Linear scaling computation of excited states

Satoshi Yokojima and GuanHua Chen

Department of Chemistry, University of Hong Kong, China

The segment-molecular-orbital representation is proposed and combined with the linear-scaling localized-density-matrix method. A segment-molecular-orbital is a molecular orbital in a segment of an entire molecule. Different cut-off lengths for different segment-molecular-orbitals are naturally introduced, which results in a drastic reduction of the computational cost.

Introduction

Ab initio molecular orbital calculations are usually limited to small- and medium-size molecular systems. The obstacle lies in the rapid increase of computational cost as the systems become larger and more complex. For the ground state, a number of linear scaling computational methods are proposed.[1,2] Here the locality of the ground state density matrix was the key to achieve the linear scaling computation. However, the linear scaling computation for excited states is a much more difficult task, since the excited state density matrix is more delocalized compared with the ground state one. Several linear scaling calculations based on the noninteracting electron models have been proposed.[3]

Recently we proposed linear scaling computational method, the localized-density-matrix (LDM) method,[4] which solves the equation of motion of the single electron reduced density matrix with explicit inclusion of electron-electron Coulomb interaction. This method has also been applied to the ground state calculation[5] and as well as excited states. The largest LDM calculation for excited state properties has been performed for a polyacetylene oligomer containing 33000 carbon atoms by employing the PPP Hamiltonian.[6]

We further improved the efficiency of the LDM method by introducing segment molecular orbital (SMO) representation. The LDM/SMO method allows us to investigate much larger systems than original LDM method.[7] We here give the detail of the idea of SMO representation and its application for the computation of optical response of very large systems.

Model

Although the LDM/SMO method may employ different models, we use here the CNDO/S Hamiltonian. In the presence of an external electric field $\mathcal{E}(t)$, the total Hamiltonian is given by the following expression:

$$H = H_{\text{CNDO/S}} + H_{\text{ext}}. \tag{1}$$

Here the Nishimoto-Mataga formula[8] is employed for two-electron integrals in $H_{\text{CNDO/S}}$. H_{ext} describes the interaction between the electrons and the external electric field $\vec{\mathcal{E}}(t)$, the dipole matrix elements $\vec{\mu}_{ij}$ in H_{ext} are evaluated using the zero differential overlap approximation. Within the time-dependent Hartree-Fock (TDHF) approximation,[9] a closed nonlinear self-consistent equation of motion is yielded for the reduced single-electron density matrix $\rho(t)$,

$$i(\hbar\frac{d}{dt} + \gamma)\rho(t) = [h(t) + f(t), \rho(t)]. \tag{2}$$

Here $h(t)$ is the Fock matrix:

$$h_{nm}(t) = t_{nm} + 2\delta_{nm}\sum_l v_{nl}\rho_{ll}(t) - v_{nm}\rho_{nm}(t), \tag{3}$$

where t_{nm} is the hopping matrix element between orbital m and n, and v_{nm} is the Coulomb repulsion between two electrons at AOs m and n, respectively. γ is a phenomenological dephasing constant. $f(t)$ describes the interaction between an electron and the external field $\vec{\mathcal{E}}(t)$, $f_{ij}(t) = e\vec{\mu}_{ij} \cdot \vec{\mathcal{E}}(t)$.

LDM

We denote the ground state density matrix and Fock matrix by $\rho^{(0)}$ and $h^{(0)}$, respectively. The i-th order of induced density matrix or Fock matrix in $\vec{\mathcal{E}}(t)$ is indicated by the superscript (i), *i.e.*, $\delta\rho^{(i)}$ or $\delta h^{(i)}$, respectively. Thus the equation for the linear optical response of the density matrix ρ is given by inserting $\rho = \rho^{(0)} + \delta\rho^{(1)} + \delta\rho^{(2)} + \delta\rho^{(3)} + \cdots$ into Eq. (2), retaining the first order in $\vec{\mathcal{E}}(t)$:

$$
\begin{aligned}
i\hbar\delta\dot{\rho}_{ij}^{(1)} = & \sum_l (h_{il}^{(0)}\delta\rho_{lj}^{(1)} - \delta\rho_{il}^{(1)}h_{lj}^{(0)}) \\
& + \sum_l (\delta h_{il}^{(1)}\delta\rho_{lj}^{(0)} - \delta\rho_{il}^{(0)}\delta h_{lj}^{(1)}) \\
& + \sum_l e\vec{\mathcal{E}}(t) \cdot (\vec{\mu}_{il}\rho_{lj}^{(0)} - \rho_{il}^{(0)}\vec{\mu}_{lj}).
\end{aligned}
\tag{4}
$$

Because of the locality of the density matrix, we can solve Eq. (4) with the following approximation when the system is large enough:

- $h_{i,j}^{(0)} = 0; \rho_{i,j}^{(0)} = 0$ if $|\vec{r}_i - \vec{r}_j| > l_0$.
- $\rho_{i,j}^{(1)} = 0$ if $|\vec{r}_i - \vec{r}_j| > l_1$.

In addition, the FMM[6,10] is employed to evaluate Coulomb interaction in $h^{(1)}$, which leads to the linear-scaling of computational time.

SMO representation

The following procedure is adopted to give the SMO representation. First the entire system is divided into many

segments, and the bonds between adjacent segments are severed. Each dangling bond is assigned two electrons. Secondly, the Hartree-Fock solution is determined for each segment by employing CNDO/S Hamiltonian. Denote respectively the Hartree-Fock molecular orbital coefficients and energy as S_{ij}^α and E_j^α for the j-th MO of the segment α, where i is the index of the i-th AO. Finally, all the resulting SMOs are taken as the basis functions, and transformation matrix S_{ij} from the AO representation to the SMO representation for the entire molecule is constructed as follows

$$S_{ij} = \begin{cases} S_{ij}^\alpha & i,j \in \alpha \\ 0 & i \in \alpha,\ j \in \beta\,(\alpha \neq \beta). \end{cases} \tag{5}$$

We denote the SMO representation by bar. For example, the density matrix $\bar\rho$ in the SMO representation is thus expressed as

$$\bar\rho_{ij} = \sum_{i_1 j_1} S_{ii_1}^\dagger \rho_{i_1 j_1} S_{j_1 j}. \tag{6}$$

Note that Eq. (6) is merely a change of the representation, and thus no approximation has been made here.

Computationally, there are two advantages for employing the SMO representation for the LDM calculation.
(1) In the AO representation many orbitals have a long cutoff length for the density matrix, whereas in the SMO representation the SMOs far from the Fermi energy or chemical potential have much shorter cutoff length. By introduction of different cutoff lengths for different pairs of SMOs, we may reduce drastically the number of density matrix elements to be considered explicitly, and consequently save much of the computational time.
(2) Those SMOs far away from the Fermi energy or chemical potential are optically inactive the visible range, and may be simply cut off from the density matrix and the equation of motion.

In our calculation we keep all SMOs but adopt different cutoff lengths for different SMO pairs, $i.e.$, employing only the approximation 1.

LDM/SMO method

By transforming the Eq. (4) from AO to SMO we have

$$i\hbar\delta\dot{\bar\rho}_{ij}^{(1)} = \sum_l (\bar h_{il}^{(0)} \delta\bar\rho_{lj}^{(1)} - \delta\bar\rho_{il}^{(1)} \bar h_{lj}^{(0)})$$
$$+ \sum_l (\delta\bar h_{il}^{(1)} \delta\bar\rho_{lj}^{(0)} - \delta\bar\rho_{il}^{(0)} \delta\bar h_{lj}^{(1)})$$
$$+ \sum_l e\vec{\mathcal{E}}(t) \cdot (\vec\mu_{il}\bar\rho_{lj}^{(0)} - \bar\rho_{il}^{(0)}\vec\mu_{lj}), \tag{7}$$

where

$$\delta\bar h_{ij}^{(1)} = 2 \sum_{i_1 j_1 mn} S_{ii_1}^\dagger S_{i_1 j} v_{i_1 j_1} S_{j_1 m} \delta\bar\rho_{mn}^{(1)} S_{nj_1}^\dagger$$
$$- \sum_{i_1 j_1 mn} S_{ii_1}^\dagger v_{i_1 j_1} S_{i_1 m} \delta\bar\rho_{mn}^{(1)} S_{nj_1}^\dagger S_{j_1 j}. \tag{8}$$

A theoretical investigation of the locality of the density matrix shows that the density matrix in the spatial representation decays exponentially,[11]

$$\rho(\vec r_1, \vec r_2) \sim \exp(-\gamma|\vec r_1 - \vec r_2|), \tag{9}$$

where γ is proportional to the energy gap for semiconductors and insulators. Therefore a critical distance for $|\vec r_1 - \vec r_2|$ is introduced beyond which $\rho(\vec r_1, \vec r_2)$ is neglected. This critical distance is proportional to the inverse of the HOMO and LUMO energy gap. The above result is obtained for infinite periodic system in the weak-binding limit. We observe that different types of orbitals have different critical distances. For instance, deeply bounded orbitals or high empty orbitals rarely contribute to the optical response and thus, they have a much smaller cutoff lengths. Thus, the cutoff lengths for the ground state density matrix $\bar\rho^{(0)}$ in the SMO representation may be given as follows:

$$C_{ij} = l_0 \times \frac{E_g^\alpha}{t_0|E_i^\alpha - E^\alpha(i)| + E_g^\alpha}$$
$$\times \frac{E_g^\beta}{t_0|E_j^\beta - E^\beta(j)| + E_g^\beta}$$
$$\times \frac{\bar E_g}{t_0|f(E_i^\alpha, E_j^\beta)| + \bar E_g}, \tag{10}$$

$$f(E_i^\alpha, E_j^\beta) = (E_i^\alpha - E^\alpha(i)) - (E_j^\beta - E^\beta(j)). \tag{11}$$

Here C_{ij} is the cutoff length for orbitals i and j. i and j are in two different segments α and β, respectively. E_g^α (E_g^β) is half of the energy gap in α (β), and $\bar E_g$ is half of the averaged energy gap for the entire system. $E^\alpha(i)$ is the HOMO (LUMO) energy of segment α if E_i^α is below HOMO (above LUMO). E_j^β is similarly defined. We set the density matrix element $\bar\rho_{ij}^{(0)} = 0$ in Eq. (7) when the distance between the centers of mass for α and β is longer than C_{ij}. The first term on the right hand side (R.H.S.) of Eq. (10), l_0, is simply the cutoff length when i and j are either HOMO or LUMO of respective segments. The second and third terms on the R.H.S. of Eq. (10) take into account that further the orbital i or j from the HOMO or LUMO, the shorter the critical length. $|E_i^\alpha - E^\alpha(i)|$ and $|E_j^\beta - E^\beta(j)|$ measure the energy difference of i and j with respect to their segment HOMO or LUMO. t_0 is a scaling constant which is introduced to control the variation of critical length for orbitals other than segment HOMO or LUMO. The fourth term accounts for the following fact: the larger the energy difference between i and j, the smaller their critical length.

The cutoff lengths for excited states are given by replacing (l_0, t_0) with (l_1, t_1). To take into account the bonding information between segments correctly, we include all the density matrix elements $\bar\rho_{ij}$ for i and j which reside on the same segment or the nearest neighbors. When we take $t_0 = t_1 = 0$, LDM/SMO method is almost equivalent to the original LDM except that the LDM/SMO method uses the distance between the centers of mass of two segments instead of the distance between two atoms.

Result and discussion

The absorption spectra of the 8-unit 4-chain PPV aggregates,

which contain 456 atoms (1220 orbitals), is shown in Fig. 1. The calculation is carried out in time-domain with SMO representation. It is an important issue whether the excitations in PPV is intra- or inter-chain excitations. However, calculating an optical response of a system with multiple chains is a difficult task because of the increasing number of the orbitals compared with the single chain case. For example, the lowest excitation of PPV spread out about 8-unit for π-orbitals. It would have required the full TDHF calculation, if it were not for SMO. The LDM/SMO method automatically selects the relevant density-matrix elements, which results in the large reduction of the computational tasks. The number of the density-matrix elements used for obtaining Fig. 1 is 382524 with $l_0 = l_1 = 32$ Å and $t_0 = t_1 = 0.35$. This is about 1/4 of the total number of the density-matrix elements. Since there are some mixing between σ and π orbitals because of the special arrangement of the chains, some of the σ orbitals contribute to the optical response.
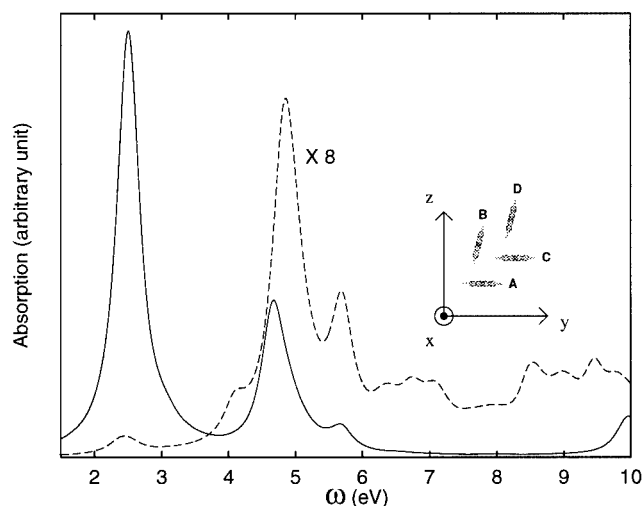
Fig. 1. Absorption spectra of the 8-unit 4-chain PPV aggregates. The electric field is polarized along the chain axis (solid line), or perpendicular to the plane of A (dashed line).

References

1) W. Yang: Phys. Rev. Lett. **66**, 1438 (1991).
2) For review, see S. Goedecker: Rev. Mod. Phys. **71**, 1085 (1999).
3) For example, T. Iitaka et al.: Phys. Rev. E **56**, 1222 (1997).
4) S. Yokojima and G. H. Chen: Phys. Rev. B **59**, 7259 (1999); Chem. Phys. Lett. **292**, 379 (1998); W. Z. Liang, S. Yokojima, and G. H. Chen: J. Chem. Phys. **110**, 1844 (1999).
5) S. Yokojima and G. H. Chen: Chem. Phys. Lett. **300**, 540 (1999); S. Yokojima, D. H. Zhou, and G. H. Chen: Chem. Phys. Lett. **302**, 495 (1999).
6) W. Z. Liang, S. Yokojima, D. H. Zhou, and G. H. Chen: J. Phys. Chem. A **104**, 2445 (2000).
7) S. Yokojima, X. J. Wang, D. H. Zhou, and G. H. Chen: J. Chem. Phys. **111**, 10444 (1999).
8) K. Nishimoto and N. Mataga: Z. Phys. Chem. N. F. **12**, 335 (1957).
9) A. Takahashi and S. Mukamel: J. Chem. Phys. **100**, 2366 (1994); P. Ring and P. Schuck: *The Nuclear Many-Body Problem*, (Springer, New York, 1980); G. H. Chen and S. Mukamel: J. Am. Chem. Soc. **117**, 4945 (1995).
10) L. Greengard: Science **265**, 909 (1994); H.-Q. Ding, N. Karasawa, and W. A. Goddard III: J. Chem. Phys. **97**, 4309 (1992).
11) S. Ismail-Beigi and T. A. Arias: Phys. Rev. Lett. **82**, 2127 (1999).