

Studies of the Electronic Structure of Very Large Non-Periodic Atomic Systems on PVP and MPP Cray Architectures

M.T. Michalewicz, CSIRO Supercomputing Support Group, Division of Information Technology, Commonwealth Scientific and Industrial Research Organization, Carlton, Victoria, Australia, and *Roger W. Brown*, Cray Research (Australia), Deakin, ACT, Australia

A program for electronic structure computations, which scales linearly ($O(N)$) with the number of atoms in a system, suitable for disordered systems, transition metal oxides, interfaces, superlattices, and in general for nanometer scale engineered structures consisting of millions of atoms, has been implemented on a vector-parallel Cray C90. It runs at a speed of 8.5 GFlop/s on 16 CPUs Cray C90. We believe that the system studied using this program (1,105,920 atoms in rutile TiO_2) is the largest quantum mechanical electronic structure calculation reported so far.

We report on our experiences of porting and optimizing this code on C90 and T3D architectures.

I. INTRODUCTION

Elucidation of the electronic structure of solids was and continues to be one of the most central problems of condensed matter physics [1]. Electronic structure of ideal, periodic solids: metals, insulators and semiconductors has been established over the last half-century through many different approaches, most of them based on applications of the Bloch Theorem. Presently considerable research efforts are devoted to development of fast (linear $O(N)$, where N is the number of atoms) algorithms, large- N methods, real space methods and their parallel implementations for the studies of electronic structure of complex systems.

In this paper we present a method which is perfectly suitable for very large atomic systems, is a real-space method, has been implemented on massively parallel SIMD, PVP and MPP architectures and scales as $O(N)$ on the Cray PVP architecture.

Why are we interested in the large- N methods? The reasons are many. Most importantly real systems in nature are almost always far from perfect. That means all approximations based on ideal symmetry, perfect stoichiometry, clean composition without impurities etc. should be treated with caution. For example, for a system with a high concentration of defects, the traditional approach based on the perturbation method is not only very difficult, but may not be applicable. Indeed, the very fact that systems are far from perfect make them

more interesting and endow them with peculiar characteristics.

Further, fabrication of artificial structures in 0, 1, 2 or quasi-2 dimensions in nano-meter scale has been achieved. Such structures have unique electronic and optical properties which can not be easily approximated by results obtained for ideal, infinite 3-D models. This calls for methods of computation of electronic structure and all related properties for *complete atomic scale models* of structures [2].

The method presented in this paper, indeed the original program on which this work is based, was recently used for elaborate studies which combined electronic structure with the dynamics of ions in dense plasmas [3].

II. ELECTRONIC STRUCTURE: EQUATION OF MOTION METHOD

The group of materials we are interested in, such as insulators or semiconductors, are reasonably well described by a model Hamiltonian which treats all outer shell electrons as rather well localized "atomic orbitals" - it is called the tight binding model. This model can in principle be also used for metals, but the delocalized nature of conduction electrons requires a very large (still impractical) basis set for each atom.

The tight binding Hamiltonian is written in a form

$$H = \sum_{i,\mu} \epsilon_{i,\mu} c_{i,\mu}^\dagger c_{i,\mu} + \sum_{i,\mu;j,\nu} (t_{i,\mu;j,\nu} c_{i,\mu}^\dagger c_{j,\nu} + h.c.) \quad (2.1)$$

where the i 's and j 's are site indices, the Greek letters μ, ν are indices labeling orbitals and the sum on (ij) is over neighbors on the lattice. The first term on the right hand side of the equation describes the "on site" energy at site i and atomic orbital μ , and the second term represents "hopping energy" of the electron "jumping" from j, ν to i, μ site/orbital state. The operators $c_{i,\mu}^\dagger$ and $c_{i,\mu}$ are called creation and annihilation operators respectively.

In this work we report application of the equation of motion method (EoM) [4,5] to computations of local, orbital projected, total, and surface electronic densities

of states on PVP and MPP Cray architectures. The equation of motion method is intuitively very simple. All computations are performed in direct space. The method effectively computes the spectrum of a Hamiltonian (Schroedinger equation) without resorting to a direct diagonalization.

Let us summarize the formal results which form the basis of the equation of motion method.

The density of states $N_\mu(\omega)$ associated with orbitals of type μ is given by:

$$N_\mu(\omega) = \sum_n \sum_i | \langle n | i, \mu \rangle |^2 \delta(\omega - \epsilon_n) \quad (2.2)$$

where $\langle n |$ are the eigenstates of the tight-binding problem in the (disordered) lattice and the $|i, \mu\rangle$ is the tight-binding state localized on site i and of the orbital type μ .

It can be shown that the above expression can be represented as

$$N_\mu(\omega) = -\frac{1}{\pi} \text{Im} \left[\int \sum_i e^{-i\phi_{i,\mu}} F_{i,\mu}(t) e^{+i\omega t} dt \right] \quad (2.3)$$

where $F_{i,\mu}(t) = \sum_{j,\nu} a_{j,\nu} G_{i,\mu;j,\nu}(t)$ is the amplitude of Green's function.

We define the Green's function in a standard way: $G_{i,\mu;j,\nu}(t) = -i\Theta(t) \langle \{c_{i,\mu}(t), c_{j,\nu}^\dagger(0)\} \rangle$.

The time evolution of the quantum system is governed by the equation of motion

$$i\hbar \partial F_{i,\mu} / \partial t = \sum_{j,\nu} H_{i,\mu;j,\nu} F_{j,\nu} \quad (2.4)$$

with the initial condition $F_{i,\mu}(0) = -ia_{i,\mu}$

Insulators or wide gap semiconductors can be described by the tight binding Hamiltonian, Eq. (2.1).

In order to compute the total electronic density of states (for the orbital of type μ) we set the coefficients in expression for $F_{i,\mu}$ to be: $a_{i,\mu} = e^{i\phi_{i,\mu}}$; with random $\phi_{i,\mu}$; $0 < \phi_{i,\mu} < 2\pi$.

The equation of motion is integrated using the formal solution

$$\mathbf{F}(t_{l+1}) = \exp(-i\mathbf{H}\delta t) \mathbf{F}(t_l) \quad (2.5)$$

with the Chebyshev polynomial fit of the exponential function expansion. More details can be found in references [6-8].

III. IMPLEMENTATION ON CRAY C90 AND T3D COMPUTERS

A. C90: PVP implementation

The massively parallel implementation of the equation-of-motion program designed for array processor SIMD architecture [7,8] was modified specifically for

MIMD (as well as SIMD) machines. The connection of physics and parallel architecture is contained in Eq. (2.4) and especially Eq. (2.1). The values of hopping integrals $t_{i,\mu;j,\nu}$ represent "exchange of information" between the neighbouring atoms. Depending on the location of an atom, it may correspond to inter- or intra-processor communication.

Most of our PVP work was done on Cray C916, serial number SN-4025. Over time, its configuration has changed. At present it is called Rain (it was Hot) and is described thus:

Environment — Unicos/Production

Mainframe — SN 4025

Model — CRAY C916

CPUs — 16

Memory Size — 512 mW

Memory banks — 512

Chip Type — MEM4M

Clock — 4.167 ns

SSD —

Model — SSDE

Size — 1024

IOS — IOSE SN 709

Depending on the size of the sample the program ran between 550 to 630 MFlops/s on a single processor C90. This means we achieved about 57% to 66% efficiency of a single vector processor of CRAY C90. The sample sizes are presented in the table below. Each unit cell contains six atoms. The dimensions of the largest sample were $18 \times 22 \times 28nm$.

TABLE I. TiO_2 model system sizes

System size $a \times b \times c$	Number of atoms
$6 \times 6 \times 6$	1296
$10 \times 10 \times 12$	7200
$12 \times 12 \times 12$	10368
$16 \times 16 \times 16$	24574
$18 \times 18 \times 18$	34992
$20 \times 20 \times 20$	48000
$24 \times 24 \times 24$	82944
$81 \times 81 \times 10$	393660
$45 \times 45 \times 42$	510300
$40 \times 48 \times 48$	552960
$40 \times 48 \times 96$	1105920

The results of benchmarks are presented in Figure 1.

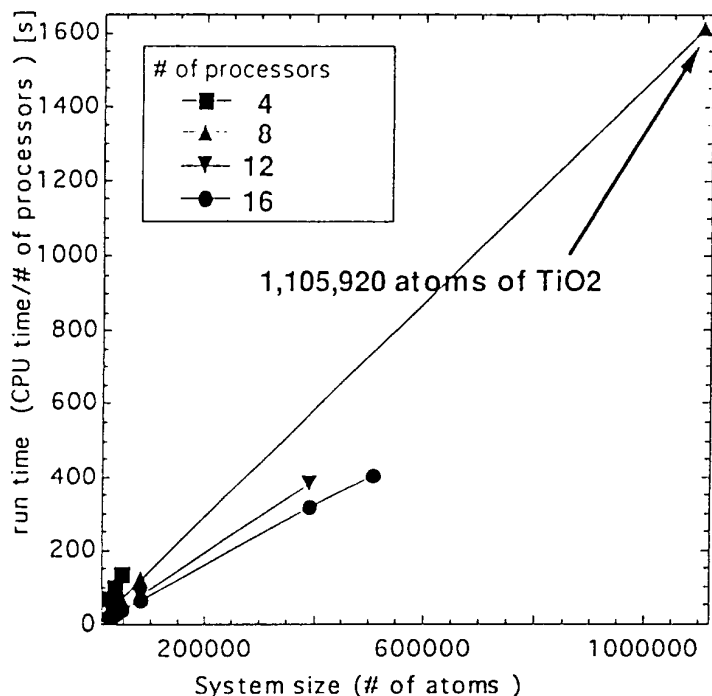


FIG. 1. Computation time as a function of system size for different numbers of Cray C90 processors employed. Linear scaling is seen for all hardware configurations.

The running time was defined as the total CPU time divided by the number of processors which ran a particular job. The running time defined this way differs slightly (up to about 15%) from the wallclock figures given on completion of each job. The discrepancy is attributed to the fact that we ran our program on a busy machine with many jobs competing for the resources. For a given sample size the slope of each curve decreases with the number of processors. This indicates deterioration of performance of each processor as the number of processors used grows. The best performance achieved was 8.5 GFlops with 16 processors (wallclock 452s, system size 552960 atoms).

The linear scaling of the equation of motion algorithm can be expressed as $T = a_{ncpus} \times N_{atoms}$, where T is the run time defined above, N_{atoms} is the number of atoms in the sample and a_{ncpus} is the linear coefficient depending on the number of processors $ncpus$. The coefficients a_{ncpus} normalized to unity ($a_1 = 1$) are plotted on Figure 2. The solid triangles are the results for our algorithm and the open squares are fractions 1, 1/4, 1/8, 1/12 and 1/16 corresponding to the best possible speed-up on a vector-parallel machine when increasing the number of processors from 1 to 4 to 8 to 12 to 16 respectively. Our results very closely follow the relationship $a_{ncpus} = a_1 \times ncpus^{-1}$.

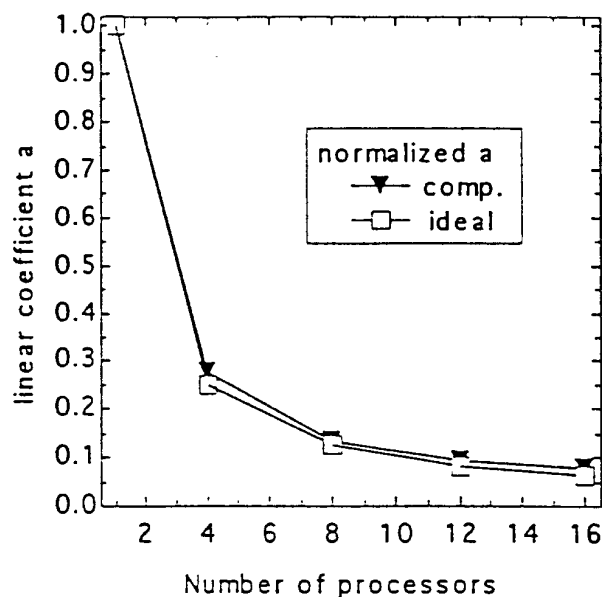


FIG. 2. Normalized coefficients a for the equation of motion program on a Cray C90 (solid triangles) and the best theoretical speed-up (open squares).

B. T3D: MPP implementation

The Cray T3D machine which we used has the serial number SN 6216. The standard configuration was the following:

MPP — SN 6216
 Model — T3D MC/AC128-8
 PEs — 128
 MW/nodes — 8

It must be expected that the performance will drop, when an algorithm that has been successively refined for a particular architecture, is relocated to a different computing environment. This is the current situation with our effort to transfer the computation to the Cray T3D. The basic structure of the algorithm is very well suited to MPP architectures. However, at present we see a ratio of about 1:40 between execution speeds per C90 CPU, to the speed per T3D PE. These figures suggest that it will take a 320 node T3D (640 PEs) to rival a C90/16. It's not that bad, as we see some performance degradation due to memory conflicts, with more than 12 C90 CPUs. So a 256 node T3D should provide a worthwhile contest. Further refinement of the MPP version of the code should result in another doubling of performance. Also

the better caching capabilities and faster PEs of a T3E will enable the latest MPP line, to provide a significant challenge to the current supremacy of the PVP machines, for this type of work.

Some explanation is in order, as to why the equation of motion technique should be so well suited to MPP computation. The calculation is iterative with respect to time-step (Eq. 2.5). Each iteration revises a number of field quantities that are calculated at each atomic site of the solid being considered. The revision of each field quantity requires only values of the set of field quantities at the site being considered and some (few) neighbours (Eq. 2.4).

Implementation is simple. The atomic sites of the model solid are divided into a number of hypercells. One hypercell per PE. Arrays are allocated to hold the field quantities. In fact, the arrays are oversized, the local field quantities are stored in the middle of the arrays while the borders hold the field values for those near neighbours of local sites, that are located in adjacent hypercells. An important point is that for problems with more than, say 350 sites per hypercell, the purely local sites outnumber the sites that require special consideration because they have neighbours on adjacent hypercell. Once past this threshold, inter PE data transfer times cease to be a significant component of the total computation time. This is a very low threshold so, for problems of practical interest, inter PE data transfer times present no burden. The sequence of computational steps is the following:

1. Set up the initial data.
2. Perform an iteration on each PE.
3. Barrier. (To ensure that each PE completes step 2 before proceeding)
4. Inter PE data transfers, to fill the border areas of the field arrays.
5. Barrier. (Ensure that the update is complete before the next iteration)
6. Loop back to step 2, as often as required.

At the outset of this project it was decided to explicitly code the synchronization primitives and to perform all inter PE data transfers via routines from the Shared Memory library (SHMEM). We have had no cause to regret this decision. These simple tools are more than sufficient to meet the requirements of our algorithm. It does mean, however, that we must maintain separate MPP and PVP versions of the code.

IV. RESULTS FOR TEST MODELS

During the last four years a number of refined experimental studies of the rutile $TiO_2(100)$ surface have

appeared in the literature [9-13]. With sophisticated experimental methods such as STM, glancing angle X-ray diffraction and LEED, the researchers were able to observe surface reconstruction and the appearance of $TiO_2(100)1 \times 3$ microfacets. The surface electronic valence-band structure of $TiO_2(100)$ and (110) faces was studied recently using angle-resolved photoemission [14]. It is hoped that by understanding the atomic and electronic structure of rutile surface, and especially the effects of surface steps and microfacets, the mechanism of photocatalytic decomposition of water observed by Fujisima and Honda in 1972 could be better understood [15,9]

The crystallographic structure of ideal rutile TiO_2 [16,17] and its most stable surfaces (110) and (100) [18] is well known. The tetragonal unit cell consists of two Ti and four O atoms. The titanium atoms occupy the positions $(0, 0, 0)$ and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ whereas oxygens are at the positions $\pm(x, x, 0)$, and $\pm(\frac{1}{2} + x, \frac{1}{2} - x, \frac{1}{2})$, where $x = 0.306 \pm 0.001$ [16].

Based on data derived from LEED symmetry and photoemission spectra Muryn *et al.* (1991) proposed the missing row model (MR) for the reconstructed $TiO_2(100)$ surface. Soon after, Zschack *et al.* [11] found that the microfacet model (MM) (Figure 3) was in better agreement with glancing angle X-ray diffraction and LEED measurements than the missing row model.

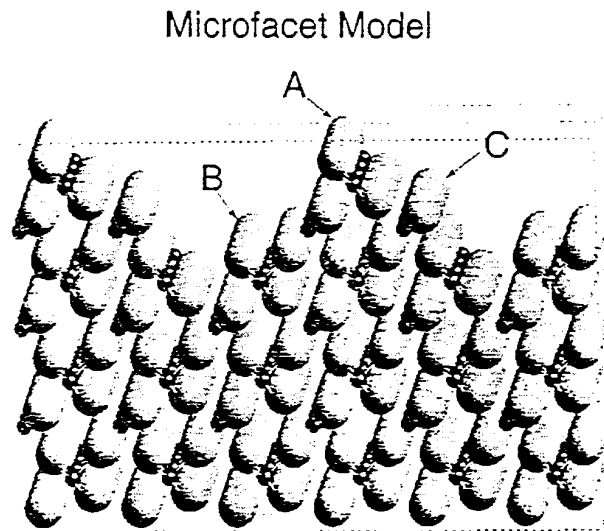


FIG. 3. Microfacet Model of the $TiO_2(100)1 \times 3$ reconstruction (after Ref. [5]). Oxygen is represented by large and titanium by small circles. The oxygen vacancy sites studied in this work are labelled by letters A, B and C.

We used all five models: missing row (MR), microfacet model (MM) and microfacet models A, B and C (MMA, MMB and MMC respectively) to study the effect of the local environment on the local electronic density of states (LDOS).

STM images of reconstructed and reduced surfaces of $TiO_2(100)1 \times 3$ were obtained by a number of groups [10,13]. In all studies the images are obtained at positive sample bias, *i.e.* the electrons tunnel from the tip into unoccupied Ti 3d states. It was reported that the images at negative tip bias could not be obtained. This means that the STM probes the positions of the Ti atoms and cannot distinguish between different reduced surface models (MMA, MMB or MMC). The results of our calculations attempt to answer the question: which model is physically realised?

The local electronic density of states for up to thirteen unequivalent atoms on the surface of microfacet for five models of microfacet (MR, MM, MMA, MMB and MMC) were computed at positions indicated in Figure 4.

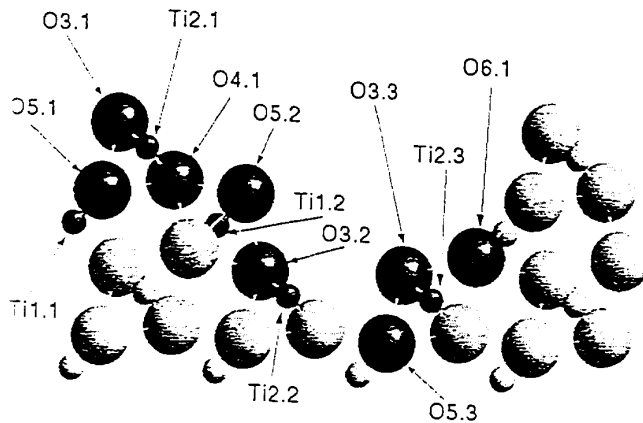


FIG. 4. Five Ti (small black circles) and eight O (large black circles) atom sites at which the LDOS was computed.

The typical results are presented in Figure 5.

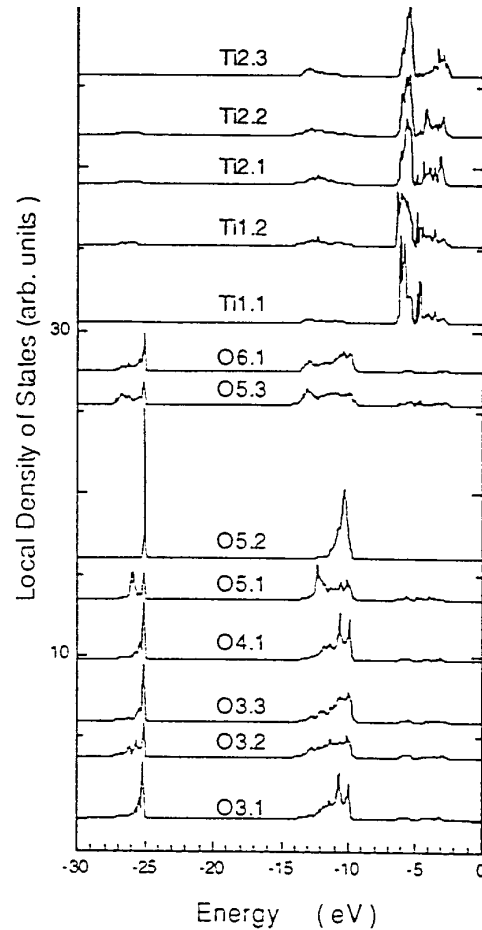


FIG. 5. The Local Electronic Density of States for surface atoms in Microfacet Model (MM) at positions indicated in Figure 2. Note complete lack of covalent mixing and spread of bands at site $O5.2$.

The LDOS for all four models derived from Microfacet Model are very similar. One striking feature is complete lack of covalent mixing and narrow s and p bands at oxygen site $O5.2$. This feature is present in all three models (MM, MMA and MMB). We interpret this as a lack of bonding of $O5.2$ atom at a microfacet. It would mean that Microfacet Model C is the most likely to represent the reduced reconstructed $TiO_2(100)1 \times 3$ surface. This result is different from the conclusions of Hardman *et al.*, [13] and needs to be investigated further.

V. CONCLUSIONS

There is a need for very fast and efficient codes to calculate the electronic structure of novel complex materials. The search for electronic structure algorithms which scale linearly is an intense field of research [19–21]. Defects engineering might lead to exciting new materials and devices. The equation-of-motion method, as imple-

mented in this work, can be used to study disordered transition metal oxides, amorphous semiconductors, the electronic structure of random point and extended defects and their influence on the electronic properties of materials. The current versions of the program, developed specifically for PVP and MPP parallel architectures scale identically in all three spatial dimensions of a sample. It is possible to manipulate a sample, rotate it, and study various crystallographic faces and extended surface defects.

With the program reported in this work we studied the local electronic structure of Ti and O atoms in the vicinity of $TiO_2(100)1 \times 3$ microfacets. The samples were up to 1,105,920 atoms or $18 \times 22 \times 28nm$ in size. The peak speed achieved in benchmarks was 8.5 GFlops/s.

ACKNOWLEDGMENTS

The CRAY part of this project was developed on the CSIRO Supercomputing Facility CRAY Y-MP4/464 in Melbourne. The computations were done on CRAY C916/16256, CRAY C916/12512, CRAY C916/16512 and CRAY T3D located at Cray Headquarters in Egan, MN. These machines are part of Cray Computer Network. Some preliminary results presented in this paper were communicated in the Proceedings *High Performance Computing 1995*. The fragments which are duplicated are reprinted by permission from *High Performance Computing 1995*, pp.101-106, 1995 Simulation Councils, Inc.

- [8] M.T. Michalewicz and M. Priebatsch, *Parallel Computing* **21**, 853 (1995).
- [9] C.A. Muryn, P.J. Hardman, J.J. Crouch, G.N. Raiker, G. Thornton and D.S.-L. Law, *Surface Science* **251/252**, 747 (1991).
- [10] G.W. Clark and L.L. Kesmodel, *Ultramicroscopy* **41**, 77 (1992).
- [11] P. Zschack, J.B. Cohen and Y.W. Chung, *Surface Science* **262**, 395 (1992).
- [12] P.M. Murray, F.M. Leibsle, C.A. Muryn, H.J. Fisher, C.F.J. Flipse and Thornton, G. *Physical Review Letters* **72**, 689 (1994).
- [13] P.J. Hardman, N.S. Prakash, C.A. Muryn, G.N. Raiker, A.G. Thomas, A.F. Prime, G. Thornton and R.J. Blake, *Physical Review B* **47**, 16056 (1993).
- [14] P.J. Hardman, G.N. Raiker, C.A. Muryn, G. van der Laan, P.L. Wincott, G. Thornton, D.W. Bullett and P.A.D.M.A. Dale, *Physical Review B* **49**, 7170 (1994).
- [15] A. Fujishima and K. Honda, *Nature* **238**, 37 (1972).
- [16] F.A. Grant, *Reviews of Modern Physics* **31**, 646 (1959).
- [17] C.J. Howard, T.M. Sabine and F. Dickson, *Acta Crystallographica B* **47**, 462 (1991).
- [18] V.E. Henrich and P.A. Cox, *The Surface Science of Metal Oxides*. Cambridge University Press, Cambridge (1993).
- [19] F. Mauri, G. Galli and R. Carr, *Physical Review B* **47**, 9973 (1993).
- [20] M. Krajcí and J. Hafner, *Physical Review Letters* **74**, 5100 (1995).
- [21] D.A. Drabold and O.F. Sankey, *Physical Review Letters* **70**, 3631 (1993).

-
- [1] Euro-Network for Electronic Structure Calculations: <http://www.dl.ac.uk/TCSC/HCM/PSIK/main.html>
 - [2] M. T. Michalewicz, "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, Email: hpc@nchc.gov.tw
 - [3] Z. Zinamon and E. Nardi, "Electronic and Ionic Structure of Dense Plasmas", preprint, (1995)
 - [4] R. Alben, M. Blume, H. Krakauer and L. Schwartz, *Physical Review B* **12**, 4090 (1975).
 - [5] D. Beeman and R. Alben, *Advances in Physics*, **26**, 339 (1977).
 - [6] M.T. Michalewicz, H. Shore, N. Tit and J.W. Halley, *Comp. Phys. Comm.* **71** 222-234 (1992)
 - [7] M.T. Michalewicz, *Computer Physics Communications* **79**, 13 (1994).