB}}{\partial Q_A} \right|_{Q_A=0} \tag{2.133}$$

In these methods, the energy is determined by a self-consistent charge.

The DFTB method is several hundred times faster than DFT. However, DFTB parameters are prepared for each element pair. Therefore, the number of elemental species that can be handled is small, and there is a problem with versatility.

2.8.4 The extended Tight-Binding method

The DFTB adopt the elecment-pair parameter. Thus, the determination of these parameter is difficult and its application have been limited only 10-30 elements. Grimme *et al.* have developed the extended Tight-Binding (xTB) method as general-purpose semi-empirical method[28, 7, 6]. The xTB method adopt element-specific

parameter at the expense of the target limitation. The xTB parameter are determined for geometry, frequencies, and non-covalent interaction. The xTB specializes these GFN property , not energy. So far, 86 elements are available in the GFN-xTB method. There are several version of the GFN-xTB. The GFN1/2-xTB are explicitly considered charge fluctuation though self-consistently-charge (SCC). In the GFN0-xTB, the diagonalization of Hamiltonian matrix is performed only once. In addition, the GFN-FF has been developed for Force-Field calculation. The GFN1-xTB method adopts monopole approximated charge and minimal basis except for hydrogen, which is Double-Zeta basis for description of hydrogen bond. In addition, the halogen bond correction term is incorporated and the Grimme's D3 dispersion correction method is adopted in the GFN1-xTB method. In the GFN2-xTB method, the multipole approximated charge is adopted, which enables to use minimal basis for all elements. In addition, Grimme's D4 dispersion correction method is adopted, which depends on the charge.

Chapter 3

Cb-TDHFB method: the excited state calculation including static correlation in black-box manner

3.1 Preface

The multi-configurational method such as the CASSCF is popular in the accurate excited state including the static correlation. However, there are some drawbacks that requirement for setting the active space properly and its high computational cost. Thus, black-box and low-cost methods for the excited calculation including the static correlation are desired. I applied the HFB method, the ground state calculation method with a black-box treatment of the static correlation and low cost, to the excited state calculation. I adopted the time-dependent (TD) method, obtaining the information of the excited states from the response against the external fields, as the excited states calculation method. In this study, the TD equation is solved by the real-time (RT) method, tracking the

change of the electronic states over time. However, the original HFB method has the defect of the electron number failure, which is worse than the ground state. Thus, I introduced the "canonical-basis" approximation form of the TDHFB method for preserving the average electron number. However, the TDHFB wave function remains not an eigenfunction of the number operator.

An accurate evaluation of excitation energies for molecules is challenging because it requires the description of both ground and excited states having different correlation energies with similar accuracy. If the ground state of the target system has a large static correlation energy (e.g., bond-dissociated molecules and transition-metal compounds), a multi-configurational wave function is required for the qualitative description. The most popular method for such a system is the active space-based self-consistent field (SCF) method, which is typified by state-averaged complete active space SCF (SA-CASSCF)[92, 91]. However, the scope of its application is limited because its computational cost increases with the size of the active space. In addition, users must set an appropriate active space to maintain accurate results at affordable computational costs. The density matrix renormalization group has been developed[73, 126, 5] and has enabled the treatment of a large active space; however, black-box treatment remains difficult[106]. Another approach is geminal-based wave function theories[108, 109]. In this method, the N-electron wave function is represented as the antisymmetric product of two-electron wave functions called gem-

inals, rather than the one-electron wave functions called orbitals. This method can cover a large static correlation energy because each geminal can fully describe the intra-geminal electron correlation. In addition, the computational scaling of practical geminal-based methods is $O(N^{3-5})$[111, 99]. Thus, this method can efficiently construct the N-electron wave function by incorporating static electron correlation with a lower computational cost than the active-space scheme. Several variants of the geminal-based methods have been proposed both for ground[84, 87, 85, 86, 38] and excited[75, 8, 77, 31] state calculations. Here, I focus on the Hartree–Fock–Bogoliubov (HFB) method[89, 88]. The HFB method originates from the fusion of the Hartree–Fock (HF) theory and Bardeen–Cooper–Schrieffer theory[98] in superconductivity. It describes the static correlation by considering quasiparticles, which are formally related to geminals. Number-projection associates the HFB theory with the geminal-based wave function, which is called antisymmetrized geminal power[13, 56, 62]. In the HFB scheme, both the electron and quasiparticle describe the electronic distribution and Hamiltonian. The dimension of the Hamiltonian that must be diagonalized is only twice that of HF. Thus, the HFB method enables the treatment of the static correlation with mean-field computational cost. The HFB method has been highly successful in nuclear physics and superconductivity. In 2002, Staroverov and Scuseria applied this to molecular systems[105]. They found that the HFB energy functional was identical to the corrected HF density mat-

rix functional proposed by Csányi and Arias[14], except for the sign of the correlation term. It indicates that the HFB method is also related to the density matrix functional theory (DMFT)[78]. They demonstrated that the HFB wave function can qualitatively describe strongly correlated systems. However, the HFB wave function breaks the electron number conservation, which can be avoided by introducing a chemical potential in the sense that the expectation value of the electron number can be conserved. Although several electron correlation theories based on the HFB concept have been developed and applied to molecular systems[99, 116, 100, 117, 114, 19], their extension to the excited state remains a frontier[75, 8], whereas the time-dependent version of related DMFT has been proposed[24, 25]. The excited-state calculation based on the HFB method, which is the time-dependent HFB (TDHFB) method[89], was also formulated in superconductivity and nuclear dynamics. In this study, I formulate the TDHFB method for molecular systems as a simple method for calculating excited states that can consider static electron correlations. The two methods for solving the TD equation are time- and frequency-domains. The frequency-domain method, based on the Casida equation[10], has been popular in quantum chemistry because the excitation energy and associated transition moment are the most significant properties. Meanwhile, the time-domain method, called the real-time (RT) method[124, 123, 4, 3, 26], is popular in nuclear physics because it enables observation of the time evolution of a state. In addition, its compu-

tational cost is formally lower than that of the frequency-domain method[125]. Almost all previous studies in nuclear physics have solved the TDHFB equation using the RT method. Therefore, I applied the RT method and developed RT-TDHFB to obtain excited-state information. The RT-TDHFB method has the disadvantage that the HFB wave function is not necessarily an eigenfunction of the number operator; therefore, the electron number can change during time evolution. To avoid this change, I focus on applying the canonical basis (Cb)-TDHFB method developed in nuclear physics to nuclear dynamics calculations[18, 17, 16]. This method introduces a TD canonical basis that diagonalizes the density matrix at any time and approximates the pairing field matrix in diagonal form. Accordingly, the expectation value of the electron number was conserved.

3.2 Real-Time Time-Dependent method

3.2.1 RT-TDHFB method

The TD HF/KS equation for a given TD potential for an external field, $V^{\mathrm{ext}}(r,t)$, is expressed as the following:

$$i\frac{\partial}{\partial t}\psi_m(r,t) = \left(\hat{F}(t) + \hat{V}^{\mathrm{ext}}(r,t)\right)\psi_m(r,t) \tag{3.1}$$

where $\hat{F}(t)$ is the TD Fock operator and $\hat{V}^{\mathrm{ext}}$ represents the interaction between the electron and the external field. This equation can be reformulated using the TD density matrix $\boldsymbol{D}(t)$:

$$i\frac{\partial}{\partial t}\boldsymbol{D}(t) = \left[\boldsymbol{F}(t) + \boldsymbol{V}^{\mathrm{ext}}(t), \boldsymbol{D}(t)\right]. \tag{3.2}$$

In the RT method, this equation was solved by directly integrating the time variable. Here, I applied the RT method to solve the TDHFB equation[21], which can be expressed using the density matrix form as

$$i\frac{\partial}{\partial t}\bar{\boldsymbol{R}}(t) = \left[\boldsymbol{H}'(t), \bar{\boldsymbol{R}}(t)\right]. \tag{3.3}$$

$$\boldsymbol{H}'(t) = \boldsymbol{H}(t) + \boldsymbol{V}^{\mathrm{ext}}(t) \tag{3.4}$$

where $\bar{\boldsymbol{R}}(t)$ and $\boldsymbol{H}(t)$ are the generalized density matrix and quasiparticle Hamiltonian, respectively, at time t. The TD generalized density matrix $\bar{\boldsymbol{R}}(t)$ comprises static and induced parts in terms of the external field.

$$\bar{\boldsymbol{R}}(t) = \bar{\boldsymbol{R}}^{(0)} + \delta\bar{\boldsymbol{R}}(t) = \bar{\boldsymbol{R}}^{(0)} + \sum_{n=1}^{\infty}\delta\bar{\boldsymbol{R}}^{(n)}(t) \tag{3.5}$$

Similarly, the quasiparticle Hamiltonian is expressed as the sum of the static part corresponding to $\bar{\boldsymbol{R}}^{(0)}$ and the parts induced by $\delta\bar{\boldsymbol{R}}(t)$ as

$$\boldsymbol{H}(t) = \boldsymbol{H}^{(0)} + \delta\boldsymbol{H}(t) = \boldsymbol{H}^{(0)} + \sum_{n=1}^{\infty}\delta\boldsymbol{H}^{(n)}(t) \tag{3.6}$$

In this study, I focus on the linear-response case:

$$i\frac{\partial}{\partial t}\delta\bar{\boldsymbol{R}}^{(1)}(t) = \left[\boldsymbol{H}^{(0)}, \delta\bar{\boldsymbol{R}}^{(1)}(t)\right] + \left[\delta\boldsymbol{H}^{(1)}(t), \bar{\boldsymbol{R}}^{(0)}\right] + \left[\boldsymbol{V}^{\mathrm{ext}}(t), \bar{\boldsymbol{R}}^{(0)}\right] \tag{3.7}$$

Note that the nonlinear equation is also soluble using the RT-TDHFB method.

3.3 Canonical-basis TDHFB method

Because the ground-state HFB wave function is not the eigenfunction of the number operator, the RT-TDHFB method suffers from the fluctuation of the electron number during time evolution. Ebata *el al.* [44] introduced a TD canonical basis into the RT-TDHFB theory in nuclear physics. In the resulting Cb-TDHFB method, TD canonical orbitals are obtained each time by diagonalizing the TD density matrix $\mathbf{D}(t)$. This method can avoid the fluctuation of the electron number during time evolution. Note that this definition of the "canonical orbital" does not mean the orbitals that diagonalize the effective Hamiltonian such as canonical HF orbitals, but is similar to the "natural orbitals" in quantum chemical terminology. The TD canonical orbital is defined by

$$i\frac{\partial}{\partial t}\psi'_m(\mathbf{r},t) = \left(\hat{F}(t) + \hat{V}^{\mathrm{ext}}(\mathbf{r},t)\right)\psi'_m(\mathbf{r},t). \tag{3.8}$$

where $\psi'_m(r,t)$ is the TD canonical molecular orbital and is represented by the MO expansion coefficient $\boldsymbol{C}(t)$

$$\psi'_m(\mathbf{r},t) = \sum_j C_{mj}(t)\chi_j(\mathbf{r}) \tag{3.9}$$

This coefficients matrix diagonalizes the density matrix:

$$\boldsymbol{C}^\dagger(t)\boldsymbol{D}(t)\boldsymbol{C}(t) = \begin{pmatrix} n_1(t) & \cdot & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdot & n_k(t) \end{pmatrix} \tag{3.10}$$

The time evolution of the TD canonical orbitals is given by

$$i\frac{\partial}{\partial t}\boldsymbol{C}(t) = \boldsymbol{F}(t)\boldsymbol{C}(t) - \boldsymbol{C}(t)\boldsymbol{\eta}(t) \tag{3.11}$$

where $\boldsymbol{\eta}(t)$ is the gauge that makes the density matrix invariant to the orbital transformation and $\boldsymbol{\eta}(t)$ is the diagonal matrix with diagonal elements of the Fock matrix:

$$\eta_k(t) = \langle \phi_k | \hat{F}(t) | \phi_k \rangle \tag{3.12}$$

I rewrite the TDHFB equation in the form of a time evolution equation for the density and pair matrices:

$$i\frac{\partial}{\partial t}\bar{\boldsymbol{D}}(t) = [\boldsymbol{F}(t), \bar{\boldsymbol{D}}(t)] + \bar{\boldsymbol{K}}(t)\boldsymbol{\Delta}^*(t) - \Delta(t)\bar{\boldsymbol{K}}^*(t), \tag{3.13}$$

$$i\frac{\partial}{\partial t}\bar{\boldsymbol{K}}(t) = \boldsymbol{F}(t)\bar{\boldsymbol{K}}(t) + \bar{\boldsymbol{K}}(t)\boldsymbol{F}(t)^* + \boldsymbol{\Delta}^*(t)[\mathbf{1} - \bar{\boldsymbol{D}}^*(t)] - \bar{\boldsymbol{D}}(t)\boldsymbol{\Delta}(t). \tag{3.14}$$

Note that the pairing matrix at $t = 0$ is also diagonal according to Eq. (12). In addition, by approximating the pair potential to be diagonal:

$$\Delta_k(t)\delta_{kl} = \Delta_{kl}(t) \tag{3.15}$$

The equations for the diagonal elements of the density and pair matrices were derived from Eqs. (33) and (34).

$$i\frac{\partial}{\partial t}\bar{D}_k(t) = \bar{K}_k(t)\Delta_k^*(t) - \bar{K}_k^*(t)\Delta_k(t), \tag{3.16}$$

$$i\frac{\partial}{\partial t}\bar{K}_k(t) = 2\eta_k(t)\bar{K}_k(t) + \Delta_k(t)[2\bar{D}_k(t) - 1] \tag{3.17}$$

Eq. (36) is the time evolution equation for the occupation number of time-dependent natural orbitals. In the usual Cb-TDHF equation, the number of occupancies does not change with time

evolution and is either two or zero (occupied or unoccupied). In the TD-DMFT, it is known that this time variation of the occupation number is necessary to describe correctly states in which two-electron excitations are predominant[25]. Since the right-hand side of Eq. (36) is the product of the diagonal components, the sum over k of the first and second terms is equal to the trace of the product of the pairing matrix and the pairing field matrix, respectively.

$$\sum_k \bar{K}_k(t)\Delta_k^*(t) = \mathrm{Tr}\left(\boldsymbol{K}\boldsymbol{\Delta}^*\right) \tag{3.18}$$

$$\sum_k \bar{K}_k^*(t)\Delta_k(t) = \mathrm{Tr}\left(\boldsymbol{K}^*\boldsymbol{\Delta}\right) \tag{3.19}$$

Assuming that the pairing energy remains constant during time evolution, the sum over k on the right-hand side of Eq. (36) is zero.

$$\mathrm{Tr}\left(\boldsymbol{K}(t)\boldsymbol{\Delta}^*(t)\right) = \mathrm{Tr}\left(\boldsymbol{K}^*(t)\boldsymbol{\Delta}(t)\right) \tag{3.20}$$

Hence, the trace of the density matrix, i.e., the number of electrons, is also constant during time evolution. In addition to Eqs. (36) and (37), the time evolution equation of the canonical orbital in Eq. (31) is solved by numerical integration.

3.4 Implementation

I implemented RT-TDHFB and Cb-TDHFB codes in the GAMESS program[97]. The following is the procedure of the Cb-TDHFB method[68]:

1. The HFB wave function in the ground state is calculated to obtain the static generalized density and Hamiltonian.

2. The canonical basis set is constructed by diagonalizing the density matrix. The pairing matrix was transformed using a canonical basis. Off-diagonal elements of the pairing-field matrix is discarded.

3. The new canonical basis function, diagonal density, and pairing matrices under an external field are updated using Eqs. (31), (36), and (37). I adopted the fourth-order Runge–Kutta method for numerical integration. The potential of the applied external field is given by

$$V_{ij}^{\mathrm{ext}}(t) = \varepsilon(t)\boldsymbol{e}^{\mathrm{ext}} \cdot \langle \phi_i | \boldsymbol{r} | \phi_j \rangle \,, \tag{3.21}$$

   where $\boldsymbol{e}^{\mathrm{ext}}$ is a unit vector of the external field, $\langle \phi_i | \boldsymbol{r} | \phi_j \rangle$ is the dipole integral, and $\varepsilon(t)$ represents the TD part of the external field. In this study, I adopted a delta-function-type external field:

$$\varepsilon(t) = \delta(t) \tag{3.22}$$

4. The induced polarization of the electrons at each time step is calculated:

$$\delta\boldsymbol{P}(t) = -\sum_{ij} \delta\bar{D}_{ij}(t) \langle \phi_i | \boldsymbol{r} | \phi_j \rangle \tag{3.23}$$

5. The absorption spectrum is obtained from the time variation of the induced polarization through Fourier transform (FT) analysis.

Note that only states with non-zero dipole transition moments are obtained as excited states since this study employed a dipole-type operator.

3.5 Numerical Application

In this section, I present the numerical results obtained using the TDHFB method. I applied the RT/Cb-TDHFB method to H_2 molecules. In the TDHFB calculation, I set ζ, which is a parameter of the HFB method, to 1.0, in all cases. In the time evolution, I adopted the direction of the external fields, $\boldsymbol{e}^{\mathrm{ext}} = (1, 1, 1)$ in all calculations. All time evolutions were run over $T = 8.0$ fs with a time interval of $\Delta t = 0.48$ as (0.02 a.u. of time). Therefore, the resolution of the absorption spectrum ΔE is 0.52 eV from the uncertainty principle $\Delta E \sim h/T$, where h is the Planck constant and T is the simulation time. Throughout this study, I adopted the correlation-consistent polarized valence double-ζ (cc-pVTZ) basis set unless otherwise stated.

3.5.1 Ground State

Figure 3.1 shows the potential energy curves for the H_2 molecule. Around the equilibrium bond length where the static correlation is sufficiently small, the HFB energy coincides with the RHF energy. In contrast, the HFB results qualitatively describe the bond dissociation curve. However, because the HFB method overestimates the static correlation, the HFB energy becomes lower than the full CI

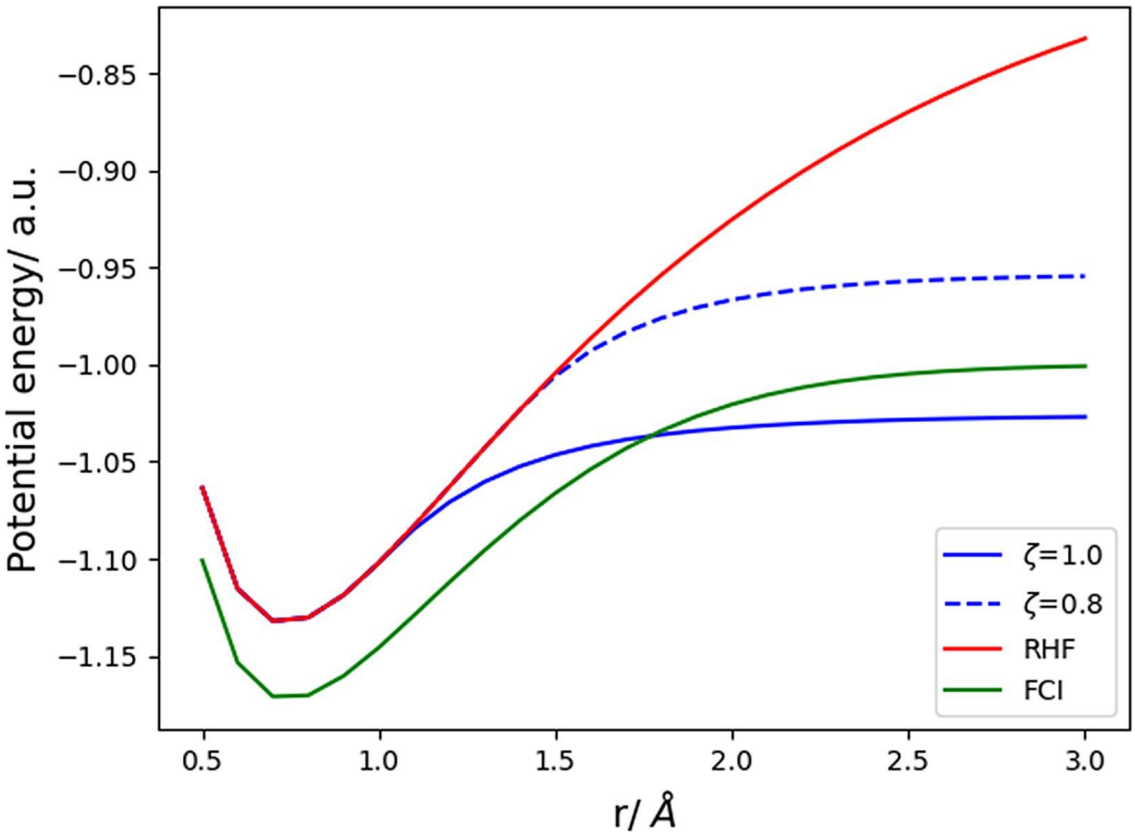

Figure 3.1: The potential energy curves for the H_2 molecule.

energy. If the parameter ζ was adopted as 0.8, the HFB potential curve was almost parallel to the full CI potential curve.

3.5.2 RT-TDHFB calculation

First, I performed the RT-TDHFB calculation of the H_2 molecule at the equilibrium bond length ($r = 0.75$ Å), where the static correlation is sufficiently small. Figure 3.2 (upper) shows the time evolution of the induced polarization $\delta\boldsymbol{P}^{(1)}(t)$ and the absorption spectrum obtained from the FT analysis of $\delta\boldsymbol{P}^{(1)}(t)$. For comparison, the results of the RT-TDHF calculations are shown in Figure 3.2 (lower). In this simple weak-correlation case, the RT-TDHFB res-

ults coincided with the RT-TDHF results. As a stretched bond ($r = 2.6$ Å), the results changed, as shown in Figure 3.3. Note that the RT-TDHF results are qualitatively inadequate because it is a spin-restricted formalism. The oscillation of the induced polarization by the RT-TDHFB calculation is approximately one order of magnitude larger than that by the RT-TDHF calculation. Accordingly, many unphysical peaks appear in the absorption spectrum, especially in the high-energy regions. Figure 3.4 (a) shows the fluctuation of the expectation value of the electron number during the time evolution in the RT-TDHFB calculation compared with the RT-TDHF calculation. Figure 3.4 (b) shows the FT analysis of the electron number fluctuation. Strong peaks are observed in the high-energy regions, which correspond to the absorption spectrum. Therefore, the change in the electron number affects the induced polarization in the RT-TDHFB calculation. Furthermore, dependence of the fluctuation on the time step used in the time evolution (Δt) is investigated in Figure 3.5.

3.5.3 Canonical-basis TDHFB calculation

I performed the Cb-TDHFB calculations for the H_2 molecule. In the equilibrium bond length ($r = 0.75$ Å), the Cb-TDHFB results depicted in Figure 3.6coincide with the RT-TDHF results as well as the RT-TDHFB ones. As the H-H bond is stretched ($r = 2.6$ Å), the results changed, as shown in Figure 3.8. As for the electron number, the Cb-TDHFB method could suppress the change in the

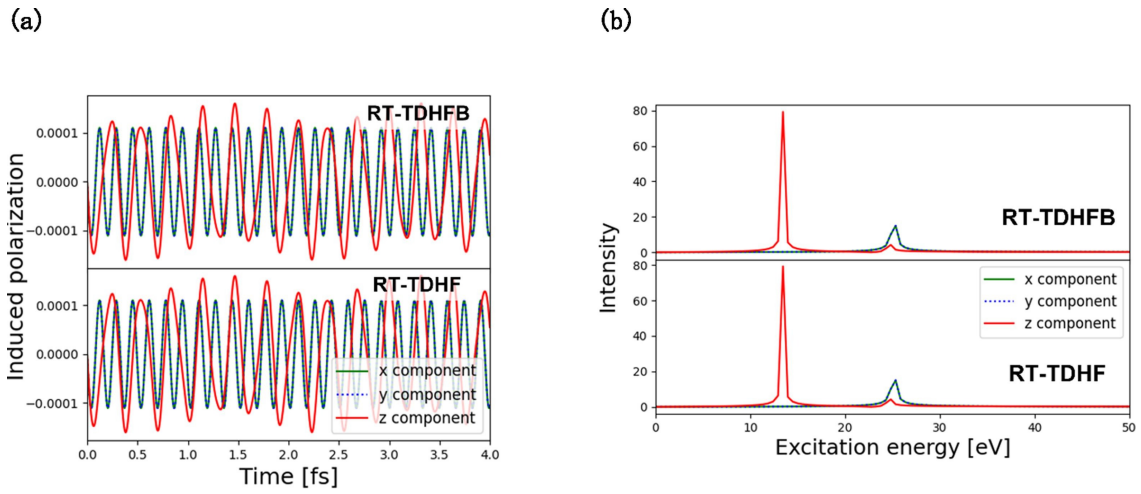

Figure 3.2: (a) Time variation of the induced polarization in H_2 ($r = 0.75$ Å) and (b) absorption spectrum with (Upper) RT-TDHFB and (Lower) RT-TDHF methods.

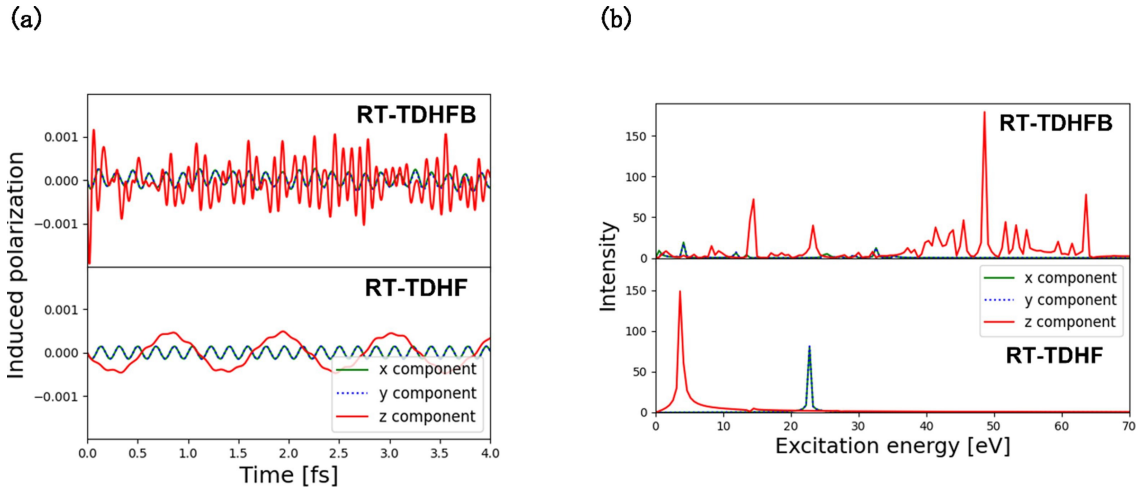

Figure 3.3: (a) Time variation of the induced polarization in H_2 ($r = 2.6$ Å) and (b) absorption spectrum with (Upper) RT-TDHFB and (Lower) RT-TDHF methods.

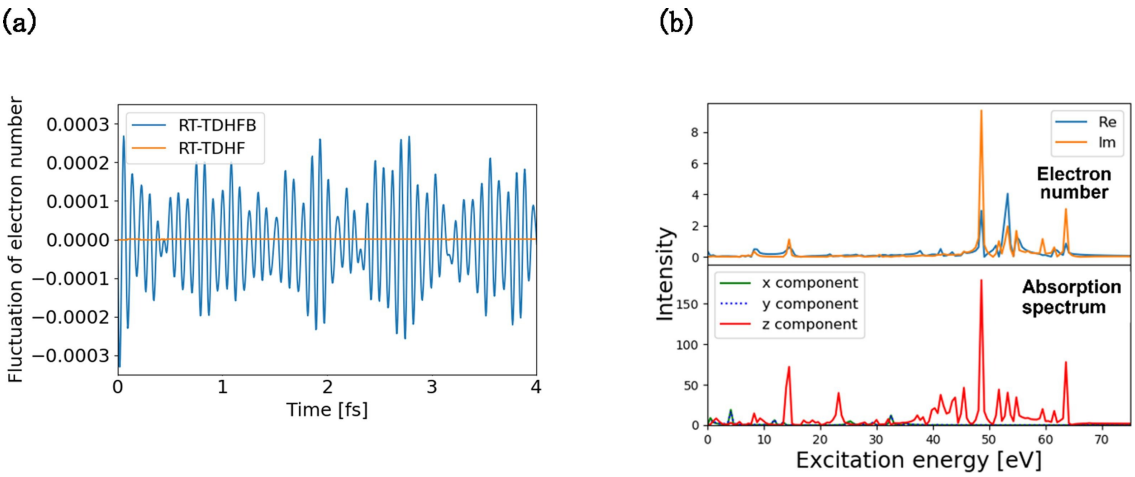

Figure 3.4: (a) Fluctuation of the electron number during the time evolution in the RT-TDHFB or RT-TDHF calculation of H_2 (r = 2.6 Å) and (b) FT analysis of (Upper) the electron number fluctuation and (Lower) absorption spectrum in the RT-TDHFB calculation of H_2 (r = 2.6 Å).

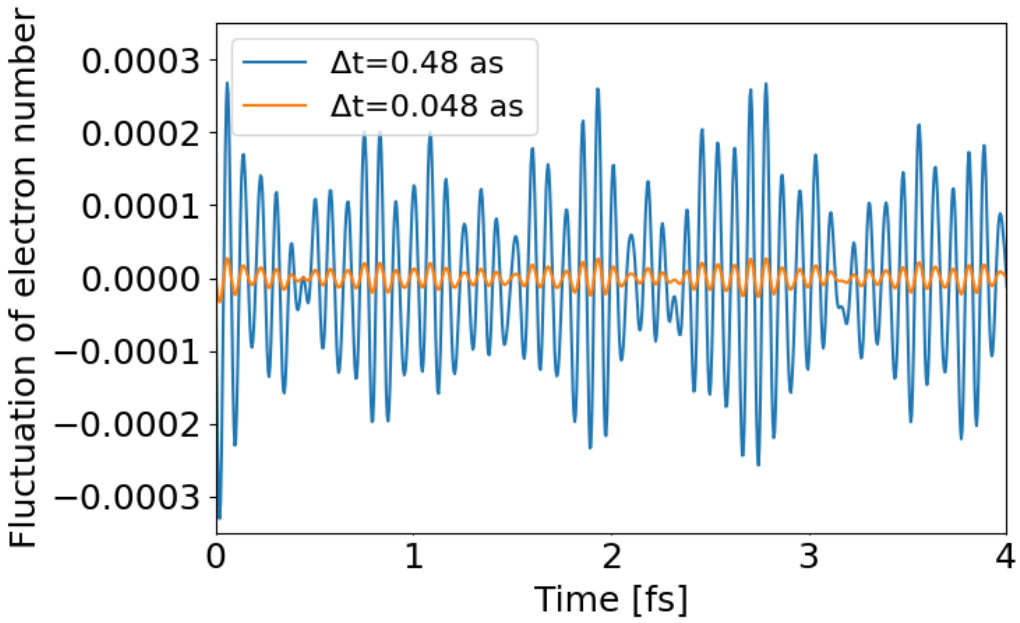

Figure 3.5: The time step dependency of the fluctuation of electron number in the RT-TDHFB calculation of H_2 (r=2.6 Å).

electron number during the time evolution (Figure 3.7). This indicates that the polarization and absorption spectra in the Cb-TDHFB calculation can eliminate the effect of the electron number change. Table 3.1 compares the excitation energies of the Cb-TDHFB, RT-TDHF, and full CI calculations for the H_2 molecule (r = 2.6 Å). The Σ_u^+ states appear as excitations in the z direction in the Cb-TDHFB calculation. In the low energy region ($\leqslant$10 eV), the Cb-TDHFB results are not consistent with those of full CI. However, at high energy region (>15 eV), the Cb-TDHFB energies accurately reproduce the full-CI ones. The Π_u state appears as an excitation in the x and y directions, and the excitation energy and oscillator strength by the Cb-TDHFB method also correspond to those of full CI. The other states, which are not dipole-allowed, do not appear in excited state calculations using real-time methods. I also performed Cb-TDHFB calculations for an LiH molecule. I found as well as for the H_2 molecule, i.e., the high-energy region is described correctly, whereas the low-energy region is described poorly (Figure 3.9). The TDHFB method equips with a time evolution equations not only for the natural orbitals but also for the occupation numbers, suggesting a connection with TD-DMFT. Previous studies have shown that TD-DMFT can correctly describe two-electron excited states missing in the TD-DFT[35,36]. In a hydrogen molecule, this state is a totally symmetric representation. However, this state does not appear in the Cb-TDHFB calculation because it does not have a transition moment. I consider that these states can be treated if

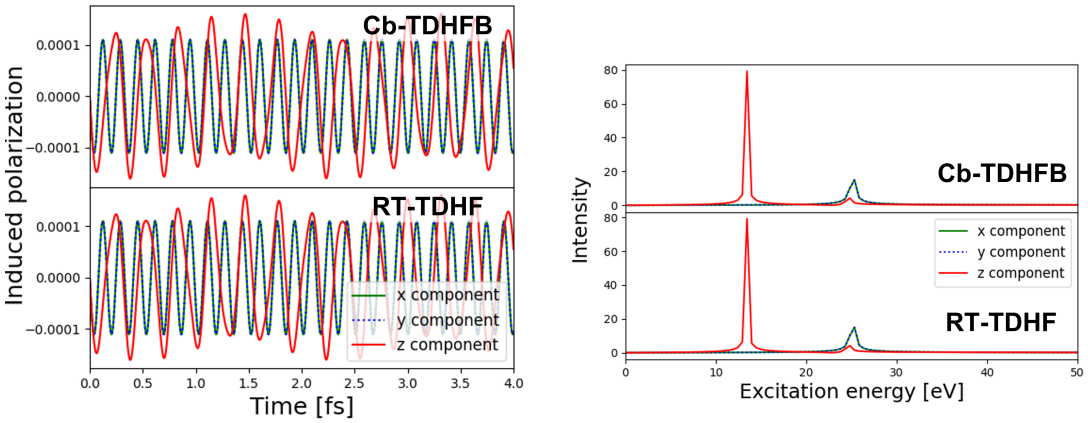

Figure 3.6: (left) Time variation of $\delta\mathbf{P}^{(1)}(t)$ and (right) absorption spectrum of H_2 (r=0.75 Å). (Upper) Cb-TDHFB calculation. (Lower) RT-TDHF calculation.

I formulate the Cb-TDHFB method in the Casida-like equation as well as the TD-DMFT, but the present Cb-TDHFB calculation includes doubly excited configurations in the ground state, and it can obtain one-electron excited states from such configurations.

3.5.4 Hartree–Fock orbital analysis

In this study, I analyzed the absorption spectrum using canonical HF orbitals. First, the TDHFB density matrix is projected onto a static canonical HF orbital basis. Note that the canonical HF orbitals were constructed separately from the HFB calculations for analysis. Although the occupation numbers of the canonical HF orbitals are zero or two for the RHF density matrix, these canonical orbitals are fractionally occupied for the HFB density matrix. Therefore, excitation from the orbital, which is virtual in the RHF, but partially occupied in the HFB, can occur. The time variation of the density matrix in the canonical HF orbital basis provides a

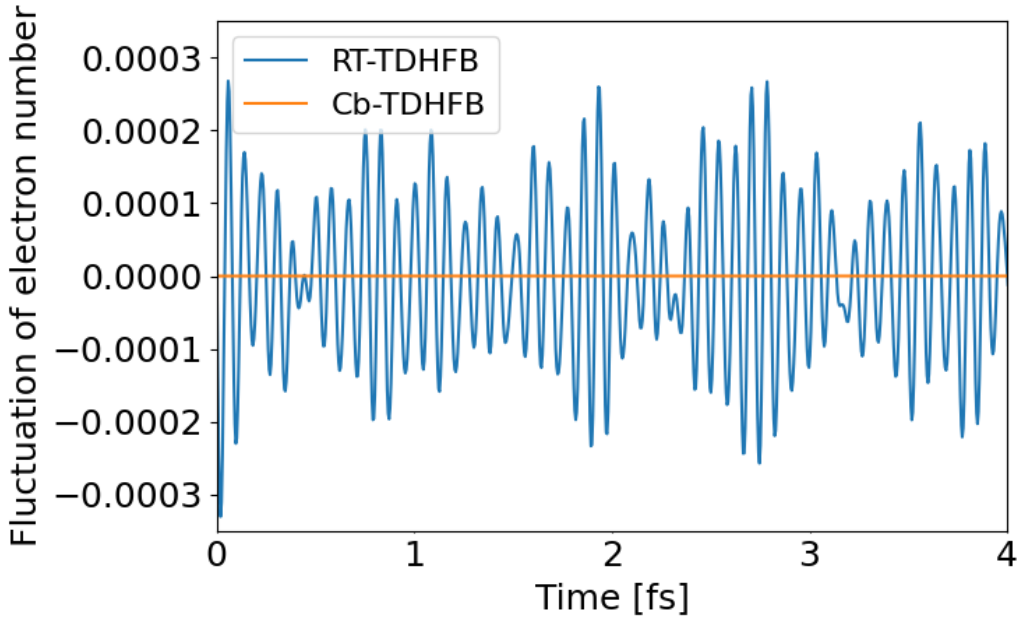

Figure 3.7: Fluctuation of the electron number during the time evolution of H_2 (r=2.6 Å). (Blue) RT-TDHFB calculation. (Orange) Cb-TDHFB calculation.

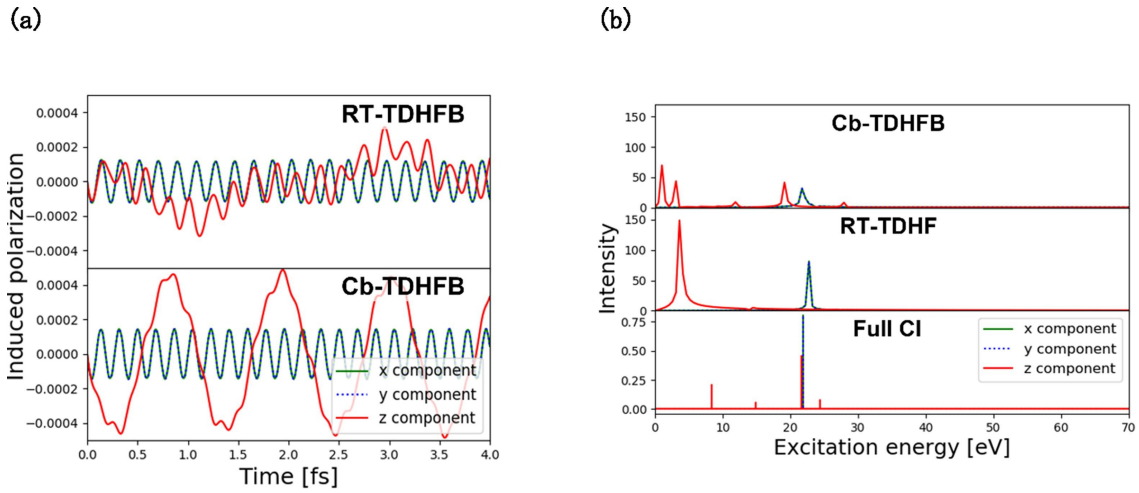

Figure 3.8: (a) Time evolution of the induced polarization of H_2 (r = 2.6 Å) with (Upper) Cb-TDHFB and (Lower) RT-TDHF methods. (b) Absorption spectrum with (Upper) Cb-TDHFB, (middle) RT-TDHF, and (Lower) full CI methods.

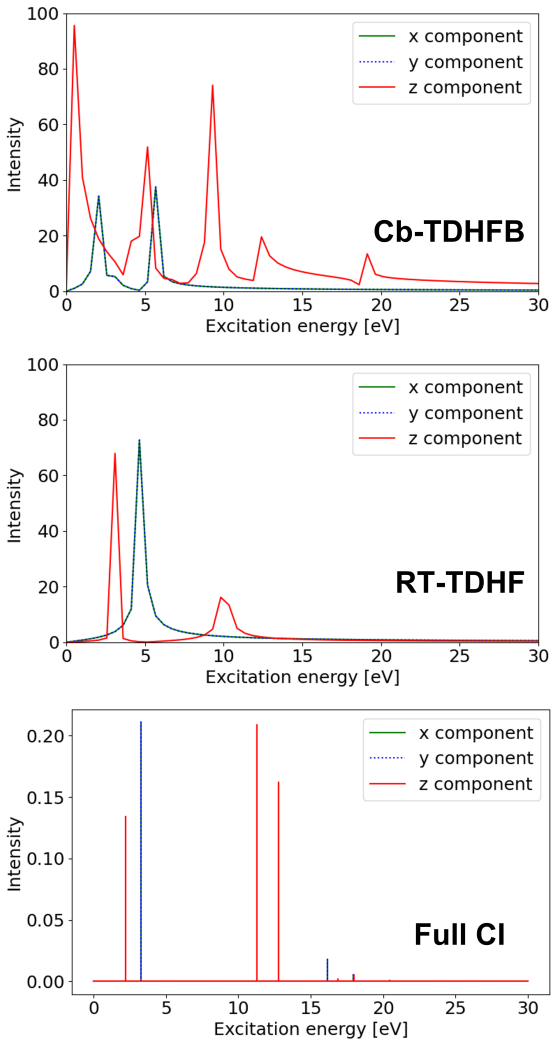

Figure 3.9: Absorption spectra of LiH molecule (r=2.5 Å) using STO-3G. (Upper) Cb-TDHFB, (middle) RT-TDHF, and (Lower) full CI calculation.

transition based on the HF canonical orbital. For example, matrix elements $D_{ij}(t)$ correspond to the TD interaction between canonical HF orbitals i and j. FT analysis of the time variation $D_{ij}(t)$ gives us a spectrum

$$D_{ij}(\omega) = \int_{\infty}^{\infty} D_{ij}(t) e^{-i\omega t} dt. \tag{3.24}$$

Then, the spectrum was compared with the absorption spectrum to assign the nature of the transition based on the HF canonical orbitals. I applied this analysis to the results of the Cb-TDHFB calculation of the stretched H_2 molecule with $r = 2.6$ Å. Figure 3.10 (b) shows the spectrum of the canonical HF orbital analysis of Eq. (44). For simplicity, I used the 6-31G basis set. Figure 3.10 (a) shows the canonical HF orbitals and their occupation numbers in the ground-state HFB wave function. The spectrum of the RT-TDHF results is shown in Figure 3.10 (c). For parity symmetry, D_{13} and D_{24} were zero. The Cb-TDHFB result has non-zero components that are not relevant to MO1, which do not appear in the RT-TDHF result where only MO1 is occupied. In the first and second excited states (i.e., 1.6 and 2.6 eV), the transition of MO1 and MO2 (D_{12}) is dominant, but that between MO1 and MO4 (D_{14}) also exists. A similar contribution was observed for the first excited state (4.1 eV) of the RT-TDHF sample. Note that, in the case of the Cb-TDHFB result, the D_{12} contribution may include the electron transition from MO2 to MO1, as well as that from MO1 to MO2. In addition, transitions between MO2 and MO3 (D_{23}) and between MO3 and MO4 (D_{34}) exist in the Cb-TDHFB results. It

is because that the ground state wave function of the HFB method has a multi-configurational character. For comparison, the full CI wave functions and excitation energies are summarized in Table 3.2. The first excitation energy by the full CI method is 10.05 eV, indicating that the energy by the Cb-TDHFB method is estimated to be lower. S1← S0 excitation in the full CI method can be approximated as transitions from MO1 to MO2 ($\Phi_{11} \rightarrow \Phi_{22}/\Phi_{21}$), MO2 to MO1 ($\Phi_{22} \rightarrow \Phi_{12}/\Phi_{21}$), MO1 to MO4 ($\Phi_{11} \rightarrow \Phi_{14}/\Phi_{41}$), and MO2 to MO3 ($\Phi_{22} \rightarrow \Phi_{23}/\Phi_{32}$). Therefore, the transition analysis for the Cb-TDHFB results was consistent with the full-CI result. In the third excited state (27.4 eV), the transition of D_{23} mainly contributed to the spectrum, although the D_{14} and D_{34} transitions also contributed to this state. However, only the D_{14} transition contributed to the second state (30.5 eV) in the RT-TDHF result. In the full CI result, this state corresponds to S4← S0 excitation (26.38 eV) and can be approximated as transitions from MO2 to MO3 ($\Phi_{22} \rightarrow \Phi_{23}/\Phi_{32}$) and from MO1 to MO4 ($\Phi_{11} \rightarrow \Phi_{14}/\Phi_{41}$). This result is consistent with the Cb-TDHFB results, indicating that the Cb-TDHFB method can qualitatively describe the static correlation in the excited state. However, the accuracy of low-energy excited states is poor. This is probably because these peaks originate from the breakdown of number symmetry, called the Nambu–Goldstone mode in nuclear physics[49,50].

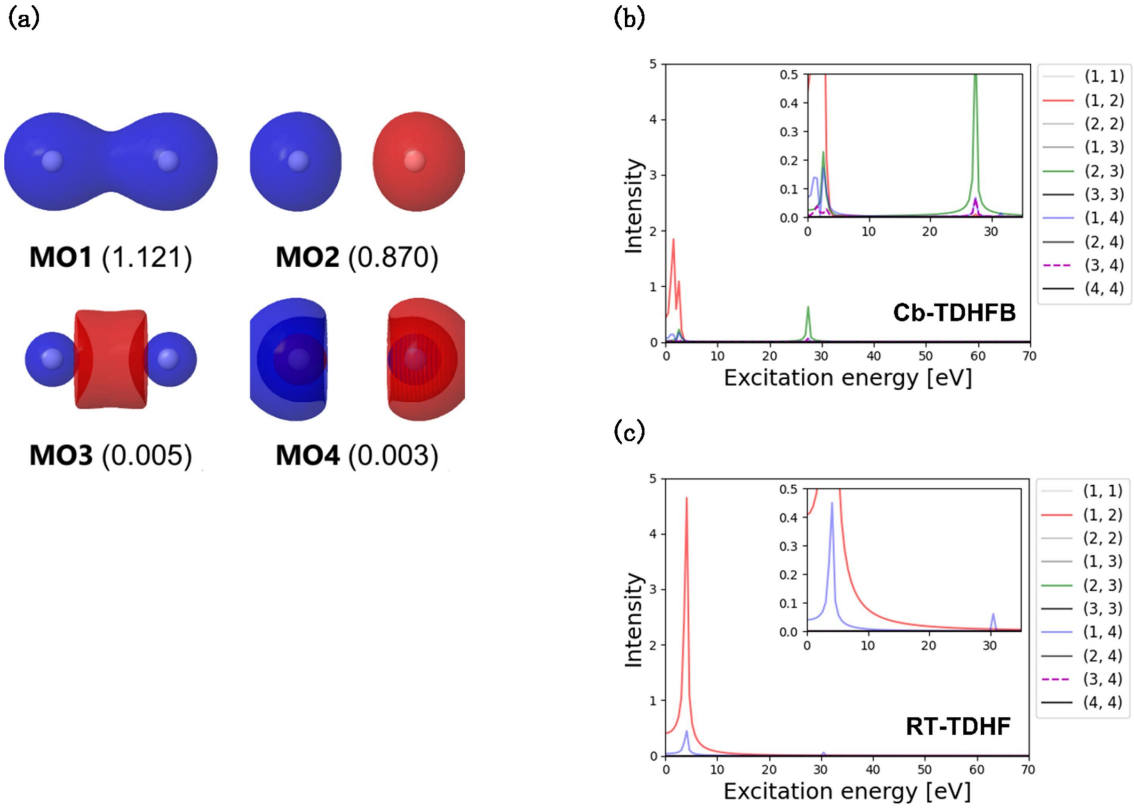

Figure 3.10: HF orbital analysis of the spectrum in H_2 ($r = 2.6$ Å). (a) Canonical HF orbital and occupation number in the parentheses. (b) Cb-TDHFB. (c) RT-TDHF.

3.5.5 Conclusion

In this study, the TDHFB method was applied to molecular systems to treat excited states with large static correlations at a reasonable computational cost. I first developed the RT-TDHFB method by combining the real-time propagation method with the HFB wave function, which can describe static correlation. Around the equilibrium structure of the H_2 molecule, where the static correlation is weak, the RT-TDHFB result is identical to the RT-TDHF result because the HFB ground state coincides with the HF ground state. In the structure with stretched bond, the RT-TDHFB result indicate many unphysical peaks in its absorption spectrum that originate from violation of the total electron number. To avoid this, I adopted the Cb-TDHFB, which introduces the TD canonical basis and the diagonal approximation of the pairing field matrix to preserve the electron number. The Cb-TDHFB method successfully suppressed the fluctuation of the electron number and provided an absorption spectrum that was free from electron number changes. Projection onto the canonical HF orbitals provides the origin of the absorption peaks. I confirmed that the Cb-TDHFB method can qualitatively treat the ground and excited states with static electron correlation. However, the Cb-TDHFB results underestimated the excitation energies and generated spurious peaks around the low-energy region. Although the Cb-TDHFB method conserves the average electron number, it is not an eigenfunction of the number operator. To recover the number symmetry, I will attempt the

projection scheme[99] in the future.

Table 3.1: Excitation energies $\hbar\omega$ of H_2 molecule ($r = 2.6$ Å) obtained by the RT-TDHF, Cb-TDHFB, and full CI calculation using the cc-pVTZ basis.

Direction	**RT-TDHF (eV)**	**Cb-TDHFB (eV)**	**Full CI (eV)**
z		1.00	
z	3.60	3.10	$8.42/\Sigma_u^+$
z	14.50	11.90	$14.91/\Sigma_u^+$
z		19.10	$21.57/\Sigma_u^+$
x, y	22.70	21.70	$21.93/\Pi_u$
z	27.90	27.90	$24.38/\Sigma_u^+$

Table 3.2: The full CI calculation in H_2 molecule with $r = 2.6$ Åusing the 6-31G basis.

State	**Excitation energy (eV)**	**Wave function**
S_0	0.0	$0.787\Phi_{11} - 0.606\Phi_{22}$
S_1	10.05	$0.702\Phi_{12} + 0.702\Phi_{21}$
S_2	10.27	$0.780\Phi_{22} + 0.612\Phi_{11}$
S_3	25.91	$0.557\Phi_{31} + 0.557\Phi_{13} + 0.416\Phi_{42} + 0.416\Phi_{24}$
S_4	26.38	$0.515\Phi_{32} + 0.515\Phi_{23} + 0.482\Phi_{41} + 0.482\Phi_{14}$

Chapter 4

DC-PUHF: large-scale calculation method including static correlation in black-box manner

4.1 Preface

Accurate calculations, including electron correlation, require a high computational cost. Attempts have been made to develop a large-scale electron correlation calculation method based on linear scaling methods [129]. The dynamic correlation is the repulsion effect between electrons in the proximity region, so it is the local effect. Thus, many large-scale calculation methods based on the "nearsightedness" principle [51, 82] have correctly described this type of correlation [22, 39, 20, 44, 47]. On the other hand, large-scale calculations including static correlation are challenging because it is long-range effects in some cases. In 2005, Fedrov and Kitaura developed the FMO-MCSCF combining the FMO method with the MCSCF method [21]. However, this method has the draw-

back that only one fragment can be treated with MCSCF and others are treated with RHF. Thus, it is necessary to understand which fragment has the largest static correlation in advance. In addition, the method cannot properly handle systems with large static electron correlations in multiple fragments. Kobayashi et al. developed the DC-HFB method [48] combining the DC method [64] with the HFB method [105] for treating all fragments as multiconfigurational. In the DC-HFB method, the HFB equations are solved within the localization region the same as in the normal DC method. They utilized the fact that the fermi level in the DC method and the chemical potential in the HFB method are physically identical. This method enables the large-scale static correlation calculation in a black-box manner. However, the contamination problem for the number of electrons in the original HFB method remains. I focused on the projection scheme to recover the symmetry and applied the projection scheme to the large-scale calculation using the broken symmetry wave function [99, 37]. In practice, I developed the DC-PUHF method applying the projection scheme to the DC-UHF method. In the PUHF calculation, there are three bottlenecks. The first is the diagonalization of the effective Fock matrix. The second is the calculation of the inverse matrix in the construction of the transition density matrix. The last is the determinant calculation of the overlap matrices. All three bottlenecks have $O(N^3)$ costs. This study has the significance of revealing the relationship between the projection and the

wave function approximated by the local wave functions.

4.2 Accelaration scheme for diagonalization of effective Fock matrix based on the DC method

4.2.1 Theory

In the DC method, the total Fock matrix is constructed from the approximate total density matrix.

$$F_{\mu\nu} = H_{\mu\nu}^{\mathrm{core}} + \sum_{\lambda\sigma}^{N_{\mathrm{basis}}} D_{\lambda\sigma}^{\mathrm{DC}} \langle \mu\sigma || \nu\lambda \rangle \tag{4.1}$$

The approximated total density matrix is constructed from local density matrices as follows,

$$D_{\mu\nu}^{\mathrm{DC}} = \sum_{\alpha}^{N_{\mathrm{subsystem}}} p_{\mu\nu}^{\alpha} D_{\mu\nu}^{\alpha}. \tag{4.2}$$

where $p_{\mu\nu}^{\alpha}$ is the element of the partition matrix. In this study, this density matrix is applied to the projection scheme. The natural orbital (NO) basis is adopted in the projection scheme. Thus, diagonalizing the α, β spin density matrices,

$$\mathbf{D}_{\mathrm{NO}}^{\sigma} = \mathbf{C}^{\sigma\dagger} \mathbf{D}^{\mathrm{DC}(\sigma)} \mathbf{C}^{\sigma} \qquad (\sigma = \alpha/\beta) \tag{4.3}$$

where $\mathbf{C}^{\sigma}$ is the spin-natural orbital coefficient matrix. Using the DC-UHF wave function, the occupation number is fractional, unlike the normal UHF method. In this study, orbitals whose occupation number is greater than 0.1 are regarded as occupied, and smaller ones are considered unoccupied. In addition, the density matrix is

constructed for each spin and occupied-unoccupied block.

$$\mathbf{D}_{\mathrm{NO}} = \begin{pmatrix} \mathbf{D}_{oo}^{\alpha} & \mathbf{0} & \mathbf{D}_{ov}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{oo}^{\beta} & \mathbf{0} & \mathbf{D}_{ov}^{\beta} \\ \mathbf{D}_{vo}^{\alpha} & \mathbf{0} & \mathbf{D}_{vv}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{vo}^{\beta} & \mathbf{D}_{vv}^{\alpha} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{oo}^{\alpha} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{oo}^{\beta} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.4}$$

Here, I used that occ-vir and vir-occ, and vir-vir block are also zero matrices,

$$\mathbf{D}_{ov}^{\sigma} = \mathbf{D}_{vo}^{\sigma} = \mathbf{D}_{vv}^{\sigma} = \mathbf{0}. \tag{4.5}$$

Only occ-occ block has non-zero components,

$$\mathbf{D}_{oo} = \begin{pmatrix} \mathbf{D}_{oo}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{oo}^{\beta} \end{pmatrix} \tag{4.6}$$

Note that occ-occ block matrices are diagonal,

$$\mathbf{D}_{oo}^{\sigma} = \begin{pmatrix} n_1^{\sigma} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & n_k^{\sigma} \end{pmatrix} \tag{4.7}$$

Next, the transformation matrix is constructed as follows,

$$\mathbf{U}_{\mathrm{NO}} = \begin{pmatrix} \mathbf{C}_{\mathrm{occ}}^{\alpha} & \mathbf{0} & \mathbf{C}_{\mathrm{vir}}^{\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\mathrm{occ}}^{\beta} & \mathbf{0} & \mathbf{C}_{\mathrm{vir}}^{\beta} \end{pmatrix} \tag{4.8}$$

The dimension of the matrix is $2N_{\mathrm{basis}}$. Transforming the rotation matrices into the natural orbital basis,

$$\mathbf{R}_{\mathrm{NO}}(\Omega) = \mathbf{U}_{\mathrm{NO}}^{\dagger} \mathbf{R}(\Omega) \mathbf{U}_{\mathrm{NO}} \tag{4.9}$$

The overlap between the rotated states and normal ones in the NO

basis follows as,

$$\mathbf{N}(\Omega) = \left[\begin{pmatrix} \mathbf{D}_{oo} & \mathbf{D}_{ov} \end{pmatrix} \mathbf{R}_{\mathrm{NO}}(\Omega) \begin{pmatrix} \mathbf{D}_{oo} \\ \mathbf{D}_{vo} \end{pmatrix} \right]^{-1} \tag{4.10}$$

Since $\mathbf{D}_{\mathrm{oo}}$ is diagonal matrix and $\mathbf{D}_{\mathrm{ov}}$, $\mathbf{D}_{\mathrm{vo}}$ are zero matrices, $\mathbf{N}(\Omega)$ follows as,

$$\mathbf{N}(\Omega) = \mathbf{R}_{\mathrm{NO}}^{-1}(\Omega) \tag{4.11}$$

Finally,

$$\langle \Phi | \hat{R}(\Omega) | \Phi \rangle = \frac{1}{\det(\mathbf{N}_{\mathrm{NO}}(\Omega)\mathbf{D}_{oo})}. \tag{4.12}$$

In the present DC-PUHF method, the total overlap/transition density matrices are constructed from the total density matrix. The rotated Fock matrices are constructed at each angle. The integral of rotated Fock is calculated by grid integral as the normal PUHF method.

The advantage of this study is that it diagonalizes local effective Fock matrices instead of diagonalizing the total one. In other words, as in the usual DC method, the total effective Fock matrix is divided into local matrices.

$$\mathcal{F}_{\mu\nu} = \sum_{\alpha} \mathcal{F}_{\mu\nu}^{\alpha} \tag{4.13}$$

The eigenvalue equations for each local system are then solved to obtain the local orbital coefficients and eigenvalues.

$$\mathcal{F}^{\alpha}\mathbf{C}^{\alpha} = \mathbf{S}^{\alpha}\mathbf{C}^{\alpha}\epsilon^{\alpha} \tag{4.14}$$

This SCF cycle iterates until energy convergence [Figure 4.1].

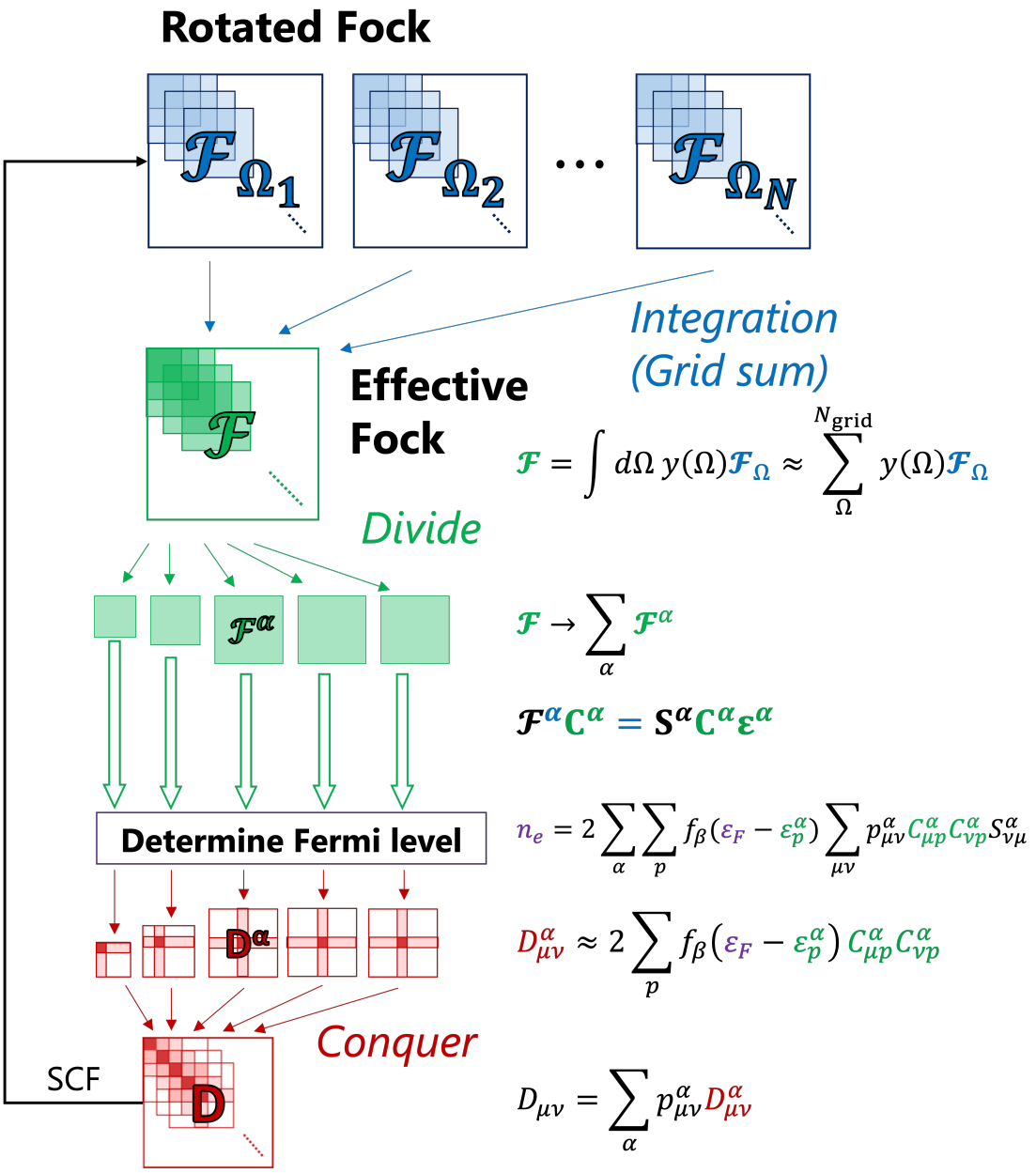

Figure 4.1: DC-PUHF scheme

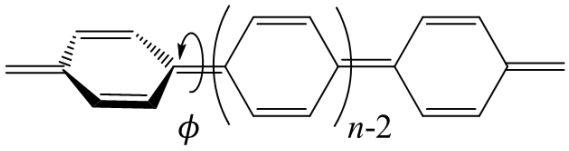

Figure 4.2: poly(phenoquinodimethane) (p-PQDM)

4.2.2 Result

I test the DC-PUHF method. Poly(phenoquinodimethane) (p-PQDM) [Figure 4.2] has a diradical character, so it is a static correlation system. I investigated the potential energy curve (PEC) for singlet p-PQDM against the rotation of C=C bond[48]. First, I tested the performance of the PUHF method for p-PQDM($n = 2$). Figure 4.5 shows PECs against the rotation of C=C bond. RHF, UHF, and CASSCF calculations are also performed for comparison. The active space of the CASSCF calculation is set to 14 electrons and 14 orbitals [Figure 4.3,4.4]. Figure 4.5 shows the absolute energy for PECs. CASSCF energy is the lowest energy, followed by PUHF, UHF, and RHF. Figure 4.6 shows the relative energy for PECs. RHF PEC differs significantly from the other three methods. The UHF and PUHF results are qualitatively consistent with the CASSCF one. However, $\langle \hat{S}^2_{\mathrm{UHF}} \rangle = 2.931$ at $\phi = 0$ degree, while $\langle \hat{S}^2_{\mathrm{PUHF}} \rangle = 1.235 \times 10^{-5}$, indicating that the PUHF wave function is in the correct spin state. The PUHF wave function is superior to the UHF wave function.

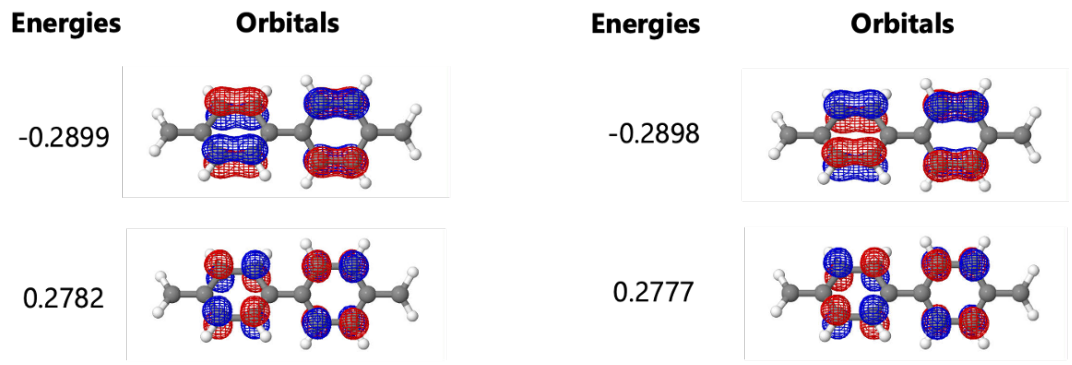

Figure 4.3: Active orbitals of the CASSCF calculation. (left) A symmetry, (right) B1 symmetry.

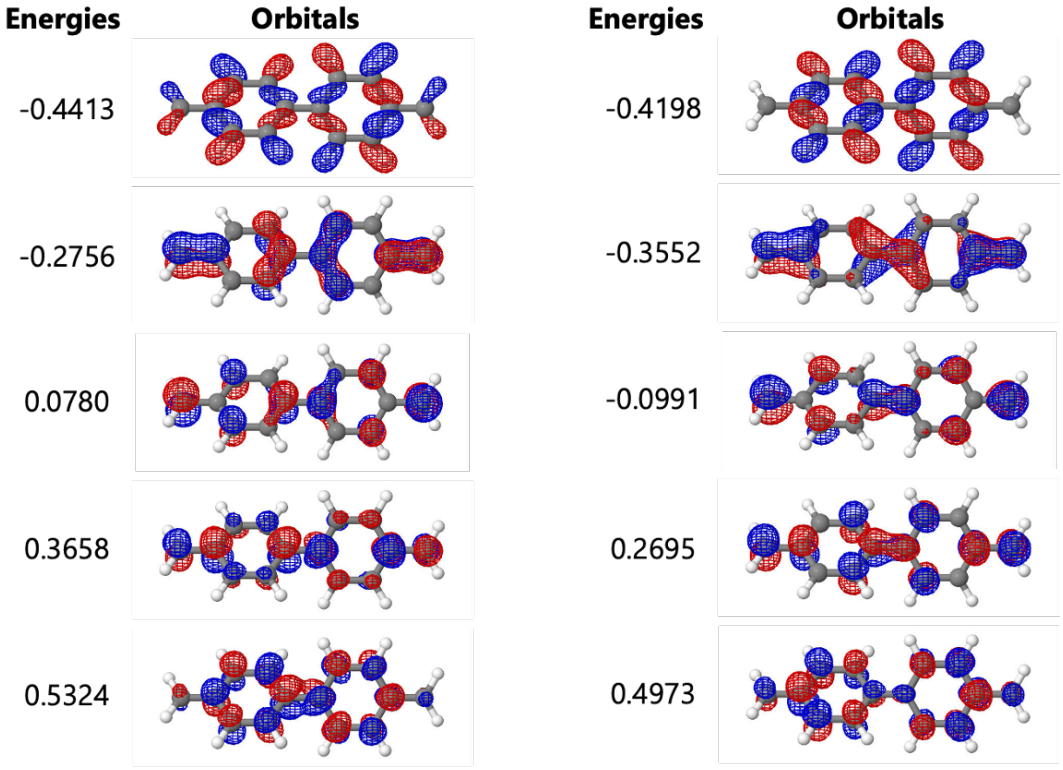

Figure 4.4: Active orbitals of the CASSCF calculation. (left) B2 symmetry, (right) B3 symmetry.

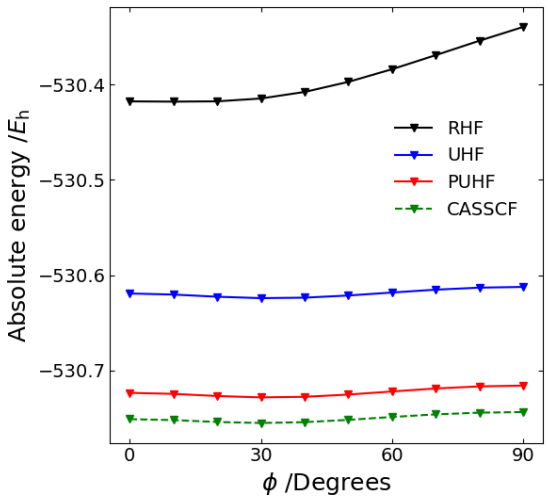

Figure 4.5: The absolute potential energy curve for p-PQDM($n = 2$).

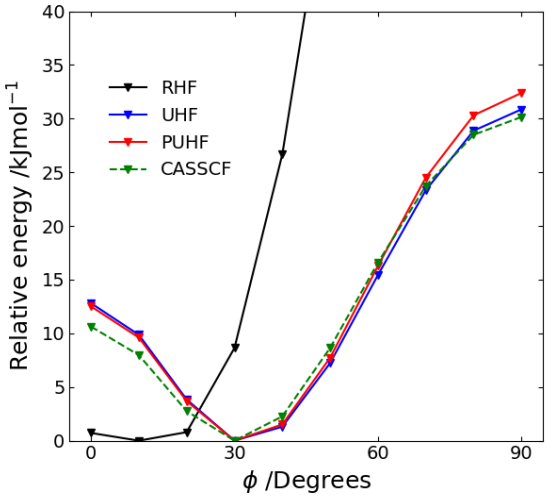

Figure 4.6: The relative potential energy curve for p-PQDM($n = 2$).

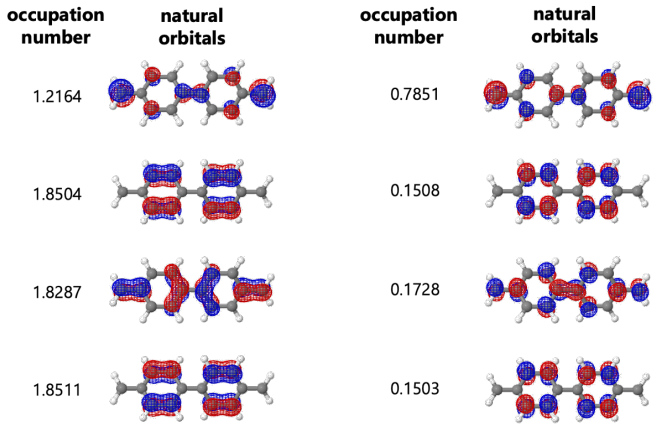

Figure 4.7: Natural orbitals and its occupation number of CASSCF wave function for p-PQDM($n = 2$) molecule at $\phi = 40$ degree.

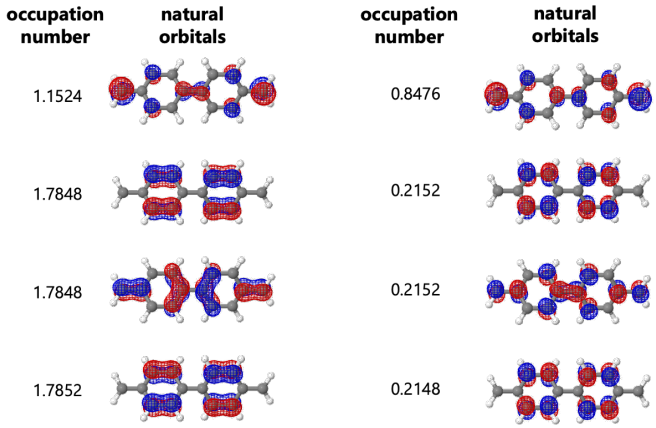

Figure 4.8: Natural orbitals and its occupation number of PUHF wave function for p-PQDM($n = 2$) molecule at $\phi = 40$ degree.

Next, I examined the PUHF wave function in terms of NOs and occupation numbers. Figure 4.7 shows the NOs and the occupation number of the CASSCF results. Radical orbitals at both ends are the most fractional and contribute to static electron correlation. NOs of the PUHF wave function are consistent with CASSCF ones, and the occupation number is qualitatively agreement with CASSCF ones [Figure 4.8].

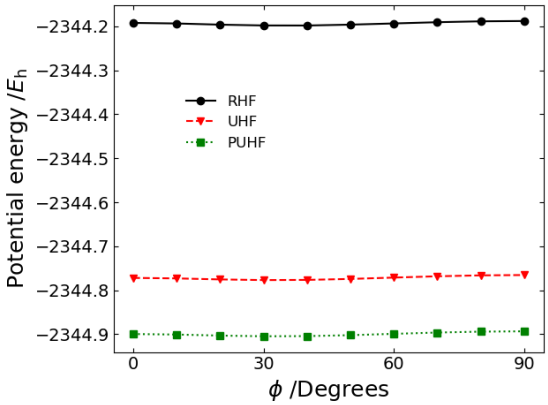

Figure 4.9: PEC for p-PQDM($n = 10$).

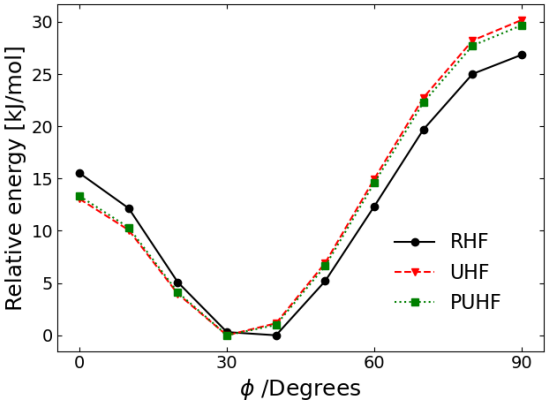

Figure 4.10: Relative energy curve for p-PQDM($n = 10$).

Next, p-PQDM($n = 10$) molecules are calculated. Figure 4.9 shows the absolute energy diagram and shows that the PUHF energy is the lowest. It can be seen that the PUHF method is able to incorporate more electron correlations than the UHF method on the basis of the variational principle. Figure 4.10 shows the relative energy. RHF gives a different energy minimum for the other two methods, and the RHF energy barrier is also inconsistent with the UHF and PUHF ones.

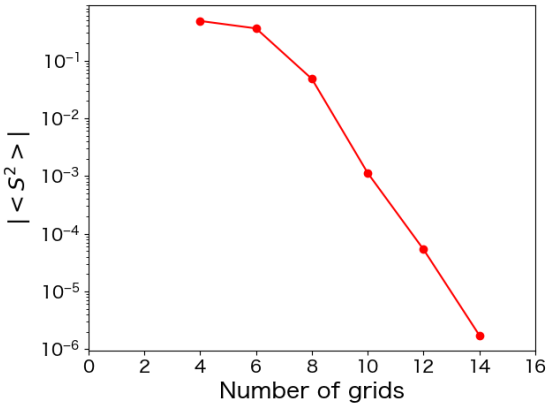

Figure 4.11: The grid dependency of $\langle \hat{S^2}_{\mathrm{PUHF}} \rangle$ for p-PQDM($n = 10$) molecule at $\phi = 40$ degree.

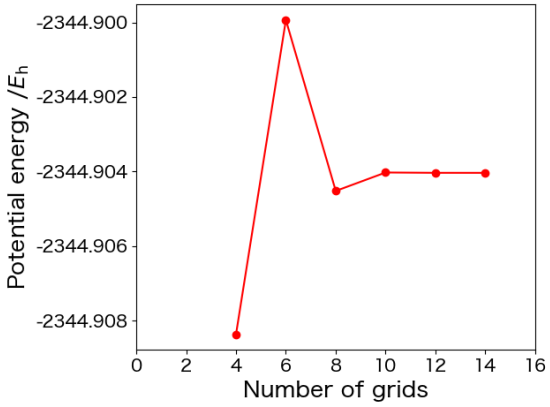

Figure 4.12: The grid dependency of PUHF energy for p-PQDM($n = 10$) molecule at $\phi = 40$ degree.

In addition to the energy, I investigated the PUHF wave function. First, I investigated the expectation value of $\hat{S^2}$. In PUHF, $\langle \hat{S^2} \rangle$ is not strictly zero ($\langle \hat{S}^2_{\mathrm{PUHF}} \rangle = -0.04796$ at $\phi = 40$ degree). This is due to the lack of grid points. In the current calculation, I adopted 8 grid points. In fact, the grid dependence of $\langle \hat{S^2} \rangle$ is examined in Figure 4.11. It can be seen that $\langle \hat{S^2} \rangle$ rapidly approaches zero as the grid increases. The PUHF energy was also found to converge when the number of grids was also increased [Figure 4.12].

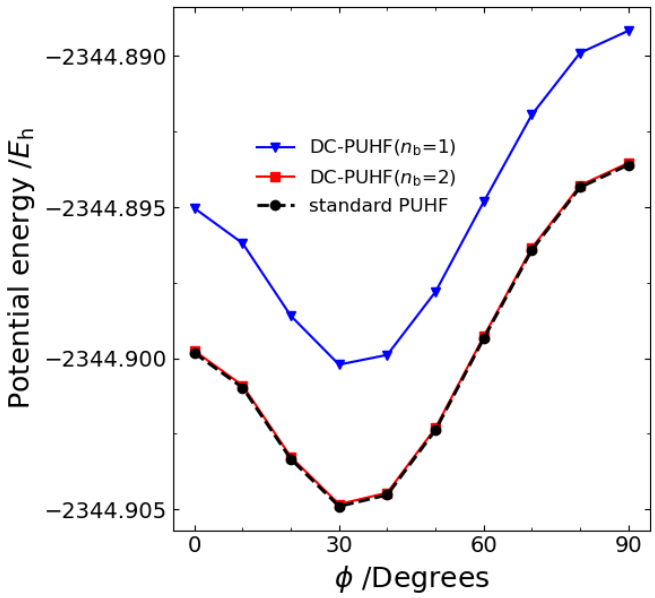

Figure 4.13: The DC-PUHF PECs for p-PQDM($n = 10$).

Next, I show the DC-PUHF results. Each C6 ring is assigned to each subsystem, so there are 10 subsystems. The buffer regions are defined as the number of adjoint C6 rings: n_b. Figure 4.13 shows DC-PUHF PECs for p-PQDM($n = 10$) compared to standard PUHF for comparison. The energy error exists in $n_\mathrm{b} = 1$ case. On the other hand, the DC-PUHF results are consistent with standard PUHF results in $n_\mathrm{b} = 2$ case. Therefore, the DC-PUHF method can reproduce standard PUHF results by adopting appropriate buffer regions. This result indicates that the PUHF effective Fock matrix has sparsity. Thus, linear-scaling PUHF calculations based on locality may be possible.

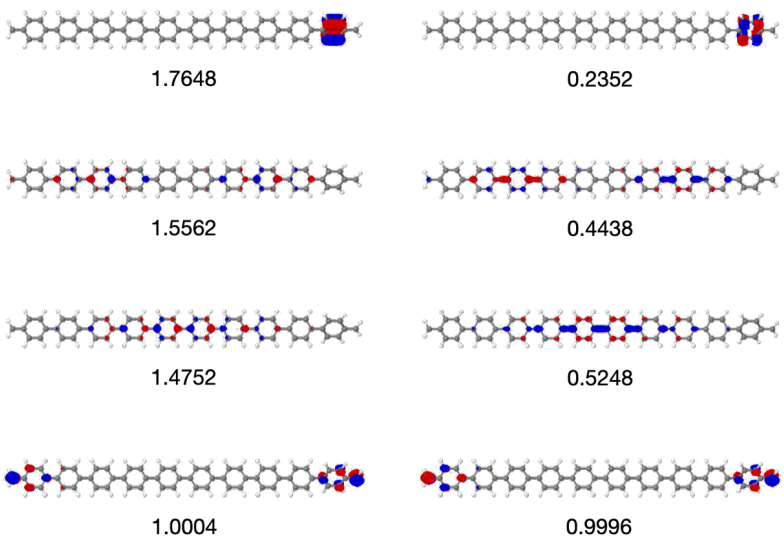

Figure 4.14: Natural orbitals and its occupation number of PUHF wave function for p-PQDM($n = 10$) at $\phi = 40$ degree.

The effect of dividing in the DC-PUHF method was investigated in terms of wave functions. Figure 4.14 shows the NOs and occupation number obtained by the PUHF method. First, there are localized π and $\pi*$ orbitals in the twisted C6 ring. Next, there are many delocalized π and $\pi*$ orbitals in the molecular plane. Finally, there are two radical orbitals at both ends that have the largest contribution to static electron correlations. Note that there are about 50 orbitals whose occupation numbers differ by more than 0.2 from 0 and 2. The fractional occupation is due to multi-configurational effect. In the CASSCF method, only active orbitals have fractional occupation numbers, while the occupation numbers of other orbitals are 0 or 2. This means that the PUHF method is able to include more than 50 electron correlations; it is currently impossible to calculate such a large active space with the CASSCF

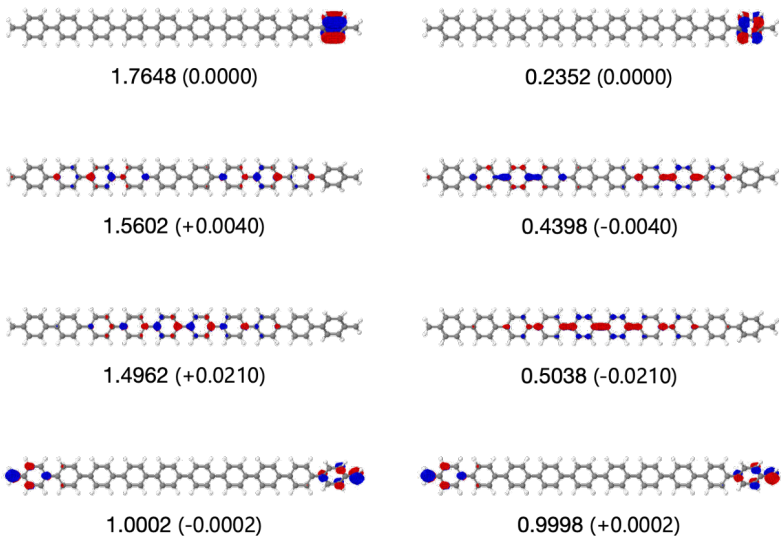

Figure 4.15: Natural orbitals and its occupation number of DC-PUHF($n_\mathrm{b} = 1$) for p-PQDM($n = 10$) at $\phi = 40$ degree. The difference between DC-PUHF occupation number and PUHF ones are given in parentheses.

method.

I investigated DC-PUHF wave function in terms of NOs. Figure 4.15 shows DC-PUHF($n_\mathrm{b} = 1$) NOs and values in parentheses indicate the difference of the occupation number from the PUHF ones. Its NOs are consistent with the PUHF ones. The occupation numbers have small errors in the localized π, $\pi*$, and radical orbitals. On the other hand, for the delocalized π and $\pi*$ orbitals, the errors are relatively large. Figure 4.16 shows DC-PUHF($n_\mathrm{b} = 2$) NOs. The error of occupation numbers in delocalized π and $\pi*$ orbitals has decreased by one order of magnitude. From this, it can be seen that the DC-PUHF method can also accurately describe wave function by expanding the buffer region.

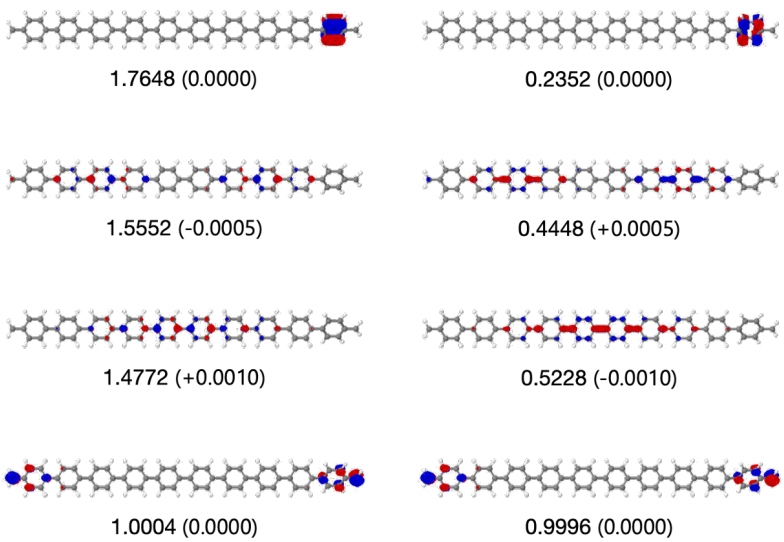

Figure 4.16: Natural orbitals and its occupation number of DC-PUHF($n_\mathrm{b} = 2$) for p-PQDM($n = 10$) at $\phi = 40$ degree. The difference between DC-PUHF occupation number and PUHF ones are given in parentheses.

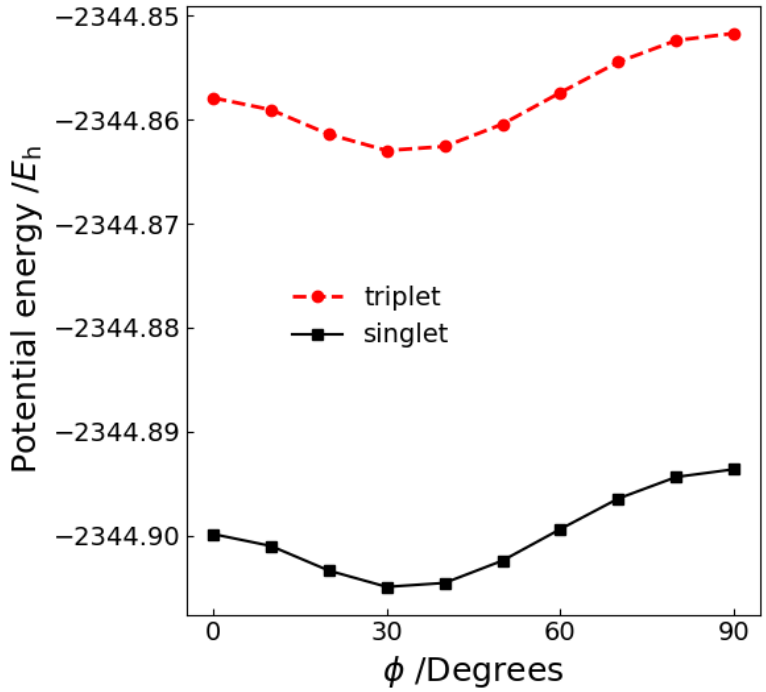

Figure 4.17: PUHF PECs for singlet/triplet p-PQDM($n = 10$).

I also investigated other spin states and performed the triplet state. Figure 4.17 shows the PUHF PECs for singlet and triplet p-PQDM($n = 10$). From this, it can be seen that the shapes of the singlet and triplet PECs are almost the same, but the singlet state is more stable. Figure 4.18 shows the DC-PUHF PECs. It indicates that PUHF results could be reproduced by expanding the buffer region as in the singlet state [Figure 4.18].

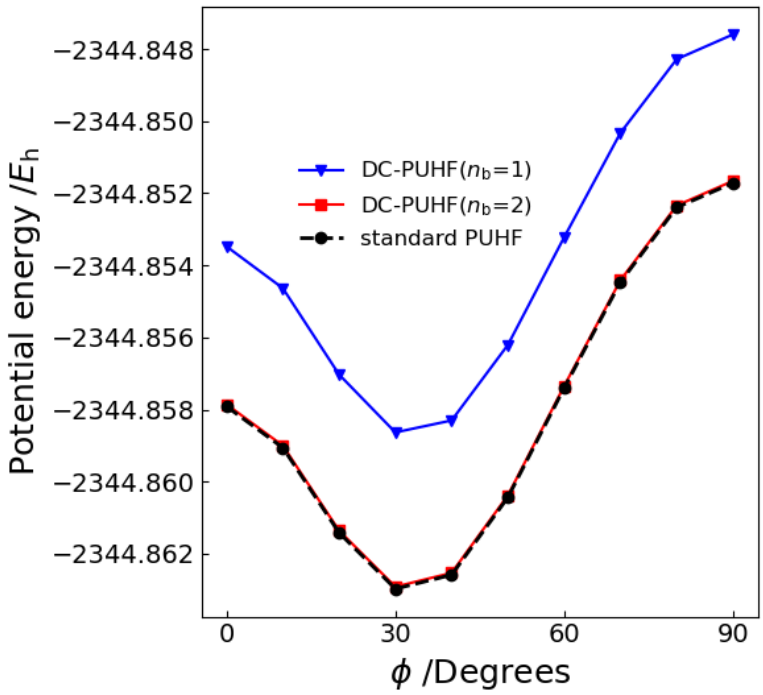

Figure 4.18: DC-PUHF PECs for triplet p-PQDM($n = 10$).

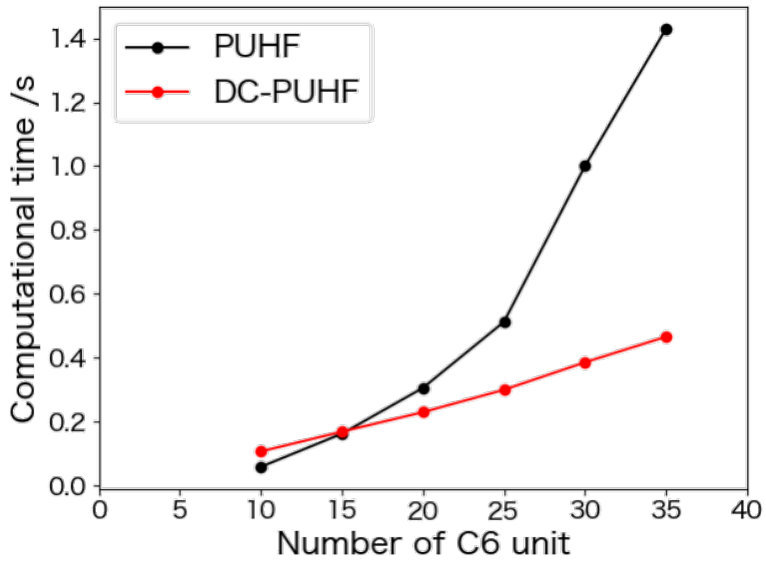

Figure 4.19: Computational time for diagonalization of effective Fock matrix

Finally, I investigated computational time. In this case, the time required to diagonalize the effective Fock matrix was compared between the PUHF and DC-PUHF methods. The computational time here was measured using a single node computer equipped with two Intel(R) Xeon(R) Gold 6326 processors. Figure 4.19 shows the relationship between the computational time and the number of C6 rings: n_{C6}. The PUHF method scales $O(n_{\mathrm{C6}}^{2.5})$, while the DC-PUHF method scales $O(n_{\mathrm{C6}}^{1.2})$, indicating that almost linear scaling was achieved.

4.3 Acceleration scheme for construction of projection matrix based on the DC method

4.3.1 Theory

Next, I developed the projection scheme that approximates the overlap and transition density matrices $\mathbf{M}(\Omega)$, $\mathbf{D}(\Omega)$ from the local ones $\mathbf{M}^{\alpha}(\Omega)$, $\mathbf{D}^{\alpha}(\Omega)$ as follows,

$$M^{\mathrm{DC}}_{\mu\nu}(\Omega) = \sum_{\alpha}^{N_{\mathrm{subsystem}}} p^{\alpha}_{\mu\nu} M^{\alpha}_{\mu\nu}(\Omega) \tag{4.15}$$

$$D^{\mathrm{DC}}_{\mu\nu}(\Omega)) = \sum_{\alpha}^{N_{\mathrm{subsystem}}} p^{\alpha}_{\mu\nu} D^{\alpha}_{\mu\nu}(\Omega) \tag{4.16}$$

where $p^{\alpha}_{\mu\nu}$ is the element of partition matrix in DC method. $\mathbf{F}(\Omega)$ are calculated using the transition density matrix approximated by Eq 4.16

$$F_{\mu\nu}(\Omega) = h_{\mu\nu} + \frac{1}{2} \sum_{\lambda\sigma}^{N_{\mathrm{basis}}} \langle \mu\lambda || \nu\sigma \rangle \, D^{\mathrm{DC}}_{\lambda\sigma}(\Omega) \tag{4.17}$$

However, the determinant calculation for the overlap matrices $\mathbf{M}(\Omega)$ remains $O(N^3)$ cost. Currently, this scheme is available for only PAV scheme.

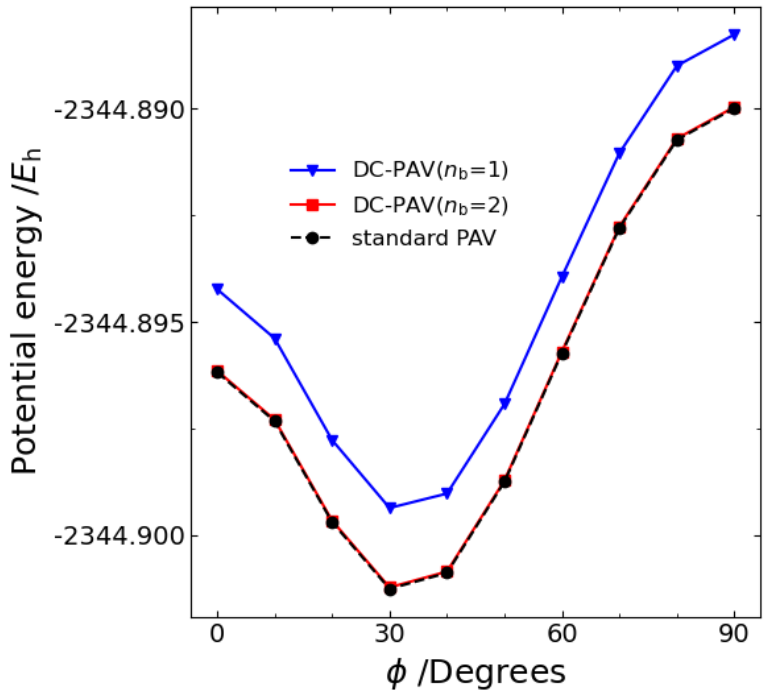

Figure 4.20: DC-PAV PECs for p-PQDM($n = 10$).

4.3.2 Result

I test this projection scheme for p-PQDM($n = 10$) molecules. Figure 4.20 shows the DC-PAV PECs compared to the standard PAV ones. This means that the transition density $\mathbf{M}(\Omega)$ and the overlap matrices $\mathbf{D}(\Omega)$ are also sparse, suggesting that they can be divisible.

The DC-PAV wave function was examined. Figure 4.21 shows the NOs and occupation number obtained from the PAV wave function. Although these orbitals are not optimized, the PAV NOs and occupation numbers are consistent with the VAP ones. In DC-PAV($n_b = 1$), the NOs of the radicals are localized on one side of the molecule, unlike those of PAV [Figure 4.22]. In addition, the occupation number is also not "corresponding". In other words, the sum

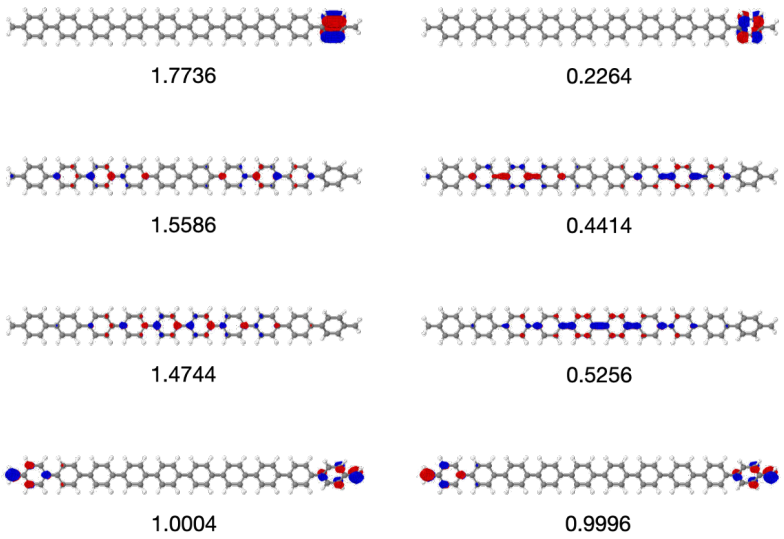

Figure 4.21: Natural orbitals and its occupation number of PAV wave function for p-PQDM($n = 10$) at $\phi = 40$ degree.

of occupation number of the paired orbitals is not 2. This suggests that the buffer region is small and does not fully incorporate the static electron correlation. Figure 4.23 shows the DC-PAV($n_{\mathrm{b}} = 2$) results. NOs are consistent with VAP ones, and occupation number remains not "corresponding" but the errors are smaller. Thus, it was found that the wave function can be reproduced by setting up an appropriate buffer region even if the transition density matrix and the overlap matrix are divided.

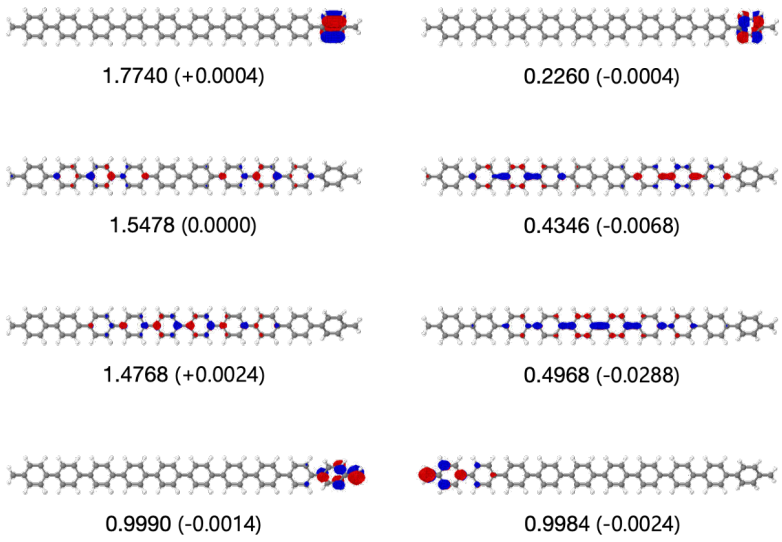

Figure 4.22: Natural orbitals of DC-PAV($n_\mathrm{b} = 1$) wave function for p-PQDM($n = 10$) at $\phi = 40$ degree. The difference between DC-PAV occupation number and PAV ones are given in parentheses.

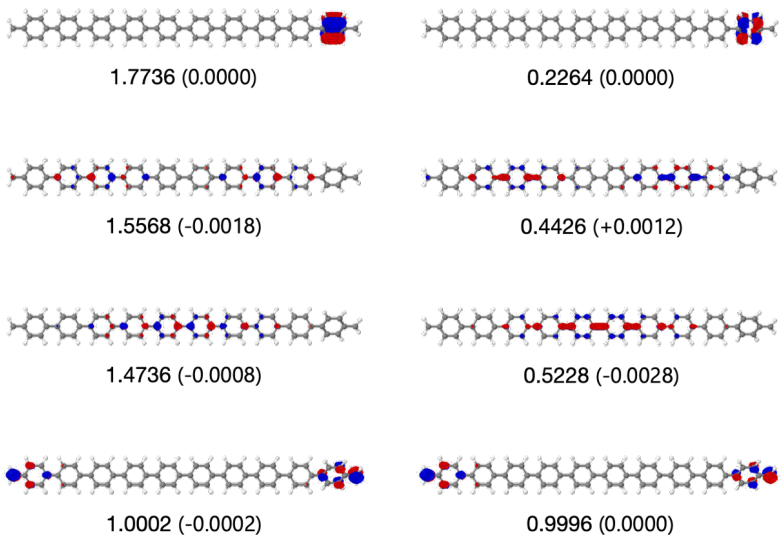

Figure 4.23: Natural orbitals of DC-PAV($n_\mathrm{b} = 2$) wave function for p-PQDM($n = 10$) at $\phi = 40$ degree. The difference between DC-PAV occupation number and PAV ones are given in parentheses.

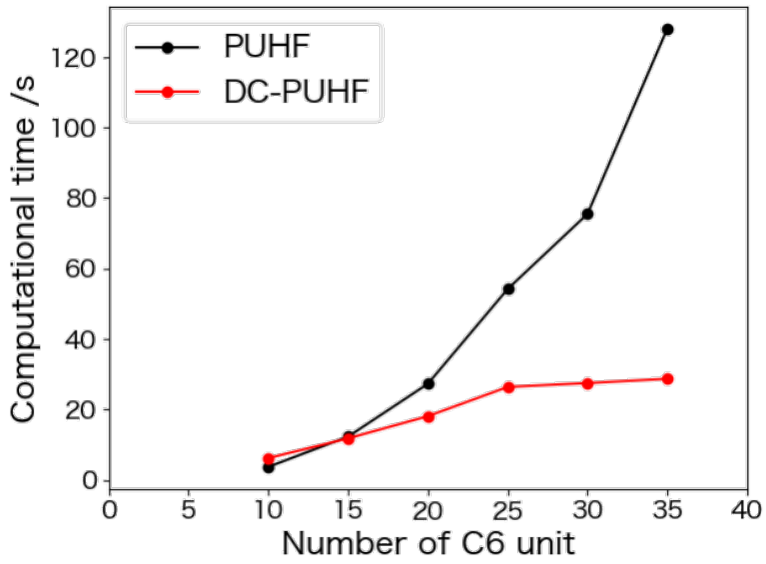

Figure 4.24: The computational time for constructing overlap and transition density matrices.

Finally, I examine the computational time. Figure 4.24 shows the computational time for constructing the overlap and transition density matrices, which include the calculation of the inverse matrix. The computational time here was measured using a single node computer equipped with two Intel(R) Xeon(R) Gold 6326 processors. The PUHF method scales $O(n_{\mathrm{C6}}^{2.8})$, while the DC-PUHF method scales $O(n_{\mathrm{C6}}^{1.3})$, indicating that almost linear scaling was achieved. The time required for the remaining bottleneck, the determinant calculation, is also shown in Figure 4.25. The PUHF method scales $O(n_{\mathrm{C6}}^{3.8})$ and the DC-PUHF method scales $O(n_{\mathrm{C6}}^{3.9})$. This step is the most time-consuming throughout the entire calculation.

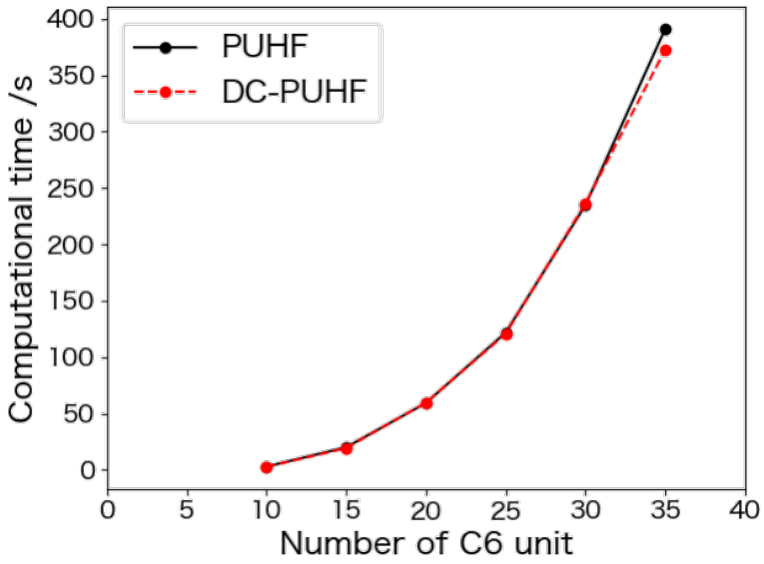

Figure 4.25: The computational time for determinants.

4.4 Conclusion

I developed the DC-PUHF method for the large-scale static correlation calculation. The test calculation shows that the DC-PUHF energy is consistent with the standard PUHF results if appropriate buffer regions are adopted. In addition, $\langle \hat{S}^2 \rangle$ and NOs are also consistent using appropriate buffer regions. This indicates that the effective Fock, overlap, and transition density matrices , which includes the projection effect, are sparse. This fact suggests that it can be computed by division. The diagonalization of the effective Fock matrix and the construction of the overlap and transition density matrices, which were accelerated in this study, were confirmed to achieve linear scaling. In the future, the last bottleneck, the determinant calculations, will be approximated from the subsystem calculations. I look forward to the future development of

large-scale static electron correlation methods.

Chapter 5

DC-xTB: A general-purpose, very large-scale calculation method

5.1 Preface

Due to significant improvements in computational efficiency and the development of advanced electronic structure theories, the scope of computational quantum chemistry has been extended to a wide variety of systems. In particular, the large-scale electronic structure theories[130] have made it possible to apply the method to the system with thousands of atoms. Even for typical large-scale targets such as proteins, the consideration of surrounding water molecules is essential, which leads to the model with tens of thousands of atoms[33]. In these systems, the dynamic geometrical change may also affect not only reactions directly but also through the bunch of conformations by the entropic terms. Although classical molecular dynamics (MD) can handle such systems, it cannot treat chemical bond rearrangement, which is essential both in the amorphous formation and protein reactions. For enabling large-scale quantum

MD simulations, much effort has been devoted so far[130]. Note that an alternative way of treating chemical rearrangement in cheap computational cost is the machine learning potential (MLP)[121]. Even when the accuracy of MLP is sufficiently improved, the dynamic quantum characters such as the electron distribution and density of states are still important, which has to be obtained through the quantum chemical calculations.

The semi-empirical method can be a balanced option for large systems while preserving the quantum nature. Since the invention of Pariser–Parr–Pople (PPP) method[74, 79], many traditional semi-empirical Hamiltonians have been proposed in quantum chemistry[112]. The accuracy of the semi-empirical methods was drastically improved with the advent of the density-functional tight-binding (DFTB) method[81, 101], which is based on the density functional theory (DFT). However, only adopting semi-empirical Hamiltonian cannot make very large-scale calculations possible due to the diagonalization of the Hamiltonian matrix taking the computational time of $O(N^3)$, where N refers to the system size. Previous study has combined the DFTB Hamiltonian with the divide-and-conquer (DC) method[71, 69]. The DC method, originally proposed by Yang and coworkers[127], is a linear-scaling quantum chemical method that avoids explicit diagonalization of the full Hamiltonian matrix based on its sparsity. The DC-based self-consistent field methods, including DFTB, can produce the energy gradient[131, 45], enabling molecular dynamics simulation of large systems[65, 94, 11, 54, 122,

53, 70]. However, the DFTB method requires element-pair parameters, and its available elements are limited. Thus, it cannot treat enormously large systems containing various atom species, such as doped amorphous and protein with trace metals.

The GFN-xTB method, called xTB hereafter, has recently been developed and attracted attention as a general-purpose semi-empirical quantum chemical calculation[28, 7, 6]. This method adopted the DFTB-type Hamiltonian while using element-specific parameters, basically non-element-pair parameters, although the target molecular properties were limited to geometry, vibrational frequencies, and non-covalent interaction. The default parameters are provided up to $Z = 86$, which makes the method applicable to any systems consisting of up to period 6 elements. Recently, the xTB has been implemented in the CP2K package[95], where the linear-scaling density matrix minimization optimizer can be adopted in addition to the conventional self-consistent scheme[59, 57].

This study combined the DC scheme with the xTB method to develop a general-purpose, huge-scale quantum chemical MD simulation program.

5.2 The DC-xTB method

Several versions of GFN-xTB have been developed so far[6]. In current DC implementation, I only focused on the GFN1-xTB Hamiltonian. Thus the formulae in this manuscript is based on the GFN1-xTB method, while its extensions to other xTB Hamilto-

nians are straightforward. The GFN-xTB energy is composed of the electronic energy E_{el}, the repulsion energy E_{rep}, the dispersion energy E_{disp}, and the halogen bond energy E_{XB}:

$$E = E_{\mathrm{el}} + E_{\mathrm{rep}} + E_{\mathrm{disp}} + E_{\mathrm{XB}} \tag{5.1}$$

In the DC-xTB method, the electronic energy E_{el} is given by

$$\begin{aligned} E_{\mathrm{el}} = &\sum_{\mu\nu}^{N_{\mathrm{AO}}} D_{\mu\nu}^{\mathrm{DC}} h_{\nu\mu} + \frac{1}{2} \sum_{AB}^{N_{\mathrm{atom}}} \sum_{l}^{N_{\mathrm{shell}}(A)} \sum_{l'}^{N_{\mathrm{shell}}(B)} \gamma_{AB,ll'} \Delta p_l^A \Delta p_{l'}^B \\ &+ \frac{1}{3} \sum_{A}^{N_{\mathrm{atom}}} \Delta Q_A^3 \Gamma_A \end{aligned} \tag{5.2}$$

where $h_{\nu\mu}$ represents the one-electron Hamiltonian matrix element and $D_{\mu\nu}^{\mathrm{DC}}$ is the density matrix element approximated by the DC scheme. N_{AO}, N_{atom}, and $N_{\mathrm{shell}}(A)$ refer to the number of AOs, the number of atoms, and the number of atomic shells in atom A, respectively. ΔQ_A and Δp_l^A are atomic charge on atom A and shell charge on shell l of atom A, respectively,

$$\Delta Q_A = \sum_{\nu}^{N_{\mathrm{AO}}} \sum_{\mu \in A} D_{\mu\nu}^{\mathrm{DC}} S_{\nu\mu} - Z_A \tag{5.3}$$

$$\Delta p_l^A = \sum_{\nu}^{N_{\mathrm{AO}}} \sum_{\mu \in l(A)} D_{\mu\nu}^{\mathrm{DC}} S_{\nu\mu} - p_l^{A_0} \tag{5.4}$$

where $S_{\nu\mu}$, Z_A, and $p_l^{A_0}$ are the overlap matrix element, the neutral nuclear charge of atom A, and the reference shell charge of shell l in atom A, respectively. The second and third terms of Eq. (5.2) represent the second-order electronic Coulomb repulsion and diagonal of the third-order Coulomb interaction terms, respectively.

The second-order Coulomb repulsive function, $\gamma_{AB,ll'}$, is the function of atomic distance R_{AB}, while parametrized by the chemical hardness, $\eta_{AB,ll'}$, and other parameters, k_g and κ_A^l, as

$$\gamma_{AB,ll'} = \left(\frac{1}{R_{AB}^{k_g} + \eta_{AB,ll'}^{-k_g}} \right)^{1/k_g} \tag{5.5}$$

$$\eta_{AB,ll'} = 2 \left(\frac{1}{(1+\kappa_A^l)\eta_A} + \frac{1}{(1+\kappa_B^{l'})\eta_B} \right)^{-1} \tag{5.6}$$

The third-order Coulomb term, Γ_A, is element-specific atomwise parameter. The one-electron Hamiltonian is given as the following:

$$h_{\mu\nu} = \frac{1}{2} K_{AB}^{ll'} (h_A^l + h_B^{l'}) S_{\mu\nu} (1 + k_{\mathrm{EN}} \Delta EN_{AB}^2) \Pi(R_{AB,ll'}) \quad (\mu \in l(A), /\nu \in l'(B)) \tag{5.7}$$

Here, $K_{AB}^{ll'}$ is the shell-specific parameter, and it is only the element-pair parameter referred to as the scaling constant. $\Delta EN_{AB} = EN_{\mathrm{A}} - EN_{\mathrm{B}}$ is defined with the Pauling's electronegativity of element A, EN_A, and k_{EN} is a global parameter. k_l is the Hückel constant of the shell l, $\Pi(R_{AB,ll'})$ is a distance and l-dependent function, and h_A^l is a shell- and element-specific parameter. The h_A^l depends on the D3 coordination number of atom A, CN_A;

$$h_A^l = H_A^l (1 + k_{CN,l} CN_A) \ (l \in A) \tag{5.8}$$

where H_A^l and $k_{CN,l}$ are the atom energy level of atom A and l-dependent global scaling parameter. The repulsion energy is as follows,

$$E_{\mathrm{rep}} = \sum_{AB} V_{AB}^{\mathrm{rep}} = \sum_{AB} \frac{Z_A^{\mathrm{eff}} Z_B^{\mathrm{eff}}}{R_{AB}} \exp\left[-(\alpha_A \alpha_B)^{1/2} \right] (R_{AB})^{k_f}, \tag{5.9}$$

where Z_A^{eff} is the effective nuclear charge on atom A, k_f is a global parameter, and α_A is the element-specific parameter. The dispersion energy, E_{disp}, is computed by Grimme's D3 method[27]. In GFN1-xTB, the Lennard–Jones type halogen bonding interaction function, E_{XB}, is added for every halogen bond[28]. The minimal basis set is adopted except for hydrogen, where a double-zeta set is adopted for the description of hydrogen bonding.

In the DC scheme, the system of interest is divided into disjoint subsystems. Each disjoint subsystem is called the *central region* and the set of basis functions belonging to the central region of subsystem α is represented by $\mathcal{S}(\alpha)$. Note that $\mathcal{S}(\alpha) \cap \mathcal{S}(\beta) = \varnothing^{\forall}\alpha \neq \beta$ by definition. To improve the accuracy, the *buffer region*, defined as the neighbor of the central region, is included in the subsystem calculation. The basis functions belonging to the buffer region of subsystem α is denoted by $\mathcal{B}(\alpha)$. The combined region of the central and buffer regions is referred to as the *localization region* and the basis functions of the localization region of subsystem α is denoted by $\mathcal{L}(\alpha)$, namely, $\mathcal{L}(\alpha) = \mathcal{S}(\alpha) \sqcup \mathcal{B}(\alpha)$. In the DC method, the error originating from the fragmentation can be systematically reduced by enlarging the buffer region. Fujimori *et al.* developed a scheme to automatically determine the appropriate buffer region from single energy criterion value[42, 23, 43]. In each localization region, the local Hamiltonian matrix is constructed as

following,

$$\begin{aligned} H^{\alpha}_{\mu\nu} =& h_{\mu\nu} + \frac{1}{2} S_{\mu\nu} \sum_{C}^{N_{\mathrm{atom}}} \sum_{l"}^{N_{\mathrm{shell}}(C)} (\gamma_{AC,ll"} + \gamma_{BC,l'l"}) \Delta p^{C}_{l"} \\ &+ \frac{1}{2} S_{\mu\nu} (\Delta Q^{2}_{A} \Gamma_{A} + \Delta Q^{2}_{B} \Gamma_{B}) \\ &(\mu \in l(A), \nu \in l'(B), \mu, \nu \in \mathcal{L}(\alpha)) \end{aligned} \tag{5.10}$$

Then, the following local eigenvalue equation is solved for every localization region:

$$\mathbf{H}^{\alpha} \mathbf{C}^{\alpha}_{p} = \varepsilon^{\alpha}_{p} \mathbf{S}^{\alpha} \mathbf{C}^{\alpha}_{p}, \tag{5.11}$$

where $\mathbf{S}^{\alpha} = [S_{\mu\nu}]_{\mu\nu \in \mathcal{L}(\alpha)}$ is the local overlap matrix, and $\mathbf{C}^{\alpha}_{p}$ and ε^{α}_{p} are local MO coefficient vector and energy of p-th orbital, respectively. The local density matrix is calculated from the local MOs as,

$$D^{\alpha}_{\mu\nu} = 2 p^{\alpha}_{\mu\nu} \sum_{p} f_{\beta}(\varepsilon_{\mathrm{F}} - \varepsilon^{\alpha}_{p}) C^{\alpha}_{\mu p} C^{\alpha *}_{\nu p}. \tag{5.12}$$

ε_{F} and $f_{\beta}(x)$ are the Fermi level and the Fermi function, respectively. $p^{\alpha}_{\mu\nu}$ is a partition matrix to avoid overlap between localization regions as follows,

$$p^{\alpha}_{\mu\nu} = \begin{cases} 1 & \text{if } \mu \in \mathcal{S}(\alpha) \text{ and } \nu \in \mathcal{S}(\alpha), \\ 1/2 & \text{if } (\mu \in \mathcal{S}(\alpha) \text{ and } \nu \in \mathcal{B}(\alpha)) \text{ or vice versa}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.13}$$

The total density matrix is constructed from the local density matrices and partition matrix,

$$D^{\mathrm{DC}}_{\mu\nu} = \sum_{\alpha} p^{\alpha}_{\mu\nu} D^{\alpha}_{\mu\nu}. \tag{5.14}$$

The Fermi level is determined so that the number of electrons is conserved in the entire system,

$$N_{\mathrm{e}} = \sum_{\mu} (\mathbf{D}^{\mathrm{DC}}\mathbf{S})_{\mu\mu}. \tag{5.15}$$

The charges defined by Eqs. (5.3) and (5.4) are calculated from the density matrix that is defined by Eqs. (5.12) and (5.14) with Fermi level, the solution of Eq. (5.15), and the subsystem MOs, the solutions of Eq. (5.11). The subsystem Hamiltonian matrix in Eq. (5.11) is defined by Eq. (5.10) with atomic and shell charges. Therefore, in the standard self-consistent charge (SCC) procedure, these charges are determined self-consistently from the iterative calculation.

5.3 Implementation of the DC-xTB program

I have developed the present DC-xTB energy and gradient calculation codes by incorporating the Hamiltonian construction routines in xtb package, an open-source xTB calculation program, into the DCDFTBK program[71]. DCDFTBK is the DC-DFTB calculation program that was originally tuned for Japan's previous-generation supercomputer K, which is oriented toward the massively parallel calculation. Since the original DCDFTBK program can only handle the single-zeta quality basis set, I have modified the program to handle the double-zeta basis for hydrogen atoms.

Figure 5.1 shows the computational procedure of the DC-xTB calculation. The steps highlighted in yellow are constructed from

the xTB parameters and those in red are processed in the subsystem loop in the DC method. Initially, the effective nuclear repulsion and halogen bond energies, E_{rep} and E_{XB}, are constant for fixed geometry and are pre-calculated. The electron repulsion parameter, $\gamma_{AB,ll'}$, also becomes a constant for each combination of AB pair and ll' shells, while the third-order parameter, Γ_A, is constant by definition. Although these pre-calculation steps require $O(N^2)$ computational time, they are calculated only once and their computational time is relatively small compared to the SCC part. The overlap and one-electron Hamiltonian matrix elements, $S^{\alpha}_{\mu\nu}$ and $h^{\alpha}_{\mu\nu}$, are also precomputed for each subsystem, not for the entire system to avoid huge memory requirements and to reduce the computational scaling to $O(N)$. The local hamiltonian matrix element, $H^{\alpha}_{\mu\nu}$, is then constructed in the subsystem loop according to Eq. (5.10) and is diagonalized to obtain the subsystem MOs. The routines to determine the Fermi level and to compute the charges are from the DCDFTBK program. The efficient implementation of these routines is reported in detail in Ref. [71]. I have compiled the present DCxTBMD program on the Japanese supercomputer "Fugaku" to handle more than tens of thousands of atoms containing various atom species.

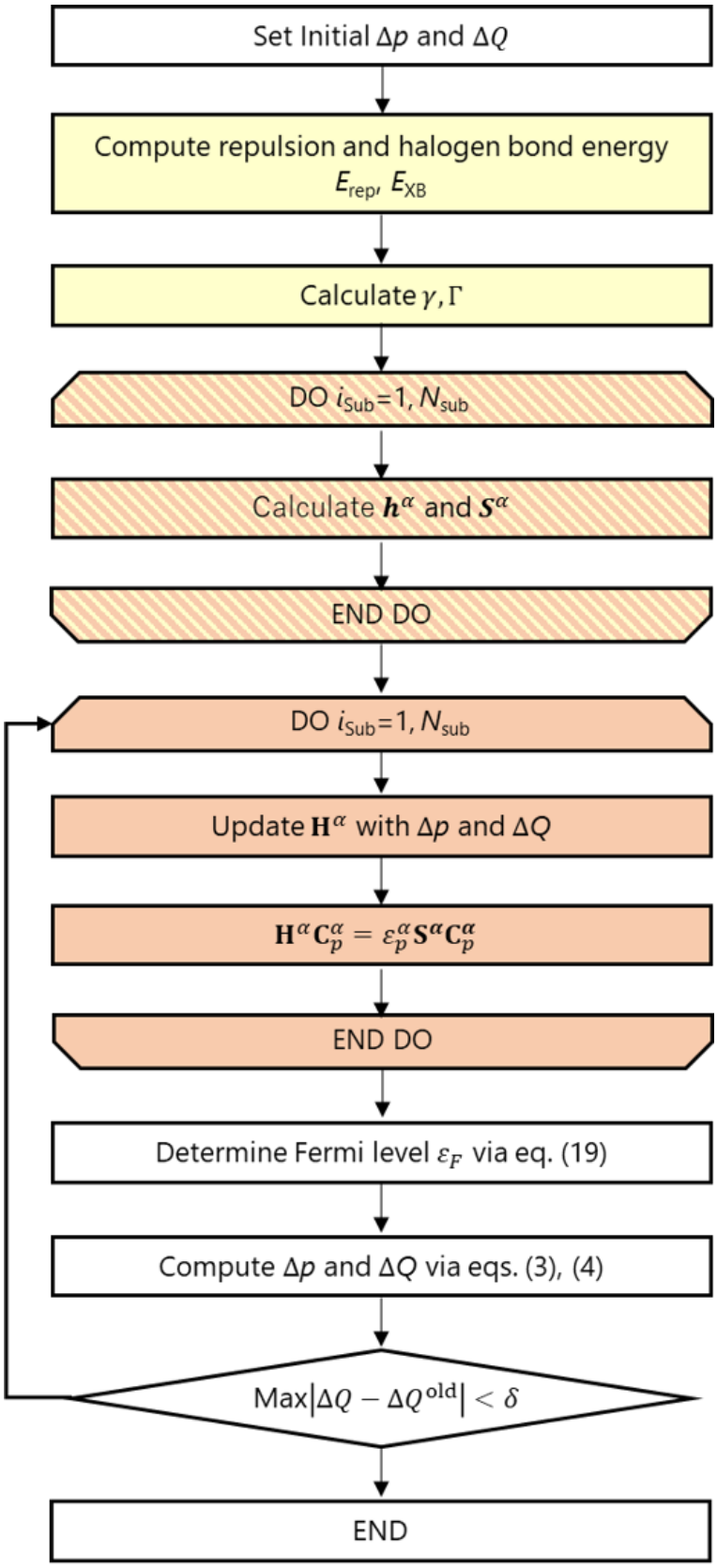

Figure 5.1: The DC-xTB calculation scheme.

5.4 The DC-xTB calculation result

I here assessed the computational cost, accuracy, and parallelization efficiency of the present DC-xTB method. Throughout this study, I have defined the central region as a cube of d_{cent} per side and the buffer region as atoms within r_{buf} radius of the atoms in the central region, unless otherwise stated.

First, the computational times for the standard xTB and DC-xTB calculations of n_{water} (= 500 –4500) water molecules were measured using a single node of the Fugaku computer. In the DC calculations, I set $d_{\mathrm{cent}} = 7.0$ Å and $r_{\mathrm{buf}} = 5.0$ Å. Figure 5.2 summarizes the system-size dependence of the computational time per SCC iteration for the standard and DC-xTB calculations. The time for the DC-xTB calculation becomes shorter than that for the standard xTB for $n_{\mathrm{water}} \geqslant 1500$. According to the order estimation based on the double logarithmic plot, the time for the DC and standard xTB calculations scale as $O(n_{\mathrm{water}}^{1.23})$ and $O(n_{\mathrm{water}}^{2.71})$, respectively.

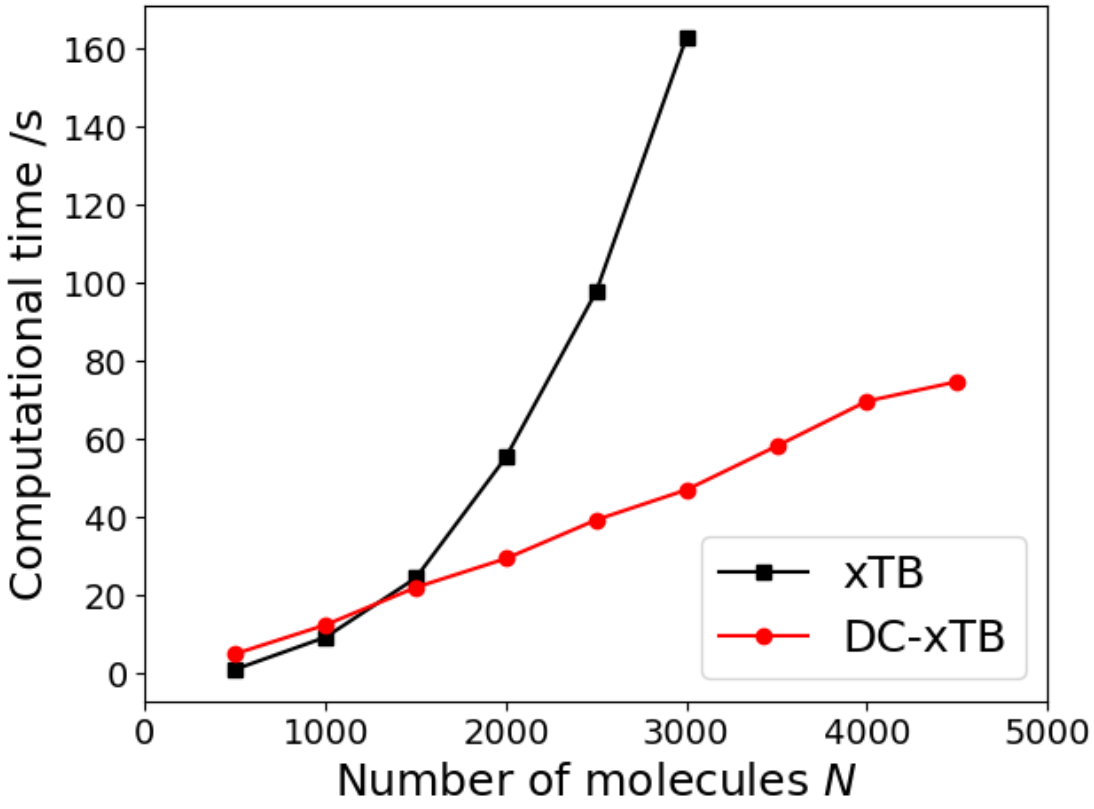

Figure 5.2: Computational time per one SCC iteration as a function of the number of water molecules (n_{water}) for the standard and DC-xTB calculations.

Next, the parallel efficiency of the DCxTBMD program was investigated in calculation of 3000 water molecule system using the supercomputer "Fugaku". Here, I set $d_{\mathrm{cent}} = 5.0$ Å and $r_{\mathrm{buf}} = 6.0$ Å. Table 5.1 shows the dependence of the computational time (t) on the number of nodes (n_{nodes}) used in the calculation. The parallel efficiency (PE), which was defined as $PE = (t_{\mathrm{ref}} \cdot n_{\mathrm{nodes}}^{\mathrm{ref}})/(t \cdot n_{\mathrm{nodes}})$ with $n_{\mathrm{nodes}}^{\mathrm{ref}} = 12$, was also tabulated. The number of subsystems in the DC-xTB calculation of this system was 775. Up to $n_{\mathrm{nodes}} = 192$, where each node is responsible for an average of ~ 4 subsystems, PE is over 90%, showing high efficiency. Although the PE deteriorates to $\sim$ 65% for $n_{\mathrm{nodes}} = 384$, the main reason is the load imbalance. Therefore, the efficiency would improve as the system size increases.

Table 5.1: Computational time (s) and parallel efficiency (%) for the DC-xTB calculations of 3000 water molecule system. The number of nodes, n_{nodes}, was varied between 12 and 384.

Number of nodes (n_{nodes})	**Time** (t) **[s]**	**Parallel efficiency (PE) [%]**
12	25.59	-
24	12.82	99.8
48	6.43	99.5
96	3.27	98.0
192	1.76	91.0
384	1.23	64.8

Table 5.2: The DC-xTB-MD calculation results using massively parallel computer "Fugaku".

System	**# nodes**	**# processes**	**# threads**	**Average time (s/step)**
Myoglobin	48	192	12	46.8
Myoglobin+3000H_2O	384	768	24	226.2

Figure 5.3 shows the buffer-size dependence of the energy error from the standard xTB calculation and the computational time for the DC-xTB calculation of the 3000 water molecule system. d_{cent} was fixed at 7.0 Å. The computational time here was measured using a single node computer equipped with two Intel® Xeon® Gold 6326 processors. For $r_{\mathrm{buf}} = 3.5$ Å, the energy error is more than 15 kcal mol^{-1}. Increasing the buffer size increases the computational time slightly, but the energy error decreases rapidly. Actually, the energy error becomes less than 1 kcal mol^{-1} for $r_{\mathrm{buf}} = 5.0$ Å, while the time increase from $r_{\mathrm{buf}} = 3.5$ Å is 47.8%.

Note that the time for the standard xTB calculation was 163.0 s, which is more than three times longer than the DC-xTB time for $r_{\mathrm{buf}} = 5.0$ Å.

Finally, I performed the DC-xTB-MD simulation of myoglobin protein, containing 2226 atoms, on Fugaku. Here, I set $d_{\mathrm{cent}} = 5.0$ Å and $r_{\mathrm{buf}} = 7.0$ Å. The initial structure was taken from the protein data bank (PDBID: 5MBA). The solution-phase simulation was additionally performed by placing 3000 water molecules around the myoglobin protein and running under the cubic confined potential. The arrangement of the water molecules was carried out using the Winmostar program connected to the GROMACS package. Table 5.2 summarizes the computational time per one MD step together with the calculation conditions.

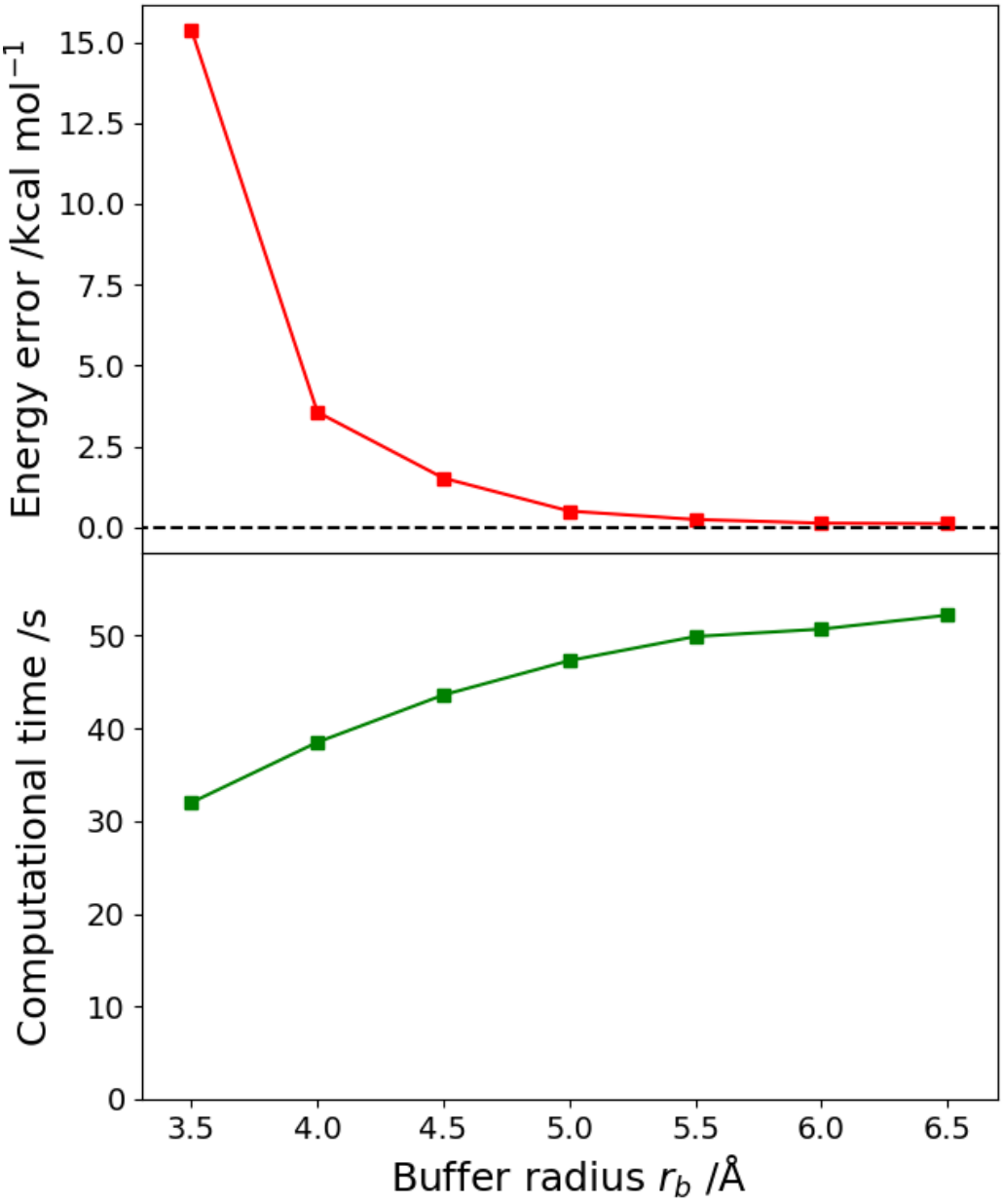

Figure 5.3: (Upper) The DC-xTB energy error with buffer radius for water 3000 molecules. The reference value is the xTB calculation. (Lower) The relationship between the buffer radius and its computational time per one SCC iteration for water 3000 molecules.

5.5 Conclusion

Tremendous-scale systems such as amorphous and viruses are untractable in quantum chemical calculation. Semi-empirical methods such as the DFTB are insufficient for exceeding tens of thousands of atom systems. Thus, for further speed up, the previous study combined the DFTB with the DC scheme, one of the linear scaling methods, and developed the DC-DFTB method. The DC-DFTB method enables very large-scale calculations. However, its parameter format is complicated, so available elements are limited. I focused on the GFN-xTB method, which has developed as a general-purpose semi-empirical method. I combined the xTB method with the DC scheme and developed the DC-xTB method as a general-purpose, enormously large-scale quantum calculation method.

I implemented the DC-xTB energy and its gradient for quantum MD calculation. I confirmed the DC-xTB can make computational cost linearly, and its accuracy is tunable by the size of the buffer region.

The massively parallel computation can draw out the maximum power of the DC method because the DC scheme can handle each subsystem independently. Thus, I investigated the DC-xTB calculation with massively parallel computation in a supercomputer. I confirmed its parallel efficiency in strong scaling achieved a high score. Finally, the DC-xTB-MD calculation for myoglobin molecule

surrounding water, which has more than tens of thousands of atoms, indicated that very large-scale quantum MD simulation with a few minutes per MD step is possible. In the future, I will extend to periodic boundary conditions for amorphous.

Chapter 6

General Conclusion

In this thesis, I worked on the black-box accurate calculation and the general-purpose very large-scale quantum chemical calculation method.

In Chapter 3, I applied the HFB method to the excited state calculation, which includes static correlation in a black-box manner. In this thesis, the RT method is adopted as the solution of the TDHFB equation. However, the number of electrons is varied in time propagation. Thus, the "canonical orbital" diagonalizing the electron density matrix avoids the change of the electron number. Although the Cb-TDHFB calculation gave the excited state, including the static correlation, the unphysical states also appeared. That is why the TDHFB wavefunction is not an eigenfunction of the electron number operator. The projection scheme recovers the symmetry for such broken symmetry wavefunctions. In the future, I will apply the projected HFB wave function to the excited state calculation.

In Chapter 4, the PUHF method, which removes spin-contaminated

states from the UHF wave function, is applied to the DC scheme. First, I developed methods to divide the effective Fock matrix into local ones and to construct the total transition density matrix from local ones. The test calculations for biradical molecules showed that the DC-PUHF method reproduces the PUHF calculation results if the appropriate buffer regions are adopted. I also measured the computation time for diagonalization and inverse matrix computation and found that the cost is about $O(N^3)$ for the PUHF method, but almost linear for the DC-PUHF method.

In Chapter 5, I applied the GFN-xTB method, a general-purpose semi-empirical method, to the DC scheme. The DC-xTB energy calculation achieved linear scaling. In addition, the larger the buffer region, the more accurate energies are obtained. Since the DC method can treat each subsystem independently, the parallel computation can accelerate the DC calculation. The massively parallel DC-xTB calculations using the supercomputer "Fugaku" are performed. Using more than 100 nodes, high parallel efficiency is achieved. Finally, I performed the large-scale quantum MD calculations using the DC-xTB energy gradient. For protein molecules, DC-xTB-MD calculations using a massively parallel computer can achieve the fast quantum MD, whose cost per step is within several minutes.

In this thesis, I have described large-scale and highly accurate electronic structure calculations. I hope that this research will contribute to the future development of theoretical chemistry.

Bibliography

[1] Carlo Adamo and Vincenzo Barone. 'Toward reliable density functional methods without adjustable parameters: The PBE0 model'. In: *The Journal of chemical physics* 110.13 (1999), pp. 6158–6170 (cit. on p. 10).

[2] Tomoko Akama, Atsuhiko Fujii, Masato Kobayashi and Hiromi Nakai. 'Is the divide-and-conquer Hartree–Fock method valid for calculations of delocalized systems?' In: *Molecular Physics* 105.19-22 (2007), pp. 2799–2804 (cit. on p. 34).

[3] Tomoko Akama, Yutaka Imamura and Hiromi Nakai. 'Application of real-time time-dependent density functional theory with the CVB3LYP functional to Core excitations'. In: *Chemistry letters* 39.4 (2010), pp. 407–409 (cit. on p. 46).

[4] Tomoko Akama and Hiromi Nakai. 'Short-time Fourier transform analysis of real-time time-dependent Hartree–Fock and time-dependent density functional theory calculations with Gaussian basis functions'. In: *The Journal of chemical physics* 132.5 (2010) (cit. on p. 46).

[5] Alberto Baiardi and Markus Reiher. 'The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges'. In: *The Journal of Chemical Physics* 152.4 (2020) (cit. on p. 44).

[6] Christoph Bannwarth, Eike Caldeweyher, Sebastian Ehlert, Andreas Hansen, Philipp Pracht, Jakob Seibert, Sebastian Spicher and Stefan Grimme. 'Extended tight-binding quantum chemistry methods'. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 11.2 (2021), e1493 (cit. on pp. 41, 97).

[7] Christoph Bannwarth, Sebastian Ehlert and Stefan Grimme. 'GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions'. In: *Journal of chemical theory and computation* 15.3 (2019), pp. 1652–1671 (cit. on pp. 41, 97).

[8] Christoph Bannwarth, Jimmy K Yu, Edward G Hohenstein and Todd J Martnez. 'Hole–hole Tamm–Dancoff-approximated density functional theory: A highly efficient electronic structure method incorporating dynamic and static correlation'. In: *The Journal of Chemical Physics* 153.2 (2020) (cit. on pp. 45, 46).

[9] Mark E Casida and Miquel Huix-Rotllant. 'Progress in time-dependent density-functional theory'. In: *Annual review of physical chemistry* 63.1 (2012), pp. 287–323 (cit. on p. 12).

[10] Mark E Casida and Miquel Huix-Rotllant. 'Progress in time-dependent density-functional theory'. In: *Annual review of physical chemistry* 63.1 (2012), pp. 287–323 (cit. on p. 46).

[11] Chien-Pin Chou, Aditya Wibawa Sakti, Yoshifumi Nishimura and Hiromi Nakai. 'Development of Divide-and-Conquer Density-Functional Tight-Binding Method for Theoretical Research on Li-Ion Battery'. In: *The Chemical Record* 19.4 (2019), pp. 746–757 (cit. on p. 96).

[12] Ji ek. 'On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods'. In: *The Journal of Chemical Physics* 45.11 (1966), pp. 4256–4266 (cit. on p. 2).

[13] AJ Coleman. 'Structure of fermion density matrices. II. Antisymmetrized geminal powers'. In: *Journal of Mathematical Physics* 6.9 (1965), pp. 1425–1431 (cit. on p. 45).

[14] Gábor Csányi and TA Arias. 'Tensor product expansions for correlation in quantum many-body systems'. In: *Physical Review B* 61.11 (2000), p. 7348 (cit. on p. 46).

[15] Paul Adrien Maurice Dirac. 'Quantum mechanics of many-electron systems'. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 123.792 (1929), pp. 714–733 (cit. on p. 1).

[16] Shuichiro Ebata. 'Studies on nuclear structure and nuclear dynamics using Cb-TDHFB'. In: *Frontiers in Physics* 8 (2020), p. 102 (cit. on p. 47).

[17] Shuichiro Ebata, Takashi Nakatsukasa and Tsunenori Inakura. 'Systematic investigation of low-lying dipole modes using the canonical-basis time-dependent Hartree-Fock-Bogoliubov theory'. In: *Phys. Rev. C* 90 (2 Aug. 2014), p. 024303. DOI: 10.1103/PhysRevC.90.024303. URL: https://link.aps.org/doi/10.1103/PhysRevC.90.024303 (cit. on p. 47).

[18] Shuichiro Ebata, Takashi Nakatsukasa, Tsunenori Inakura, Kenichi Yoshida, Yukio Hashimoto and Kazuhiro Yabana. 'Canonical-basis time-dependent Hartree-Fock-Bogoliubov theory and linear-response calculations'. In: *Phys. Rev. C* 82 (3 Sept. 2010), p. 034306. DOI: 10.1103/PhysRevC.82.034306. URL: https://link.aps.org/doi/10.1103/PhysRevC.82.034306 (cit. on p. 47).

[19] Jason K Ellis, Carlos A Jiménez-Hoyos, Thomas M Henderson, Takashi Tsuchimochi and Gustavo E Scuseria. 'Constrained-pairing mean-field theory. V. Triplet pairing formalism'. In: *The Journal of chemical physics* 135.3 (2011) (cit. on p. 46).

[20] Dmitri G Fedorov and Kazuo Kitaura. 'Coupled-cluster theory based upon the fragment molecular-orbital method'. In: *The Journal of chemical physics* 123.13 (2005) (cit. on p. 69).

[21] Dmitri G Fedorov and Kazuo Kitaura. 'Multiconfiguration self-consistent-field theory based upon the fragment molecular orbital method'. In: *The Journal of chemical physics* 122.5 (2005) (cit. on p. 69).

[22] Dmitri G Fedorov and Kazuo Kitaura. 'Second order Møller-Plesset perturbation theory based upon the fragment molecular orbital method'. In: *The Journal of Chemical Physics* 121.6 (2004), pp. 2483–2490 (cit. on p. 69).

[23] Toshikazu Fujimori, Masato Kobayashi and Tetsuya Taketsugu. 'Energy-based automatic determination of buffer region in the divide-and-conquer second-order Møller–Plesset perturbation theory'. In: *Journal of computational chemistry* 42.9 (2021), pp. 620–629 (cit. on pp. 36, 100).

[24] KJH Giesbertz, EJ Baerends and OV Gritsenko. 'Charge Transfer, Double and Bond-Breaking Excitations<? format?> with Time-Dependent Density Matrix Functional Theory'. In: *Physical review letters* 101.3 (2008), p. 033004 (cit. on p. 46).

[25] KJH Giesbertz, K Pernal, OV Gritsenko and EJ Baerends. 'Excitation energies with time-dependent density matrix functional theory: Singlet two-electron systems'. In: *The Journal of chemical physics* 130.11 (2009) (cit. on pp. 46, 51).

[26] Joshua J Goings, Patrick J Lestrange and Xiaosong Li. 'Real-time time-dependent electronic structure theory'. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 8.1 (2018), e1341 (cit. on p. 46).

[27] Stefan Grimme, Jens Antony, Stephan Ehrlich and Helge Krieg. 'A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu'. In: *The Journal of chemical physics* 132.15 (2010) (cit. on p. 100).

[28] Stefan Grimme, Christoph Bannwarth and Philip Shushkov. 'A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements (Z= 1–86)'. In: *Journal of chemical theory and computation* 13.5 (2017), pp. 1989–2009 (cit. on pp. 41, 97, 100).

[29] Stefan Grimme, Jan Gerit Brandenburg, Christoph Bannwarth and Andreas Hansen. 'Consistent structures and interactions by density functional theory with small atomic orbital basis sets'. In: *The Journal of chemical physics* 143.5 (2015) (cit. on p. 10).

[30] George G Hall. 'The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials'. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 205.1083 (1951), pp. 541–552 (cit. on p. 8).

[31] Micha Hapka, Katarzyna Pernal and Hans Jørgen Aa Jensen. 'An efficient implementation of time-dependent linear-response theory for strongly orthogonal geminal wave function models'. In: *The Journal of Chemical Physics* 156.17 (2022) (cit. on p. 45).

[32] John E Harriman. 'Natural Expansion of the First-Order Density Matrix for a Spin-Projected Single Determinant'. In: *The Journal of Chemical Physics* 40.10 (1964), pp. 2827–2839 (cit. on p. 30).

[33] Ryo Hatada, Koji Okuwaki, Yuji Mochizuki, Yuma Handa, Kaori Fukuzawa, Yuto Komeiji, Yoshio Okiyama and Shigenori Tanaka. 'Fragment molecular orbital based interaction analyses on COVID-19 main protease- inhibitor N3 complex (PDB ID: 6LU7)'. In: *Journal of chemical information and modeling* 60.7 (2020), pp. 3593–3602 (cit. on p. 95).

[34] Thomas M Henderson and Gustavo E Scuseria. 'Linearized Jastrow-style fluctuations on spin-projected Hartree-Fock'. In: *The Journal of chemical physics* 139.23 (2013) (cit. on pp. 23, 28).

[35] Roald Hoffmann. 'An extended Hückel theory. I. hydrocarbons'. In: *The Journal of Chemical Physics* 39.6 (1963), pp. 1397–1412 (cit. on p. 37).

[36] Erich Hückel. 'Zur quantentheorie der doppelbindung'. In: *Zeitschrift für Physik* 60.7 (1930), pp. 423–456 (cit. on p. 37).

[37] Carlos A Jiménez-Hoyos, Thomas M Henderson, Takashi Tsuchimochi and Gustavo E Scuseria. 'Projected hartree–fock theory'. In: *The Journal of chemical physics* 136.16 (2012) (cit. on pp. 23, 27, 70).

[38] Paul A Johnson, Paul W Ayers, Peter A Limacher, Stijn De Baerdemacker, Dimitri Van Neck and Patrick Bultinck. 'A size-consistent approach to strongly correlated systems using a generalized antisymmetrized product of nonorthogonal geminals'. In: *Computational and Theoretical Chemistry* 1003 (2013), pp. 101–113 (cit. on p. 45).

[39] Michio Katouda. 'Application of resolution of identity approximation of second-order Møller–Plesset perturbation theory to three-body fragment molecular orbital method'. In: *Theoretical Chemistry Accounts* 130 (2011), pp. 449–453 (cit. on p. 69).

[40] Kazuo Kitaura, Eiji Ikeo, Toshio Asada, Tatsuya Nakano and Masami Uebayasi. 'Fragment molecular orbital method: an approximate computational method for large molecules'. In: *Chemical Physics Letters* 313.3-4 (1999), pp. 701–706 (cit. on pp. 4, 32).

[41] Masato Kobayashi. 'Gradient of molecular Hartree–Fock–Bogoliubov energy with a linear combination of atomic orbital quasiparticle wave functions'. In: *The Journal of Chemical Physics* 140.8 (2014) (cit. on p. 18).

[42] Masato Kobayashi, Toshikazu Fujimori and Tetsuya Taketsugu. 'Automated error control in divide-and-conquer self-consistent field calculations'. In: *Journal of computational chemistry* 39.15 (2018), pp. 909–916 (cit. on pp. 36, 100).

[43] Masato Kobayashi, Toshikazu Fujimori and Tetsuya Taketsugu. 'Automatic Determination of Buffer Region in Divide-anc-Conquer Quantum Chemical Calculations'. In: *Journal of Computer Chemistry, Japan* 20.2 (2021), pp. 48–59 (cit. on pp. 36, 100).

[44] Masato Kobayashi, Yutaka Imamura and Hiromi Nakai. 'Alternative linear-scaling methodology for the second-order Møller-Plesset perturbation calculation based on the divide-and-conquer method'. In: *The Journal of chemical physics* 127.7 (2007) (cit. on pp. 34, 69).

[45] Masato Kobayashi, Tomotaka Kunisada, Tomoko Akama, Daisuke Sakura and Hiromi Nakai. 'Reconsidering an analytical gradient expression within a divide-and-conquer self-consistent field approach: Exact formula and its approximate treatment'. In: *The Journal of Chemical Physics* 134.3 (2011) (cit. on pp. 34, 96).

[46] Masato Kobayashi and Hiromi Nakai. 'An effective energy gradient expression for divide-and-conquer second-order Møller–Plesset perturbation theory'. In: *The Journal of Chemical Physics* 138.4 (2013) (cit. on p. 34).

[47] Masato Kobayashi and Hiromi Nakai. 'Extension of linear-scaling divide-and-conquer-based correlation method to coupled cluster theory with singles and doubles excitations'. In: *The Journal of chemical physics* 129.4 (2008) (cit. on pp. 34, 69).

[48] Masato Kobayashi and Tetsuya Taketsugu. 'Divide-and-Conquer Hartree–Fock–Bogoliubov Method and Its Application to Conjugated Diradical Systems'. In: *Chemistry Letters* 45.11 (2016), pp. 1268–1270 (cit. on pp. 17, 70, 75).

[49] Masato Kobayashi, Takeshi Yoshikawa and Hiromi Nakai. 'Divide-and-conquer self-consistent field calculation for open-shell systems: Implementation and application'. In: *Chemical Physics Letters* 500.1-3 (2010), pp. 172–177 (cit. on p. 34).

[50] Nobuaki Koga, Koichi Yamashita and Keiji Morokuma. 'On incorrect behavior of single annihilation equations of spin-projected UHF and UMP energies'. In: *Chemical physics letters* 184.4 (1991), pp. 359–362 (cit. on p. 22).

[51] Walter Kohn. 'Density functional and density matrix method scaling linearly with the number of atoms'. In: *Physical Review Letters* 76.17 (1996), p. 3168 (cit. on pp. 31, 69).

[52] Walter Kohn and Lu Jeu Sham. 'Self-consistent equations including exchange and correlation effects'. In: *Physical review* 140.4A (1965), A1133 (cit. on pp. 4, 9).

[53] Nana Komoto, Takeshi Yoshikawa, Yoshifumi Nishimura and Hiromi Nakai. 'Large-scale molecular dynamics simulation for ground and excited states based on divide-and-conquer long-range corrected density-functional tight-binding method'. In: *Journal of chemical theory and computation* 16.4 (2020), pp. 2369–2378 (cit. on p. 97).

[54] Nana Komoto, Takeshi Yoshikawa, Junichi Ono, Yoshifumi Nishimura and Hiromi Nakai. 'Development of large-scale excited-state calculations based on the divide-and-conquer time-dependent density functional tight-binding method'. In: *Journal of chemical theory and computation* 15.3 (2019), pp. 1719–1727 (cit. on p. 96).

[55] Jörg Kussmann, Matthias Beer and Christian Ochsenfeld. 'Linear-scaling self-consistent field methods for large molecules'. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 3.6 (2013), pp. 614–636 (cit. on p. 34).

[56] Werner Kutzelnigg. 'On the validity of the electron pair approximation for the Beryllium ground state'. In: *Theoretica chimica acta* 3.3 (1965), pp. 241–253 (cit. on p. 45).

[57] Michael Lass, Robert Schade, Thomas D Kühne and Christian Plessl. 'A submatrix-based method for approximate matrix function evaluation in the quantum chemistry code CP2K'. In: *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE. 2020, pp. 1–14 (cit. on p. 97).

[58] X-P Li, Ricardo W Nunes and David Vanderbilt. 'Density-matrix electronic-structure method with linear system-size scaling'. In: *Physical Review B* 47.16 (1993), p. 10891 (cit. on p. 34).

[59] X-P Li, Ricardo W Nunes and David Vanderbilt. 'Density-matrix electronic-structure method with linear system-size scaling'. In: *Physical Review B* 47.16 (1993), p. 10891 (cit. on p. 97).

[60] Per-Olov Löwdin. 'Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects'. In: *Physical review* 97.6 (1955), p. 1509 (cit. on p. 10).

[61] István Mayer. 'The spin-projected extended Hartree-Fock method'. In: *Advances in quantum chemistry*. Vol. 12. Elsevier, 1980, pp. 189–262 (cit. on p. 22).

[62] David A Mazziotti. 'Energy functional of the one-particle reduced density matrix: a geminal approach'. In: *Chemical physics letters* 338.4-6 (2001), pp. 323–328 (cit. on p. 45).

[63] Chr Møller and Milton S Plesset. 'Note on an approximation treatment for many-electron systems'. In: *Physical review* 46.7 (1934), p. 618 (cit. on pp. 2, 11).

[64] Hiromi Nakai, Masato Kobayashi, Takeshi Yoshikawa, Junji Seino, Yasuhiro Ikabata and Yoshifumi Nishimura. 'Divide-and-Conquer Linear-Scaling Quantum Chemical Computations'. In: *The Journal of Physical Chemistry A* 127.3 (2023), pp. 589–618 (cit. on p. 70).

[65] Hiromi Nakai, Aditya Wibawa Sakti and Yoshifumi Nishimura. 'Divide-and-conquer-type density-functional tight-binding molecular dynamics simulations of proton diffusion in a bulk water system'. In: *The Journal of Physical Chemistry B* 120.1 (2016), pp. 217–221 (cit. on p. 96).

[66] H Nakatsuji and K Hirao. 'Cluster expansion of the wavefunction. Symmetry-adapted-cluster expansion, its variational determination, and extension of open-shell orbital theory'. In: *The Journal of Chemical Physics* 68.5 (1978), pp. 2053–2065 (cit. on p. 2).

[67] Eric Neuscamman. 'Size Consistency Error in the Antisymmetric Geminal Power Wave Function<? format?> can be Completely Removed'. In: *Physical review letters* 109.20 (2012), p. 203001 (cit. on pp. 23, 28).

[68] Masatsugu Nishida, Tomoko Akama, Masato Kobayashi and Tetsuya Taketsugu. 'Time-dependent Hartree–Fock–Bogoliubov method for molecular systems: An alternative excited-state methodology including static electron correlation'. In: *Chemical Physics Letters* 816 (2023), p. 140386 (cit. on p. 51).

[69] Yoshifumi Nishimura and Hiromi Nakai. 'Parallel implementation of efficient charge–charge interaction evaluation scheme in periodic divide-and-conquer density-functional tight-binding calculations'. In: *Journal of Computational Chemistry* 39.2 (2018), pp. 105–116 (cit. on p. 96).

[70] Yoshifumi Nishimura and Hiromi Nakai. 'Quantum chemical calculations for up to one hundred million atoms using DCDFTBMD code on supercomputer Fugaku'. In: *Chemistry Letters* 50.8 (2021), pp. 1546–1550 (cit. on p. 97).

[71] Hiroaki Nishizawa, Yoshifumi Nishimura, Masato Kobayashi, Stephan Irle and Hiromi Nakai. 'Three pillars for achieving quantum mechanical molecular dynamics simulations of huge systems: Divide-and-conquer, density-functional tight-binding, and massively parallel computation'. In: *Journal of computational chemistry* 37.21 (2016), pp. 1983–1992 (cit. on pp. 96, 102, 103).

[72] 5_R W Nunes and David Vanderbilt. 'Generalization of the density-matrix method to a nonorthogonal basis'. In: *Physical Review B* 50.23 (1994), p. 17611 (cit. on p. 34).

[73] Roberto Olivares-Amaya, Weifeng Hu, Naoki Nakatani, Sandeep Sharma, Jun Yang and Garnet Kin Chan. 'The ab-initio density matrix renormalization group in practice'. In: *The Journal of chemical physics* 142.3 (2015) (cit. on p. 44).

[74] Rudolph Pariser and Robert G Parr. 'A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Molecules. I.' In: *The Journal of Chemical Physics* 21.3 (1953), pp. 466–471 (cit. on pp. 38, 96).

[75] Degao Peng, Helen van Aggelen, Yang Yang and Weitao Yang. 'Linear-response time-dependent density-functional theory with pairing fields'. In: *The Journal of Chemical Physics* 140.18 (2014) (cit. on pp. 45, 46).

[76] John P Perdew, Kieron Burke and Matthias Ernzerhof. 'Generalized gradient approximation made simple'. In: *Physical review letters* 77.18 (1996), p. 3865 (cit. on p. 10).

[77] Katarzyna Pernal, Koushik Chatterjee and Piotr H Kowalski. 'How accurate is the strongly orthogonal geminal theory in predicting excitation energies? Comparison of the extended random phase approximation and the linear response theory approaches'. In: *The Journal of Chemical Physics* 140.1 (2014) (cit. on p. 45).

[78] Katarzyna Pernal and Klaas JH Giesbertz. 'Reduced density matrix functional theory (RDMFT) and linear response time-dependent RDMFT (TD-RDMFT)'. In: *Density-Functional Methods for Excited States* (2016), pp. 125–183 (cit. on p. 46).

[79] John A Pople. 'Electron interaction in unsaturated hydrocarbons'. In: *Transactions of the Faraday Society* 49 (1953), pp. 1375–1385 (cit. on pp. 38, 96).

[80] John A Pople and Robert K Nesbet. 'Self-consistent orbitals for radicals'. In: *The Journal of Chemical Physics* 22.3 (1954), pp. 571–572 (cit. on p. 15).

[81] Dirk Porezag, Th Frauenheim, Th Köhler, Gotthard Seifert and R Kaschner. 'Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon'. In: *Physical Review B* 51.19 (1995), p. 12947 (cit. on pp. 39, 96).

[82] Emil Prodan and Walter Kohn. 'Nearsightedness of electronic matter'. In: *Proceedings of the National Academy of Sciences* 102.33 (2005), pp. 11635–11638 (cit. on pp. 31, 69).

[83] Yiheng Qiu, Thomas M Henderson, Jinmo Zhao and Gustavo E Scuseria. 'Projected coupled cluster theory'. In: *The Journal of Chemical Physics* 147.6 (2017) (cit. on p. 31).

[84] Vitaly A Rassolov. 'A geminal model chemistry'. In: *The Journal of chemical physics* 117.13 (2002), pp. 5978–5987 (cit. on p. 45).

[85] Vitaly A Rassolov and Feng Xu. 'Geminal model chemistry III: Partial spin restriction'. In: *The Journal of chemical physics* 126.23 (2007) (cit. on p. 45).

[86] Vitaly A Rassolov and Feng Xu. 'Geminal model chemistry. IV. Variational and size consistent pure spin states'. In: *The Journal of chemical physics* 127.4 (2007) (cit. on p. 45).

[87] Vitaly A Rassolov, Feng Xu and Sophya Garashchuk. 'Geminal model chemistry II. Perturbative corrections'. In: *The Journal of chemical physics* 120.22 (2004), pp. 10385–10394 (cit. on p. 45).

[88] P. Ring and P. Schuck. *The Nuclear Many-Body Problem*. Physics and astronomy online library. Springer, 2004. ISBN: 9783540212065. URL: https://books.google.co.jp/books?id=PTynSM-nMA8C (cit. on pp. 22, 45).

[89] S.R.P.G. Ripka, J.P. Blaizot and G. Ripka. *Quantum Theory of Finite Systems*. MIT Press, 1986. ISBN: 9780262022149. URL: https://books.google.co.jp/books?id=s_xlQgAACAAJ (cit. on pp. 45, 46).

[90] Pablo Rivero, Carlos A Jiménez-Hoyos and Gustavo E Scuseria. 'Entanglement and polyradical character of polycyclic aromatic hydrocarbons predicted by projected Hartree–Fock theory'. In: *The Journal of Physical Chemistry B* 117.42 (2013), pp. 12750–12758 (cit. on p. 28).

[91] B.O. Roos, R. Lindh, P. Malmqvist, V. Veryazov and P.O. Widmark. *Multiconfigurational Quantum Chemistry*. Wiley, 2016. ISBN: 9780470633465. URL: https://books.google.co.jp/books?id=KFaQuAAACAAJ (cit. on p. 44).

[92] Björn O Roos et al. 'The complete active space self-consistent field method and its applications in electronic structure calculations'. In: *Advances in chemical physics* 69 (2007), pp. 399–445 (cit. on p. 44).

[93] C. C. J. Roothaan. 'New Developments in Molecular Orbital Theory'. In: *Rev. Mod. Phys.* 23 (2 Apr. 1951), pp. 69–89. DOI: 10.1103/RevModPhys.23.69. URL: https://link.aps.org/doi/10.1103/RevModPhys.23.69 (cit. on p. 8).

[94] Aditya Wibawa Sakti, Yoshifumi Nishimura and Hiromi Nakai. 'Divide-and-conquer-type density-functional tight-binding simulations of hydroxide ion diffusion in bulk water'. In: *The Journal of Physical Chemistry B* 121.6 (2017), pp. 1362–1371 (cit. on p. 96).

[95] Robert Schade, Tobias Kenter, Hossam Elgabarty, Michael Lass, Ole Schütt, Alfio Lazzaro, Hans Pabst, Stephan Mohr, Jürg Hutter, Thomas D Kühne et al. 'Towards electronic structure-based ab-initio molecular dynamics simulations with hundreds of millions of atoms'. In: *Parallel Computing* 111 (2022), p. 102920 (cit. on p. 97).

[96] H Bernhard Schlegel. 'Møller-Plesset perturbation theory with spin projection'. In: *The Journal of Physical Chemistry* 92.11 (1988), pp. 3075–3078 (cit. on p. 22).

[97] Michael W Schmidt, Kim K Baldridge, Jerry A Boatz, Steven T Elbert, Mark S Gordon, Jan H Jensen, Shiro Koseki, Nikita Matsunaga, Kiet A Nguyen, Shujun Su et al. 'General atomic and molecular electronic structure system'. In: *Journal of computational chemistry* 14.11 (1993), pp. 1347–1363 (cit. on p. 51).

[98] J Robert Schrieffer. *Theory of superconductivity*. CRC press, 2018 (cit. on p. 45).

[99] Gustavo E Scuseria, Carlos A Jiménez-Hoyos, Thomas M Henderson, Kousik Samanta and Jason K Ellis. 'Projected quasiparticle theory for molecular electronic structure'. In: *The Journal of chemical physics* 135.12 (2011) (cit. on pp. 23, 30, 45, 46, 66, 70).

[100] Gustavo E Scuseria and Takashi Tsuchimochi. 'Constrained-pairing mean-field theory. II. Exact treatment of dissociations to nondegenerate orbitals'. In: *The Journal of chemical physics* 131.16 (2009) (cit. on p. 46).

[101] Gotthard Seifert and Jan-Ole Joswig. 'Density-functional tight binding—an approximate density-functional theory method'. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2.3 (2012), pp. 456–465 (cit. on pp. 39, 96).

[102] Javid A Sheikh, J Dobaczewski, Peter Ring, Luis Miguel Robledo and Constantine Yannouleas. 'Symmetry restoration in mean-field approaches'. In: *Journal of Physics G: Nuclear and Particle Physics* 48.12 (2021), p. 123001 (cit. on p. 22).

[103] Javid A Sheikh and Peter Ring. 'Symmetry-projected hartree–fock–bogoliubov equations'. In: *Nuclear Physics A* 665.1-2 (2000), pp. 71–91 (cit. on p. 22).

[104] Oktay Sinanolu. 'Many-electron theory of atoms, molecules and their interactions'. In: *Advances in chemical physics* (1964), pp. 315–412 (cit. on pp. 2, 10).

[105] Viktor N Staroverov and Gustavo E Scuseria. 'Optimization of density matrix functionals by the Hartree–Fock–Bogoliubov method'. In: *The Journal of chemical physics* 117.24 (2002), pp. 11107–11112 (cit. on pp. 18, 45, 70).

[106] Christopher J Stein and Markus Reiher. 'autoCAS: A program for fully automated multiconfigurational calculations'. In: *Journal of computational chemistry* 40.25 (2019), pp. 2216–2226 (cit. on p. 44).

[107] Péter R Surján. 'An introduction to the theory of geminals'. In: *Correlation and localization* (1999), pp. 63–88 (cit. on p. 17).

[108] Péter R Surján. 'An introduction to the theory of geminals'. In: *Correlation and localization* (1999), pp. 63–88 (cit. on p. 44).

[109] PR Surján. 'Geminal approach'. In: (2016) (cit. on p. 44).

[110] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 1996 (cit. on p. 7).

[111] Moto Tarumi, Masato Kobayashi and Hiromi Nakai. 'Accelerating convergence in the antisymmetric product of strongly orthogonal geminals method'. In: *International Journal of Quantum Chemistry* 113.3 (2013), pp. 239–244 (cit. on p. 45).

[112] Walter Thiel. 'Semiempirical quantum–chemical methods'. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 4.2 (2014), pp. 145–157 (cit. on p. 96).

[113] Tsuguki Touma, Masato Kobayashi and Hiromi Nakai. 'Time-dependent Hartree–Fock frequency-dependent polarizability calculation applied to divide-and-conquer electronic structure method'. In: *Chemical Physics Letters* 485.1-3 (2010), pp. 247–252 (cit. on p. 34).

[114] Takashi Tsuchimochi, Thomas M Henderson, Gustavo E Scuseria and Andreas Savin. 'Constrained-pairing mean-field theory. IV. Inclusion of corresponding pair constraints and connection to unrestricted Hartree–Fock theory'. In: *The Journal of chemical physics* 133.13 (2010) (cit. on p. 46).

[115] Takashi Tsuchimochi, Yuto Mori and Seiichiro L Ten-no. 'Spin-projection for quantum computation: A low-depth approach to strong correlation'. In: *Physical Review Research* 2.4 (2020), p. 043142 (cit. on p. 31).

[116] Takashi Tsuchimochi and Gustavo E Scuseria. 'Strong correlations via constrained-pairing mean-field theory'. In: *The Journal of chemical physics* 131.12 (2009) (cit. on p. 46).

[117] Takashi Tsuchimochi, Gustavo E Scuseria and Andreas Savin. 'Constrained-pairing mean-field theory. III. Inclusion of density functional exchange and correlation effects via alternative densities'. In: *The Journal of chemical physics* 132.2 (2010) (cit. on p. 46).

[118] Takashi Tsuchimochi and Seiichiro Ten-no. 'Communication: Configuration interaction combined with spin-projection for strongly correlated molecular electronic structures'. In: *The Journal of Chemical Physics* 144.1 (2016) (cit. on p. 31).

[119] Takashi Tsuchimochi and Seiichiro L Ten-No. 'Second-order perturbation theory with spin-symmetry-projected Hartree–Fock'. In: *Journal of Chemical Theory and Computation* 15.12 (2019), pp. 6688–6702 (cit. on p. 31).

[120] Takashi Tsuchimochi and Troy Van Voorhis. 'Time-dependent projected Hartree-Fock'. In: *The Journal of Chemical Physics* 142.12 (2015) (cit. on p. 31).

[121] Oliver T Unke, Stefan Chmiela, Huziel E Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T Schütt, Alexandre Tkatchenko and Klaus-Robert Müller. 'Machine learning force fields'. In: *Chemical Reviews* 121.16 (2021), pp. 10142–10186 (cit. on p. 96).

[122] Hiroki Uratani, Chien-Pin Chou and Hiromi Nakai. 'Quantum mechanical molecular dynamics simulations of polaron formation in methylammonium lead iodide perovskite'. In: *Physical Chemistry Chemical Physics* 22.1 (2020), pp. 97–106 (cit. on p. 96).

[123] K Yabana, T Nakatsukasa, J-I Iwata and GF Bertsch. 'Real-time, real-space implementation of the linear response time-dependent density-functional theory'. In: *physica status solidi (b)* 243.5 (2006), pp. 1121–1138 (cit. on p. 46).

[124] K. Yabana and G. F. Bertsch. 'Time-dependent local-density approximation in real time'. In: *Phys. Rev. B* 54 (7 Aug. 1996), pp. 4484–4487. DOI: 10.1103/PhysRevB.54.4484. URL: https://link.aps.org/doi/10.1103/PhysRevB.54.4484 (cit. on p. 46).

[125] ChiYung Yam, Satoshi Yokojima and GuanHua Chen. 'Linear-scaling time-dependent density-functional theory'. In: *Phys. Rev. B* 68 (15 Oct. 2003), p. 153105. DOI: 10.1103/PhysRevB.68.153105. URL: https://link.aps.org/doi/10.1103/PhysRevB.68.153105 (cit. on p. 47).

[126] Takeshi Yanai, Masaaki Saitow, Xiao-Gen Xiong, Jakub Chalupský, Yuki Kurashige, Sheng Guo and Sandeep Sharma. 'Multistate complete-active-space second-order perturbation theory based on density matrix renormalization group reference states'. In: *Journal of chemical theory and computation* 13.10 (2017), pp. 4829–4840 (cit. on p. 44).

[127] Weitao Yang and Tai-Sung Lee. 'A density-matrix divide-and-conquer approach for electronic structure calculations of large molecules'. In: *The Journal of Chemical Physics* 103.13 (1995), pp. 5674–5678 (cit. on pp. 4, 34, 96).

[128] Takeshi Yoshikawa, Masato Kobayashi, Atsuhiko Fujii and Hiromi Nakai. 'Novel approach to excited-state calculations of large molecules based on divide-and-conquer method: application to photoactive yellow protein'. In: *The Journal of Physical Chemistry B* 117.18 (2013), pp. 5565–5573 (cit. on p. 34).

[129] Robert Zaleny, Manthos G Papadopoulos, Paul G Mezey and Jerzy Leszczynski. *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*. Vol. 13. Springer Science & Business Media, 2011 (cit. on p. 69).

[130] Robert Zaleny, Manthos G Papadopoulos, Paul G Mezey and Jerzy Leszczynski. *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*. Vol. 13. Springer Science & Business Media, 2011 (cit. on pp. 95, 96).

[131] Qingsheng Zhao and Weitao Yang. 'Analytical energy gradients and geometry optimization in the divide-and-conquer method for large molecules'. In: *The Journal of chemical physics* 102.24 (1995), pp. 9598–9603 (cit. on pp. 34, 96).

Chapter 7

Acknowledgement

I would like to thank my supervisor, Prof. Tetsuya Taketsugu. He watched over me with a warm heart even when my research was not going well. I was also very impressed with his attitude towards science and nature. He was the most caring teacher I have ever met. I want to express my gratitude to Assoc. Prof. Masato Kobayashi, who is the actual supervisor for me. He is tolerant of my laziness and inadequate research skills and gave me appropriate advice when I struggled with research. He also allowed me to experience various research topics and helped me grow as a researcher. I would like to express my appreciation to Asst. Prof. Tomoko Akama. She patiently taught me how to code the quantum chemistry calculation program, even though I was not very good at remembering things. Other staff members, Asst. Prof. Takeshi Iwasa, Assoc. Prof. Andley Lyalin, Asst. Prof. Kenichiro Saita, Asst. Prof. Yuriko Ono, Visiting Prof. Takao Tsuneda, Dr. Sonu Kumar, Dr. Keisuke Tashiro, and Dr. Satoi Wada made appropriate comments at the seminar and always encouraged me.

In Chapter 3, I developed and implemented the Cb-TDHFB method. Assoc. Prof. Masato Kobayashi and Asst. Prof. Tomoko Akama taught me about the HFB method and RT method. In addition, they gave me a way to code the GAMESS program.

In Chapter 4, I developed the DC-xTB method and implemented its program. At that time, I developed the program based on the DCDFTBK program developed by the Nakai laboratory at Waseda University. I would like to thank all those involved in the development of the DCDFTBK program. In addition, I appreciate my junior colleague, Kotaro Fujiwara, with whom I worked on program development. His very unique personality always kept me excited in the laboratory. He was the first student I directly supervised and taught me many important things about working with colleagues. The experience of working with him has been useful for my subsequent research with younger students.

Mr. Gen Ogawa worked together on many projects, including the development of a method for calculating electron correlation in large-scale periodic systems, the development of a method for calculating large-scale electronic states in solvents, and joint research with a company. He has a very broad range of interests, which exposed me to a variety of methods and theories. He is very passionate about research and is always able to do more than what I taught him. I think the experience of working with him has helped me grow as a researcher.

In Chapter 5, I developed the DC-PUHF method. Mr. Sousei

Kasaya worked with me on the development and implementation of the DC-PUHF program. He was the first student to work on my proposed theme, and I think I improved my teaching skills while working with him.

I am indebted to many senior colleagues, especially Dr. Toshikazu Fujimori, Mr. Yuta Oba, and Dr. Shuichi Ebisawa. Dr. Fujimori showed me the behavior as the head of the research group to make the students come alive and do research! Mr. Oba kindly taught me about research and the laboratory when I was a fourth-year undergraduate student. Dr. Ebisawa was a very talented senior student and I have worked hard to become a doctoral student like him. If I have been able to give back to my junior students what they have given me, even a little, I consider my doctoral life to be fortunate.

I am very grateful to the IYASAKA Foundation and the Kawamura Scholarship Foundation for the financial support they provided me when I am in my master course. Without this support, I would not have been able to pursue the doctoral course. I sent my thanks to the Ambitious program for smart materials science (SMatS) and the EXEX Doctoral Fellowship. These programs not only supported me financially, but also allowed me to gain a lot of experience and improve my skills.

Finally, I would like to thank my family. My paternal grandparents continued to support me in my life despite their old age. My younger sister and brother provided emotional support when I was

sick of everything. I am proud to be their older brother.