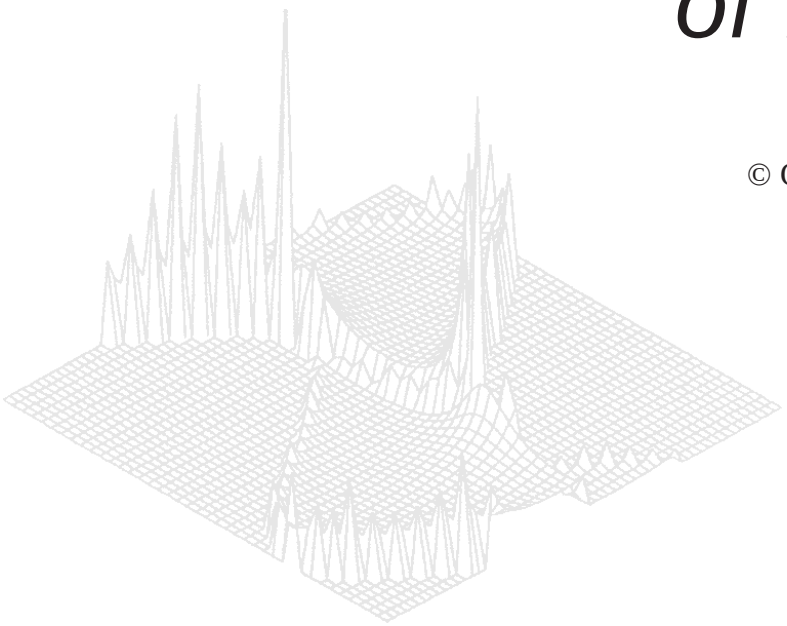
CSIRO PUBLISHING

Australian Journal of Physics

Volume 52, 1999
© CSIRO Australia 1999



A journal for the publication of
original research in all branches of physics

www.publish.csiro.au/journals/ajp

All enquiries and manuscripts should be directed to
Australian Journal of Physics
CSIRO PUBLISHING
PO Box 1139 (150 Oxford St)
Collingwood Telephone: 61 3 9662 7626
Vic. 3066 Facsimile: 61 3 9662 7611
Australia Email: peter.robertson@publish.csiro.au

Published by **CSIRO** PUBLISHING
for CSIRO Australia and
the Australian Academy of Science

Aust. J. Phys., 1999, **52**, 919–27

Equation-of-Motion $O(N)$ Electronic Structure Studies of Very Large Systems ($N \sim 10^7$)*

Marek T. Michalewicz[A] *and Per Nyberg*[B]

[A]Division of Molecular Science, CSIRO,
343 Royal Parade, Parkville, Vic. 3052, Australia.
email: Marek.Michalewicz@molsci.csiro.au

[B]NEC Australia Pty Ltd, 635 Ferntree Gully Road,
Glen Waverley, Vic. 3150, Australia.

Abstract

Extremely fast parallel implementation of the equation-of-motion method for electronic structure computations is presented. The method can be applied to non-periodic, disordered nanocrystalline samples, transition metal oxides and other systems. It scales linearly, $O(N)$, runs with a speed of up to 43 GFLOPS on a NEC SX-4 vector-parallel supercomputer with 32 processors and computes electronic densities of states (DOS) for multi-million atom samples in mere minutes. The largest test computation performed was for the electronic DOS for a TiO_2 sample consisting of 7,623,000 atoms. Mathematically, this is equivalent to obtaining the spectrum of an $n \times n$ Hermitian operator (Hamiltonian) where $n = 38,115,000$. We briefly discuss the practical implications of being able to perform electronic structure computations of this great speed and scale.

1. Introduction

Detailed knowledge of the electronic structure of materials is crucial to a deeper understanding of a vast range of their microscopic and macroscopic physical properties. Over the years, the refinement of theoretical methods, improvements in computational techniques and algorithms and the rapid increase in computational power of new machines has led to an optimistic conviction in the theoretical condensed matter physics community that a very detailed understanding of 'real materials' derived theoretically from the fundamental principles of quantum mechanics is within our reach.† By 'real materials' we mean solids which can be crystalline (single or poly) or amorphous, they may possess any type of defects (point or extended), they may be bounded by surfaces (usually are!), can be doped, or have impurities, dirt, voids, imperfections, etc. Understanding at the quantum level, in the crudest sense, means understanding of the properties derived through a solution of the Schrödinger equation, or related formalism, such as the density functional theory.

* Refereed paper based on a contribution to the Eighth Gordon Godfrey Workshop on Condensed Matter Physics held at the University of New South Wales, Sydney, in November 1998.

† See the Ψ_k network: European Union program: 'Ab initio (from electronic structure) calculation of complex processes in materials': http://www.dl.ac.uk/TCSC/HCM/PSIK/main.html

© CSIRO 1999
10.1071/PH99002 0004-9506/99/050919$05.00

From a practical point of view it is desirable that any method for studying 'real materials' should satisfy the following criteria:

(i) it should be able to account for very large systems, consisting of $\sim 10^2 - 10^6$ atoms or more,
(ii) it should exhibit linear computational complexity (scaling) $O(N)$, i.e. computational time depends linearly on the number of atoms in a system,
(iii) the computer implementation should have a good parallel performance, and
(iv) computations should be very fast.

These criteria are complementary since (i) can hardly be achieved without (ii)–(iv).

This paper is concerned exclusively with computational and high performance aspects of the equation-of-motion method for studying the electronic structure of disordered systems. We report on the performance of our code on the NEC SX-4 vector parallel supercomputers. Our computations satisfy all four criteria above and, we believe, exceed any previous benchmark results and 'high water marks' for code performance in this class of computational problems. We ran a test computation for the electronic DOS of a TiO_2 sample consisting of 7,623,000 atoms. The equation-of-motion method exhibits linear scaling, $O(N)$. The peak computational speed achieved with our electronic structure code on 32 processors was nearly 43 billion floating point operations per second (43 GFLOPS).

The work reported here is the latest step in a series of improvements of the code and implementations on various computer architectures; from vector implementation on the Cray 2 by Michalewicz *et al.* (1992), through massively parallel implementation on the SIMD MasPar-1 and MasPar-2 by Michalewicz (1994) and Michalewicz and Priebatsch (1995), to vector-parallel implementation on the Cray C90 by Michalewicz and Brown (1997). This evolution is represented in Table 1. Over the four generations of implementation on different machines, the size of the system studied was increased by a factor of nearly 2000, and the speed of computation increased by a factor of 428.

Table 1. System size increase and performance improvement of the electronic structure code on different machines 1990–98

Machine and number of processors	Year	Maximum number of atoms	Size scale ratio	Peak performance (GFLOPS)	Peak performance ratio
Cray 2 (1CPU)	1990	3,840	1	$\leq 0{\cdot}1$	1
MasPar-1, MasPar-2 (16,000 PE)	1992	491,520	128	$\sim 0{\cdot}5$–$0{\cdot}7$	~ 7
Cray C90 (16 CPUs)	1995	1,105,920	288	$8{\cdot}5$	85
NEC SX-4 (32 CPUs)	1998	7,623,000	1985	$42{\cdot}8$	428

Our 'brute force' approach of using the most powerful computers available and trying to achieve the highest possible performance serves two distinct purposes. Firstly, for each new computer architecture we can gain a good understanding of hardware, software and compilers—we can 'test drive' the machine and then apply this knowledge to other computational problems; secondly, for each new machine we extend the limits of possible computations. We will discuss possible extensions and applications of this work in Section 4 of our paper.

2. Electronic Structure of Disordered Systems: Equation-of-Motion Method

One of the physical quantities of interest which we compute is the electronic density of states (DOS). Depending on the initial conditions and the mode of computation we can obtain total DOS for a disordered nanocrystallite, surface DOS for an arbitrary exposed surface of a nanocrystallite, or local DOS for a selected atom in a sample or per orbital resolved DOS (both local and total). Symbolically the total DOS is expressed as

$$N(\omega) = \sum_{n} \sum_{i,\mu} |\langle n|i,\mu\rangle|^2 \delta(\omega - \epsilon_n), \tag{1}$$

$$= -\frac{1}{\pi} \mathrm{Im}\{TrG^{+}(\omega)\}, \tag{2}$$

where $\langle n|$ is the eigenvector and ϵ_n the eigenvalue of the Hamiltonian for the disordered lattice, and $|i,\mu\rangle$ is the tight-binding state localised on site i and of the orbital type μ, and i and μ are the site and the orbital indexes respectively. In equation (2), $G^{+}(\omega)$ is Green's function (in the energy domain). We use the tight binding Hamiltonian, although the equation-of-motion method is not limited to this model only. Our model Hamiltonian is non-trivial; it takes into account the s- and p-states on the oxygen atoms, five d-states on the Ti atoms, where each oxygen atom has 14 neighbours and each titanium atom has 16 neighbours.

In the equation-of-motion method, as implemented in our program, the trace summation is done through the use of random phases, and $G^{+}(\omega)$ is evaluated as a Fourier transform of the time dependent Green function. The required Green function time series is obtained by integrating the equation-of-motion obeyed by the Green function; hence the name of the method. More details on the method and numerical implementation can be found in our previous articles and references cited therein (Halley *et al.* 1990; Michalewicz *et al.* 1992; Michalewicz and Priebatsch 1995).

Since the equation-of-motion method is solved in direct space, it can be used in situations where the Bloch Theorem is not satisfied. The Hamiltonian we construct (use) can possess both diagonal and off-diagonal disorder, reflecting substitutional disorder (e.g. random alloys) or positional disorder (e.g. glassy or amorphous phases). The DOS is only one of the quantities which can be computed using the equation-of-motion method. Further examples will be discussed in the last section.

3. Computational Performance, Benchmarks

(3a) Machines and System Sizes

The computer code was run on two machines; one located at the NEC research and training centre in Fuchu, Japan was an SX-4/16A with 32 GBytes SDRAM main memory and 16 CPUs. All reported computations where we utilised 1, 2, 4, 8 and 16 processors were run on that machine. Computations requiring 24 or 32 processors were run on the SX-4/32 with 8 GBytes SSRAM main memory machine with 32 processors owned and operated by the Joint Bureau of Meteorology/CSIRO High Performance Computing and Communications Centre

and located in Melbourne, Australia. Both machines have 8 ns clock, 8 vector pipelines per processor and their theoretical peak speed is 2 GFLOPS per CPU. Table 2 lists all test runs for different system sizes and number of CPUs used.

Table 2. TiO_2 model system sizes and number of processors used
The meaning of symbols in the table is: x—run on SX4/16A; •—run on SX4/32

System size $n_i \times n_j \times n_k$	Number of atoms	Number of processors used 1	2	4	8	16	24	32
$40 \times 40 \times 40$	384000	x	x	x	x	x	•	•
$40 \times 50 \times 40$	480000	x	x	x	x	x	•	•
$40 \times 50 \times 50$	600000	x	x	x	x	x	•	•
$50 \times 50 \times 50$	750000	x	x	x	x	x	•	•
$70 \times 70 \times 70$	2058000	x	x	x	x	x		
$80 \times 80 \times 80$	3072000					x		
$90 \times 90 \times 90$	4374000					x		
$100 \times 100 \times 100$	6000000					x		
$105 \times 105 \times 105$	6945750					x		
$105 \times 110 \times 110$	7623000					x		

The TiO_2 model system sizes were defined by $n_i \times n_j \times n_k$, where the size of the sample is $n_i a \times n_j b \times n_k c$ and a, b, c are the rutile lattice parameters ($a = b = 4 \cdot 59$ Å, $c = 2 \cdot 96$ Å) (Howardine and Dickson 1991). The largest sample studied had dimensions of 48 nm × 50 nm × 32 nm. With this scale of computation we enter into the domain of nanosystems observable under the electron microscope and having many interesting and important properties (Cao *et al.* 1995; Shklover *et al.* 1997)—we will return to this point briefly in the last section.

The results of our electronic DOS for the largest sample consisting of 7,623,000 atoms ($105 \times 110 \times 110$ unit cells) for TiO_2 rutile are presented in Fig. 1. They are compared with experimental photoemission spectra (Zimmermann *et al.* 1998) and inverse photoemission spectroscopy data (Taverner *et al.* 1993). Our computed DOS agrees well with a number of other experimental and computed data from photoelectron spectroscopy (Zimmermann *et al.* 1999) and inverse photoelectron spectroscopy (See *et al.* 1993; Tezuka *et al.* 1994).

(3b) Timing and Performance

The simplest and perhaps most valuable measure of computer code performance is the elapsed time. In the case of parallel machine implementation, elapsed time is defined as the difference between the time when the process begins to run and the time when it ends. It includes system time, sequential sections time (unparallelisable sections) and the time of running parallel sections.

The elapsed time for runs on 1, 2, 4 and 8 processors for the systems from 384,000 to 2,058,000 atoms is plotted in Fig. 2. Two important observations are: the scaling is linear, i.e. the time of computation grows linearly with the number of atoms, $O(N)$, and the total elapsed time varies like $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$, the reciprocal of the number of processors.

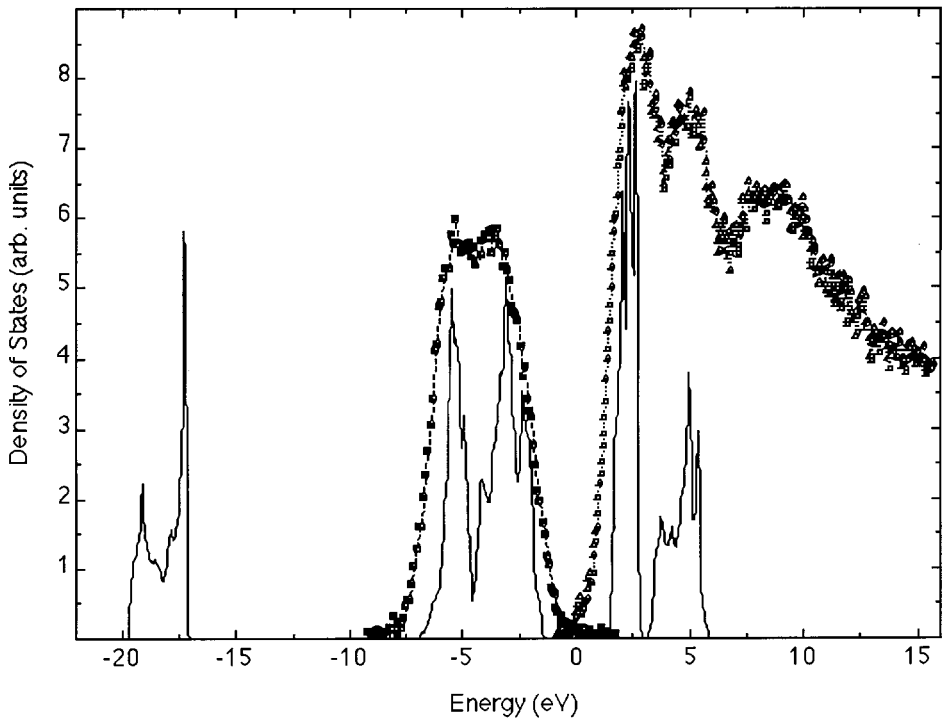


Fig. 1. Electronic density of states (DOS) for the TiO_2 sample consisting of 7,623,000 atoms (solid line) compared with experimental data obtained from photoemission spectroscopy [solid squares, Zimmermann *et al.* (1998) with permission from Elsevier Science] and from inverse photoemission spectroscopy [triangles, Taverner *et al.* (1993) with permission from Elsevier Science].

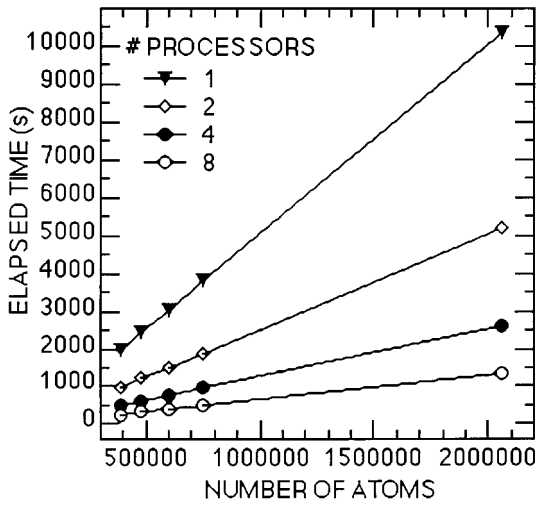


Fig. 2. Elapsed time versus system sizes of up to 2,058,000 atoms for runs on 1, 2, 4 and 8 CPUs.

The other performance measure is the number of floating point operations per second (FLOPS). The code ran at 1·5 GFLOPS on a single processor when we used the FORTRAN77 compiler. This represents 75% of the theoretical peak speed performance of a single processor. However, all runs were performed using the FORTRAN90 compiler, since this allowed us to study larger systems. The performance of the FORTRAN90 compiled program on a single CPU is about 6% slower.

Fig. 3 depicts the timing results for runs on 24 and 32 processors. The model system sizes ranged from 384,000 to 750,000. The downward bending of the timing lines represents 'superlinear' scaling—for a system size of more than 600,000 on this particular architecture, the performance is better than linear and can be attributed to a marginally higher vector operation ratio and average vector length. The elapsed time for a sample of 750,000 atoms on 32 CPUs was less than 2 minutes and the parallel performance was 42·8 GFLOPS. This represents 66·8% of the theoretical peak speed of this machine—an impressive feat, since achieving anything more than 50% on real applications is considered a very efficient utilisation of an architecture. This is one of the fastest real applications on this type of machine reported to date.

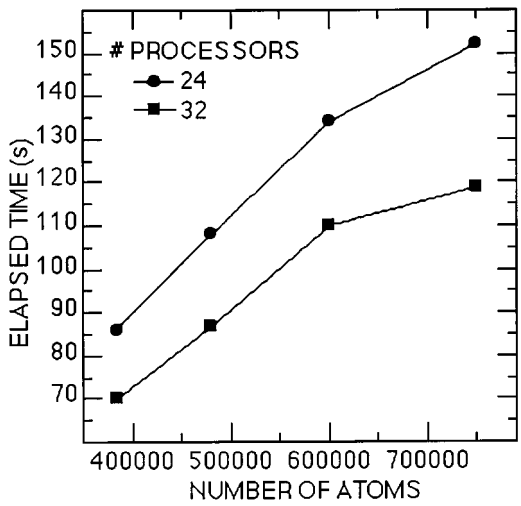


Fig. 3. Elapsed time versus system sizes of up to 750,000 atoms for runs on 24 and 32 CPUs.

Fig. 4 depicts our timing results for runs on the 16 CPU SX4/16A machine. The most important point to note is the very large system sizes we were able to study. The largest system studied had 7,623,000 atoms and it took only 41 minutes to compute the electronic DOS. The scaling is linear, $O(N)$, with one computation for 6,945,750 atoms being slightly above the straight line, with a vector operation ratio that was marginally smaller than for the other points (99·736% versus 99·744%). The peak parallel performance for the largest system size was 21 GFLOPS. This test computation, to the authors' knowledge, exceeds by some 2–3 orders of magnitude the size of the studied sample for all previous computations of a similar kind (quantum level, semi-empirical Hamiltonian).

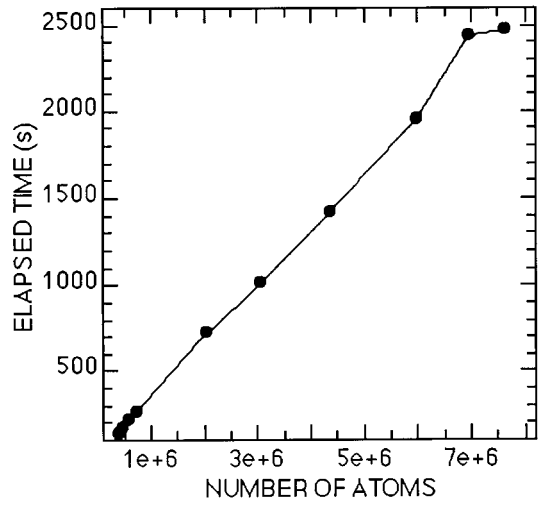


Fig. 4. Elapsed time versus system sizes of up to 7,623,000 atoms for runs on 16 CPUs.

(3c) Parallel Efficiency

Tuning and subsequent measurement of a code performance on a parallel architecture is a far from trivial task (Van de Velde 1994). The speed-up of an n-node (n-CPU) computation with execution time T_n is given by $S_n = T_1/T_n$, where T_1 is the best single-CPU time obtained for the same problem on the same computer.

In Fig. 5 we plot speed-up of our code as a function of the number of CPUs on which it was run. The various lines represent runs for a fixed system size,

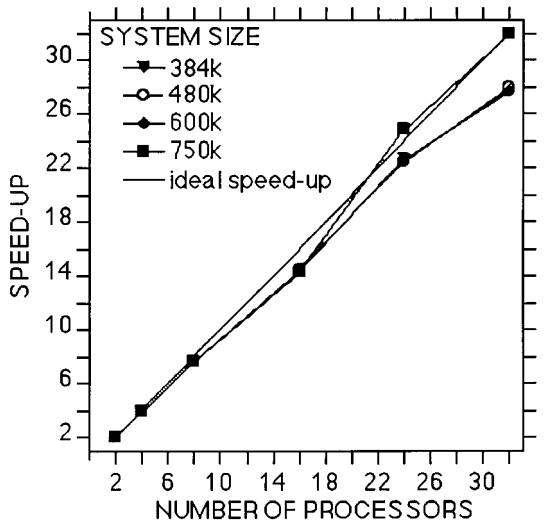


Fig. 5. Speed-up of the equation-of-motion code on the SX-4 supercomputer.

from 384,000 atoms up to 750,000 atoms. The $y = x$ line represents the ideal speed-up. There is one point at 750,000 for 24 CPUs where we observe superlinear speed-up.

It should be emphasised that for most real applications a degradation of performance is observed when some usually moderate number of processors is exceeded. Here the situation is very close to optimal, since we achieve nearly ideal speed-up for the sample consisting of 750,000 atoms. In other words we 'compress' the entire hour of computation (on a single CPU) into less than two minutes of real time computations on 32 CPUs!

4. Future Directions and Opportunities

The equation-of-motion method has been used previously for computations of the electronic conductivity via the Kubo formula, localisation, spectral functions, interband linear and nonlinear optical properties, the Hall coefficient, the electronic structure of hydrogenated a-Si and the effective mass of electron and hole in amorphous silicon; for detailed references see Michalewicz and Priebatsch (1995). A host of other properties in complex systems can be explored computationally with programs based on the equation-of-motion method and the tight-binding Hamiltonian. These include total energy computations (Cohen *et al.* 1994) or the electronic structure of elemental materials, transition and noble metals (Papaconstantopoulos 1986; Mehl and Papaconstantopoulos 1996).

It is likely that modelling of entire monolithic electronic devices (Michalewicz 1995) will be possible with the use of extremely fast computer codes. In materials research we may envisage a very fast computational materials 'discovery' process (a 'virtual' counterpart to the continous composition-spread method of van Dover *et al.* 1998). In solid state physics and materials sciences the method described here and its extensions could be invaluable for studying systems for which very large supercells make ab-initio methods (still) impractical. Such systems include (modulated) superlattices, surfaces, atomic steps, islands, interfaces, extended surface defects, random alloys, sol-gel systems and sintered nanocrystalline materials (Cao *et al.* 1995; Shklover *et al.* 1997), or carbon nanostructures (Menon *et al.* 1996).

Some of the immediate extensions and enhancements worth considering are: positional relaxation, tight-binding molecular dynamics (TBMD) (Menon and Subbwaswamy 1997), non-orthogonal tight-binding schemes (Menon and Subbwaswamy 1994) and a combination of methods described here with ab-initio schemes (Turek *et al.* 1997).

Acknowledgments

MTM wishes to thank R. Bell and R. Thurling from the Joint CSIRO–Bureau of Meteorology High Performance Computing and Communication Centre for their support and assistance during this work. We wish to thank NEC for generously granting us time on their NEC SX4/16A in Fuchu, Japan.

References

Cao, F., Oskam, G., Searson, P. C., Stipkala, J. M., Heimer, T. A., Farzad, F., and Meyer, G. J. (1995). *J. Phys. Chem.* **99**, 11974.

Cohen, R. E., Mehl, M. J., and Papaconstantopoulos, D. A. (1994). *Phys. Rev.* B **50**, 14694.

Halley, J. W., Michalewicz, M. T., and Tit, N. (1990). *Phys. Rev.* B **41**, 10165
Howardine, C. J., and Dickson, F. (1991). *Acta Crystallogr.* B **47**, 462.
Mehl, M. J., and Papaconstantopoulos, D. A. (1996). *Phys. Rev.* B **54**, 4519.
Menon, M., Richter, E., and Subbwaswamy, K. R. (1996). *J. Chem. Phys.* **104**, 5875.
Menon, M., and Subbwaswamy, K. R. (1994). *Phys. Rev.* B **50**, 11577.
Menon, M., and Subbwaswamy, K. R. (1997). *Phys. Rev.* B **55** 9231.
Michalewicz, M. T. (1994). *Comput. Phys. Commun.* **79**, 13.
Michalewicz, M. T. (1995). 'Rational nanoelectronic device design—A definition of another Grand Challenge problem'. High Performance Computing Asia '95 Electronic Proceedings, National Center For High-Performance Computing, Taiwan (hpc@nchc.gov.tw)
Michalewicz, M. T., and Brown, R. W. (1997) *In* 'Computational Chemistry and Chemical Engineering' (Eds G. Cisneros *et al.*), p. 301 (World Scientific: Singapore).
Michalewicz, M. T., and Priebatsch, M. (1995). *Parallel Comp.* **21**, 853.
Michalewicz, M. T., Shore, H., Tit, N., and Halley, J. W. (1992). *Comput. Phys. Commun.* **71**, 222.
Papaconstantopoulos, D. A. (1986). 'Handbook of the Band Structure of Elemental Solids' (Plenum: New York).
See, A. K., Thayer, M., and Bartynski, R. A. (1993). *Phys. Rev.* B **47**, 13722.
Shklover, V., Nazeeruddin, M. K., Zakeeruddin, S. M., Barbe, C., Kay, A., Haibach, T., Steurer, W., Hermann, R., Nissen, H. U., and Gratzel, M. (1997). *Chem. Mater.* **9**, 430.
Taverner, A. E., Hollamby, P. C., Aldridge, P. S., and Edgell, R. G. (1993). *Surf. Sci.* **287/88**, 653.
Tezuka, Y., Shin, S., Ishii, T., Ejima, T., Suzuki, S., and Sato, S. (1994). *J. Phys. Soc. Jpn* **63**, 347.
Turek, I., Drchal, V., Kudrnovsky, J., Sob, M., and Weinberger, P. (1997). 'Electronic Structure of Disordered Alloys, Surfaces and Interfaces' (Kluwer Academic: Boston).
Van de Velde, E. F. (1994). 'Concurrent Scientific Computing' (Springer: New York).
van Dover, R. B., Schneemeyer, L. F., and Fleming, R. M. (1998). *Nature* **392**, 162.
Zimmermann, R., Steiner, P., Claessen, F., Reiner, F., and Hüfner, S. (1998). *J. Elect. Spectrosc. Rel. Phenom.* **96**, 179.
Zimmermann, R., Steiner, P., Claessen, F., Reiner, F., Hüfner, S., Blaha, P., and Dufek, P. (1999). *J. Phys.* C **11**, 1657.

Manuscript received 5 January, accepted 28 June 1999