Polarizable Density Embedding Coupled Cluster Method

Hršak, Dalibor; Olsen, Jógvan Magnus Haugaard; Kongsted, Jacob

Published in: Journal of Chemical Theory and Computation

DOI: 10.1021/acs.jctc.7b01153

Publication date: 2018

Document version: Accepted manuscript

Citation for published version (APA):

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk
The Polarizable Density Embedding Coupled Cluster Method
Dalibor Hršak, Jógván Magnus Haugaard Olsen, and Jacob Kongsted

J. Chem. Theory Comput., Just Accepted Manuscript • DOI: 10.1021/acs.jctc.7b01153 • Publication Date (Web): 14 Feb 2018
Downloaded from http://pubs.acs.org on February 16, 2018

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.
The Polarizable Density Embedding Coupled
Cluster Method

Dalibor Hršak, Jógván Magnus Haugaard Olsen, and Jacob Kongsted*

Department of Physics, Chemistry and Pharmacy, University of Southern Denmark
Campusvej 55, 5230 Odense M, Denmark

E-mail: kongsted@sdu.dk
Phone: +45 65502304

Abstract

We present the theory and implementation of the polarizable density embedding (PDE) model in combination with coupled cluster (CC) theory (PDE-CC). This model has been implemented in the Dalton quantum chemistry program by adapting the CC code to the polarizable embedding library (PElib). In the PDE-CC method, the smaller, but chemically important core region is described with a high-level CC method. The environment surrounding the core region can be separated into two levels of description: an inner and an outer region. The effect of the inner region on the core region is described by an embedding potential consisting of a set of fragment densities obtained from calculations on isolated fragments with a quantum-chemistry method such as Hartree–Fock (HF) or Kohn–Sham density functional theory (KS-DFT) and dressed with a set of atom-centered anisotropic dipole-dipole polarizabilities. The outer region consists of distributed atom-centered multipoles and polarizabilities, i.e. in the same way as in the polarizable embedding (PE) model. The PDE-CC method contains embedding potential operators that account for the electrostatic and polarization interactions between the core region and the environment, as well as for non-electrostatic
(also known as Pauli and exchange) repulsion. All environmental effects are included through one-electron operators and account very efficiently for the response of the environment due to the change in the electron density of the core region, e.g., upon an electronic transition.

1 Introduction

One of the important topics of theoretical and computational chemistry is an accurate description of spectroscopic properties of large molecular systems, such as proteins or solute-solvent systems. Hybrid quantum mechanics/molecular mechanics (QM/MM) methods\textsuperscript{1–5} offer efficient calculations on such systems, where interactions between the chromophore and its environment modulate the molecular properties. The main advantage of such methods is the focused effort for accurate description of the chemically important part, while still retaining an atomistic description of the environment and thereby the capability of describing explicit interactions of the molecule with its nearest environment. Another, less refined, but computationally less expensive environment model is the polarizable continuum model (PCM), which emulates the bulk properties of the surrounding, but lacks the explicit interactions between the two subsystems.\textsuperscript{6,7}

Use of polarizable force fields enables an advanced and flexible modeling of solvents and biochemical environments.\textsuperscript{8–10} One of the models able to describe polarization effects from a discrete environment is the polarizable embedding (PE) model,\textsuperscript{11,12} which has been developed in recent years. The PE model has a convenient mathematical structure that allows for addition of simple one-electron effective operators to the vacuum Hamiltonian and thereby accounts for electrostatic interactions, as well as mutual polarization between the quantum region and its environment in an efficient fashion. This is particularly amenable for the description of molecular response properties, where the environment needs to be able to adapt to changes in the electron density of the molecule in the quantum region, e.g., upon an electronic excitation. The environment in the PE model is described by atom-centered mul-
tipoles and polarizabilities up to a given order. Such embedding potentials are here denoted as \( M_x P_y \), where \( x \) is an integer denoting the order of multipoles (0: charges, 1: charges and dipoles, etc.) and \( y \) denotes the type of polarizabilities (0: no polarizabilities, 1: isotropic polarizabilities, 2: anisotropic polarizabilities, etc.). The multipoles and polarizabilities are calculated for isolated fragment molecules (e.g. solvent molecules or amino acids) using a given quantum chemistry method.

The mathematical structure of the PE operators stems from classical electrostatic expressions that describe permanent electrostatics and polarization. In addition to Hartree–Fock (HF) and Kohn–Sham density functional theory (KS-DFT), PE has also been combined with coupled cluster (CC), multi-configurational self-consistent field (MCSCF), as well as a number of other theoretical methods and response theory approaches. These approaches have been shown to be able to model the effects of the environment on various spectroscopic properties in different types of systems and in general achieve good agreement with experiments. Examples of such applications include solvent effects on one- and two-photon absorption spectra, phosphorescence lifetimes of chromophores in solutions, spectroscopic properties of proteins, as well as UV spectra of molecules embedded into DNA environments.

The promising performance of the PE model calls for further advances in the field, one of which being a more refined description of the surrounding environment in the nearest vicinity of the target molecule. This effort has recently resulted in the development of a three-layered QM/QM/MM model termed polarizable density embedding (PDE). The total system in this model can be separated into three regions; 1) the core region consisting of the target molecule (solute or chromophore) fully described by a quantum-chemistry method, 2) an inner region where each fragment (solvent molecule) is described semi-classically and in isolation, and 3) an outer region consisting of the rest of the environment where each fragment is described classically in the same fashion as in the PE model. The polarization interactions in the inner region retains the same form as in the PE model. The embedding
potential that describes the inner region is here denoted as FDP\_y\_R, where FD denotes the description of electrostatic interactions with fragment electron densities, y denotes the type of the polarizabilities and R denotes the inclusion of the non-electrostatic repulsion in the model. This latter component is brought about by another layer of complexity needed for an accurate description of explicit interactions in cases where the electron density of the chromophore extends into the environment and overlaps with the electron densities of the environment molecules.\textsuperscript{29,30} The PDE model has recently been successfully applied in a case where solvent confinement effect was the main solvent contribution.\textsuperscript{31}

Until now, the PDE model has only been combined with HF and KS-DFT. In this work, we report the theory and implementation of the PDE model combined with CC methods (PDE-CC). Conventional CC is known to be a very accurate method, but characterized by a steep scaling in computational cost with respect to the size of the system, and this feature prohibits a full CC treatment of larger molecular systems. Here, we will harness both the high accuracy of CC and the ability of PDE to accurately represent intermolecular interactions using relatively simple one-electron interaction operators. A crucial feature of the PDE model is to include polarization effects classically via polarizabilites just as in the PE model. This allows us to use the same density-driven strategy as outlined in Ref. 14. This polarization plays an important role when addressing molecular response-property calculations.

Coupled cluster has previously been combined with other embedding type models for polarizable force fields. One such implementation is the combination of equation-of-motion (EOM) CC and the effective fragment potential\textsuperscript{32} (EFP) model.\textsuperscript{33–36} Embedding using EFP treats the environment in a similar way as PE and the parameters for the embedding potential are also obtained through first-principles calculations on isolated fragments. Equation-of-motion CCSD has also been combined with a two-level fluctuating charges / polarizable continuum model\textsuperscript{37} (EOM-CCSD-FQ-PCM), where polarization is accounted for through the redistribution of the point charges according to the electronegativity equalization principle.
A multilevel QM/QM analogue of the frozen density embedding (FDE) is the CC-in-DFT approach, that has been implemented for the calculation of the molecular response properties. The latter approach bears resemblances to PDE-CC due to the CC-in-DFT formalism, while the difference lies in the treatment of the polarization effects, where CC-in-DFT applies the freeze-and-thaw procedure while PDE-CC introduces polarization through classical electrostatics.

2 Theory

2.1 The PDE model

The first step in the derivation of the PDE-CC method is to introduce the relevant interactions between the core part and its environment. The contributions specific for the PDE model are detailed in the original work (at the SCF level of theory), but will be described here along with contributions that are shared with the PE model. For an in-depth description of PE-HF/DFT or PE-CC, we refer the reader to Refs. 11, 12, and 14. The polarizable environment is divided into two layers that surround the central molecule: an inner region that consists of the fragments in the nearest vicinity and an outer region that consists of the rest of the environment. The embedding potential of the inner region is defined by a set of fragment electronic densities, point nuclei, and a set of atom-centered polarizabilities, all obtained through calculations on isolated fragments. The embedding potential of the outer region is defined by fragment-based atom-centered multipoles and polarizabilities. The expansion points used for the atom-centered multipoles and polarizabilities will be referred to as the classical sites.

The interactions with the environment are described through one-electron interaction operators accounting for electrostatic interaction of the outer-region multipoles, electrostatic interaction of the inner-region fragment densities, non-electrostatic repulsion and polarization terms. The operator describing the electrostatic interaction of the core wave function...
with the outer-region multipoles, $\hat{V}^{\text{mul}}$, is equivalent to the electrostatic interaction operator in the PE model. This operator has the form

$$\hat{V}^{\text{mul}} = \sum_{s=1}^{S} \sum_{|k|=0}^{K_s} \frac{(-1)^{|k|}}{k!} M^{(k)}_s \sum_{pq} v^{(k)}_{s,pq} \hat{E}_{pq}$$  \hspace{1cm} (1)$$

where the summation over $s$ runs over the $S$ sites in the classical region, $k$ is a multi-index, $K_s$ is the truncation level of the multipole expansion, and $M^{(k)}_s$ is a component of a multipole moment located at site $s$, and $v^{(k)}_{s,pq}$ is a one-electron interaction integral which involves partial derivatives of the $\frac{1}{r}$ operator where $r$ is the distance between an electron and a multipole site. The superscript $(k)$ notation is used to indicate order and Cartesian component. For example, in the case of multipoles, $k = (0, 0, 0)$ specifies a monopole (i.e. charge), $k = (0, 1, 1)$ corresponds to the $yz$-component of a quadrupole, and so on. The electrostatic interactions between the core and the fragments in the inner region are described by the electrostatic fragment density operator

$$\hat{V}^{\text{fd}} = -\sum_{a=1}^{N_a} \sum_{pq} v^a_{pq} \hat{E}_{pq} + \sum_{a=1}^{N_a} \sum_{pq} \sum_{ij \in a} v^a_{pq,ij} D^a_{ij} \hat{E}_{pq}$$  \hspace{1cm} (2)$$

where the summations over $a$ run over the $N_a$ fragments in the inner region, $v^a_{pq}$ is a nuclear attraction integral involving orbitals $p$ and $q$ belonging to the core and the nuclei in the inner-region fragment $a$, $v^a_{pq,ij}$ is an inter-fragment two-electron repulsion integral involving orbitals $p$ and $q$ belonging to the core region and orbitals $i$ and $j$ belonging to inner-region fragment $a$, and $D^a_{ij}$ is an element of the density matrix of the inner-region fragment $a$. This set of operators describe the intermolecular interactions under the assumption that the wave functions of the two interacting fragments do not overlap. To model non-electrostatic exchange repulsion, another operator is introduced that is based on the Huzinaga-Cantu
projection operator method\textsuperscript{41} and has the following form in the MO basis

\[
\hat{V}^{\text{rep}} = -\sum_{a=1}^{N_a} \sum_{i \in a} \sum_{pq} \varepsilon_i S_{pi} S_{iq} \hat{E}_{pq}
\]  

(3)

Here, \(S_{pi}\) and \(S_{iq}\) are elements of the intermolecular overlap matrix involving the occupied orbitals of the inner-region fragments and \(\varepsilon_i\) is a fragment MO energy. Transformation of this operator from the MO to the AO basis gives rise to the presence of an energy-weighted density matrix as given in Ref. 28. The projection operator method simulates orthogonality between the wave functions of the inner-region fragments and the core-region wave function.

In order to derive the equations for optimization of the cluster amplitudes and multipliers, we proceed by adding the above one-electron operators to the vacuum Hamiltonian. This results in a time-independent PDE-CC Lagrangian which given as

\[
\mathcal{L}^{\text{PDE-CC}}(t, \bar{t}) = \left\langle \Lambda \right| \hat{H}^{\text{vac}} + \hat{V}^{\text{mul}} + \hat{V}^{\text{fd}} + \hat{V}^{\text{rep}} \right| \text{CC} \rangle - \frac{1}{2} \sum_{s=1}^{S} \sum_{s'=1}^{S} E_s(D)^\dagger B_{ss'} E_{s'}(D)
\]  

(4)

based on the usual definitions of the CC states.\textsuperscript{42} The last part of Eq. (4) is the energy associated with polarization of the environment. The \(B_{ss'}\) matrix is a \(3 \times 3\) subblock of the classical linear response matrix (also known as the relay matrix) in which the diagonal subblocks \((s = s')\) contain the inverse of the polarizabilities and the off-diagonal subblocks \((s \neq s')\) contain dipole–dipole interaction tensors. The electric field at a site \(s\), i.e. \(E_s(D)\), contains the field from the multipoles and fragment densities in the environment, and in addition the fields from the nuclei and electrons in the core region which thus depends on the CC density.

By enforcing stationarity of the Lagrangian with respect to both amplitudes and multipliers, we obtain the zeroth-order CC amplitudes and multipliers. The resulting PDE-CC optimization equations exhibit the exact same structure as the corresponding PE-CC opti-
mization equations

\[ \frac{\partial \mathcal{L}^{\text{PDE-CC}}}{\partial \mu_i} = \left( \bar{\mu}_i \right| \hat{H}^{\text{vac}} + \hat{V}^{\text{emb}} \right| \text{CC} \rangle = 0 \] (5a)

\[ \frac{\partial \mathcal{L}^{\text{PDE-CC}}}{\partial \bar{\mu}_i} = \left( \Lambda \left| \hat{H}^{\text{vac}} + \hat{V}^{\text{emb}}, \hat{r}_{\mu_i} \right| \text{CC} \rangle = 0 \] (5b)

The embedding operator \( \hat{V}^{\text{emb}} \) is defined in a density-driven polarization fashion through the equation

\[ \hat{V}^{\text{emb}} = \hat{V}^{\text{mul}} + \hat{V}^{\text{fd}} + \hat{V}^{\text{rep}} - \sum_{s=1}^{S} \mu_{s}^{\text{ind}} (D) \hat{E}^{\text{el}}_{s}, \] (6)

where \( \hat{E}^{\text{el}}_{s} \) is the electric-field operator, due to the electrons at site \( s \), and \( \mu_{s}^{\text{ind}} \) is the induced dipole at the site \( s \). As for the case of PE-CC, the PDE-CC amplitudes and multipliers equations are coupled because the embedding operator contains both sets of parameters (amplitudes and multipliers), which calls for an iterative optimization procedure. This problem is solved as outlined in Ref. 43, i.e. by performing the so-called inner and outer iterations and rely on a full solution of the coupled set of equations.

### 2.2 Polarizable Density Embedding CC response theory

To describe molecular response properties and excited states, we apply response theory as defined in Ref. 44 and formulated through Fourier component perturbation theory applied to the quasi-energy in Ref. 42. We refer to these papers for more details on the formalism. Frequency-dependent molecular response properties are calculated from response functions, defined as the derivatives of the time-averaged quasi-energy Lagrangian with respect to the perturbation strength parameters. For this work, the linear and quadratic response functions are considered and their final expressions after applying the chain-rule for partial derivatives
and the $2n + 1$ rule\textsuperscript{45} for the amplitudes and multipliers are

$$\langle\langle X; Y \rangle\rangle_{\omega Y} = \frac{1}{2} \hat{C}^{\pm \omega} \hat{P}^{XY} \left\{ \eta^{X}_t t^Y(\omega_Y) + \frac{1}{2} F^{X}_t (\omega_X) t^Y(\omega_Y) - \frac{1}{2} t^{X}_t (\omega_X) \bar{t}^Y(\omega_Y) J \right\} \quad (7a)$$

$$\langle\langle X; Y, Z \rangle\rangle_{\omega Y, \omega Z} = \frac{1}{2} \hat{C}^{\pm \omega} \hat{P}^{XY} \left\{ \frac{1}{2} F^{X} t^Y(\omega_Y) + \frac{1}{6} G^{X} t^Y(\omega_Y) + \frac{1}{2} \bar{t}^X(\omega_X) \bar{t}^Y(\omega_Y) A^Y + \frac{1}{2} \bar{t}^X(\omega_X) B t^Y(\omega_Y) + \frac{1}{2} \bar{t}^X(\omega_X) \bar{t}^Y(\omega_Y) K \right\} t^Z(\omega_Z) \quad (7b)$$

where $\omega_X$, $\omega_Y$, and $\omega_Z$ are the frequencies associated with the external perturbations, $\hat{C}^{\pm \omega}$ is the frequency symmetrizer, and $\hat{P}$ is the permutation operator. The response matrices and vectors entering into the expression for the linear and quadratic response functions are defined in Table 1. Enforcing the stationarity of the second-order time-averaged quasi-energy

Table 1: General expressions for the matrices and vectors in the PDE-CC response theory.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Vacuum contribution</th>
<th>PDE contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta^{X}_t$</td>
<td>$\langle \langle \hat{X}, \hat{t} \rangle \rangle_{CC}$</td>
<td>-</td>
</tr>
<tr>
<td>$F^{X}_{\mu, \nu}$</td>
<td>$\langle \langle \hat{H}<em>0, \hat{t} \rangle \rangle</em>{CC}$</td>
<td>$\langle \langle \hat{V}^{emb}, \hat{t} \rangle \rangle_{CC}$</td>
</tr>
<tr>
<td>$J^{X}_{\mu, \nu}$</td>
<td>$\langle \langle \hat{Y} \rangle \rangle_{CC}$</td>
<td>$\langle \langle \hat{V}^{emb}, \hat{t} \rangle \rangle_{CC}$</td>
</tr>
<tr>
<td>$G^{X}_{\mu, \nu, \sigma_k}$</td>
<td>$\langle \langle \hat{H}<em>0, \hat{t} \rangle \rangle</em>{CC}$</td>
<td>$\langle \langle \hat{V}^{emb}, \hat{t} \rangle \rangle_{CC}$</td>
</tr>
<tr>
<td>$A^Y_{\mu, \nu}$</td>
<td>$\langle \langle \hat{Y} \rangle \rangle_{CC}$</td>
<td>-</td>
</tr>
<tr>
<td>$B^{X}_{\mu, \nu, \sigma_k}$</td>
<td>$\langle \langle \hat{H}<em>0, \hat{t} \rangle \rangle</em>{CC}$</td>
<td>$\langle \langle \hat{V}^{emb}, \hat{t} \rangle \rangle_{CC}$</td>
</tr>
<tr>
<td>$K^{X}_{\mu, \nu, \sigma_k}$</td>
<td>-</td>
<td>$\hat{P}^{X}<em>{\nu} \langle \langle \hat{V}^{emb}, \hat{t} \rangle \rangle</em>{CC}$</td>
</tr>
</tbody>
</table>

Lagrangian with respect to the first-order amplitudes and multipliers results in the first-order
CC equations, which enable us to obtain the first-order parameters

\[ \xi^Y + (A - \omega_Y I)t^Y(\omega_Y) + \bar{t}^Y(\omega_Y)J = 0 \]  
\[ \eta^Y + Ft^Y(\omega_Y) + \bar{t}^Y(\omega_Y)(A + \omega_Y I) = 0 \]

These equations show the same structure as for the case of PE-CC and, exactly as for the case of PE-CC, these equations are coupled in the amplitudes and multipliers and are therefore solved based on a similar strategy as the one outlined for the ground state parameters. As seen from Table 1, the form of the PDE response matrices are similar in structure to the form of the PE response matrices, with the only difference being the new contributions to the effective environment operator entering in the definition of the \( \hat{V}_{\text{emb}} \) operator. The single and double partial derivatives of the effective environment operator with respect to the amplitudes (\( \hat{V}_{\text{emb},\mu} \) and \( \hat{V}_{\text{emb},\mu_1\nu} \)) and with respect to the multipliers (\( ^{\nu}V_{\text{emb}} \)) have exactly the same form as in the PE-CC theory because only the polarization part depends on the cluster amplitudes and Lagrange multipliers, while the new contributions specific for PDE do not. The exact forms of these operators can be found in Ref. 14.

3 Implementation Aspects

The PDE-CC method has been implemented in a development version of Dalton 2016\textsuperscript{46} by coupling the existing CC routines to the polarizable embedding library, \textit{PElib}\textsuperscript{47}. The available CC wave-function methods in the new PDE-CC implementation are CC2,\textsuperscript{48} CCSD and CCSDR(3)\textsuperscript{49}. In addition to the PDE model, \textit{PElib} provides a number of new features and functionalities not present in the previous PE-CC implementation, such as for example inclusion of multipole moments up to fifth order (M5). In order to efficiently solve the response equations, the matrices are transformed with left or right trial vectors.\textsuperscript{50} The expressions for
the transformed matrices are listed in the following equation

\[
A^{PDE_c} = \sum_{\mu_i \nu_j} \left\langle \hat{\mu}_i \left| \hat{V}^{emb, \hat{\tau}_{\nu_j}} + \hat{V}^{emb, \nu_j} \right| CC \right\rangle c_{\nu_j} = \sum_{\mu_i} \left\langle \hat{\mu}_i \left| \hat{V}^{emb, C} \right| CC \right\rangle + \xi^{emb, C} \tag{9a}
\]

\[
cA^{PDE} = \sum_{\mu_i \nu_j} c_{\mu_i} \left\langle \hat{\mu}_i \left| \hat{V}^{emb, \nu_j} \right| CC \right\rangle = \sum_{\nu_j} \left\langle C \left| \hat{V}^{emb, \hat{\tau}_{\nu_j}} \right| CC \right\rangle + \eta^{C^{emb}} \tag{9b}
\]

\[
cJ = \sum_{\mu_i \nu_j} c_{\mu_i} \left\langle \hat{\mu}_i \left| \hat{V}^{emb} \right| CC \right\rangle = \xi^{C^{emb}} \tag{9c}
\]

\[
F^{PDE_c} = \sum_{\mu_i \nu_j} \left\langle A \left| \left[ \hat{V}^{emb, \hat{\tau}_{\mu_i}}, \hat{V}^{emb, \nu_j} \right] + \left[ \hat{V}^{emb, \hat{\tau}_{\mu_i}}, \hat{\tau}_{\nu_j} \right] \right| CC \right\rangle c_{\nu_j} = \sum_{\mu_i} \left\langle A \left| \left[ \hat{V}^{emb, \hat{\tau}_{\mu_i}}, C \right] \right| CC \right\rangle + \eta^{emb, C} \tag{9d}
\]

\[
G^{PDE_{bc}} = \sum_{\mu_i \nu_j \sigma_k} \left\langle A \left| \left[ \hat{V}^{emb, \hat{\tau}_{\mu_i}}, \hat{\tau}_{\nu_j}, \hat{\tau}_{\sigma_k} \right] + \left[ \hat{V}^{emb, \nu_j, \hat{\tau}_{\mu_i}}, \hat{\tau}_{\sigma_k} \right] + \left[ \hat{V}^{emb, \hat{\tau}_{\mu_i}, \nu_j}, \hat{\tau}_{\sigma_k} \right] \right| CC \right\rangle b_{\mu_i, c_{\nu_j}}
\]

\[= F^{emb, C} b + F^{emb, B} c + \eta^{emb, BC} \tag{9e}
\]

\[
B^{PDE_{bc}} = \sum_{\mu_i \nu_j \sigma_k} \left\langle \hat{\sigma}_k \left| \left[ \hat{V}^{emb, \hat{\tau}_{\nu_j}}, \hat{\tau}_{\mu_i}, \hat{\tau}_{\sigma_k} \right] + \left[ \hat{V}^{emb, \nu_j, \hat{\tau}_{\mu_i}}, \hat{\tau}_{\sigma_k} \right] + \left[ \hat{V}^{emb, \hat{\tau}_{\mu_i}, \nu_j}, \hat{\tau}_{\sigma_k} \right] + \hat{V}^{emb, \nu_j, \nu_j} \right| CC \right\rangle b_{\mu_i, c_{\nu_j}}
\]

\[= \sum_{\sigma_k} \left\langle \hat{\sigma}_k \left| \left[ \hat{V}^{emb, B}, C \right] \right| CC \right\rangle + A^{emb, B} c + A^{emb, C} b + \xi^{emb, BC} \tag{9f}
\]

\[
bcK = \sum_{\mu_i \nu_j \sigma_k} b_{\mu_i, c_{\nu_j}} \left( \left\langle \hat{\mu}_i \left| \hat{V}^{emb, \hat{\tau}_{\sigma_k}} \right| CC \right\rangle + \left\langle \nu_j \left| \hat{\mu}_i \hat{V}^{emb, \hat{\tau}_{\sigma_k}} \right| CC \right\rangle \right) = C^{B^{emb}} + B^{C^{emb}} \tag{9g}
\]

Matrices J and K represent relatively small contributions and they can be neglected when calculating the residues of the response functions needed for the transition moments. This approximation facilitates the calculations and produces only small errors as previously observed.51,52 The matrices and vectors in the above equation have been constructed in the density-driven response polarization fashion, in which first the transformed densities (D^B, D^C, B_D, C_D and D^{BC}) are calculated and used to construct the effective environment operators. The difference from the previous PE-CC implementation is in the response calculation, whereby PELib allows to pass multiple densities simultaneously for the construction of the \( \hat{V}^{emb} \) operators instead of passing only one density at a time. This has the potential to save computational time when calculating the response as it avoids repeated calculations of electric-field integrals.
4 Calculations

4.1 Formamide in water

To illustrate the performance of the PDE-CC method – and compare the results to PE-CC – we first consider the case of formamide in water solution. The electronic spectrum of formamide is characterized by a number of Rydberg transitions in the region around the most intense $\pi \rightarrow \pi^*$ transition,\textsuperscript{53–56} and such Rydberg states are destabilized by the solvent through non-electrostatic repulsion.\textsuperscript{57,58} The spectra of aqueous formamide have already been a target of PE-CC studies,\textsuperscript{14,59,60} but the aforementioned destabilization of the Rydberg states has not been clearly analyzed.

For our analysis, 120 configurations (snapshots) of aqueous formamide were extracted from an MD simulation as previously reported.\textsuperscript{59} Each snapshot includes approximately 345 water molecules corresponding to a buffer size of 12 Å from the solute. The water solvent was described with one of two types of embedding potentials; for PDE-CC the FDP2R potential was used for all solvent molecules and for PE-CC the potential was M2P2.\textsuperscript{61} Local properties (atom-centered multipole moments and polarizabilities) were calculated based on B3LYP\textsuperscript{62–65}/aug-cc-pVDZ.\textsuperscript{66–69} Before the local property calculations the basis set has been recontracted to atomic natural orbital (ANO) form. The local property calculations were performed in Molcas\textsuperscript{70} using the LoProp formalism,\textsuperscript{71} whereas the fragment densities were calculated in the Dalton program\textsuperscript{46} also with B3LYP/aug-cc-pVDZ. The non-electrostatic repulsion was scaled by a factor of 0.7 to remedy the overestimation of short-range effects that arises from the fact that the fragment densities are optimized for isolated solvent molecules.\textsuperscript{72} For the vacuum calculations the formamide molecule was optimized at the level of B3LYP/aug-cc-pVTZ using the Gaussian 09 program.\textsuperscript{73}

The 10 lowest singlet excited states were calculated with either CC2, CCSD or CCSDR(3) in combination with PDE or PE and with the aug-cc-pVDZ basis set. From experiments in vacuum a series of transitions have been observed consisting of one $n \rightarrow \pi^*$ transition,
one Rydberg transition, an intense $\pi \rightarrow \pi^*$ transition, followed by a Rydberg and a Q-type band.\textsuperscript{53,54,57} Here, the calculated electronic spectrum of formamide in isolation is characterized by a weak $n \rightarrow \pi^*$ transition at $\sim 5.75$ eV for the three CC models, followed by a series of weak Rydberg-type excitations, an intense $\pi \rightarrow \pi^*$ transition, and another set of weak transition. For CC2 the intense $\pi \rightarrow \pi^*$ transition corresponds to the seventh excited state located at 7.62 eV. For CCSD and CCSDR(3) it is the sixth excited state located at 7.73 eV and 7.66 eV, respectively. The distribution – based on the 120 configurations – of the excited states of formamide in aqueous solution for the given embedding models are shown in Fig. 1. The differences between PDE and PE for all three CC methods are clearly seen. The PE-CC results exhibit a substantial mixture of Rydberg and valence excited states, whereas for PDE-CC valence states and Rydberg states are clearly separated, where all of the configurations have the intense transition with clear $\pi \rightarrow \pi^*$ character as the second excited state, and a weak transition with clear $n \rightarrow \pi^*$ character as the lowest excited state. We have further investigated the non-electrostatic repulsion effect on Rydberg states by changing the repulsion factor. The stick spectra are shown in Fig. S1 in the Supplementary Information. We see that the positions and intensities of $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ are not visibly affected by

Figure 1: Distribution of the 10 lowest-lying singlet excited states across the 120 snapshots of formamide solvated in water, calculated with PDE-CC or PE-CC methods. The insets show in more detail the region where the lowest singlet excited state ($S_1$) can be found (the units are same as in the main plots).
either reduction or increase in non-electrostatic repulsion. However, the Rydberg states are more affected. For the low repulsion factor (0.5), the mixing with the valence states as seen in PE-CC start to occur (especially for CC2), and for the high repulsion factor (0.9), the Rydberg states are shifted towards higher energies than for the optimal repulsion factor of 0.7. This confirms that the Rydberg states are strongly affected by non-electrostatic repulsion, which is expected because these states are diffuse in nature.

4.2 Acrolein in water

The next example is the case of acrolein in a water solution. These calculations are based on 120 configurations produced by an MD simulation already reported in a previous publication. The embedding potentials were obtained in the same fashion as for the case of formamide and they can be found in Ref. 75. The target molecular property was the UV/vis spectrum, or more precisely, the dependence of the excitation energies and solvent shifts of $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions on the embedding model. The acrolein molecule was solvated in a sphere consisting of approximately 330 water molecules reaching up to 12.0 Å from the solute.

Absorption spectra of acrolein in water were based on the 10 lowest singlet excited states, with either PDE or PE in combination with CCSDR(3), CCSD and CC2 methods. The spectra were convoluted using a Lorentzian line shape function with full-width-at-half-maximum (FWHM) set at 0.2 eV (for further details on how to produce UV/vis absorption spectra see Ref. 22) and averaged over the snapshots. The resulting spectra can be seen in Fig. 2. The intense peaks located at $\sim 6.5$ eV for the different methods correspond to the $\pi \rightarrow \pi^*$ electronic excitation. While there are no significant differences between the PDE- or PE-based results in that part of the spectrum, this is not the case for the higher transitions. For PDE-CC, the peaks corresponding to the higher excited states are well separated from the intense peak and located at $\sim 8.5 - 9$ eV, while for PE-CC the higher transitions lie closer to the intense peak, namely at $\sim 7.5 - \sim 8.5$ eV. These higher transitions in aqueous acrolein
Figure 2: Absorption spectra of a water-solvated acrolein molecule described by the PDE or PE model in combination with CCSDR(3), CCSD or CC2.
are not readily available with the experiment due to the overlap with the solvent excitations, so it is not clear which results are more accurate in this case. However, these excited states show similar behavior as the Rydberg states in formamide.

The results for the averaged excitation energies (based directly on the values for the calculated transition energies and not extracted from the positions of the peaks in the absorption spectrum) and solvatochromic shifts of the two lowest transitions, as well as the comparison with the experiments are given in Table 2. The given standard errors of the mean (SEM) are calculated as $\sigma / \sqrt{n}$, where $\sigma$ is the variance of the excitation energies through the snapshots and $n$ is the number of snapshots. Excitation energies as a function of snapshot are plotted in Fig. S3. The difference in the performance between PDE-CC and PE-CC is in general small compared to other errors involved, when describing the excitation energies of both $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions. It has previously been observed$^{28,72}$ that for some systems PE performs well in case of valence excitations based on cancellation of errors due to lack of non-electrostatic repulsion causing charge-leaks into the environment. PDE, on the contrary, provides a physically more correct picture of the short-range interactions, but this may not necessarily be reflected in a significantly better agreement with the experimental data. When comparing calculated solvent shifts with experimental solvent shifts, we see, however, that PDE-CC performs marginally better for all three CC models, except in case

<table>
<thead>
<tr>
<th></th>
<th>$n \rightarrow \pi^*$</th>
<th>$\pi \rightarrow \pi^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta E$</td>
<td>$\Delta \Delta E$</td>
</tr>
<tr>
<td>CC2</td>
<td>PDE 4.184 ± 0.013</td>
<td>0.285</td>
</tr>
<tr>
<td></td>
<td>PE 4.212 ± 0.016</td>
<td>0.313</td>
</tr>
<tr>
<td>CCSD</td>
<td>PDE 4.207 ± 0.013</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>PE 4.247 ± 0.016</td>
<td>0.315</td>
</tr>
<tr>
<td>CCSDR(3)</td>
<td>PDE 4.119 ± 0.013</td>
<td>0.280</td>
</tr>
<tr>
<td></td>
<td>PE 4.161 ± 0.016</td>
<td>0.323</td>
</tr>
<tr>
<td>experiment</td>
<td>3.94</td>
<td>0.25</td>
</tr>
</tbody>
</table>
of the $\pi \to \pi^*$ transition calculated with CC2.

### 4.3 4-nitrophenolate in water

The last system considered is the anionic 4-nitrophenolate ($p$-nitrophenolate or PNP$^-$) molecule solvated in water. We used 120 configurations consisting of one PNP$^-$ molecule, one Na$^+$ ion, and 400 water molecules, that were extracted from a QM/MM Born-Oppenheimer molecular dynamics (BOMD) trajectory.\(^{76}\)

Table 3: Average excitation energies (in eV) of the $S_1$ excited state of aqueous PNP$^-$ calculated with PDE or PE in combination with CCSDR(3), CCSD or CC2.

<table>
<thead>
<tr>
<th></th>
<th>CC2</th>
<th>CCSD</th>
<th>CCSDR(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDE</td>
<td>2.916 ± 0.018</td>
<td>3.147 ± 0.019</td>
<td>2.996 ± 0.019</td>
</tr>
<tr>
<td>PE</td>
<td>2.851 ± 0.020</td>
<td>3.089 ± 0.021</td>
<td>2.946 ± 0.022</td>
</tr>
</tbody>
</table>

The FDP2R and M2P2 embedding potentials\(^{77}\) for the PDE- and PE-CC calculations were parametrized in the same manner as in the case of formamide and acrolein, but this time using the 6-31+G* basis set\(^{78-81}\) also recontracted to ANO-type basis representation as required by the LoProp approach. The same scaling factor of 0.7 was included for the non-electrostatic repulsion. We calculated the excitation energies and oscillator strengths of the two lowest singlet excited states with either PDE or PE in combination with CCSDR(3), CCSD and CC2 methods using the 6-31+G* basis set.

The intense band in the spectrum (Fig. 3) corresponds to the HOMO $\to$ LUMO transition occurring upon excitation to the lowest singlet excited state ($S_1$). The corresponding experimental peak for PNP$^-$ in water has been observed at 3.09 eV (i.e. 401 nm).\(^{82}\) The peaks for PDE are located at somewhat higher energies than the PE peaks and closer to experimental peaks, however the differences between the embedding models are most likely not as significant as errors associated with other approximations involved in the overall computational procedure such as, e.g., the quality of the geometrical parameters of the solute and configurational sampling. We also observe that there is no systematic convergence towards...
the experimental value when including more electron correlation, i.e. the PDE-CCSD result is very close to the experimental result, while PDE-CCSDR(3) is closer to PDE-CC2.

![Absorption spectra](image)

**Figure 3:** Absorption spectra of a water solvated 4-nitrophenolate described using CCSDR(3), CCSD and CC2 in combination with either PDE or PE model.

The averaged excitation energies are given along with SEM in Table 3, and excitation energies as function of snapshot is plotted in Fig. S3. The lower excitation energies from PE-CC can be explained by the lack of non-electrostatic repulsion in PE-CC which leads to over-stabilization of the excited state. Because of the more diffuse character of the excited state it will be stabilized more than the ground state, thus, leading to a lower excitation energy. It is expected that such phenomena are more important for anionic solutes, where a more significant contribution of the diffuse functions can be found, which makes them prone to electron spill-out effects.
5 Conclusion

We have presented the theory and implementation of the PDE-CC method designed for calculation of molecular response properties of molecules in solutions. Specifically, the PDE model is combined with the CC2, CCSD, and CCSDR(3) methods. The PDE model is an extension of the previous PE model. In PDE the electrostatic effect of each solvent molecule on the core quantum part is modeled through the electron density and point nuclei in contrast to PE which relies on a multipole expansion. This provides a more accurate description of the electrostatic interactions between the solute and the solvent molecules in its nearest environment. Short-range non-electrostatic repulsion effects are described through a projection operator, which is not present in the PE model, and which effectively prevents electron spill-out and constitutes a significant improvement for cases where confinement of the solute wave function by the solvent plays an important role. Another advantage of PDE is that it retains the same efficient procedure for inclusion of environment polarization effects as the PE model.

The performance of the PDE-CC method was exemplified on three different molecules, namely aqueous solutions of formamide, acrolein and 4-nitrophenolate, and compared to the corresponding PE-CC results. The case of formamide has clearly shown the advantages of PDE-CC over PE-CC where we have demonstrated that non-electrostatic repulsion plays a crucial role for a correct description of the aqueous UV spectrum of formamide. For acrolein, PDE-CC is somewhat better at predicting the solvent shifts of the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions when comparing to the experiment. The overall spectrum of acrolein is not heavily affected by non-electrostatic repulsion in the commonly observed part (where the first two excited states are located), but the higher transitions are more strongly affected. The PDE-CC method predicts higher excitation energies for the $\pi \rightarrow \pi^*$ transition in 4-nitrophenolate compared to PE-CC which is explained by the difference in the stabilization of the involved states. The PE model is prone to over-stabilization effects due to the lack of non-electrostatic repulsion which plays an important role especially for anionic solutes.
In conclusion, we find PDE-CC to be a valuable extension of the PE-CC model that introduces significant refinements in the description of especially short-range interactions in systems with a CC core and a polarizable environment.

Acknowledgement

The authors thank Erik Rosendahl Kjellgren for providing the structures of 4-nitrophenolate in water solution. Computation/simulation for the work described in this paper was supported by the DeIC National HPC Centre, SDU. J. M. H. O. acknowledges financial support from the Carlsberg Foundation (grant id. CF15-0823). J.K. acknowledges financial support from the Danish e-Infrastructure Cooperation (DeIC), the Villum Foundation, and the Danish Council for Independent Research (DFF).

Supporting Information Available

This material is available free of charge via the Internet at http://pubs.acs.org/.

References


(2) Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? Theor. Chem. Acc. 2006, 117, 185–199.


(72) Hršak, D.; Olsen, J. M. H.; Kongsted, J. Optimization and transferability of non-


Graphical TOC Entry