

COMMUNICATION

Efficient Implementation of Time-Dependent Density Functional-Based Tight-Bind Method on Multi-core and GPU System

Guo-Hong Fan, Ke-Li Han^{*1}

¹State Key Laboratory of Molecular Reaction Dynamics and Center for Theoretical and Computational Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Science, Dalian 116023, P.R.China.

Received 29 Mar 2013; Accepted (in revised version) 15 April 2013

Special Issue: Guo-Zhong He Festschrift

Abstract: This communication focuses on the implementation of the time-dependent density functional-based tight-bind (TD-DFTB) method on the Multi-core and GPU system to be used in excited state calculations of large system. Sparse matrix and OpenMP multithreading are used to speed up the program. The most time-consuming part of the DFTB ground state is implemented on the graphical processing units (GPUs) with double precision. The block Davidson algorithm is used for finding the lowest eigenvalue and eigenvector of the large time-dependent density-functional (TDDFT) matrix. Compared with the first-principle CIS and full TDDFT calculation, the implementation of the fast TD-DFTB code on serials of organic molecules shows that the fast TD-DFTB code can obtain reasonable result with a much cheaper computational requirement.

AMS subject classifications: 68N19, 68U01, 68U20

Keywords: Density functional theory, tight-bind method, DFT, TDDFT, DFTB, TD-DFTB, GPU

SCC-DFTB is derived using a Taylor series expansion of the DFT [1] total energy with respect to the charge density fluctuations around a given reference charge density. From the

^{*}Corresponding author. E-mail address: klhan@dicp.ac.cn (Ke-li Han).

full Kohn-Sham DFT expressed on a local basis, SCC-DFTB undergoes several main approximations [2]. With the approximations the DFTB total energy and eigenvalue equation can be briefly written as: [3]

$$E = \sum_i \sum_{\mu\nu}^{\text{occ}} c_{\mu i}^* c_{\nu i} H_{\mu\nu}^0 + \frac{1}{2} \sum_{AB} q_A \gamma_{AB} q_B + E_{\text{rep}}$$

$$\sum_{\nu} \left(H_{\mu\nu} c_{\nu i} - \epsilon_i S_{\mu\nu} c_{\nu i} \right) = 0 \quad (1)$$

The time-dependent linear response extension of DFT is used for the excited states. The resulting TDDFT eigenvalue equation is: [4, 5]

$$\sum_{jb} \left[\omega_{ia}^2 \delta_{ij} \delta_{ab} + 4 \sqrt{\omega_{ia}} K_{ia,jb}^{\Sigma} \sqrt{\omega_{jb}} \right] F_{jb}^{\Sigma} = \Omega_{\Sigma}^2 F_{ia}^{\Sigma} \quad (2)$$

To ignore the sophisticated integration, the same gamma approximation in the ground state is introduced to coupling matrix $K_{ia,jb}$ [6]. Eq. 2 is diagonalized to obtain the singlet and triplet excited state energies.

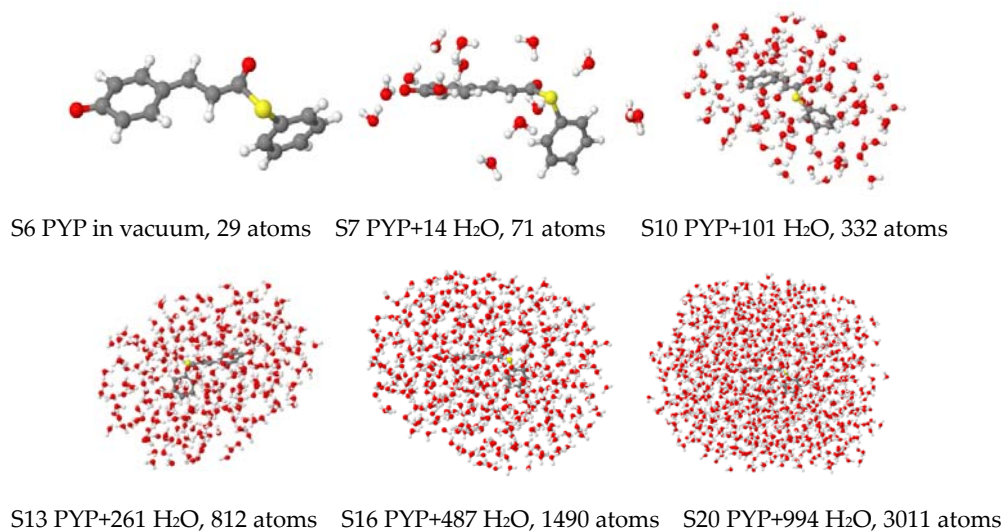


Figure 1: Structure of a PYP chromophore in vacuum (S6) and with an increasing number of solvated waters. Only five solvated PYP structure are shown (S7, S10, S13, S16, and S20).

The Hamiltonian matrix and the overlap matrix in Eq.1 are built and updated in a compressed sparse-row format. The OpenMP multi-thread technique is used to build up the sparse matrix. This reduces the computational cost of building the Hamiltonian matrix element into a small portion of the total ground-state procedure. The main computational time is then spent on the diagonalization of the eigenvalue Eq. 1. The MAGMA[7] library is a dense linear algebra library similar to LAPACK, but used for heterogeneous/hybrid

"Multi-core + GPU" architectures. The diagonal LAPACK routine DSYGVD is replaced with the MAGMA GPU diagonal routine MAGMA_DSYGVD. Double precision is used throughout the program and all the properties can be preserved. An iterative method such as the Davidson algorithm [8-10] is used to calculate the eigenvectors without explicitly constructing the TDDFT matrices. The block Davidson algorithm finds several lowest eigenvalues of the TDDFT matrix by starting with an arbitrary eigenspace with dimension l ($l \cong N$) and constructing the orthogonal basis ($K_1, K_2, \dots, K_l$). The trial vector and the eigenspace at some iteration are expanded in a linear combination of the orthogonal vectors. The linear combination coefficients in the equation are found by the Ritz variational procedure, leading to the projected eigenvalue equation. The dimension of the eigenvalue equation is much lower than the original TDDFT eigenvalue problem which can be solved more efficiently. New subspace vectors orthogonalized to the previous subspace vectors are added to the subspace. The whole procedure is repeated until convergence is reached.

At each iteration, the algorithm requires the formation of the matrix-vector product HF_l which is the most time-consuming part of the whole procedure. The matrix vector-product of the eigenvalue equation 2 can be further written as:

$$\omega_{jb}^2 F_{jb} + 4\omega_{ia}^2 \sum_A q_A^{ia} \left(\sum_B \gamma_{AB} \left(\sum_{jb} q_B^{jb} \omega_{jb}^2 F_{jb} \right) \right). \quad (3)$$

The OpenMP implementation of the second part is conducted as follows. Step 1: The temp product with the initial eigenvector for each atom is calculated; Step 2: The product of the gamma matrix with temp vector is produced; Step 3: The dot-like product is conducted for each ia . The first step for the OpenMP can be parallelized by adding the OpenMP parallel reduction directive before the jb loop. The parallel of the matrix vector product of Step 2 is implemented by the OpenMP version of the LAPACK routine DSYMV in the Intel Math Kernel Library. The parallelization of the last step of the matrix vector product is implemented using the OpenMP parallel do loop directive. We have also tried to utilize GPU to accelerate the matrix vector product.

Table 1: TD-DFTB excited state energy and the time spent on the CPU with the OpenMP and block Davidson methods obtained on a platform with an 8-core Intel Xeon E5530 CPU. The full GPU-TDDFT results and molecule labels are taken from Ref. 11 for comparison.

	Time (second)				Excite energy S1-S0			
	CIS	BLYP	B3LYP	TDDFTB	CIS	BLYP	B3LYP	TDDFTB
S1(44)	12.1	7.8	23.6	0.3	3.95	2.57	3.13	2.54
S2(128)	97.1	42.6	92.5	1.1	4.11	2.28	2.98	2.40
S3(296)	711.8	94.4	425.6	7.2	4.12	2.10	2.93	2.26

S4(636)	3508.2	230.2	1369.8	46.1	3.97	1.76	2.76	2.14
S5(112)	41.9	19.4	50.0	0.7	4.71	2.81	3.52	3.00
S6(29)	7.1	4.7	8.7	0.3	3.75	2.32	2.95	2.06
S7(71)	13.7	9.1	19.7	0.4	3.70	2.15	2.90	2.29
S8(149)	38.2	19.3	54.6	1.2	3.67	0.83	2.52	2.20
S9(218)	71.9	32.1	83.2	2.9	3.66	0.73	2.40	2.16
S10(332)	133.2	54.8	111.9	8.3	3.64	0.08	1.65	1.88
S11(467)	217.5	77.7	173.4	15.5	3.64	0.03	1.27	2.059
S12(605)	318.1	-	217.7	28.5	3.64	-	1.14	1.83
S13(812)	493.2	-	335.2	35.9	3.65	-	0.86	1.80
S14(962)	656.3	-	266.9	40.5	3.65	-	0.60	1.62
S15(1220)	894.0	-	910.3	62.1	3.65	-	0.09	0.71
S16(1490)	1221.2	-	1218.7	96.9	3.65	-	0.05	0.89
S17(1751)	1583.5	-	833.6	159.2	3.66	-	-0.05	1.16
S18(2078)	2118.2	-	958.0	266.1	3.65	-	0.57	1.48
S19(2399)	4504.1	-	1113.7	397.5	3.65	-	0.80	1.76
S20(3011)	5644.3	-	-	773.5	3.65	-	-	1.61

We then implement the TD-DFTB code on the excited state energy of large molecules shown in **Table 1**. The full TDDFT result is obtained from Isborn et al[11] which use GPU to accelerate the full TDDFT calculation. Their calculations are conducted with a GPU-based software package using the configuration interaction singles (CIS) and the adiabatic linear response of Tamm-Dancoff time-dependent density functional theory [12]. Their calculations are run on platform with dual quad-core Intel Xeon X5570 CPUs and 8 Tesla C1060 GPUs. The tests[11] shown in **Figure 1** include four generations of oligothiophene dendrimers[13] (S1-S4), the entire photoactive yellow protein (PYP)[14] solvated by TIP3P[15] water molecules (S5), and an analog of the PYP chromophore called deprotonated trans-thiophenyl-p-coumarate[16] solvated with an increasing number of QM waters (S6-S20). **Table 1** shows that the CIS result is generally higher than the TDDFT with the BLYP functional and the hybrid B3LYP functional[17, 18]. Some of the GPU-TDDFT results from Ref. 11 are unknown, because the ground-state SCF procedure does not converge to reach a reasonable result. The TD-DFTB result does not show the problems. All ground-state SCC run with the charge error of 1.0×10^{-6} converge in a number of circles less than 20. However, some molecules such as the test case S15 and S16 have an over underestimated excited state energy which may be caused by the general problem of TDDFT to unable to correctly describe the charge transfer characteristic of these transitions[17]. The TD-DFTB

result generally lies between the result of the CIS and TD-B3LYP. It can reach an excited state energy close to that of the TD-B3LYP, with much less time and memory requirement.

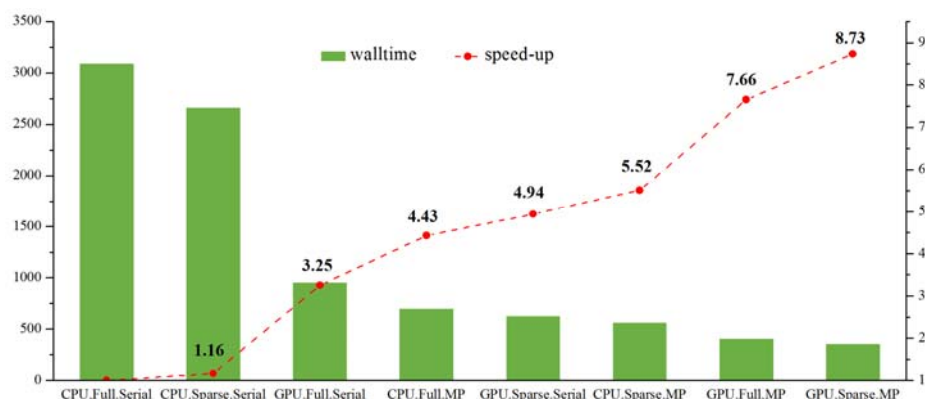


Figure 2: Time and speed of the ground state with different programming skills. The calculation is three self-consistency charge steps of the ground-state run for a 270 angstrom long carbon nanotube with 2560 atoms. The tests are conducted on a machine with an 8-core Intel Xeon E5620 CPU and an NVIDIA Tesla C2050 GPU card.

Figure 2 shows the program performance on a machine with an 8-core Intel Xeon E5620 CPU and an NVIDIA Tesla C2050 GPU card. The calculation is three self-consistency steps of the ground-state run on a 270 angstrom long carbon nanotube with 2560 atoms. With the GPU acceleration the best speed-up is 8.73 compared with the serial code run on CPU. The figure also shows that each programming technique contributes to the total speed-up of 8.73. **Figure 3(a)**, shows the timing information of the first three excited-state calculations for the block Davidson method and the ARPACK package which is a commonly used package for computation of few eigenvalues of large matrices by the Implicitly Restarted Arnoldi/Lanczos Method [19]. As the systems size increases, the time for the block Davidson increases slowly, whereas the time for the ARPACK solver increases rapidly. The iterations taken and the number of matrix vector product with ARPACK and the block Davidson method is shown in **Figure 3(b)**. Compared with the commonly used ARPACK package, the block Davidson method uses fewer iterations and matrix vector products. **Figure 4** shows the scaling property of the program as the basis size increases. The time for the excited state increase in the quadratic trend as the number of basis size increases. The TD-DFTB code with the block Davidson method can obtain excited energy within tens of minutes for systems with thousands of atoms, much faster than the GPU-CIS calculation with the minimal 6-31G basis set.

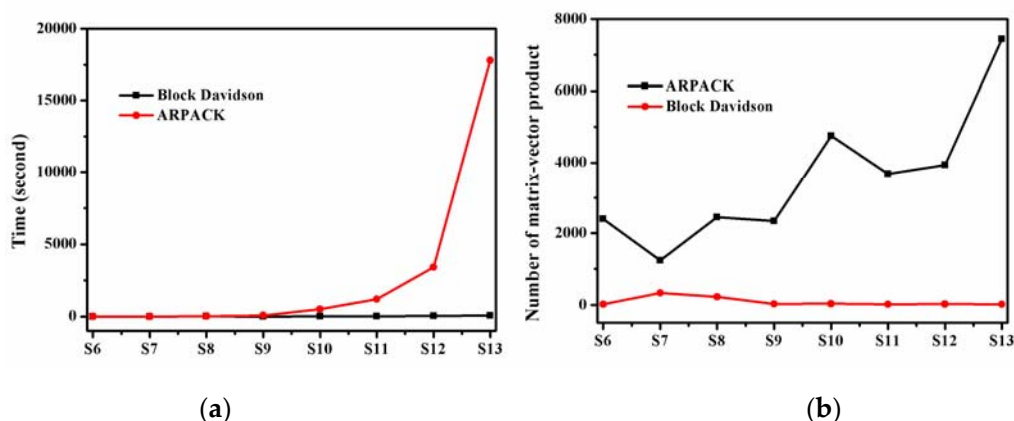


Figure 3: (a) Timing report of the first three excited-state calculations from S6 to S13 with different iterative methods. (b) The corresponding number of matrix vector products taken for TD-DFTB excited states. The convergence threshold (norm of the residual vector) of both methods is set to 10^{-5} .

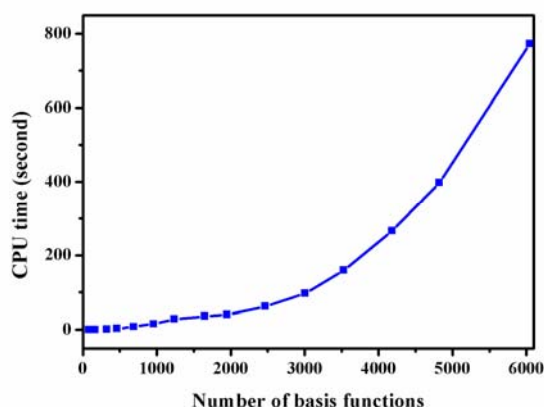


Figure 4: Time spent for the TD-DFTB computation of the lowest excited state energy S1 as the number of basis function size increases. The test system is the above test case S6 to S20, with 29 to 3011 atoms.

In summary, we have presented the implementation of the TD-DFTB method and the performance of the code in large systems. The ground state calculation on a 270A carbon nanotube with 2560 atoms shows that the sparse matrix and the OpenMP multithreading accelerate the program by 5.52 times with an 8-core CPU while the GPU acceleration can further accelerate the code to 8.73. Compared with the commonly used ARPACK package the block Davidson iterative solver requires fewer numbers of iterations. The whole scale

property of the program increases in the quadratic trend as the number of basis size increases. The program can obtain reasonable excited energy compare to TDDFT within tens of minutes for systems of thousands of atoms.

References

- [1] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, *Phys. Rev.*, 136 (1964) B864.
- [2] M. Rapacioli, R. Barthel, T. Heine, G. Seifert, Car-Parrinello treatment for an approximate density-functional theory method, *J. Chem. Phys.*, 126 (2007) 124103.
- [3] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties, *Phys. Rev. B*, 58 (1998) 7260.
- [4] M.E. Casida, K.C. Casida, D.R. Salahub, Excited-state potential energy curves from time-dependent density-functional theory: A cross section of formaldehyde's 1A1 manifold, *Int. J. Quantum Chem.*, 70 (1998) 933.
- [5] M.E. Casida, D.R. Salahub, Asymptotic correction approach to improving approximate exchange-correlation potentials: Time-dependent density-functional theory calculations of molecular excitation spectra, *J. Chem. Phys.*, 113 (2000) 8918.
- [6] T.A. Niehaus, S. Suhai, F. Della Sala, P. Lugli, M. Elstner, G. Seifert, T. Frauenheim, Tight-binding approach to time-dependent density-functional response theory, *Phys. Rev. B*, 63 (2001) 085108.
- [7] S. Tomov, R. Nath, P. Du, J. Dongarra, *MAGMA Users' Guide*, 2010.
- [8] C.W. Murray, S.C. Racine, E.R. Davidson, Improved algorithms for the lowest few eigenvalues and associated eigenvectors of large matrices, *J. Comput. Phys.*, 103 (1992) 382.
- [9] D. Ernest R, The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices, *J. Comput. Phys.*, 17 (1975) 87.
- [10] V. Chernyak, M.F. Schulz, S. Mukamel, S. Tretiak, E.V. Tsiper, Krylov-space algorithms for time-dependent Hartree-Fock and density functional computations, *J. Chem. Phys.*, 113 (2000) 36.
- [11] C.M. Isborn, N. Luehr, I.S. Ufimtsev, T.J. Martinez, Excited-State Electronic Structure with Configuration Interaction Singles and Tamm-Dancoff Time-Dependent Density Functional Theory on Graphical Processing Units, *J. Chem. Theory Comput.*, 7 (2011) 1814.
- [12] S. Hirata, M. Head-Gordon, R.J. Bartlett, Configuration interaction singles, time-dependent Hartree-Fock, and time-dependent density functional theory for the electronic excited states of extended systems, *J. Chem. Phys.*, 111 (1999) 10774.
- [13] M.R. Harpham, O.z.n. Suñer, C.-Q. Ma, P. Bañerle, T. Goodson, Thiophene Dendrimers as Entangled Photon Sensor Materials, *J. Am. Chem. Soc.*, 131 (2009) 973-979.
- [14] S. Yamaguchi, H. Kamikubo, K. Kurihara, R. Kuroki, N. Niimura, N. Shimizu, Y. Yamazaki, M. Kataoka, Low-barrier hydrogen bond in photoactive yellow protein, *Proc. Natl. Acad. Sci. U.S.A.*, 106

(2009) 440.

- [15] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *The Journal of Chemical Physics*, 79 (1983) 926-935.
- [16] I.B. Nielsen, S. Boye-Peronne, M.O.A. El Ghazaly, M.B. Kristensen, S. Brondsted Nielsen, L.H. Andersen, Absorption spectra of photoactive Yellow Protein chromophores in vacuum, *Biophys. J.*, 89 (2005) 2597.
- [17] A. Dreuw, J.L. Weisman, M. Head-Gordon, Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange, *J. Chem. Phys.*, 119 (2003) 2943.
- [18] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A*, 38 (1988) 3098.
- [19] D.C. Sorensen, in: *Implicitly restarted Arnoldi/Lanczos methods for large scale eigenvalue calculations*, Institute for Computer Applications in Science and Engineering (ICASE), 1996.