

ON PARALLEL CALCULATION OF ENERGY BAND STRUCTURE

JANUSZ ANDRZEJEWSKI, JAN MISIEWICZ

*Institute of Physics, Wrocław University of Technology,
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland*

Abstract: We propose a parallel method for calculating the energy band structure. Our method can be easily implemented and does not require high speed of communication between the processors. In practical application our method allows for linear increase of the speed of calculations with the number of processors. We also analyse the optimization algorithms for energy band structure. We conclude that the most suitable optimization methods for nonlocal empirical pseudopotential method (NEPM) is adaptive simulated annealing (ASA).

1. INTRODUCTION

The determination of the electronic structure and related properties of condensed matter is an important field in modern solid state physics, and the progress made in it during the last two decades is truly impressive. The progress is in part due to the breathtaking development in computer hardware and software technologies.

There are two kinds of methods of calculation - the empirical one and the first principle band structure methods. Up to recently, first principle theories, even as sophisticated as they can be [1], could not accurately predict the band structure of semiconductors. Most of the understanding of these materials was obtained from less accurate descriptions [2]. Among these, empirical theories have played (and still play) very important role since they allow us to simulate the true energy bands in terms of restricted number of adjustable parameters. There is at least one essentially distinct method of achieving this goal - the empirical pseudopotential method.

2. NONLOCAL EMPIRICAL PSEUDOPOTENTIAL METHOD (NEPM)

The basic concepts of pseudopotential formalism in crystals were formulated by Philips and Kleinman [3]. The fundamental concept involved in a pseudopotential calculation is that the ion core can be omitted. The pseudopotential approximation can be seen from different perspectives. Pseudopotentials can be constructed either theoretically, using Hartree-Fock atomic calculations, or semi-empirically, by means of spectroscopically known excited states of valence electrons [4]. They can show an angular momentum and an eigenvalue dependence. Typically used pseudopotentials are nonlocal ones.

Our program is written in FORTRAN77 and it is based on the theory described elsewhere [2, 4, 5]. Here, we only want to outline the most essential features.

For the plane-wave representation the solutions of Schrödinger equation are given by secular equation

$$\det | H(\mathbf{k}, \mathbf{G}, \mathbf{G}', E) - E(\mathbf{k})\mathbf{I} | = 0, \quad (1)$$

where \mathbf{k} is wave vector, \mathbf{G} is reciprocal-lattice vector and $E(\mathbf{k})$ is the energy of bands at \mathbf{k} -point. The one-electron Hamiltonian is given by

$$H(\mathbf{k}, \mathbf{G}, \mathbf{G}', E) = -\left[\frac{\hbar^2 k^2}{2m} \right] \delta_{\mathbf{G}, \mathbf{G}'} + V(\mathbf{k}, \mathbf{G}, \mathbf{G}', E)S(\mathbf{G} - \mathbf{G}'). \quad (2)$$

$V(\mathbf{k}, \mathbf{G}, \mathbf{G}', E)$ is the pseudopotential form factor, and $S(\mathbf{G})$ is the structure factor. The pseudopotential form factor is

$$V(\mathbf{k}, \mathbf{G}, \mathbf{G}', E) = V^{loc}(\mathbf{G} - \mathbf{G}') + V^{nl}(\mathbf{k}, \mathbf{G}, \mathbf{G}', E), \quad (3)$$

where $V^{loc}(\mathbf{G} - \mathbf{G}')$ is a local pseudopotential and $V^{nl}(\mathbf{k}, \mathbf{G}, \mathbf{G}', E)$ represents nonlocal corrections to the pseudopotential. For the local part of the pseudopotential we used a function suggest by Falicov and Lin [6].

The nonlocal corrections to the pseudopotential are caused by d states within the ion core which modify the conduction-band structure. The expression for V^{nl} can be found in [4, 5].

With respect to the computational details, the energy cutoffs in the matrix in equation (1) were determined so that plane waves with $\hbar^2/(2m) (\mathbf{k} + \mathbf{G})^2 < E_{cut1}$ were treated directly; plane waves with $E_{cut1} < \hbar^2/(2m) (\mathbf{k} + \mathbf{G})^2 < E_{cut2}$ were treated using Löwdin-Brust [7, 8] method. For calculating the eigenvalue and eigenvectors it was used the standard numerical library - LAPACK [9], for which the source code can be found at <http://www.netlib.org>.

The standard procedure for obtaining theoretically the optical constants of a crystal is to evaluate expression [4]

$$\varepsilon_2(\omega) = \frac{4\pi^2 e^2 \hbar}{3m^2 \omega^2} \sum_{cv} \int_{BZ} \delta(\omega_{cv}(\mathbf{k}) - \omega) \cdot |\mathbf{e} \mathbf{M}_{cv}(\mathbf{k})|^2 d\mathbf{k}, \quad (4)$$

where ε_2 is the imaginary part of the dielectric function; $\hbar\omega_{cv}(\mathbf{k}) = E_c(\mathbf{k}) - E_v(\mathbf{k}) = E_{cv}$ is the interband transition energy between the valence (v) and conduction (c) band at the \mathbf{k} -point in the Brillouin zone.

Using Kramers-Kröning relations the real part of the dielectric function can be obtained and thus, the reflectivity; $|\mathbf{e} \cdot \mathbf{M}_{cv}(\mathbf{k})|^2$ is the matrix element between a given couple of the valence and the conduction band.

Equation (4) can be written in the form

$$\varepsilon_2(\omega) = \frac{4\pi^2 e^2 \hbar}{3m^2 \omega^2} \sum_{cv} \frac{2\hbar}{(2\hbar)^3} J_{cv}(\omega), \quad (5)$$

where

$$J_{ij}(\omega) = \int_{E=E_{cv}} \frac{dS}{|\nabla_{\mathbf{k}} E|} |\mathbf{e} \mathbf{M}_{cv}(\mathbf{k})|^2 \quad (6)$$

Table I. The time of calculations of the objective function as dependency on the number of processors

Number of processors	Time of calculation
1	222 min.
2	112 min.
3	80 min.
4	64 min.
5	54 min.
6	44 min.

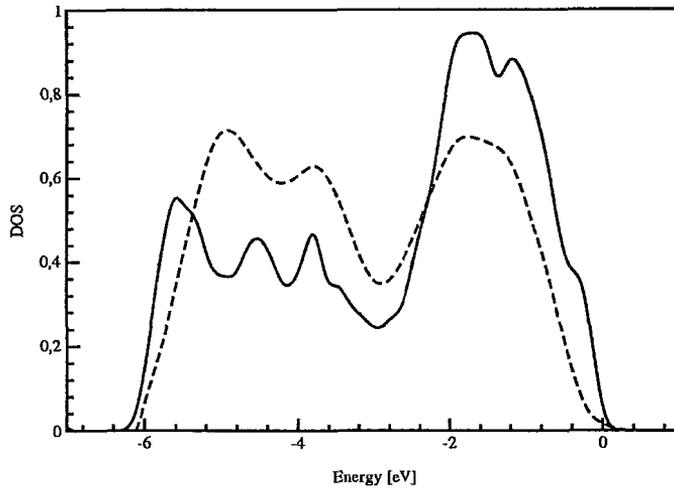


Fig. 1. The density of states of Zn_3P_2 . Solid line - theory, dashed line - experiment. The energy zero is at the top of the valence band.

is the total density of states. For calculating these integrals in \mathbf{k} space over the Brillouin zone tetrahedron method [10, 11] was used.

The word "empirical" in the name of the method means, that one adjust the form factors in such a way that there should be some resemblance between theory and experiment in some kind of optical spectrum. Hand adjust form factors are impossible because calculations take too much time. So we construct some objective function - this function describes the agreement between the theory and experiments. The lower values of the objective function the better agreement is achieved. Thus, once we have the objective function, the process of adjusting the form factors to the experimental data can be done by some optimization algorithm.

Although there may be several ways to classify optimization algorithms from different points of view [12], one natural candidate is a classification based on the deterministic or non-deterministic nature of the search algorithm. The deterministic algorithms (like gradient methods or Newton methods) are very sensitive to the starting point and they are often trapped in local

minima. For our task these systematic methods fail and we must resort to non-systematic techniques i.e. random search ones: So, we try two different kinds of stochastic methods: genetic algorithm and simulated annealing.

Genetic algorithm (GA) is an attempt to simulate the phenomenon of natural evolution which, in biology, was first observed by Darwin. Any GA [13] emphasizes the role of the representation and the interaction between the representation and perturbation operators. The GA use the representation to implicitly divide the search space into several non-overlapping classes. The algorithm works on a population of samples. We use one of the modifications of the GA called messy genetic algorithms (mGA) [14]. Unlike simple GA's, the mGA emphasize the search for appropriate relations among the members of the search space.

As its name implies, the Simulated Annealing (SA) method exploits an analogy between the way in which a metal cools and freezes into minimum energy crystalline structure (the annealing process) and the search for a minimum in more general system. The algorithm is based upon that of Metropolis et al. [15] (which is also known as Monte Carlo method). SA's major advantage with respect to other methods is the ability to avoid becoming trapped at local minima and it has a proof for asymptotic convergence to the optimal solution [16]. In this work we use the algorithm called ASA (Adaptive Simulated Annealing) developed by Ingber [17],

The construction of the objective function is a separate problem. In case of Zn_3P_2 we even do not know what is the kind of the energy band gap [18]. So, in this case the situation is very difficult and we concentrate only on that particular compound. For the construction of the objective function we consider and use the following experimental data:

- 1) Direct energy band gap is in the range of $1.5 \div 1.6$ eV.
- 2) Maximum of valence band is located at the Γ point.
- 3) The width of the valence density of states is $\sim (5.5 \div 6.5)$ eV.
- 4) Any indirect band gap and direct band gap at points other than Γ point are greater than 1.3 eV.

Constructing the objective function we also used the experimental reflectivity [19] in the range of $0 \div 8$ eV and valence density of states [20]. In order to achieve better agreement between the theory and experiment we use the quantity defined as $(\partial R_{cal}/\partial E) \cdot (1/R_{cal})$ and quantity $(\partial(DOS_{cal})/\partial E) \cdot 1/(1 + DOS_{cal})$, respectively, for calculated reflectivity and density of states. These two quantities have been fitted in the sense of the least square approximation. This way of fitting has some advantages:

- 1) Calculating derivatives one gets good fitting to the optical transitions.
- 2) Dividing by reflectivity or density of states, one gets good fitting to the absolute values of these ones.

3. PARALLEL IMPLEMENTATION

Parallel processing, the method of having many small tasks to solve one large problem, has emerged as a key-enabling technology in modern computing. In general, there are two kinds of parallel computing:

- massively parallel processors (MPP). In MPP every processor is exactly like every other in capability, resources, software, and communication speed.
- distributed computing - a set of computers connected by a network; the computers may be made by different vendors, may have different compilers, computational speed, etc.

MPP computers are very expensive. In contrast, for distributed computing one can use many computers available at a university or even in a city. Distributed computing also has some other advantages - the computational speed can be enhanced by adding new computers.

The most popular libraries for distributed computing are message passing interface (MPI) [21] and parallel virtual machine (PVM) [22], Both implementations are open source and can run on a variety of computers and operating systems. We choose PVM, because MPI does not include necessities such as process management (the ability to start a task). This is a very important feature, because sometimes connections between computers are lost (because of the network problems or shut down a computer) and it is necessary to start a new task on computers when the connection can be established again.

Our parallel implementation of computing energy band structure is based on PVM. Its source code is available in the Internet at <http://www.netlib.org> and we successfully compiled it on computers such as the HP Series 700, SGI Power Challenge XL and PC with running Linux (both with kernel 2.0 and 2.2).

In distributed computing, where sometimes one computer is far from the other ones, it is important to take care how much data are being sent between them. If one wants to calculate the energy band structure using NEPM, it is necessary to calculate the objective function. The main computational effort in calculating the objective function is the calculation of the density of states (DOS) and the reflectivity. The time spent in calculation of the DOS and the reflectivity is greater than 95% of total time. In order to calculate this quantity one must calculate the expression (6), but before that secular equations (1) on grids of \mathbf{k} -points should be solved. We use master-slave model of calculation. The slave process computes the matrix (equation 2), solve the secular equation (equation 1) and calculate the matrix elements. The most demanding task in the slave process is calculating the eigenvalue and eigenvectors of secular equation.

The master process is responsible for sending a new task to the slave and for collecting the calculated data. When for all \mathbf{k} -points the energy of bands and matrix elements are calculated, the master process calculates joint density of states (equation 6) and finally calculates the objective function. Further, master process performs the next step in optimization and parallel computation begins.

This method of parallel computation of energy band does not send big amount of data between the master process and slave one and can use as many slave processes as the number of different \mathbf{k} -points in the Brillouin zone (typically 70-100 points). The network load generated by the communication between processes are small because the times needed to complete the whole task are big (they depend of the size of matrix; for very big matrix this can be several hours on SGI). If distributed computing is done on heterogeneous computers the master process also must take care about the connections because sometimes connections between master and slave

processes are lost due to various reasons like: network problems, shut down of computers on which slave task is running etc.

In order to save the CPU time the process of computation of the objective function can be truncated to the calculation of the eigenvalues and eigenvectors at the points of high symmetry. Next, for instance if the energy of transitions at these points are reasonable, the process of calculation is continued - the DOS and the reflectivity are calculated.

4. EXAMPLE RESULT

Because the scope of this article is not a calculation of the energy band of a certain structure, we only present a result of example calculation - Fig. 1. From our experience, we find out that deterministic algorithms are not helpful because of their sensitivity to starting point and their local character. The mGA algorithm gives much better results, but still not satisfactory. As it was stated in [23] the ASA is at least an order of magnitude superior to the GA in the convergence speed and it is more likely to find a global minimum during a time limited search. This is an important fact if one takes into consideration that the calculation of the objective function sometimes is very expensive. Our numerical experiments show that the ASA method is more suitable than the mGA one for this kind of calculations.

In the table 1 we present the time of calculation of the objective function as dependent on the number of processors used in the calculations. It can be seen that the dependence is almost linear with number of processors. These calculations were done on the SGI Power Challenge XL computer with 12 processors. The results presented in this table concern the situation that there are no other running programs on the computer.

5. CONCLUSION

For adjusting the form factors the stochastic algorithms like the mGA and the ASA were used. These methods require much more calculations of objective function than deterministic algorithms. The deterministic algorithms are useless, however our numerical experience shows that the ASA method gives much better results than mGA. The construction of the objective function is also important, and we suggest the way of construction of the objective function.

The algorithm described in this paper is efficient and, in our opinion, can be rather easily implemented, especially if one has a working sequential program for energy band structure. Another advantage is that in practice this algorithm allows for linear increase of the efficiency with number of processors. Moreover, it does not require high communication speed. Because the PVM library is an open source and it can be compiled on a variety of computers, the same source program can be used both on supercomputers, on workstations, and PC computers.

The numerical calculations were done on various computers in Technical University of Wrocław and in the Poznań Supercomputing and Networking Center.

References

- [1] Singh D. J., *Planewaves, Pseudopotentials and the LAPW method*, Kluwer Academic Publishers (1994).
- [2] Cohen M. L., Chelikowsky J. R., *Electronic Structure and Optical Properties of Solids*, Springer Verlag, New York (1988).
- [3] Phillips J. C., Kleinman L., Phys. Rev., **116**, 287 (1959); Phys. Rev., **116**, 800 (1959).
- [4] Ehrenreich H., Seitz F., Turnbull D., *Solid State Physics*, vol. 24, Academic Press, New York and London (1970).
- [5] Chelikowsky J. R., Cohen M. L., Phys. Rev., **B14**, 556 (1974).
- [6] Falicov L. M., Lin P. J., Phys. Rev., **141**, 562 (1966).
- [7] Löwdin P., J. Chem. Phys., **19**, 1396 (1951).
- [8] Brust D., Phys. Rev., **134**, A1337 (1964).
- [9] Anderson E., Bai Z., Bischof C., Demmel J., Dongarra J., Du Croz J., Greenbaum A., Hammarling S., McKenney A., Ostrouchov S., Sorensen D., *LAPACK Users' Guide - Release 2.0*, SIAM Philadelphia (1994).
- [10] Lehmann G., Taut M., phys. stat. sol. (b), **54**, 469 (1972).
- [11] Baldereschi A., Phys. Rev., **B7**, 5212 (1973).
- [12] Törn A., Zilinskas A., *Global optimization*, Springer Verlag, Berlin 1989.
- [13] Goldberg D. E., *Genetic algorithms in search, optimization, and machine learning*, Addison Wesley, New York (1989).
- [14] Goldberg D. E., Deb K., Korb B., Complex Systems, **4**, 415 (1990).
- [15] Metropolis N., Rosenbluth A. W., Rosenbluth M. M., Teller A. H., Teller E., J. Chem. Phys., **21**, 1087(1958).
- [16] Kirpatrick S, Gelatt C. D., Vecchi M. P., Science, **220**, 671 (1983).
- [17] Ingber L., J. Mathem. Comput. Model., **12**, 967 (1989).
- [18] Misiewicz J., J. Phys.: Condens. Matter, **2**, 2053 (1990).
- [19] Misiewicz J., Jezierski K., Solid State Commun., **70**, 465 (1989).
- [20] Hupfer A., Hirsch D., Schulze S., phys. stat. sol. (b), **152**, 505 (1989).
- [21] Gropp W., Lusk E., Skjellum A., *Using MPI: Portable Parallel Programming with the Message-Passing Interface*, MIT Press, Cambridge (1994).
- [22] Geist A., Beguelin A., Dongarra J., Jiang W., Manchek R., Sunderam V., *PVM: Parallel Virtual Machine. A Users' Guide and Tutorial for Networked Parallel Computing*, MIT Press, Cambridge (1994).
- [23] Ingber L., Rosen B., Mathem. Comput. Model., **16**, 87 (1992).