Lagrangian of extended multiconfigurational self-consistent field second-order quasidegenerate perturbation theory combined with reference interaction site model self-consistent field constraint spatial electron density

Naoki Negishi[1] and Daisuke Yokogawa[1, a)]

Department of Arts and Sciences, University of Tokyo

(Dated: 5 January 2024)

Lagrangians of the SA multiconfigurational self-consistent field(SA-MCSCF) and multi-state extended second-order quasidegenerate perturbation theory (MS-XMCQDPT2) coupled with the reference interaction site model self-consistent field constraint spatial electron density (RISM-SCF-cSED) are defined. In addition, variational equations were derived to calculate the excitation energies of target molecules dissolved in various solvents. The theory was applied to a phenol molecule in various solutions and the gradients and Hessian matrices were calculated to evaluate the absorption spectral lines including the broadening bandwidth. Numerical calculations revealed fine structures in any solvent surroundings. The main intramolecular vibrational modes related to such fine structures were stretching vibrations of the aromatic ring and the oxygen atom of the phenol molecule. The present theory plays an important role when predicting the structure of potential energy surfaces, such as Hessian matrices for various solvent types, during the photoexcitation process.

[a)]Electronic mail: c-d.yokogawa@mail.ecc.u-tokyo.ac.jp

I. INTRODUCTION

It is important to evaluate the free energy and electron density distribution of a molecule to improve the numerical performance of the condensed matter phase. Multiscale theories, such as quantum mechanics molecular mechanics (QM/MM)[1–3],frequently require a considerably more accurate electron density distribution of a target system to simulate the chemical phenomena with the chemical accuracy which are very sensitive to the chemical properties of the bath. For example, solvatochromic effects on the chemical reactions,[4,5] photoabsorption spectra,[6] and photoemission spectra[7] are typically observed experimentally. The solvent type occasionally controls such chemical properties.

Previous computational studies on multiscale theories provided significant evidence that the methodology of *ab initio* calculation and the theoretical modeling of bath surroundings are essential only in terms of the single point energy.[8–11] During photoexcitation processes, for example, charge transfer after electronic excitation typically leads to destabilization by electrostatic interaction between photoexcited molecules and solvents.[12] Thus, it is necessary to accurately the charge transfer calculated by *ab initio* and the electrostatic interaction field by theoretical modeling the solvent surroundings. In addition, energy fluctuations causing the peak broadenings of absorption spectral lines are dependent on the electron density.[13,14]

To simulate the photoexcitation process of target molecules, excited state calculations must be performed. Time-dependent density functional theory (TDDFT) is one of the most famous methods and is completed in combination with various solvation models such as the QM/MM, polarizable continuum model (PCM)[10,11], and reference interaction site model (RISM)[15–18]. TDDFT is an effective method in terms of the balance between its computational costs and accuracy. However, the quantitatively detailed analysis with chemical accuracy relative to experimental data appears to require a higher-level theory.

The multiconfigurational self-consistent field second-order quasidegenerate perturbation theory (XMCQDPT2) via the SA complete-active-space self-consistent field (SA-CASSCF)[19–21] is a powerful method in terms of precision. The computational frameworks of the SA-CASSCF and multistate XMCQDPT2 (MS-XMCQDPT2) combined with the RISM have already developed and exhibited good performances.[9,14,22,23] However, variational approaches have not yet been proposed. Thus, we present a theoretical framework for introducing the SA-CASSCF and the MS-XMCQDPT2 combined with the RISM self-consistent field constraint spatial electron density

(RISM-SCF-cSED) to satisfy the accurate descriptions of the electronic excitation process and external electrostatic field from the solvent surroundings.

The rest of the paper is organized as follows. Section II summarizes the theoretical framework and describes the variational principle of the Lagrangian of the SA-CASSCF combined with the RISM-SCF-cSED and that of the MS-XMCQDPT2 combined with its counterpart. Section III describes the computational details. Section IV describes the implementation of the proposed method applied to our previous study on the absorption spectral lines of a phenol molecule in various solutions. Finally, the numerical results are compared with the experimental data.

II. THEORY

A. EQUATIONS FOR STATE AVERAGED RISM-CASSCF

We defined the Lagrangian of the SA-CASSCF combined with the RISM-SCF-cSED to obtain variational equations that minimize the Lagrangian. We defined the SA energy in solution G_{SA} as,

$$G_{\mathrm{SA}} = \sum_{\alpha \in P} w_\alpha \{ \langle \alpha | \hat{H} | \alpha \rangle + {}^t\mathbf{V}^{(0)} (\mathbf{d}^{(\alpha)}_{\mathrm{CAS}} - \mathbf{d}^{(0)}_{\mathrm{CAS}}) \} + \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{CAS}}) \tag{1}$$

$$|\alpha\rangle = \sum_{I \in \mathrm{CSF}} C^I_\alpha |I\rangle . \tag{2}$$

G_{SA} denotes the sum of the free energy of the target state and the energies of the Franck-Condon states. The first term of Eq.(1) denotes the SA CAS energy, where α and I denote the CAS-configuration interaction (CI) state and the configuration state function (CSF), respectively.[20] Here, P in Eq.(1) denotes the CAS-CI space. The CAS-CI states comprise linear combinations of CSFs under the CI coefficient $\mathbf{C}$ in Eq.(2). w_α is the weight of the α-th SA-CASSCF wave-function and holds the sum rule as $\sum_{\alpha \in P} w_\alpha = 1$. $\hat{H}$ is the electronic Hamiltonian of the target molecule. The second term is the average of the interaction energies between the solute and solvent, which denotes the transition from the 0-th state to the α-th state. The column vector $\mathbf{V}^{(0)}$ is the electrostatic potential derived from the solvent structure as

$$\begin{aligned} V^{(0)}_\zeta &= \frac{\partial \Delta\mu}{\partial (\mathbf{d}^{(0)}_{\mathrm{CAS}})_\zeta} \\ &= -\sum_s \frac{\rho_s q_s}{\beta} \int d^3\mathbf{r} \int d^3\mathbf{r}' h_{\gamma s}(r) \frac{f_\zeta(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \ \mathrm{for}(\zeta \in \gamma) \end{aligned} \tag{3}$$

where $\mathbf{h}$ is the total correlation function between the atomic sites of the solute represented by γ and the solvent sites represented by s. f_ζ is the ζ-th auxiliary basis set (ABS) representing the electron

density distribution of the solute molecule under the RISM-SCF-cSED framework. $\boldsymbol{\rho}$ and $\mathbf{q}$ are the number density and charge of atoms in the solvent, respectively, and β is the inverse temperature. The $\Delta\mu(\mathbf{d}^{(0)})$ of the second term in Eq.(1) is the excess chemical potential in Ref.[24]. The RISM equation and closure relation are given in Eqs.(1) and (6) in Ref.[15] if using the Hyper-netted chain (HNC)[25], respectively. The interaction between the solute and solvent is given by Eqs.(4) and (6) in Ref.[26]. $\mathbf{d}_{\mathrm{CAS}}^{(\alpha)}$ is the column vector of the fitting coefficients denoting the electron density distribution of the α-th CAS-CI state under the RISM-SCF-cSED framework in Eq.(5) in Ref.[26]. $\alpha = 0$ represents the target state in which the solvent structure is relaxed.

The most important point herein is how to define the fitting coefficients $\mathbf{d}_{\mathrm{CAS}}^{(\alpha)}$. We define $\mathbf{d}_{\mathrm{CAS}}^{(\alpha)}$ as

$$\mathbf{d}_{\mathrm{CAS}}^{(\alpha)} = [\boldsymbol{\Xi} + \boldsymbol{\Gamma}(\mathbf{d}_{\mathrm{SA}})]^{-1} \sum_{pq}^{\mathrm{MO}} \mathbf{R}_{pq} \langle \alpha | \hat{a}_p^\dagger \hat{a}_q | \alpha \rangle \tag{4}$$

$$\mathbf{d}_{\mathrm{SA}} = \sum_{\alpha \in P} w_\alpha \mathbf{d}_{\mathrm{CAS}}^{(\alpha)}, \tag{5}$$

where matrices $\boldsymbol{\Xi}$ and $\boldsymbol{\Gamma}$ and the column vector $\mathbf{R}$ are defined in Ref.[24]. $\hat{a}_p^\dagger$ and $\hat{a}_p$ are creation and annihilation operators of the p-th molecular orbital (MO) respectively.

We defined the Lagrangian of the SA-CASSCF combined with the RISM-SCF-cSED $\mathscr{L}_{\mathrm{SA}}$ to minimize the SA free energy in Eq.(1).

$$\begin{aligned} \mathscr{L}_{\mathrm{SA}} &= \mathscr{L}_{\mathrm{SA}}[\mathbf{C}, \boldsymbol{\phi}, \boldsymbol{\Lambda}, \boldsymbol{\lambda}] \\ &= G_{\mathrm{SA}} + \sum_{\alpha,\beta \in P} \Lambda_{\alpha\beta} \left(\sum_{I \in \mathrm{CSF}} C_\alpha^I C_\beta^I - \delta_{\alpha\beta} \right) + \sum_{pq}^{\mathrm{MO}} \lambda_{pq} \left((p|q) - \delta_{pq} \right) \end{aligned} \tag{6}$$

where ${}^t\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots)$ denotes the vector of the MOs. $(p|q)$ denotes the overlap integral under the MO representation.[27] We assumed that the electrostatic potential $\mathbf{V}^{(0)}$ was rigid for the change in the electronic structure of the solute as

$$\frac{\partial \mathbf{V}^{(0)}}{\partial \mathbf{d}_{\mathrm{CAS}}^{(0)}} = 0. \tag{7}$$

Thereafter, we applied the variational principle to the $\mathscr{L}_{\mathrm{SA}}$ to optimize the CI coefficients and

orbitals. The differentials of the $\Delta\mu$ and $\mathbf{d}_{\mathrm{CAS}}^{(0)}$ using Eqs.(4) and (5) are given by

$$\frac{\partial \Delta\mu}{\partial C_\alpha^I} = {}^t\mathbf{V}^{(0)} \frac{\partial \mathbf{d}_{\mathrm{CAS}}^{(0)}}{\partial C_\alpha^I} \tag{8}$$

$$\frac{\partial \Delta\mu}{\partial \phi_p} = {}^t\mathbf{V}^{(0)} \frac{\partial \mathbf{d}_{\mathrm{CAS}}^{(0)}}{\partial \phi_p} \tag{9}$$

$$\frac{\partial \mathbf{d}_{\mathrm{SA}}}{\partial C_\alpha^I} = \left[\mathbf{\Xi} + (m-1)\mathbf{\Gamma}(\mathbf{d}_{\mathrm{SA}})\right]^{-1} w_\alpha \frac{\partial \langle\alpha|\hat{\mathbf{R}}|\alpha\rangle}{\partial C_\alpha^I} \tag{10}$$

$$\frac{\partial \mathbf{d}_{\mathrm{SA}}}{\partial \phi_p} = \left[\mathbf{\Xi} + (m-1)\mathbf{\Gamma}(\mathbf{d}_{\mathrm{SA}})\right]^{-1} \sum_{\alpha\in P} w_\alpha \frac{\partial \langle\alpha|\hat{\mathbf{R}}|\alpha\rangle}{\partial \phi_p}, \tag{11}$$

, respectively, where $\hat{\mathbf{R}}$ is equal to $\sum_{pq}^{\mathrm{MO}} \mathbf{R}_{pq}\hat{a}_p^\dagger \hat{a}_q$ and m is the parameter of the penalty matrix $\mathbf{\Gamma}(\mathbf{d}_{\mathrm{SA}}^{(0)})$ in Ref.[24]. Thus, the variational equations derived from the Lagrangian in Eq.(6) using Eqs.(7)–(11) are as follows.

$$\begin{aligned}\frac{\partial \mathscr{L}_{\mathrm{SA}}}{\partial C_\alpha^I} &= w_\alpha \sum_{J\in\mathrm{CSF}} \langle I|\hat{H}^{\mathrm{sol}}|J\rangle C_\alpha^J + \sum_{\beta\in P} \Lambda_{\alpha\beta} C_\beta^I + (\mathrm{c.c}) \\ &= 0 \end{aligned} \tag{12}$$

$$\begin{aligned}\frac{\partial \mathscr{L}_{\mathrm{SA}}}{\partial \phi_p} &= \sum_{\alpha\in P} w_\alpha \left\langle \frac{\partial \alpha}{\partial \phi_p} \middle| \hat{H}^{\mathrm{sol}} \middle| \alpha \right\rangle + \sum_q \lambda_{pq}\phi_q + (\mathrm{c.c.}) \\ &= 0, \end{aligned} \tag{13}$$

where

$$\hat{H}^{\mathrm{sol}} = \hat{H} + {}^t\mathbf{V}^{(0)}\left[\mathbf{\Xi} + (m-1)\mathbf{\Gamma}(\mathbf{d}_{\mathrm{SA}})\right]^{-1}\hat{\mathbf{R}} \tag{14}$$

and constraint conditions by considering the differentiation of the Lagrange multipliers $\mathbf{\Lambda}$ and $\boldsymbol{\lambda}$. Eq.(12) is the generalized eigenvalue problem for CI coefficients as follows.

$$\sum_{J\in\mathrm{CSF}} H_{\mathrm{CAS}}^{IJ} C_\alpha^J = E_{\mathrm{CAS}}^\alpha C_\alpha^I, \tag{15}$$

$$\sum_{J\in\mathrm{CSF}} C_\alpha^J C_\beta^J = \delta_{\alpha\beta}, \tag{16}$$

$$H_{\mathrm{CAS}}^{IJ} = \langle I|\hat{H}^{\mathrm{sol}}|J\rangle, \tag{17}$$

where E_{CAS}^α is the α-th CAS-CI eigenenergy. We employed the orbital gradient method to solve Eq. (13) as follows:

$$\sum_{\alpha\in P} w_\alpha \langle\alpha|[\hat{a}_p^\dagger \hat{a}_q, \hat{H}^{\mathrm{sol}}]|\alpha\rangle = 0, \tag{18}$$

$$(p|q) = \delta_{pq}, \tag{19}$$

Thereafter, we optimized the CASSCF orbitals. There are choices of representation of orbitals such as natural orbitals or canonical MOs in the SA state. This study selected canonical MOs in the SA state, and canonical Fock matrix is given by

$$\begin{aligned} f_{pq}^{\mathrm{solv}} &= h_{pq} + {}^{t}\mathbf{V}^{(0)}[\mathbf{\Xi} + (m-1)\mathbf{\Gamma}(\mathbf{d}_{\mathrm{SA}})]^{-1}\mathbf{R}_{pq} \\ &+ \sum_{\alpha\in P}\sum_{r,s}^{\mathrm{MO}} w_\alpha \langle\alpha|\hat{a}_r^\dagger \hat{a}_s|\alpha\rangle \left\{ (pq|rs) - \frac{1}{2}(pr|qs) \right\}, \\ &= \varepsilon_p \delta_{pq}, \end{aligned} \tag{20}$$

where h_{pq} and ε_p are the one-electron integral and p-th canonical orbital energy, respectively.

We describe the novel point of our introduction of the free energy G_{SA} and the difference between G_{SA} and other averaged energies in Refs.[28] and[23], respectively. Ref.[28] regards the averaged energy G'_{SA} as

$$G'_{\mathrm{SA}} = \sum_{\alpha\in P} w_\alpha \langle\alpha|\hat{H}|\alpha\rangle + \Delta\mu(\mathbf{d}_{\mathrm{CAS}}^{(0)}) \tag{21}$$

corresponding to neglecting the third term of Eq.(1). Ref.[28] assumed that the excess chemical potentials of all the states were equal. Oppositely, Ref.[23] regards the averaged energy G''_{SA} as

$$G''_{\mathrm{SA}} = \sum_{\alpha\in P} w_\alpha (\langle\alpha|\hat{H}|\alpha\rangle + \Delta\mu(\mathbf{d}_{\mathrm{CAS}}^{(\alpha)})). \tag{22}$$

$\mu(\mathbf{d}_{\mathrm{CAS}}^{(\alpha)})$ denotes the excess chemical potential determined by $\mathbf{d}_{\mathrm{CAS}}^{(\alpha)}$ and indicates that the solvent structure is relaxed in the α-th electronic structure of the solute molecule. The second term of Eq.(22) corresponds to the average of the excess chemical potentials whose solvent structures are relaxed in the electronic states of the solute molecule. The most important difference between the present theory and Refs.[23,28] is the existence of formalisms through a variational approach. It is not guaranteed that the equations solved in Ref.[23,28] to optimize the CAS states stem from the variational principle for G'_{SA} and G''_{SA}. Thus, their solutions could not be self-consistent fields and may not minimize G'_{SA} and G''_{SA}. However, the present methodology provides a self-consistent field for minimizing the G_{SA}.

B. LAGRANGIAN OF RISM-XMCQDPT2

We defined the XMCQDPT wavefunction $|i\rangle$, using linear combinations of the CASSCF wavefunctions $|\alpha\rangle$ as follows:

$$|i\rangle = \sum_{\alpha\in P} D_i^\alpha |\alpha\rangle, \tag{23}$$

D is the rotation matrix from the CAS states to the QDPT states.[21] We defined the coefficients of the ABSs of the electron density or the transition density matrices in the criteria of the RISM-SCF-cSED $\mathbf{d}_{\alpha\beta}$ as

$$\mathbf{d}_{\alpha\beta} = [\Xi + (m-1)\Gamma(\mathbf{d}_{\mathrm{SA}})]^{-1} \sum_{pq}^{\mathrm{MO}} \mathbf{R}_{pq} \langle \alpha | \hat{a}_p^\dagger \hat{a}_q | \beta \rangle . \tag{24}$$

The i-th XMCQDPT2-RISM energy $E^{(i)}_{\mathrm{QDPT}}$ was defined as

$$E^{(i)}_{\mathrm{QDPT}} = \sum_{\alpha,\beta\in P} D_i^\alpha D_i^\beta H^{\alpha\beta}_{\mathrm{gas}} + \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{QDPT}}) + \frac{\partial \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{QDPT}})}{\partial \mathbf{d}^{(0)}_{\mathrm{QDPT}}} (\mathbf{d}^{(i)}_{\mathrm{QDPT}} - \mathbf{d}^{(0)}_{\mathrm{QDPT}}), \tag{25}$$

where

$$H^{\alpha\beta}_{\mathrm{gas}} = \langle \alpha | \hat{H} | \beta \rangle + \frac{1}{2} \sum_{I\in \mathrm{CSF}} \frac{\langle \alpha | \hat{V} | I \rangle \langle I | \hat{V} | \beta \rangle}{E_0^\alpha - E_0^I} + (\alpha \leftrightarrow \beta) \tag{26}$$

$$E_0^\alpha = \sum_p \varepsilon_p \langle \alpha | \hat{a}_p^\dagger \hat{a}_p | \alpha \rangle \tag{27}$$

$$E_0^I = \sum_p \varepsilon_p \langle I | \hat{a}_p^\dagger \hat{a}_p | I \rangle \tag{28}$$

$$\mathbf{d}^{(i)}_{\mathrm{QDPT}} = \sum_{\alpha,\beta\in P} D_i^\alpha \mathbf{d}_{\alpha\beta} D_i^\beta \tag{29}$$

$$\sum_{\alpha\in P} D_i^\alpha D_i^\alpha = 1, \tag{30}$$

and $i = 0$ represents the target state of the QDPT states. Here, we assumed that the fitting coefficients of the target state of the XMCQDPT2-RISM were practically equal to those of the CAS states as $\mathbf{d}^{(0)}_{\mathrm{QDPT}} \simeq \mathbf{d}^{(0)}_{\mathrm{CAS}}$. In such a case, we approximated the $\Delta\mu(\mathbf{d}^{(0)}_{\mathrm{QDPT}})$ to

$$\Delta\mu(\mathbf{d}^{(0)}_{\mathrm{QDPT}}) = \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{CAS}}) + {}^t\mathbf{V}^{(0)} (\mathbf{d}^{(0)}_{\mathrm{QDPT}} - \mathbf{d}^{(0)}_{\mathrm{CAS}}) \tag{31}$$

$$\frac{\partial \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{QDPT}})}{\partial \mathbf{d}^{(0)}_{\mathrm{QDPT}}} = \frac{\partial \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{CAS}})}{\partial \mathbf{d}^{(0)}_{\mathrm{CAS}}} = {}^t\mathbf{V}^{(0)}. \tag{32}$$

The approximated XMCQDPT2-RISM energy was given by

$$\begin{aligned} E^{(i)}_{\mathrm{QDPT}} &= \sum_{\alpha,\beta\in P} D_i^\alpha D_i^\beta H^{\alpha\beta}_{\mathrm{gas}} + \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{CAS}}) + {}^t\mathbf{V}^{(0)} (\mathbf{d}^{(i)}_{\mathrm{QDPT}} - \mathbf{d}^{(0)}_{\mathrm{CAS}}) \\ &= \sum_{\alpha,\beta\in P} D_i^\alpha D_i^\beta \left[H^{\alpha\beta}_{\mathrm{gas}} + {}^t\mathbf{V}^{(0)} \mathbf{d}_{\alpha\beta} \right] + \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{CAS}}) - {}^t\mathbf{V}^{(0)} \mathbf{d}^{(0)}_{\mathrm{CAS}} \end{aligned} \tag{33}$$

We introduced the XMCQDPT2-RISM Lagrangian $\mathscr{L}_i$ to optimize the solvated free energy of the i-th state as

$$\begin{aligned}\mathscr{L}_i = &\sum_{\alpha,\beta\in P} D_i^\alpha D_i^\beta \left[H_{\rm gas}^{\alpha\beta} +^t \mathbf{V}^{(0)}\mathbf{d}_{\alpha\beta}\right] + \lambda \sum_{\alpha\in P}\left\{(D_i^\alpha)^2 - 1\right\} \\ &+ \sum_{\alpha,\beta\in P}\sum_{I,J\in \rm CSF}\left\{K_\alpha^I (H_{\rm CAS}^{IJ} - E_{\rm CAS}^\beta \delta_{IJ}) C_\beta^J - \kappa_{\alpha\beta}(C_\alpha^I C_\beta^J \delta_{IJ} - \delta_{\alpha\beta})\right\} \\ &+ \sum_{pq}^{\rm MO}\left\{Z_{pq}(f_{pq}^{\rm solv} - \varepsilon_p \delta_{pq}) + z_{pq}(x_{pq} - x_{qp}) + \zeta_{pq}((p|q) - \delta_{pq})\right\}, \end{aligned} \tag{34}$$

where

$$\begin{aligned} x_{pq} = &\sum_{\alpha\in P} w_\alpha \left\{\sum_r^{\rm MO}\left(h_{pr} +^t \mathbf{V}^{(0)}\left[\Xi + (m-1)\mathbf{\Gamma}(\mathbf{d}_{\rm SA})\right]^{-1}\mathbf{R}_{pr}\right)\langle\alpha|\hat{a}_q^\dagger \hat{a}_r|\alpha\rangle\right\} \\ &+ \sum_{\alpha\in P} w_\alpha \left\{\sum_{r,s,t}^{\rm MO}(pr|st)\langle\alpha|\hat{a}_q^\dagger \hat{a}_s^\dagger \hat{a}_t \hat{a}_r|\alpha\rangle\right\} \end{aligned} \tag{35}$$

using Lagrange multipliers λ, $\mathbf{K}$, $\boldsymbol{\kappa}$, $\mathbf{Z}$, $\mathbf{z}$, and $\boldsymbol{\zeta}$ to create the constraints for the equations of the SA-CASSCF. $\hat{H}$, $\hat{F}$, and $\hat{V}$ are the total electronic Hamiltonian, the SA Fock operator, and the operator of the perturbation term in the XMCQDPT2 framework, respectively. $(pq|rs)$ denotes a two-electron integral under the MO representation.[27] Considering the variations of all the Lagrange multipliers in Eq.(34), we obtained the equations of the SA-CASSCF calculations as follows:

$$x_{pq} = x_{qp}, \tag{36}$$

and we simplified the matrix $\mathbf{H}_{\rm CAS}$ as

$$H_{\rm CAS}^{IJ} = E_0^I \delta_{IJ} + \langle I|\hat{V}|J\rangle, \tag{37}$$

and

$$\langle\alpha|\hat{H}^{\rm sol}|\beta\rangle = E_{\rm CAS}^\alpha \delta_{\alpha\beta}. \tag{38}$$

The sum of the first and second terms in Eq.(34) was simplified as

$$\begin{aligned} H_{\rm gas}^{\alpha\beta} +^t \mathbf{V}^{(0)}\mathbf{d}_{\alpha\beta} &= \langle\alpha|\hat{H}^{\rm sol}|\beta\rangle \\ &\quad + \frac{1}{2}\sum_{I\in\rm CSF}\frac{\langle\alpha|\hat{V}|I\rangle\langle I|\hat{V}|\beta\rangle}{E_0^\alpha - E_0^I} + (\alpha\leftrightarrow\beta) \\ &= E_{\rm CAS}^\alpha \delta_{\alpha\beta} + \frac{1}{2}\sum_{I\in\rm CSF}\frac{\langle\alpha|\hat{V}|I\rangle\langle I|\hat{V}|\beta\rangle}{E_0^\alpha - E_0^I} + (\alpha\leftrightarrow\beta). \end{aligned} \tag{39}$$

Using the simpler formula of Eq.(39), we performed the perturbation scheme in the solution phase similarly to that of a gas phase as written in the second and third terms of the second line of the right-hand-side of Eq.(39). Finally, we obtained the free energies in the solution of XMCQDPT2 $E^{(i)}_{\mathrm{QDPT}}$ by taking the derivative of matrix $\mathbf{D}$ about the Lagrangian in Eq.(34) given by

$$\sum_{\beta\in P}\left[H^{\alpha\beta}_{\mathrm{gas}} + {}^{t}\mathbf{V}^{(0)}\mathbf{d}_{\alpha\beta}\right]D^{\beta}_{i} = \mathscr{E}^{(i)}D^{\alpha}_{i}, \tag{40}$$

where

$$E^{(i)}_{\mathrm{QDPT}} = \mathscr{E}^{(i)} + \Delta\mu(\mathbf{d}^{(0)}_{\mathrm{CAS}}) - {}^{t}\mathbf{V}^{(0)}\mathbf{d}^{(0)}_{\mathrm{CAS}}, \tag{41}$$

and solving the eigenvalue problem in Eq.(40).

III. COMPUTATIONAL DETAILS

We simulated a phenol (PhOH) molecule in various solutions using the GAMESS program package[29,30]. Acetonitrile (ACN), methanol (MeOH), and water (WAT) were selected as solvents. The CAM-B3LYP functional was employed for geometry optimization and Hessian calculations in each solvent using the RISM-SCF-cSED, as in previous studies.[26,31] In the Hessian calculations, the analytical method was employed in all simulations. After the evaluation of equilibrium geometries and Hessian matrices in the solutions, we performed MS-XMCQDPT2(7o,8e) to evaluate vertical excitation energies after the calculation of three states SA-CASSCF(7o,8e) whose averaged weights were uniform. The aug–cc–pv(D+d)Z basis sets[32] were employed for all atoms.

Here, we approximately drew the potential energy surfaces (PESs) of a PhOH molecule in solutions using the MS-XMCQDPT2 under the theoretical framework previously constructed in Ref.[14]. This enabled the evaluation of the electronic photoabsorption spectral lines including the broadening bandwidth. We modeled multimodal harmonic oscillators for drawing the PESs, which were approximated using a five-point central finite difference method with a grid width of $\delta Q = 0.01$ Bohr. The mode coordinates were set to those obtained through the analytical Hessian method. The number of grids N was 2^{20}=1048576 and the grid width in the time domain Δt was 7.49 a.u. ($\sim$ 0.18 fs), where the spectral line was given by Eqs.(16)–(20) in Ref.[14].

We applied the site—site Coulombic and Lennard—Jones potential to the intermolecular interaction between the solute and solvent using Singer and Chandler's hypernetted–chain approximation as the closure for the RISM. The atomic charge and Lennard-–Jones parameters used for the

TABLE I. Atomic charges and LJ parameters of the PhOH solute and the solvents. H(C) and H(O) denote the hydrogen atoms bonding with the carbon and oxygen atoms of phOH, respectively.

site	σ_s (Å)	ε_s (kcal/mol)	q_s
		ACN	
CH_3	3.775	0.207	0.150
C	3.650	0.150	0.280
N	3.070	0.170	-0.430
		MeOH	
CH_3	3.775	0.207	0.265
O	3.070	0.170	-0.700
H	1.000	0.056	0.435
		WAT	
O	3.166	0.155	-0.820
H	1.000	0.056	0.410
		PhOH	
C	3.550	0.070	
O	3.070	0.170	
H(C)	2.420	0.030	
H(O)	1.000	0.056	

RISM calculations were taken from OPLS-AA [33] and are provided with the solvent parameters in the Table I. In addition, we set the constraint condition for the fitting method of $\{\mathbf{d}^{\alpha}_{\mathrm{CAS}}\}$ using type II for the geometry optimization and Hessian matrices. Type III restrains were used for the excitation energies given by Eqs.(11) and (22) in Ref.[31] because of the computational stability in the RISM-SCF-cSED. Finally, we used $\mathbf{d}^{(i)}_{\mathrm{QDPT}}$ to evaluate the fluctuations in the energy during photoabsorption, as studied in Ref.[14].

IV. RESULTS AND DISCUSSION

The absorption spectral lines in the solutions are shown in FIG.1. In addition, vertical excitation energies E_{vert} and the energies of the maximum point of the spectral lines E_{max} are summarized in

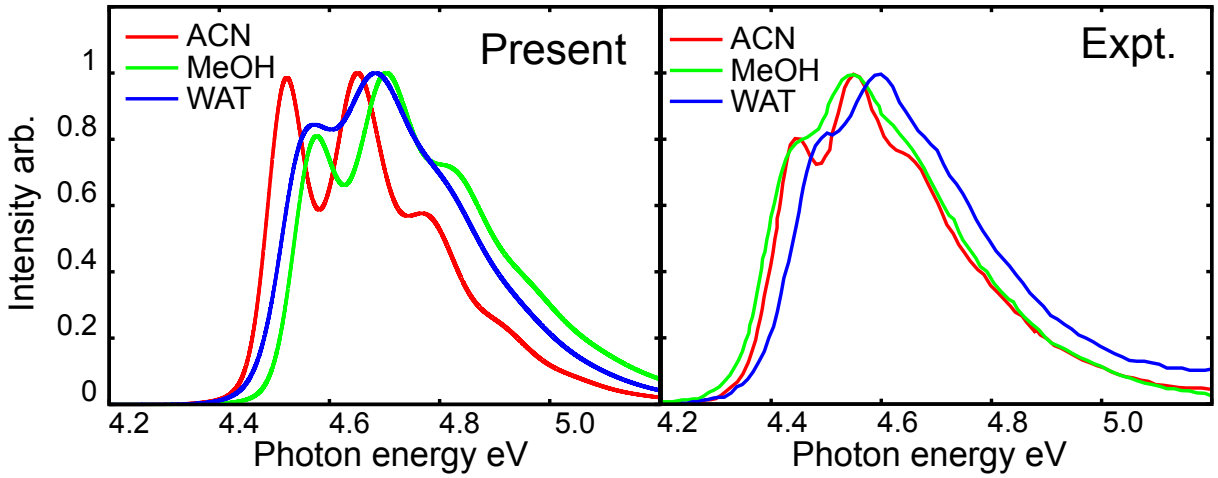

FIG. 1. Absorption spectrum of a PhOH molecule in solution. The left shows the numerical result of $P(E)$ defined in Ref.[14]. The experimental data on the right were digitized from Ref.[34]. The absorption strength was normalized to the maximum point of each spectral line.

TABLE II. We only focused on the S_0 to S_1 transition in evaluating the absorption spectral lines. The height of the spectral lines of the present numerical results and experiments[34] were normalized to the maximum point of the respective lines. In all the solvents, $\sim$0.1 eV overestimations of the maximum point of the peak from the experimental data $E_{\text{expt.}}$ were observed.

Considering the degree of freedom of molecular vibration, the maximum points of the absorption spectral lines obtained by the present method are corrected closer to the respective points of the experimental data than the vertical excitation energies obtained through this method. We verified the precision of a computational method in an excited state calculation by comparing the vertical excitation energies and the energies at the peak top in experimental data uto date. The present method might not be chemically accurate under the criterion of the gap energies between the theory and the experiment larger than $\sim$0.2 eV. However, correction by intramolecular vibration plays an important role in predicting the excitation energies of molecules, at least for a phenol molecule. The results indicate that the XMCQDPT2-RISM is not necessarily out of chemical accuracy.

In terms of the broadening bandwidth, differences exist between an aprotic solvent (ACN) and protic solvents (MeOH and WAT). The aprotic solvent shows multiple sharp peaks, probably corresponding to the fine structure of molecular vibration. However, protic solvents exhibit more broadened peaks, which appeared to indicate that the fine structures were hidden. The tendencies

TABLE II. Vertical energies, peak top energies of respective solutions, and thermal fluctuations of the electro-vibronic transition of the solvents (unit eV.)

solvent	ACN	MeOH	WAT
E_{vert}	4.75	4.76	4.81
E_{max}	4.65	4.75	4.74
E_{expt}	4.55	4.55	4.60
σ^{V}	0.042	0.044	0.037
σ^{U}	0.152	0.174	0.161

of the peak resolution of such fine structures are also observed in the experimental results. Nevertheless, fine structures were observed for all the solvents. Considering that we only considered the S_0 to S_1 transitions, it was confirmed that the fine structures of the spectral lines originated from the vibronic transition from the electronically ground state to the excited state. The broadening band originated from the thermal fluctuations of solvent σ^{V} given by

$$\sigma^{\mathrm{V}} = \left((\Delta\mathbf{d}_{10})^t \boldsymbol{\sigma} \Delta\mathbf{d}_{10}\right)^{\frac{1}{2}} \tag{42}$$

$$\Delta\mathbf{d}_{10} = \mathbf{d}^{(1)}_{\mathrm{QDPT}} - \mathbf{d}^{(0)}_{\mathrm{QDPT}}, \tag{43}$$

$$(\boldsymbol{\sigma})_{\zeta\eta} = \sum_s \beta^{-1}\rho_s q_s \int\int d\mathbf{r}d\mathbf{r}' \frac{f_\zeta(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \frac{\partial h_{\gamma s}(r')}{\partial(\mathbf{d}^{(0)}_{\mathrm{CAS}})_\eta} \ (\zeta \in \gamma), \tag{44}$$

in our previous study. Thereafter, we summarized the σ^{V} for the solvents in TABLE II to compare the width of the broadening band for each electro-vibronic transition. All σ^{V} values are $<$400 cm^{-1} typically less than the eigenfrequency of the bending or stretching intramolecular vibrational mode. Unexpectedly, WAT, which forms the strongest hydrogen bonding with the hydroxy group of a PhOH molecule, exhibited smaller thermal fluctuation σ^{V} than ACN. However, the σ^{V} was dependent on the strength of the electrostatic interaction between the solutes and solvents as shown in Eqs.(42)-(44). Thus, for PhOH, the peak resolution of the vibronic transition from the electronic ground to the excited state was sufficiently high to observe fine structures in any solvent because of the weak solvent fluctuation.

Next, we analyzed the vibrational modes to which the multiple peaks observed in the fine

structures were assigned. The absorption spectral line $P(E)$ from the S_0 to S_1 state was given by

$$
\begin{aligned}
P(E) &= EP_1(E) \\
P_1(E) &\propto \int dt \exp\left[-\frac{1}{2}(\sigma^{\mathrm{V}}t)^2 - i(E - V_1^{\min} + V_0^{\min})t\right]\left[\prod_l p_l(t)\right] \qquad (45) \\
p_l(t) &= \xi_l^{(1)}(t)\exp\left\{-\eta_l^{(1)}(t)\right\} \qquad (46)
\end{aligned}
$$

where $V_k^{\min}$ denotes the minimum free energy of the k-th PES and $\zeta_l^{(1)}$ and $\eta_l^{(1)}$ are defined in Eqs.(18)–(20) in Ref.[14]. $p_l(t)$ denotes the time correlation function of the l-th harmonic oscillator modeling the intramolecular vibration. When multiple peaks were observed, the absorption spectral line was largely distributed by the main vibrational modes characterizing the fine structure in terms of energy distribution. Here, we assumed that such vibrational modes contributed to the standard deviation of the spectral lines σ^{U} given by

$$
\begin{aligned}
(\sigma^{\mathrm{U}})^2 &= \frac{\int_{-\infty}^{\infty} dE E^2 P(E)}{\int_{-\infty}^{\infty} dE P(E)} - \left(\frac{\int_{-\infty}^{\infty} dE E P(E)}{\int_{-\infty}^{\infty} dE P(E)}\right)^2 - (\sigma^{\mathrm{V}})^2 \\
&= \sum_l \left[-\left\{\frac{\partial^2}{\partial t^2} p_l(t)\right\}_{t=0} + \left[\left\{\frac{\partial}{\partial t} p_l(t)\right\}_{t=0}\right]^2\right], \qquad (47)
\end{aligned}
$$

where the numerical results are summarized in TABLE.II. We selected the four modes whose time correlation functions $\{p_l(t)\}$ exhibited the four largest standard deviations defined as

$$
(\sigma_l^{\mathrm{U}})^2 = -\left\{\frac{\partial^2}{\partial t^2} p_l(t)\right\}_{t=0} + \left[\left\{\frac{\partial}{\partial t} p_l(t)\right\}_{t=0}\right]^2. \qquad (48)
$$

The four modes are commonly observed in all kinds of solvents. The modes are summarized in FIG.2 and their details are presented in TABLEIII. We observed that all the modes are within the plane parallel to the aromatic ring. Mode 1 and 2 in FIG.2 denote the stretching modes of the carbon atoms in the aromatic ring, whereas Mode 3 denotes the bending mode of the hydrogen atoms bonding with the carbon atoms. The vibrational transition of these three modes from the electronically ground state to the excited state was observed in many aromatic compounds like benzene, naphthalene, and anthracene. In addition, we observed that Mode 4 corresponding to the stretching vibration between the oxygen and carbon atoms is one of the main contributions of σ^{U}. All the modes in TABLEIII exhibited very weak solvatochromism in terms of the eigenfrequencies and the standard deviations. Considering the weak solvatochromism of σ^{V}, the peak resolution of the absorption spectral line was unlikely to be dependent on the kind of solvent in the case for a phenol molecule because the fine structures were observed in any solvent type in the experimental

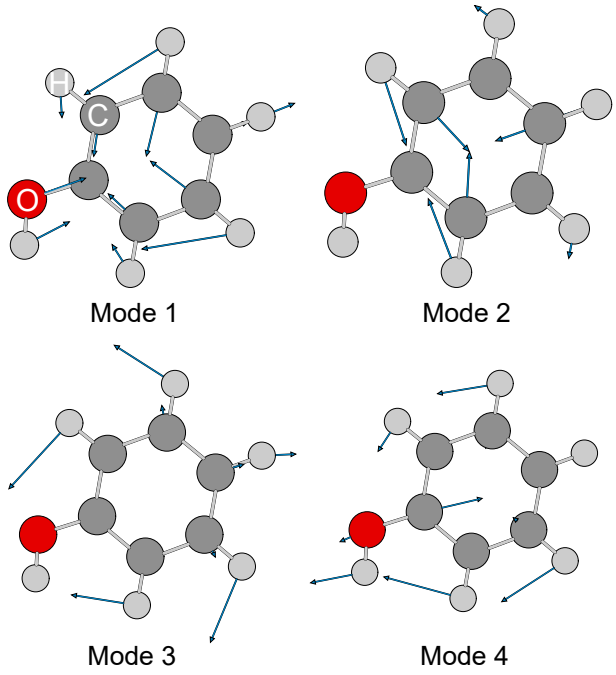

FIG. 2. Four vibrational modes with time correlation functions $\{p_l(t)\}$ showing the four largest standard deviations. The mode coordinates are shown as blue arrows obtained from the Hessian calculation in the ACN solvent surroundings. We omit the view parallel to the aromatic ring because all mode coordinates were within its plane.

TABLE III. Eigenfrequencies and standard deviations defined in Eq.(48) for a phenol molecule in solvents

l	ω_l cm^{-1}			σ_l^{U} eV		
	ACN	MeOH	WAT	ACN	MeOH	WAT
Mode 1	838	838	839	0.054	0.054	0.055
Mode 2	1017	1018	1020	0.071	0.072	0.072
Mode 3	1055	1055	1053	0.060	0.059	0.060
Mode 4	1293	1288	1283	0.076	0.076	0.075

data. The tendency of protic solvents to exhibit a broader bandwidth should be explained by the distortion of potential energy curves of other vibrational modes caused by the solvation effect.

V. CONCLUSION

We present the Lagrangian of the SA-CASSCF and the MS-XMCQDPT2 coupled with the RISM-SCF-cSED. This study derived their variational equations to calculate the excitation energies of target molecules in various solvents. After drawing the PESs and evaluating the thermal fluctuations of solvents, our previous theory was implemented for the absorption spectra of a phenol molecule dissolved in ACN, MeOH, and WAT. The numerical results exhibited that aprotic solvents (ACN) have the fine structure of the molecular vibration and protic solvents (MeOH and WAT) exhibited one or two broader peaks which appeared to be slightly hidden fine structures. We conclude that it is essential to predict the structure of PESs such as the Hessian matrices, on the kinds of solvents in terms of the photoexcitation process. Thus, the present theory is a very valuable method.

We also comfirmed that the difference in the thermal fluctuations between aprotic and protic solvents was negligible and not intrinsic in the peak resolution of the absorption spectral lines for a phenol molecule. The fine structures that were observed in any solvent surroundings stemmed from the vibrational transition of the four modes shown in FIG.2 all of which were stretching vibrations of the aromatic ring and oxygen atom of the phenol molecule. The vibrational transitions of these modes were closely related to the fine structure typically observed in many aromatic compounds.

ACKNOWLEDGMENTS

This study was supported by JST, PRESTO Grant Number JPMJPR21C9 (D.Y.), MEXT KAKENHI Grants (Green Catalysis Science) JP23H04911 (D.Y.), and Grant–in–Aid for JSPS Research Fellow No. 22KJ1094 (N.N.). The authors would like to thank Enago (www.enago.jp) for the English review.

REFERENCES

[1] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227–249 (1976).

[2] M. A. Thompson and G. K. Schenter, J. Phys. Chem. **99**, 6374–6386 (1995).

[3] J. Gao and D. G. Truhlar, Annual Review of Physical Chemistry **53**, 467–505 (2002).

[4] D. C. Rideout and R. Breslow, J. Am. Chem. Soc. **102**, 7816–7817 (1980).

[5]S. Otto, W. Blokzijl, and J. B. F. N. Engberts, J. Org. Chem. **59**, 5372–5376 (1994).

[6]S. Mai, B. Ashwood, P. Marquetand, C. E. Crespo-Hernández, and L. González, J. Phys. Chem. B **121**, 5187–5196 (2017).

[7]K. Anandhan, M. Cerón, V. Perumal, P. Ceballos, P. Gordillo-Guerra, E. Pérez-Gutiérrez, A. E. Castillo, S. Thamotharan, and M. J. Percino, RSC Adv. **9**, 12085–12096 (2019).

[8]D. Yokogawa, J. Chem. Phys. **145**, 094101 (2016).

[9]R. Y. Shimizu, T. Yanai, Y. Kurashige, and D. Yokogawa, J. Chem. Theory Comput. **14**, 5673–5679 (2018).

[10]R. Cammi and J. Tomasi, J. Chem. Phys. **100**, 7495–7502 (1994).

[11]J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999–3094 (2005).

[12]B. Carlotti, R. Flamini, I. Kikaš, U. Mazzucato, and A. Spalletti, Chem. Phys. **407**, 9–19 (2012).

[13]K. Naka, A. Morita, and S. Kato, J. Chem. Phys. **111**, 481–491 (1999).

[14]N. Negishi and D. Yokogawa, J. Phys. Chem. B **127**, 10142–10150 (2023).

[15]F. Hirata and P. J. Rossky, Chem. Phys. Lett. **83**, 329–334 (1981).

[16]S. Ten-no, F. Hirata, and S. Kato, J. Chem. Phys. **100**, 7443–7453 (1994).

[17]S. Ten-no, F. Hirata, and S. Kato, Chem. Phys. Lett. **214**, 391–396 (1993).

[18]H. Sato, F. Hirata, and S. Kato, J. Chem. Phys. **105**, 1546–1551 (1996).

[19]H. Werner and W. Meyer, J. Chem. Phys. **74**, 5794–5801 (1981).

[20]H. Nakano, J. Chem. Phys. **99**, 7983–7992 (1993).

[21]H. Nakano, K. Hirao, and M. S. Gordon, J. Chem. Phys. **108**, 5660–5669 (1998).

[22]R. Y. Shimizu, T. Yanai, and D. Yokogawa, J. Chem. Theory Comput. **16**, 4865–4873 (2020).

[23]D. Okamoto, Y. Watanabe, N. Yoshida, and H. Nakano, Chem. Phys. Lett. **730**, 179–185 (2019).

[24]D. Yokogawa, Bull. Chem. Soc. Jpn. **91**, 1540–1545 (2018).

[25]S. J. Singer and D. Chandler, Mol. Phys. **55**, 621–625 (1985).

[26]D. Yokogawa, H. Sato, and S. Sakaki, J. Chem. Phys. **126**, 244504 (2007).

[27]A. Szabo and N. S. Ostlund, *Mordern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Courier Corporation, 2012).

[28]N. Yoshida, T. Ishida, and F. Hirata, J. Phys. Chem. B **112**, 433–440 (2008).

[29]T. Mori and S. Kato, Chem. Phys. Lett. **437**, 159–163 (2007).

[30]N. Negishi and D. Yokogawa, J. Chem. Phys. **154** (2021), 154101.

[31]D. Yokogawa and K. Suda, J. Chem. Phys. **155**, 204102 (2021).

[32]J. Dunning, Thom H., J. Chem. Phys. **90**, 1007–1023 (1989).

[33]W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657–1666 (1988).

[34]S. Karmakar, D. P. Mukhopadhyay, and T. Chakraborty, J. Chem. Phys. **142**, 184303 (2015).