

A Direct INDO/SCI Method for Excited State Calculations

Aimée Tomlinson and David Yaron*

Department of Chemistry, Carnegie Mellon University, Pittsburgh, PA 15213

(Dated: (submitted to J. Comp. Chem.))

Abstract

Intermediate neglect of differential overlap (INDO) is the most commonly utilized semi-empirical technique for performing excited state calculations on large organic systems such as organic semiconductors and fluorescent dyes. The calculations are typically done at the singles configuration interaction (SCI) level. Direct methods provide a more efficient means of performing configuration interaction (CI) calculations, and the computational trade-offs associated with various approaches to direct-CI theory have been well characterized for ab initio Hamiltonians and high-order CI. However, the INDO and SCI approximations lead to a new set of trade-offs. In particular, application of the electron-electron interactions in the atomic basis leads to a computational savings that scales as the number of atomic orbitals, which for a large organic system can be two to three orders of magnitude. Furthermore, substantial memory savings are achieved by avoiding the evaluation of the two electron integrals in the molecular orbital basis.

*Electronic address: yaron@chem.cmu.edu

I. INTRODUCTION

Intermediate Neglect of Differential Overlap (INDO)[1] is a widely utilized semi-empirical model for the excited states of large organic systems, such as conjugated polymers and fluorescent dyes[2–4]. Within INDO, the excited states are typically obtained at the singles-configuration interaction (SCI) level. This work develops a direct method for INDO/SCI calculations that provides a more efficient approach to these computations. Roos and Siegbahn[5–7] first introduced the direct configuration interaction (direct-CI) method, which was instrumental in making CI applicable to large systems. The computational trade-offs associated with various approaches to direct-CI theory have been well characterized for ab initio Hamiltonians and high-order CI[8]. However, the INDO and SCI approximations lead to a new set of trade-offs. Most importantly, the evaluation of electron-electron interactions in the atomic basis provides a considerable savings in INDO theory, whereas it is not advantageous in ab initio theories[9]. For direct INDO/SCI, the computational savings relative to non-direct approaches are on the order of the number of atomic basis functions in the calculation, which can be two to three orders of magnitude for large organic systems.

In CI theory, the electronic wavefunction is written as a linear combination of many-body functions, which are obtained as excitations from the Hartree-Fock ground state. In SCI, the excited states include only singly excited electronic configurations,

$$|\psi\rangle = \sum_{r,a} c_{r,a} |\psi_a^r\rangle \quad (1)$$

For singlet excited states, $|\psi_a^r\rangle$ is given by

$$|\psi_a^r\rangle = \frac{1}{\sqrt{2}} (a_{r,\alpha}^\dagger a_{a,\alpha} + a_{r,\beta}^\dagger a_{a,\beta}) |HF\rangle \quad (2)$$

where $|HF\rangle$ is the restricted Hartree-Fock ground state, and $a_{a,\sigma}^\dagger$ ($a_{a,\sigma}$) is the creation (destruction) operator for an electron in molecular orbital a with spin σ . Throughout this paper, $a, b \dots$ are used for molecular orbitals that are occupied in the Hartree-Fock ground state, and $r, s \dots$ are used for unoccupied molecular orbitals.

The coefficients, $c_{r,a}$ of Eq. (1), are determined variationally. This is equivalent to finding the eigenfunctions of the $N \times N$ Hamiltonian matrix within the SCI basis, where N is the number of singly excited configurations. We will refer to methods that explicitly create the Hamiltonian matrix as traditional methods. For full diagonalization of this $N \times N$ matrix, the

computational effort scales as N^3 , and gives all eigenvalues and eigenstates. An alternative approach is to use an iterative diagonalization scheme that repeatedly applies the matrix to an initial starting vector,

$$|\psi'\rangle = \hat{H} |\psi\rangle \quad (3)$$

In this work, we use the implicitly restarted Arnoldi method[10–12] which, for the systems considered here, takes about 500 applications of the Hamiltonian to a vector in order to obtain the lowest few eigenstates. Since matrix-vector multiplication scales as N^2 , iterative diagonalization of the SCI Hamiltonian matrix scales as $N^2 N_{Appl}$, where N_{Appl} is the number of applications. For $N > N_{Appl}$, iterative diagonalization is less computationally demanding than full diagonalization.

In addition to the computational savings of iterative diagonalization discussed above, direct CI techniques compute $\hat{H} |\psi\rangle$ of Eq. (3) without explicitly evaluating the Hamiltonian matrix. This leads to substantial computational savings, due to the sparseness of the Hamiltonian matrix. Substantial memory savings also result, since it is no longer necessary to store the full Hamiltonian matrix.

In this work, we will demonstrate an efficient direct approach to INDO/SCI calculations. The following section introduces a benchmark computational system and discusses the computational cost and memory requirements of a traditional calculation. Section III gives an overview of the method, examines the scaling of the computational effort with system size, and presents calculations on some test systems. Section IV gives a brief summary and discusses possible extensions.

II. BENCHMARK SYSTEM

In considering the computational savings resulting from direct INDO/SCI, we will compare this method with that used in traditional INDO/SCI. For comparison, we'll assume both the traditional and the direct methods use the same iterative matrix diagonalization scheme. The numbers quoted here are for the Arnoldi method discussed above. The difference between traditional and direct methods then arises only from the way in which the Hamiltonian is applied to a trial vector, Eq. (3). In traditional methods, the Hamiltonian matrix is computed and Eq. (3) is a matrix vector multiply. The direct method uses the optimized approach developed in Section III. The differences relate only to computational

efficiency, with both approaches giving identical results (to within numerical roundoff).

To make the comparison more concrete, we will consider a system containing 100 heavy atoms and 100 hydrogen atoms such that the number of atomic basis functions, N_{AO} , is 500. In INDO/SCI calculations, the number of molecular orbitals included in the SCI basis of Eq. (1) is increased until the desired accuracy is obtained. For our benchmark system, we will assume inclusion of 100 occupied molecular orbitals, N_h , and 100 unoccupied molecular orbitals, N_e . The total number of singly excited configurations is then $N_e N_h$ (10,000) (where the number in parentheses is the numerical value for our benchmark system). Finally, we will assume that 500 applications, N_{Appl} , of $\hat{H}|\psi\rangle$ are needed to obtain the lowest lying excited states.

In the traditional method, the two most computationally demanding steps are the transformation of the two-electron integrals from the atomic orbital (AO) to the molecular orbital (MO) basis and the diagonalization of the Hamiltonian. We will report the computational effort for each process as a symbolic expression, followed in parentheses by the numerical value of this expression for our benchmark system. Transformation of the two-electron integrals to the MO basis scales as $N_e N_h N_{AO}^2$ (2.5×10^9). The memory required to store these integrals is $N_e^2 N_h^2 / 4$ floating point numbers (200 Mbytes). For iterative diagonalization, one application of the Hamiltonian to the wavefunction, Eq. (3), scales as $N_e^2 N_h^2$ (1×10^8). Therefore, the computational effort required to obtain the lowest lying excited states scales as $N_e^2 N_h^2 N_{Appl}$ (5×10^{10}). In summary, the computational effort associated with the traditional calculation scales as $N_e N_h N_{AO}^2$ for the two-electron transformation, and $N_e^2 N_h^2 N_{Appl}$ for the diagonalization.

III. DIRECT INDO/SCI

A. The Method

Direct methods can lead to substantial computational savings when a fast routine for applying the Hamiltonian to a trial wavefunction (vector) is available, see Eq. (3). The matrix elements of the Hamiltonian in SCI theory are

$$\langle \psi_b^s | \hat{H} | \psi_a^r \rangle = F_{s,r} \delta_{b,a} - f_{b,a} \delta_{s,r} - (rs|ab) + 2(ar|sb) \quad (4)$$

where F and f are subblocks of the Fock matrix within the unoccupied and occupied molecular orbital basis, respectively. The two-electron integrals are for the MOs in chemist's notation

$$(sa|rb) = \int d\vec{r}_1 d\vec{r}_2 \phi_s^*(\vec{r}_1) \phi_a(\vec{r}_1) \times \left| \vec{r}_2 - \vec{r}_1 \right|^{-1} \phi_r^*(\vec{r}_2) \phi_b(\vec{r}_2) \quad (5)$$

where ϕ_a are the molecular orbitals.

For notational convenience, the coefficients of the SCI wavefunction from Eq. (1) will be expressed as matrices, $c_{a,r}$. The Hamiltonian from Eq. (4) is then a supermatrix,

$$\hat{H}_{\{s,b\},\{r,a\}} = \langle \psi_b^s | \hat{H} | \psi_a^r \rangle \quad (6)$$

Application of the Hamiltonian to the initial starting vector may be written as

$$d_{s,b} = \sum_{r,a} \hat{H}_{\{s,b\},\{r,a\}} c_{r,a} \quad (7)$$

where $c_{r,a}$ are the expansion coefficients for the initial vector, $|\psi\rangle$ of Eq. (3), and $d_{s,b}$ are the expansion coefficients for the final vector, $|\psi'\rangle$ of Eq. (3).

The derivation begins with Eqs. (7) and (4), and examines each term of the Hamiltonian matrix elements of Eq. (4) in turn. In the supermatrix notation used here, the first two terms of Eq. (4) can be written as matrix multiplies. The first term leads to,

$$\sum_{r,a} (F_{s,r} \delta_{b,a}) c_{r,a} = \sum_r F_{s,r} c_{r,b} = Fc \quad (8)$$

and the second term leads to $-cf$. Note that these matrices are quite small ($N_e \times N_e$) and are handled efficiently in the MO basis. However, for the third and fourth terms in Eq. (4), it is advantageous to work in the atomic orbital representation. This advantage is due to the two-electron terms of INDO theory being sparse in the AO basis. To exploit this sparsity, we need to be able to transform the S-CI wavefunction of Eq. (1) between the MO and AO representation. In the AO representation, the wavefunction may be written,

$$|\psi\rangle = \sum_{j,i} \bar{c}_{j,i} |\psi_i^j\rangle \quad (9)$$

with

$$|\psi_i^j\rangle = \frac{1}{\sqrt{2}} (a_{j,\alpha}^\dagger a_{i,\alpha} + a_{j,\beta}^\dagger a_{i,\beta}) |HF\rangle \quad (10)$$

$a_{i,\sigma}^\dagger$ ($a_{i,\sigma}$) is the creation (destruction) operator for an electron in atomic orbital i with spin σ . Throughout this paper, $i, j \dots$ will be used for atomic orbitals, with summation being over all atomic orbitals. The bar in Eq. (9) indicates that the coefficients $\bar{c}_{j,i}$ are in the AO, as opposed to MO, basis. The transformation from the MO to AO basis is given by,

$$\bar{c}_{j,i} = \sum_{r,a} V_{j,r} c_{r,a} v_{a,i}^T \quad (11)$$

where $V_{j,r}$ is the amplitude of the r^{th} unoccupied MO on the j^{th} atom and $v_{i,a}$ is the amplitude of the a^{th} occupied MO on the i^{th} atom. The reverse transformation is

$$c_{r,a} = \sum_{i,j} V_{r,j}^T \bar{c}_{j,i} v_{i,a} \quad (12)$$

In order to apply the two-electron terms in the atomic orbital basis, we first express the two-electron matrix elements of Eq. (4) in terms of the AO basis. The third term of Eq. (4) becomes

$$(rs|ab) = \sum_{m,n,o,p} V_{m,r} V_{s,n}^T (mn|op) v_{a,o}^T v_{p,b} \quad (13)$$

where $(mn|op)$ are the two-electron integrals in the AO basis. The INDO Hamiltonian includes only coulomb, J,

$$\bar{J}_{ij} = (ii|jj) \quad (14)$$

and exchange, K, integrals,

$$\bar{K}_{ij} = (ij|ij) = (ij|ji) \quad (15)$$

and we take K_{ii} to be zero. The two-electron terms, $(mn|op)$, may be written as,

$$(mn|op) = \bar{J}_{n,p}(\delta_{m,n}\delta_{o,p}) + \bar{K}_{n,m}(\delta_{m,o}\delta_{n,p} + \delta_{m,p}\delta_{n,o}) \quad (16)$$

Using Eq. (13) and Eq. (16), the third term in the Hamiltonian matrix of Eq. (4) becomes:

$$-(rs|ab) = \sum_{m,n,o,p} -V_{m,r} V_{s,n}^T [\bar{J}_{n,p}(\delta_{m,n}\delta_{o,p}) + \bar{K}_{n,m}(\delta_{m,o}\delta_{n,p} + \delta_{m,p}\delta_{n,o})] v_{a,o}^T v_{p,b} \quad (17)$$

A similar transformation is performed on the fourth term in the Hamiltonian matrix, $(ar|sb)$.

There are two types of expressions which are obtained when the two-electron terms are applied to the $c_{a,r}$ matrix, Eq. (7). An example of the first type is the coulomb term from

Eq. (17),

$$\begin{aligned} \sum_{m,n,o,p,r,a} -V_{m,r} V_{s,n}^T [\bar{J}_{n,p} (\delta_{m,n} \delta_{o,p})] v_{a,o}^T v_{p,b} c_{r,a} \\ = \sum_{n,p,r,a} -V_{s,n}^T \bar{J}_{n,p} v_{p,b} V_{n,r} c_{r,a} v_{a,p}^T \end{aligned} \quad (18)$$

A comparison of Eqs. (18) and (11) leads to the following simplification,

$$\sum_{n,p} -V_{s,n}^T \bar{J}_{n,p} \bar{c}_{n,p} v_{p,b} = -V^T (\bar{J} \circ \bar{c}) v \quad (19)$$

where the “ \circ ” indicates an element by element multiply. An example of the second type is the first exchange term from Eq. (17),

$$\begin{aligned} \sum_{n,m,o,p,r,a} -V_{m,r} V_{s,n}^T [\bar{K}_{n,m} (\delta_{m,o} \delta_{n,p})] v_{a,o}^T v_{p,b} c_{r,a} \\ = \sum_{r,a,n,m} -V_{s,n}^T \bar{K}_{n,m} V_{m,r} c_{r,a} v_{a,m}^T v_{n,b} \end{aligned} \quad (20)$$

Comparison with Eq. (11) leads to the following simplification

$$\sum_{n,m} -V_{s,n}^T \bar{K}_{n,m} \bar{c}_{m,m} v_{n,b} \quad (21)$$

For convenience, we introduce the following function, which takes a matrix and returns a diagonal matrix

$$[F(\bar{K})]_{n,l} = \sum_m \bar{K}_{n,m} \bar{c}_{m,m} \delta_{n,l} \quad (22)$$

Eq. (21) then becomes

$$\sum_{n,l} -V_{s,n}^T [F(\bar{K})]_{n,l} v_{l,b} = -V^T [F(\bar{K})] v \quad (23)$$

A similar treatment of all terms in the Hamiltonian matrix element of Eq. (4) leads to,

$$\begin{aligned} d = Fc - cf + V^T [-(\bar{J} \circ \bar{c}) - [F(\bar{K})] - (\bar{K} \circ \bar{c})^T \\ + [F(2\bar{J})] + 2(\bar{K} \circ \bar{c})^T + 2(\bar{K} \circ \bar{c})] v. \end{aligned} \quad (24)$$

The first three terms in brackets arise from $-(rs|ab)$ of Eq. (4) and the last three terms in brackets arise from $2(ar|sb)$ of Eq. (4). Combining like terms, we obtain the final expression

$$\begin{aligned} d = Fc - cf + V^T \left[(2\bar{K} - \bar{J}) \circ \bar{c} \right. \\ \left. + (\bar{K} \circ \bar{c})^T + [F(2\bar{J} - \bar{K})] \right] v \end{aligned} \quad (25)$$

B. Scaling of the Method

This section discusses the computational scaling of the direct method, using the benchmark system of Section II as a concrete example for comparison with the traditional approach. The computational expense associated with Eq. (25) is dominated by the matrix multiplications. One application of the Hamiltonian to a trial vector scales as $N_e N_{AO}^2$ (2.5×10^7). Since the two-electron integrals are not transformed to the molecular orbital basis, the storage requirements are modest, N_{AO}^2 floating point numbers (1Mbyte). The diagonalization then scales as $N_e N_{AO}^2 N_{Appl}$ (1.25×10^{10}). The ratio of computation time for traditional to direct techniques scales as,

$$\frac{Trad\ INDO/SCI}{Direct\ INDO/SCI} = \mathcal{O}\left(\frac{N_e N_h^2}{N_{AO}^2}\right) \quad (26)$$

where “ \mathcal{O} ” indicates “on the order of.” Fig. (1) demonstrates the validity of the scaling factor given in Eq. (26). Each point in the graph represents a polyene with a different number of vinyl units, n , ranging from two to eight. The SCI basis includes excitations between all occupied and unoccupied orbitals. The linear relation shown in Fig. 1 confirms the predicted scaling behavior of Eq. (26).

When all singly excited configurations are included in the S-CI basis, application of the Hamiltonian to a trial vector via Eq. (25) scales as N_{AO}^3 . This is confirmed in Fig. (2).

C. Test Cases

In order to demonstrate the savings which may be achieved through the use of the direct method, two test cases are presented. These tests were performed on a Pentium III 733MHz processor with 512Mbytes of RAM running linux kernel version 2.4.18. For the first case, we wanted to directly compare the direct method to the traditional method. We computed the first five excited states of stilbene, for which N_{AO} is 68, and all singly excited configurations (1156) were included. The amount of time required by the traditional technique was 160 seconds, whereas the direct method took 2.5 seconds. This corresponds to a savings which is on the order of N_{AO} , as expected.

To illustrate the direct method on a large system, we considered a poly(paraphenylenevinylene) (PPV) oligomer with sixteen phenyl rings, where N_{AO} is 600. All singly excited configurations (90,000) were included and the calculation took 38

minutes. In order to apply the traditional INDO/SCI method to this system, it would take approximately eleven days to perform and require 300Gbytes of memory. Therefore, our direct method may be applied to systems which were previously unmanageable.

IV. CONCLUDING REMARKS

The direct approach to INDO/SCI computations reported here substantially reduces the computational resources required for excited state calculations. This enables computations to be performed on large systems at a fraction of the time and memory required by the traditional method. These savings take advantage both of the sparsity of the two-electron integrals in the INDO Hamiltonian and of the inclusion of only singly excited configurations in the CI basis. Due to these features of INDO/SCI, application of the two-electron terms in the atomic basis set leads to the observed computational savings. The degree to which this approach can lead to savings in higher order CI, or for Hamiltonians with less-sparse two-electron integrals, such as AM1[13], is left for future work.

V. ACKNOWLEDGMENTS

This work was funded by the National Science Foundation (CHE9985719) and the Dreyfus Foundation.

-
- [1] J. Ridley and M. Zerner, *Theor. Chim. Acta* **32**, 111 (1973).
- [2] M. Zerner, *Rev. Comput. Chem.* **2**, 313 (1991).
- [3] M. Zerner, *NATO ASI Series C* **500**, 249 (1997).
- [4] J. Bredas et al., *Acc. Chem. Res.* **32**, 267 (1999).
- [5] B. Roos, *Chem. Phys. Lett.* **15**, 153 (1972).
- [6] B. Roos and P. E. M. Siegbahn, *Int. J. Quan. Chem.* **17**, 485 (1980).
- [7] P. E. M. Siegbahn, *Chem. Phys. Lett.* **109**, 417 (1984).
- [8] I. Shavitt, *Mol. Phys.* **94**, 3 (1998).
- [9] V. R. Saunders and J. H. V. Lenthe, *Mol. Phys.* **100**, 167 (2002).
- [10] F. M. Gomes and D. Sorensen, *Arpack++*, URL:<http://www.ime.unicamp.br/~chico/arpac++/> accessed August 2002.
- [11] W. E. Arnoldi, *Quart. Appl. Math.* **9**, 17 (1951).
- [12] C. Lanczos, *J. Res. Nat. Bur. Stand.* **45**, 255 (1950).
- [13] M. J. S. Dewar et al., *J. Am. Chem. Soc.* **107**, 3902 (1985).

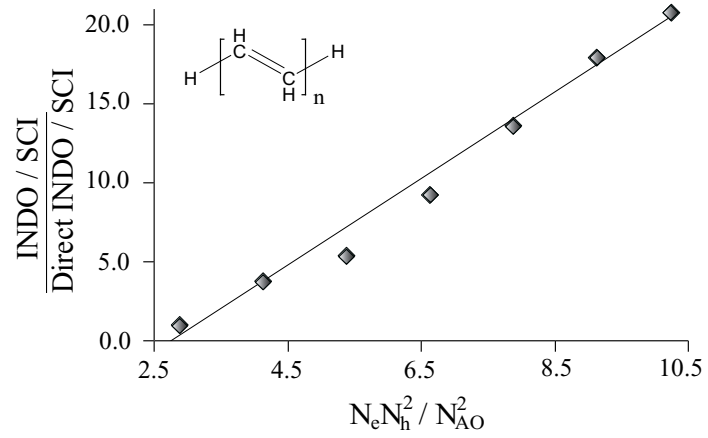


FIG. 1: Comparison of computer times required for traditional and direct INDO/SCI calculations on polyene chains containing $n=2$ to $n=8$ double bonds. The ratio of traditional to direct computational times is plotted versus the right hand-side of Eq. (26), $N_e N_h^2 / N_{AO}^2$.

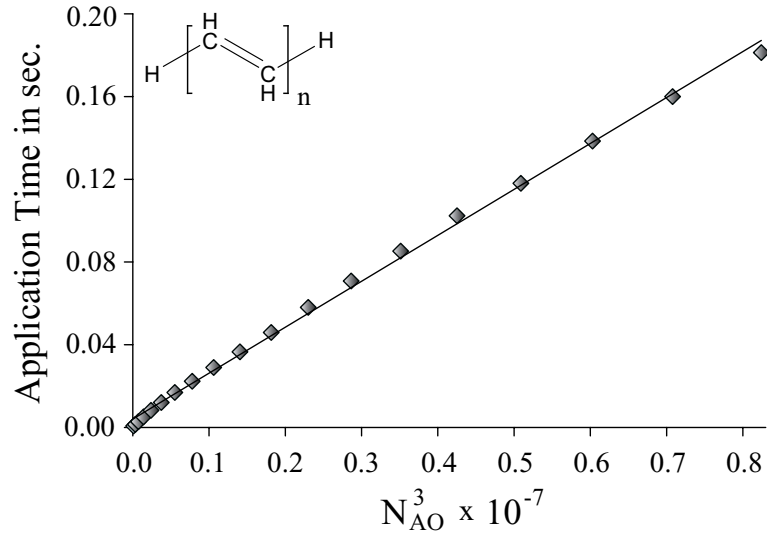


FIG. 2: Computer time required to apply the Hamiltonian to a trial vector in direct INDO/SCI, Eq. (25), for polyenes containing $n=2$ to $n=20$ double bonds. The time is plotted versus the number of atomic orbitals cubed.