

65. Sun, T. T., Shih, C. and Green, H., *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 2813.
66. Schlegel, R., Banks Schlegel, S. and Pinkus, G. S., *Lab. Invest.*, 1980, **42**, 91.
67. Shinohara, H., Yamada, K. and Tanaka, T., *J. Oral Pathol. Med.*, 1989, **18**, 133.
68. Vaidya, M. M., Borges, A. M., Pradhan, S. A., Rajpal, R. M. and Bhusey, A. N., *J. Oral Pathol. Med.*, 1989, **18**, 282.
69. Banks Schlegel, S. P. and Harris, C. C., *Cancer Res.*, 1984, **44**, 1153.
70. Fischer, H. P., Altsmannberger, M., Weber, K. and Osborn, M., *Am. J. Pathol.*, 1987, **127**, 530.
71. Kahn, H. J., Huang, S. N., Hanna, W. M., Baumal, R. and Phillips, M. J., *Am. J. Pathol.*, 1984, **81**, 184.
72. Sun, T. T., Eichner, R., Nelson, W. G., Tseng, S. C. G., Weiss, R. A., Jarvinen, M. J. and Woodcock-Mitchell, J., *J. Invest. Dermatol.*, 1983, **81**, 109s.
73. Tseng, S. C. G., Jarvinen, M. J. and Nelson, W. G., *Cell*, 1982, **30**, 361.
74. Galand, P., Lenoy, F. and Chretien, J., *J. Endocrinol.*, 1971, **49**, 243.
75. Kronenberg, M. S. and Clark, J. H., *Endocrinology*, 1985, **117**, 1469.
76. Gupta, P. D., Reha, S. B., Sudhakar, Ch. and Reddy, A. G., in *Frontiers in Reproductive Physiology* (eds. Ghosh, D. and Sengupta, J.), Wiley Eastern, New Delhi, 1991, p. 121.
77. Fuchs, E., *J. Cell Biol.*, 1990, **111**, 2807.
78. Watt, F. M., *Curr. Opin Cell Biol.*, 1989, **1**, 1107.
79. Gupta, P. D., Swarup, G. and Reddy, A. G., *FEBS Lett.*, 1990, **273**, 135.
80. Vijayasaradhi, S. and Gupta, P. D., *J. Biosci.*, 1988, **13**, 401.

RESEARCH ARTICLES

Parallelization of molecular electronic-structure calculation

Archana D. Bhusari, Vivek V. Bhate and Sourav Pal

Physical Chemistry Division, National Chemical Laboratory, Pune 411 008, India

Electronic-structure calculations are useful for understanding atomic and molecular processes. Here we present a parallel *ab initio* algorithm developed for the preliminary electronic-structure program with special emphasis on four-index transformation for the transputers at the Centre for Development of Advanced Computing (C-DAC) in Pune. We also present results obtained on a four-node machine.

DEVELOPMENTS in computational quantum chemistry have therefore been linked with advances in computer hardware and software¹. We have been developing an accurate electronic-structure package for parallel computing. The simplest *ab initio* Hartree-Fock (HF) or self-consistent field (SCF) electronic-structure calculations can be very time-consuming for large systems, and scale as N^4 (N is 2-5 times the number of electrons). Further, calculations that incorporate electron correlation like the coupled cluster (CC) calculations, are even more expensive² and scale as N^6 - N^7 . Even the first cut in correlated theory, like second-order many-body perturbation theory (MBPT 2), scales as N^5 , primarily owing to the necessity of transforming the integrals from atomic-orbital (AO) basis to molecular-orbital (MO) basis. Bender Shavitt's algorithm^{3,4} provides us with the

best of the required algorithm. For a system of reasonable size, meaningful *ab initio* computation is not possible without high-speed integrated circuits. While the development of high-speed chips and cheaper computer hardware is opening new avenues, an attractive alternative lies in parallel computation.

Much work has already been done on parallelization of the SCF code⁵⁻⁷. However, comparatively less attention has been paid on parallelization of correlation calculations as well as on the four-index transformation preceding these^{8,9}. We focus primarily on this aspect in this paper. The parallelization of four-index transformation attempted earlier was based mainly on the algorithm of two two-index transformations. However, this involves an extra sorting at the beginning and at the end of the first two-index transformation. In addition, it requires more memory. In view of the limited memory available on our hardware, we found it useful to parallelize Bender's original algorithm. We highlight the merits of the strategy used to achieve this and discuss the specific problems that can arise. We also mention parallelization of the SCF code, including the one on evaluation of the two-electron integrals. We discuss our experience with the significant parts of this code and present preliminary results on a four-node transputer machine.

Transputers

The main advantages of parallel computing are: (i) of power, from high-speed microprocessors; (ii) in performance, from running tasks in parallel; (iii) of growth, by adding more processors as and when required; and (iv) of reliability, from quality-assured components. We are using a transputer-based parallel computing system. A transputer is a single-chip computer consisting of a complete central processing unit (CPU), local memory, and four dedicated input/output links. It satisfies the criteria of power of technology, software portability, and price. The transputer can be used as a building block to form a distributed computer network, thus setting up the communicating sequential processes (CSP) model of parallel computing¹⁰. The main features of the transputer are: (i) a single transputer of the T800 range is capable of computing speed greater than 1 MFLOP (million floating-point operations) per second; (ii) each transputer has four communication links (channels) for input/output with a data transfer rate of the order of 1 M byte s⁻¹; and (iii) a single transputer can run concurrent processes.

Parallelization of SCF code

An analysis of the SCF code and the numerical evidence of others reveal the two most time-consuming processes that can be parallelized. In the case of transputers with large communication overhead and slow data-transfer rate between different channels, it is not worth trying to parallelize the parts that take insignificant time. Therefore only the two time-consuming steps, viz. the generation of two-electron integrals and their use in the construction of Fock matrices, have been parallelized. As the systems become large the CPU time for the generation of two-electron integrals scales as N^4 . We denote in the following the atomic orbitals as p, q, r, s and the integrals as $[pq|rs]$. They are categorized in different shells of orbitals according to their principal and azimuthal quantum numbers (n, l). The program has nested loops over the shells. The parallelization was done for all orbitals r, s (for different shells) for a fixed pair of shells of p, q . In the case of the smaller angular momenta (i.e. small l values) for the p th orbital, the parallelization was done for all q, r, s for a fixed shell of p th orbital. To avoid bottlenecks, we found the 'star' topology to be quite convenient for four transputers.

Table 1 shows the results of our parallelization scheme for the two-electron integral calculation on four transputers. It is seen that the speed-up factor approaches 4.0 for the larger systems. In the later part we have coded the restricted Hartree-Fock equation for molecules in which a pseudoeigenvalue equation is solved iteratively. The essential steps are generation of

Table 1. Two-electron integral generation

Input	Basis	CPU time (sec)		
		Sequential (one transputer)	Parallel (four transputers)	Speed-up factor
CH ₄ (40/18)	Huzinaga Dunning double zeta	31.5	9.3	3.4
N ₂ (48/20)	Huzinaga Dunning double zeta	25.8	7.9	3.3
C ₂ H ₆ (78/26)	Minimal basis STO-3G	220.7	56.7	3.9
C ₂ H ₄ (64/28)	Huzinaga Dunning double zeta	131.7	34.6	3.8
H ₂ O (44/32)	O9SSP1D/[5S3P1D]* H.4S1P/[3S1P]	77.8	21.0	3.7
N ₂ (60/32)	Huzinaga Dunning double zeta + polarization	168.9	44.2	3.8
C ₆ H ₆ (108/36)	Minimal basis STO-3G	776.0	196.1	3.9
CO (60/40)	C.9SSP1D/[5S3P1D]† O.9SSP1D/[5S3P1D]*	216.9	55.9	3.9
C ₃ (90/48)	Huzinaga Dunning double zeta + polarization	738.6	187.0	3.9

The numbers in parenthesis are the numbers of primitive and contracted basis functions

*Dunning's (9SSP/5S3P) + DZP's D

†Dunning's (9SSP/5S3P) + DZP's D

the Fock matrix in AO basis, transforming it to an orthogonal basis, diagonalizing the transformed Fock matrix, and continuing the iteration. The most important steps are setting up the Fock matrix and diagonalization. The 'Jacobi diagonalization' scheme is not easy to parallelize because of its complicated nature. However, the setting up of the Fock matrix is completely parallelizable.

There are several algorithms for construction of the Fock matrix, the most notable being Duke's method¹¹ and Raffinetti's method¹². In the latter the integrals are presorted into a supermatrix, which occupies a large amount of disc space and memory. We have therefore used Duke's method to set up the Fock matrix in parallel. This is done by making use of the equations

$$F = H + G, \quad (1)$$

$$G_{ij} = \sum_{k=1}^N \sum_{l=1}^N D_{kl} \{ (ij|kl) - 1/4(ik|jl) - 1/4(il|jk) \}, \quad (2)$$

where H is a matrix of one-electron integrals, and G a two-electron integral part. Duke's method suggests the use of a list of integrals categorized into various types and processing them for assembling the two-electron integral part of the Fock matrix. The method also suggests the use of different formulae according to the type of two-electron integrals. The distinct integrals are distributed randomly on different transputers and,

depending on the type of integral, the appropriate Duke's equation is used. This was done again using the 'star' network of four transputers. The integrals were fed in batches as and when the transputers were free to process them. We have tested the code for different systems, and summarize the results in Table 2. The speed-up factor is usually around 4.0 for the larger systems.

Four-index transformation and MBPT2

To perform the first step in correlation calculations like MBPT2 it is necessary to transform the two-electron integrals over the MO basis. The four-index transformation can be written as

$$[ij|kl] = \sum_{p=1}^N \sum_{q=1}^N \sum_{r=1}^N \sum_{s=1}^N \mathbf{T}_{ip}^* \mathbf{T}_{kr}^* [pq|rs] \mathbf{T}_{jq} \mathbf{T}_{ls}, \quad (3)$$

where $[ij|kl]$ are orbital integrals, \mathbf{T} the AO to MO transformation matrix, and $[pq|rs]$ a set of integrals over AOs.

This transformation of two-electron integrals requires extensive computational resources for large basis sets. The order of operation, storage organization and input/output requirements are of major importance for developing a feasible and economically viable process. A straightforward application of equation (3) involves computation of the order of N^8 . A parallel approach to this was suggested by Shavitt and coworkers⁴ by breaking up the process into a sequence of four partial summations, as shown below, using the idea of intermediate storage.

$$[iq|rs] = \sum_{p=1}^N \mathbf{T}_{ip}^* [pq|rs], \quad i, q, r, s = 1, \dots, N, \quad (4)$$

$$[ij|rs] = \sum_{q=1}^N [iq|rs] \mathbf{T}_{jq}, \quad i, j, r, s = 1, \dots, N, \quad (5)$$

$$[ij|ks] = \sum_{r=1}^N \mathbf{T}_{kr}^* [ij|rs], \quad i, j, k, s = 1, \dots, N, \quad (6)$$

$$[ij|kl] = \sum_{s=1}^N [ij|ks] \mathbf{T}_{ls}, \quad i, j, k, l = 1, \dots, N. \quad (7)$$

Thus the overall transformation scales as N^5 rather than N^8 . Four-index transformation can also be performed as two two-index transformations (equations (8) and (9) below¹):

$$[pq|kl] = \sum_{r=1}^N \sum_{s=1}^N \mathbf{T}_{kr}^* [pq|rs] \mathbf{T}_{ls}, \quad p, q, k, l = 1, \dots, N, \quad (8)$$

$$[ij|kl] = \sum_{p=1}^N \sum_{q=1}^N \mathbf{T}_{ip}^* [pq|kl] \mathbf{T}_{jq}, \quad i, j, k, l = 1, \dots, N. \quad (9)$$

It may be noted that the partially transformed integrals $[pq|kl]$ are not, in general, in the same order as required in the second two-index transformation and hence must be resorted in sequential order. It was proposed¹ that four-index transformation may be performed as two two-index transformations and that these two-index transformations can be performed in parallel. But this method requires more storage and additional sorting. In appendices 1 and 2 respectively we lay down the basic steps required for parallelizing Bender's algorithm and two two-index transformations with 'star' topology for comparison. It is clear that, from the computational viewpoint, it is desirable to parallelize the algorithm suggested by Bender³. The only core storage required for the $[pq|rs]$ integrals is an input buffer. The integrals are read only once into memory. The computation of the intermediate matrices is nested into the outer loops such that only parts of them are needed in the computer at any one time; the indices p, q and r on these matrices in equations (4) to (7) are suppressed, since they are not used in

Table 2. Fock matrix generation

Input	Basis	CPU time (sec)			SCF energy (a.u.)
		Sequential (1T)	Parallel (4T)	Speed-up factor	
CH ₂	Double zeta	9.6	3.3	3.0	-44.9450473149278
CH ₄	Double zeta	45.7	11.7	3.9	-53.6069041746763
N ₂	Double zeta	16.3	5.4	3.0	-132.504103203245
C ₂ H ₄	Double zeta	130.8	33.6	3.9	-111.232946453976
H ₂ O	O 9SSP1D / [5S3P1D]* H. 4S1P/[3S1P]	323.4	81.7	3.9	-85.2391345785609

*Dunning's (9SSP/5S3P) + DZP's D

Table 3. Four-index transformation and MBPT 2.

Input*	CPU time (sec)		Speed-up factor	Correlation energy (a.u.)
	Sequential (1T)	Parallel (4T)		
CH ₂	131	45	2.9	-0.07882858111622
CH ₄	43.9	13.4	3.3	-0.115712638152846
N ₂	70.3	22.1	3.2	-0.256174528654251
C ₂ H ₄	365.8	98.3	3.7	-0.19771086709733
H ₂ O	698.6	186.7	3.7	-0.240942557671582

*The basis is the same as in Table 2

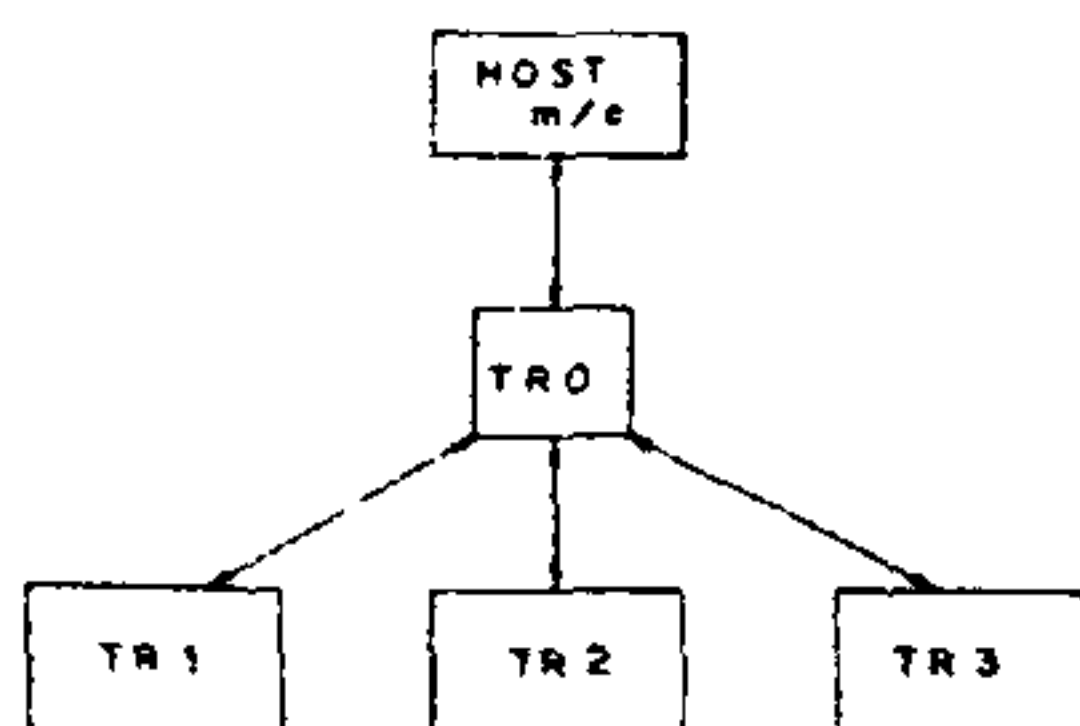


Figure 1. Four transputer network in 'star' topology.

determining the storage locations. We have set up a parallel code on a four-transputer network using the 'star' topology shown in Figure 1. It is observed that as p increases the number of integrals also increases. Hence optimum load balancing is difficult. Parallelization was done explicitly on the angular momenta.

The strategy adopted for parallelization is as follows:

For any given system: (i) individual loops of p were executed in parallel giving partial integrals on different transputers; (ii) effective load balancing was done by executing a batch p on each transputer; (iii) at the end all the partial integrals were collected and added in the host transputer.

The effectiveness of the parallelization is brought out in Table 3. The speed-up factor is typically 3.6–3.9 for a variety of systems studied by us. Once all the two-electron integrals are transformed over the MO basis the correlation energy at second-order MBPT level can be calculated easily. Since all the integrals are available in the transputer memory we directly get the MBPT2 correlation energy. This part is quite fast and needs no parallelization. The last column of Table 3 gives the MBPT2 results.

Future prospects

We intend to port our code on the 64-node machine 'PARAM' using 'PRESHAK' software developed by C-DAC. This software is independent of topology and can be used to implement the message-passing communication structure for any parallel algorithm on a network

of transputers. Hence this code works on any number of transputers with added efficiency.

Appendix 1

Steps involved in parallelization of Bender's algorithm

1. Assume that we have a list of two-electron integrals $[pq|rs]$ stored in canonical order.
2. Distribute the integrals on the different transputers for ranges of p indices, i.e. distribute the integrals for different ranges of p on different transputers for all q, r, s indices.
3. Calculate $[ij|kl]$ for a range of p on each transputer using Bender-Shavitt's method. Now we have partial $[ij|kl]$ integrals on each transputer.
4. Collect all partial $[ij|kl]$ integrals to a master transputer and add them giving the required $[ij|kl]$ integrals.

Appendix 2

Steps involved in parallelization of two two-index transformation algorithms

1. Assume that we have a list of two-electron integrals $[pq|rs]$ stored in random order.
2. Sort and distribute the integrals over various transputers with a range of p and q for all the values of r and s on a specific transputer.
3. Perform the two-index transformation using equation (8) giving all $[pq|kl]$ integrals (half-transformed integrals).
4. Sort and distribute the integrals over various transputers with a range of k and l for all the values of p and q on a specific transputer.
5. Perform the two-index transformation using equation (9) giving all $[ij|kl]$ integrals.

1. Stephen Wilson, (ed), *Methods in Computational Chemistry*. Plenum Press, New York and London, 1987, vol. 1.
2. Bartlett, R. J., *Annu. Rev. Phys. Chem.*, 1981, 32, 359, and references therein, Mukherjee, D. and Pal, S., *Adv. Quant. Chem.*, 1989, 20, 291.
3. Bender, C. F., *J. Comp. Phys.*, 1972, 9, 547.
4. Henry, F., Schaefer, (eds.), *Methods of Electronic Structure Theory*, 3rd edn, 1977, Plenum Press, New York and London.
5. Wedig, U., Burkhardt, A. and Von Schnering, H. G., *Z. Phys.*, 1989, D13, 377.
6. Ernenwein, R., Rohmer, M. M. and Benard, M., *Comp. Phys. Commun.*, 1990, 58, 305.
7. Rohmer, M. M., Denmuyneck, J., Benard, M., Wiest, R., Batchmann, C., Henriot, C. and Ernenwein, R., *Comp. Phys. Commun.*, 1990, 60, 127.

8. Watts, J. D. and Dupuis, M., *J. Comp. Chem.*, 1988, 9, 158.
9. Covick, L. A. and Sando, K. M., *J. Comp. Chem.*, 1990, 11, 151.
10. 3L Parallel FORTRAN user guide, INMOS, 1989.
11. Duke, A. J., *Chem. Phys. Lett.*, 1972, 13, 76.
12. Raffanetti, R. C., *Chem. Phys. Lett.*, 1973, 20, 335.

ACKNOWLEDGEMENTS. We thank C-DAC, Pune, for provid-

ing their parallel computation facility and for funding this venture. We used the 'MOLECULE' code (a Stockholm Institute of Physics program system) for the basic routines required in two-electron integral calculation.

Received 18 February 1991, revised accepted 18 June 1991

Fast solid-state reactions between bifluorides and vanadates at ambient temperature

B. N. Wani, S. J. Patwe and U. R. K. Rao

Applied Chemistry Division, Bhabha Atomic Research Centre, Bombay 400 085, India

We have examined solid-state reactions occurring at room temperature at a fast rate across an interface between NH_4 /alkali metal bifluorides and NH_4 /alkali metal vanadates resulting in the formation of oxyfluorocompounds in various reaction configurations from the point of view of delineating the mechanism of the reaction, nature of the products formed, and kinetics of the reaction. The reaction, initiated by moisture, always progresses from the bifluoride into the vanadate by transport of the bifluoride across an ever-growing layer of the product, and the product invariably contains the new compound with a stoichiometry of 1:2 in vanadate to bifluoride. We have also determined relative reaction rates.

MANY solid-state reactions, such as those used in the formation of ferrites, alloys, etc., have measurable reaction rates at temperatures much higher than the ambient temperature. This is because such reactions are diffusion-controlled and the diffusion rates are negligibly small at ordinary temperatures. Reactions between solids are complex and reactivity depends on several factors. For instance, the rate of reaction between two single crystals¹ of MgO and TiO_2 at 1550°C depends on the faces of the two crystals that are in physical contact with each other. KHF_2 is known to exist in two polymorphic phases, one tetragonal at $T < 198^\circ\text{C}$ and the other cubic at $T > 198^\circ\text{C}$. Recently, in some other context, we found that the tetragonal form of KHF_2 does not react with Nb_2O_5 even after three hours of heating at 180°C , whereas the cubic form reacts readily at 210°C . Hellstern and Schultz² found that Fe and Zr react to form alloys at room temperature by ball milling, but the products were amorphous. Goyal *et al.*³ reported that orthorhombic-V modification of Na_2SO_4 trans-

forms into another, orthorhombic-III, modification of Na_2SO_4 by grinding at room temperature. During the last decade several examples⁴⁻⁹ have been found by our group and by others¹⁰, in which two (or three) solids, when mixed at room temperature and on application of mechanical force such as grinding with a pestle in an agate mortar, form a new compound in a single phase. Such reactions take place in the solid state readily at ambient temperature without necessity of heating and with such ease as though they are taking place in solution. The product is crystalline. The structure of the product is totally different from that of either reactant. Recently we showed¹¹ that two crystalline solids, one a bifluoride and the other a vanadate, even without grinding, interact at room temperature by merely keeping in contact with each other at an interface and form a third crystalline solid at the interface. The structure of the product is totally different from that of either of the reactants. A preliminary report of this was made in ref. 11. It is difficult to rationalize this observation. The formation of the product involves change of symmetry, translation of atoms, breaking of bonds and making of new linkages, all taking place with almost no additional energy externally supplied above the thermal energy $k_B T$. Here we describe experiments aimed at identification of the reaction products and determination of the kinetics of these reactions. Experiments have also shown that the synthesis proceeds through catalytic initiation by traces of moisture.

Methods

The chemicals used were NH_4HF_2 (Riedel), NaHF_2 (IR grade), KHF_2 (BDH, AR grade), NH_4VO_3 (AR