ISSN 0965-5425, Computational Mathematics and Mathematical Physics, 2014, Vol. 54, No. 12, pp. 1929–1935. © Pleiades Publishing, Ltd., 2014.
Original Russian Text © K.K. Abgaryan, M.A. Posypkin, 2014, published in Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 2014, Vol. 54, No. 12, pp. 1994–2001.

Optimization Methods as Applied to Parametric Identification of Interatomic Potentials

K. K. Abgaryan and M. A. Posypkin
Dorodnicyn Computing Center, Russian Academy of Sciences, ul. Vavilova 40, Moscow, 119333 Russia
e-mail: kristal83@mail.ru, mposypkin@gmail.com

Received July 29, 2014

Abstract—Various optimization methods are applied to parametric identification of interatomic potentials. The problem of choosing parameters of the Tersoff potential in the case of covalently bonded monocrystals is considered.

DOI: 10.1134/S0965542514120021

Keywords: optimization methods, parametric identification, objective function, gradient methods, Tersoff potential, software package.

INTRODUCTION

At present, the development and application of optimization methods seems to be an especially promising approach to the important task of predictive computer simulation of crystalline structures with the use of multiscale techniques (see [1, 2]). In this approach, results of the first-principle simulation [3, 4] of atomic structures and electron properties of systems consisting of 200–1000 atoms are used as input data for the simulation of more complex structures consisting of 1000–1 000 000 atoms. Such approaches provides ample opportunities for the predictive simulation of structures with defects and make it possible to consider various dynamic (time-dependent) processes, such as diffusion and adhesion. At each scale level, the problem under study can be stated in an extremal formulation and suitable optimization methods can be applied. At the first scale level, the optimization problem of finding a configuration of basis atoms corresponding to the minimum total energy of the system is solved within the framework of the theory of electron density functionals [2, 3] by minimizing the energy functional of the system. This approach is implemented in well-known application packages, such as VASP (http://cms.mpi.univie.ac.at/vasp/) and Quantum ESPRESSO.

The number of particles in the considered system can be increased by passing to the next scale hierarchy level, namely, to molecular-dynamical simulation. Empirical interatomic potentials are used at this level. An important stage in molecular-dynamical simulation is the structural identification of potentials for a particular crystalline material. This means the selection of the form of a potential (determination of a potential structure) suitable for describing a particular crystalline structure, which is characterized by its chemical composition, geometry, and the type of chemical bonding. The parameters of the selected interatomic potential are determined by solving the optimization problem of parametric identification. Local and global optimization methods are applied (see [5, 6]). The physical parameters produced by first-principle methods can be used as reference values in solving this optimization problem (see [6]).

STATEMENT OF THE OPTIMIZATION PROBLEM

Formally, the optimization problem can be written as the minimization of the functional

$$F(\xi) = \sum_{i=1}^{m} \omega_i \left(f_i(\xi) - \hat{f}_i\right)^2 \to \min, \quad \xi \in X, \tag{1}$$

where $\hat{f}_i$ is the reference value of the ith characteristic, $f_i(\xi)$ is the value of this characteristic obtained by computations for a given set of basis atoms, $\xi \in R^n$ is the vector of parameters to be tuned, and ω_i is a weighting coefficient. The feasible set $X \subseteq R^n$ is a parallelepiped whose boundaries are chosen so as to contain the possible range of the parameters. In problem (1), the task is to find a set of parameters $\xi \in R^n$

minimizing the function $F(\xi)$. This set will ensure the minimum deviation of the computed material characteristics from the reference values, thus describing the simulated crystal properties in the most precise manner.

Below, we propose an approach based on reproducing the elastic properties of the crystal. More specifically, the approach is based on the cohesion energy value computed with the help of first-principle simulation, on experimentally determined elastic constants, and on other major characteristics of the structural properties of the material under consideration. The number of terms in (1) varies depending on the material.

As a result, we consider the function

$$F(\xi_1,\dots,\xi_n) = \sum_i \left(\omega_i(A_i(\xi) - \overline{A_i})^2\right) \to \min,$$

where $\xi_1 \dots \xi_n$ are the potential parameters to be optimized and the values A_i are obtained by ab initio computation or experimentally.

A solution is sought on the set $X \in R^n$ of all admissible values of the desired parameters, where $\xi = (\xi_1, \dots, \xi_n) \in X \in R^n$ is the vector of tuned parameters.

The form of the minimized function $F(\xi_1, \dots, \xi_n)$ is chosen depending on the chemical composition, the crystalline structure, and the type of bonding of the simulated material. Even in the simplest cubic case, $F(\xi_1, \dots, \xi_n)$ is a multiextremal function.

To solve the problem at the initial stage, we need to specify the admissible ranges of searching for potential parameters: $[\underline{\xi_i}, \overline{\xi_i}]$, $i = 1, \dots, n$, where n is the total number of potential parameters to be determined. This can be done using various methods. Below, we show how this was done in solving the parametric identification problem for the parameters of the Tersoff potential (see [7]). A point with parameter values $(\xi_1, \dots, \xi_n)$ sufficiently close to the global minimum is found using global optimization methods, such as Monte Carlo, scanning, or Granular Random Search [7]. Next, the point found is used as an initial one in a more accurate search for a local minimum, for example, by applying the gradient descent method with adaptive step size control or using granular radial search (GRS) [8].

It should be noted that the procedures described above are executed for particular, fixed positions of the basis atoms in the considered crystalline structure. As a result, after solving the parametric identification problem in this formulation, we do not know whether the positions of the basis atoms correspond to the minimum of the potential energy of the system.

Accordingly, the next step is to compute the forces acting on each atom in an elementary cell and the corresponding displacements of the atoms to positions associated with the minimum of the energy. This procedure is known as the *relaxation of the structure*. The directions of the forces are calculated according to the potential energy gradient. There is an alternative relaxation procedure, in which optimization is performed with respect to the coordinates of the particles so that the particles are placed in positions corresponding to the minimum of the total potential energy of the system.

If, with the potential parameters found, the atoms are shifted significantly from the required equilibrium positions or the material properties differ substantially from the required ones, the parametric identification procedure for the chosen potential should be repeated with the characteristics to be refined added to the set of reference values in (1).

NUMERICAL EXPERIMENT: SOLUTION OF A REAL-WORLD PROBLEM

Consider the problem of selecting parameters of the Tersoff potential [7] in the case of monocrystals. This potential is a multiparticle one based on the concept of the order of bonds: the bonding force between two atoms is not constant, but depends on the local environment. Such potentials can be used to describe the properties of covalently bonded crystals (for example, carbon, silicon, and germanium).

With the use of the Tersoff potential, the cohesion energy E_{coh} of the interaction of a group of atoms can be calculated using the formula

$$E_{\mathrm{coh}} = \frac{1}{2}\sum_{i \neq j} V_{ij},$$

$$V_{ij} = f_c(r_{ij})\left(V_R(r_{ij}) - b_{ij}V_A(r_{ij})\right),$$

$$f_c(r) = \begin{cases} 1, & r < R - R_{\text{cut}}, \\ \dfrac{1}{2}\left(1 - \sin\left(\dfrac{\pi(r-R)}{2R_{\text{cut}}}\right)\right), & R - R_{\text{cut}} < r < R + R_{\text{cut}}, \\ 0, & r > R + R_{\text{cut}}, \end{cases}$$

$$V_R(r_{ij}) = \frac{D_e}{S-1}\exp\left(-\beta\sqrt{2S}(r_{ij} - r_e)\right),$$

$$V_A(r_{ij}) = \frac{SD_e}{S-1}\exp\left(-\beta\sqrt{\frac{2}{S}}(r_{ij} - r_e)\right),$$

$$b_{ij} = \left(1 + \left(\gamma\zeta_{ij}\right)^n\right)^{-1/2n},$$

$$\zeta_{ij} = \sum_{k \neq i,j} f_c(r_{ik}) g(\theta_{ijk}) \omega_{ik},$$

$$\omega_{ik} = \exp\left(\lambda^3\left(r_{ij} - r_{ik}\right)^3\right),$$

$$g(\theta_{ijk}) = 1 + \left(\frac{c}{d}\right)^2 - \frac{c^2}{d^2 + (h - \cos\theta_{ijk})^2}.$$

The variables r_{ij} denote the distances between the atoms indexed by i and j, respectively, and θ_{ijk} is the angle made by the vectors joining atom i to atoms j and k, respectively. The Tersoff potential involves 12 parameters, which are specific to the substance modeled: D_e, r_e, β, S, n, γ, λ, c, d, h, R, R_{cut}. Note that R and R_{cut} are determined from experimentally measured geometric characteristics of the substance and do not need to be chosen.

In what follows, the parameters of the Tersoff potential are identified for isotropic crystals with a cubic crystalline lattice. According to Hooke's law, small deformations are proportional to stresses (see [9]):

$$\sigma = C\varepsilon;$$

here, σ is the stress tensor; ε is the strain tensor; and C is the elastic constant tensor, whose components are given by the expressions

$$C_{ij} = \frac{1}{V}\frac{\partial^2 E}{\partial\varepsilon_i \partial\varepsilon_j}, \quad i, j = 1, \ldots, 6, \tag{2}$$

where ε_i are the corresponding components of the strain tensor of the crystal.

Thus, $C_{ij}(\xi)$ are the elastic constants computed using formula (2) with the use of the cohesion energy calculated with the help of the chosen interatomic potential and $\hat{C}_{ij}$ are experimental values of the elastic constants. The number of independent elastic constants is determined by the crystalline structure of the modeled material. For example, in the case of a cubic crystal, there are only three independent elastic constants: C_{11}, C_{12}, and C_{44} (see [9, 10]). The other tensor components either vanish or coincide with these three. In the case under consideration, the objective function to be minimized in the parameter identification problem is given by

$$f(\xi) = \omega_1(E(\xi) - \hat{E})^2 + \omega_2(a(\xi) - \hat{a})^2 + \omega_3(B(\xi) - \hat{B})^2 + \omega_4(C'(\xi) - \hat{C}')^2$$
$$+ \omega_5(C_{44}(\xi) - \hat{C}_{44})^2 + \omega_6(\zeta(\xi) - \hat{\zeta})^2,$$

where $\xi = \left(D_e, r_e, \beta, S, n, \gamma, \lambda, c, d, h\right)$ is the vector of parameters to be identified. The objective function involves the following quantities:

E is the cohesion energy of the system divided by the number of atoms in the simulated group (specific energy);

a is the lattice constant (characteristic length);

B is the bulk modulus;

C' is the shear modulus;

C_{44} is the elasticity constant;

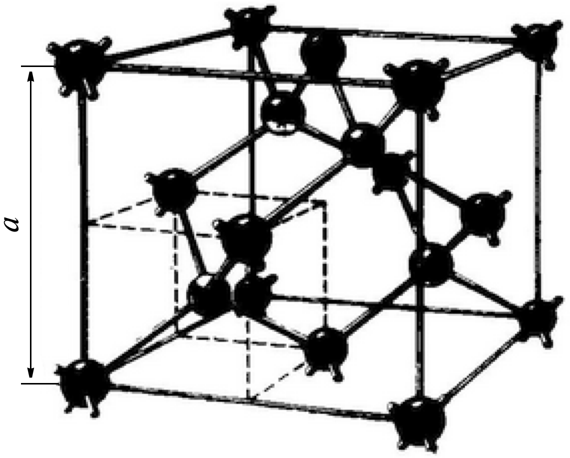

Fig. 1. Single cell of the silicon crystalline lattice.

ζ is the Kleinman parameter, which characterizes the additional change in the lattice under a shift deformation. This change is caused by the tendency of the atoms to occupy an energetically optimal position.

The parameter C_{44} is a component of the elastic constant tensor, and the bulk and shear moduli are calculated in terms of the components C_{11} and C_{12} according to the formulas

$$B = (C_{11} + 2C_{12})/3, \quad C' = (C_{11} - C_{12})/2.$$

Obtained experimentally or by quantum mechanical methods for monocrystalline silicon, the reference values of the components of the objective functions were taken from [8]:

$$\hat{E} = -4.63, \quad \hat{a} = 5.43, \quad \hat{B} = 0.97, \quad \hat{C}' = 0.51, \quad \hat{C}_{44} = 0.79, \quad \hat{\zeta} = 0.52.$$

The parallelepiped $X = [\underline{\xi}, \overline{\xi}] = \{\xi \in R^n : \underline{\xi}_i \le x_i \le \overline{\xi}_i\}$ for specifying initial approximations was chosen so that it contained all possible parameter values. We analyzed known values of the coefficients of the Tersoff potential for monocrystalline silicon. The boundaries were chosen so that they contained these values:

$$\underline{\xi} = (0.5,\ 0.5,\ 0.5,\ 0.5,\ 0.1,\ 5\times 10^{-8},\ 0.5,\ 10000,\ 1, -2),$$

$$\overline{\xi} = (10,\ 5,\ 5,\ 5,\ 2,\ 3\times 10^{-6},\ 3,\ 200000,\ 30, -0.1).$$

All computations were performed for a single silicon lattice cell, which consists of 18 atoms (Fig. 1). Not only the interactions between the cell atoms, but also the interactions with neighboring atoms within the cutoff radius R_{cut} were taken into account.

The cohesion energy was calculated using the formulas for the Tersoff potential. Then the resulting value was divided by the number of atoms in the cell to obtain specific energy. The components B and C' were calculated by approximating the ratio of the energy variation to the corresponding small lattice deformation. The deformation used for computing C_{44} is such that the some of the atoms shift from the positions determined by the deformation to occupy an energy-optimal position. This is simulated with the help of a procedure for finding the minimum of energy depending on the displacements of the group of internal atoms in the lattice determined by the Kleinman parameter. This one-dimensional minimization procedure is used to determined the lattice constant a for given potential parameters.

A software package has been implemented, which includes the following modules:

(1) a module generating a crystalline lattice, which fills up the atomic coordinate array according to the type of the crystalline lattice and its characteristic size;

(2) a module computing the basic material characteristics and the objective function $F(x)$;

(3) an optimization module searching for a minimum of $F(x)$ on the feasible set X.

The most complicated component of the software package is the optimization module, which was implemented using the BNB-Solver library (see [11]). The problem in question is of the black box class, since there is no explicit expression for the objective function and its value is produced by a complicated numerical procedure. This problem was solved by combining search globalization methods, which ensure the diversification of the search within the feasible parallelepiped, and local methods, which find a local minimum in a given neighborhood.

As a search globalization method, we used the Monte Carlo method, which generated a sequence of random initial approximations within the feasible parallelepiped. Uniformly distributed pseudorandom numbers in each coordinate were used within the parallelepiped.

Two local optimization algorithms were tested. The first was based on the gradient descent method with adaptive step size control. The gradient was numerically approximated by finite differences. The second algorithm was based on the GRS method proposed in [8]. According to this method, one of the parameters undergoes a random shift from the initial point within a given range g. If this leads to a smaller value of the objective function, then the newly found value is used as an initial one and a new shift is executed. The original initial value is chosen at random in the parallelepiped X. If the fraction of "successful" shifts leading to smaller objective function values is reduced below a given threshold, then the granularity g is decreased. The algorithm terminates when the upper bound on the number of iterations is surpassed or when the granularity becomes too small.

Since the identified parameters differ widely in characteristic values, they have to be rescaled to ensure the effective performance of the algorithm. In this case, we used linear rescaling with the coefficient multiplying the corresponding parameter equal to $1/\tilde{\xi}_i$, where $\tilde{\xi}_i = \left(\underline{\xi}_i + \bar{\xi}_i\right)/2$.

Preliminary numerical experiments showed that the gradient descent algorithm is outperformed by GRS as applied to the given problem. Under constraints on the number of iterations so that the running time was roughly identical, the GRS algorithm ensured considerably higher accuracy. In an experiment with eight runs starting from various initial approximations, the GRS method produced an optimal value of 0.0002, while the gradient descent yielded a minimal objective function value of 0.1987.

The numerical experiment was organized as follows. As many as 32768 random initial approximations uniformly distributed in X were generated. Next, the GRS algorithm was applied to these approximations. A server with two four-core Intel(R) Xeon(R) CPU E5620 2.40 GHz processors was used for the computations. To accelerate the computations, we run eight parallel processes, each working with 4096 initial approximations. The total run time was about 2 days.

Among the local minimizers found, only those with an objective function value not exceeding 10^{-5} were further processed. The local minimum condition was additionally numerically checked at these points. Specifically, the objective function was computed at some points chosen at random in a neighborhood of given radius. If its value in at least one of them was smaller than the believed minimum, the check was considered failed. The neighborhood was specified as $U(x) = \left\{y \in R^n : |x_i - y_i| \le \rho\right\}$, where $\rho = 0.1$. The number of points at which the objective function was computed was set to 1000.

The check for a local minimum was successful for 11 generated points satisfying the condition $f(\xi) \le 10^{-5}$. The deviations of the found parameters from the reference values were analyzed. For each of 10 parameters, the deviation d_i from its reference value was calculated using the formula

$$d_i = \frac{|x_i - r_i|}{|r_i|}, \quad i = 1, \ldots, 10,$$

where x_i is the value of the parameter and r_i is its reference value. The circles in Fig. 2 show the deviations for 10 identified parameters.

The results show that the parameters D_e, r_e, β, n, h were determined rather accurately with the objective function used. The other parameters were identified to a varying degree of accuracy. For some of them, the deviations were significant (1.5 times higher than the reference value).

If the last six parameters are fixed, then the distribution has the form shown in Fig. 3. In the case of 1024 initial approximations, the objective function values at 112 local minimizers were less than 0.001. Although the number of points was tenfold more than in the case of 10 parameters, while the computations were less accurate, the scatter of the parameter values was noticeably lower, being in an acceptable range.

Below are the parameter values for silicon corresponding to the point ξ^0 at which the minimum objective function value $v_0 = 5 \times 10^{-6}$ was found and the vector $\hat{\xi}$ of reference parameter values from [11]:

$$\xi^0 = (2.36, 2.34, 1.48, 1.73, 0.95, 1.34 \times 10^{-6}, 0.79, 121010, 13.91, -0.42),$$

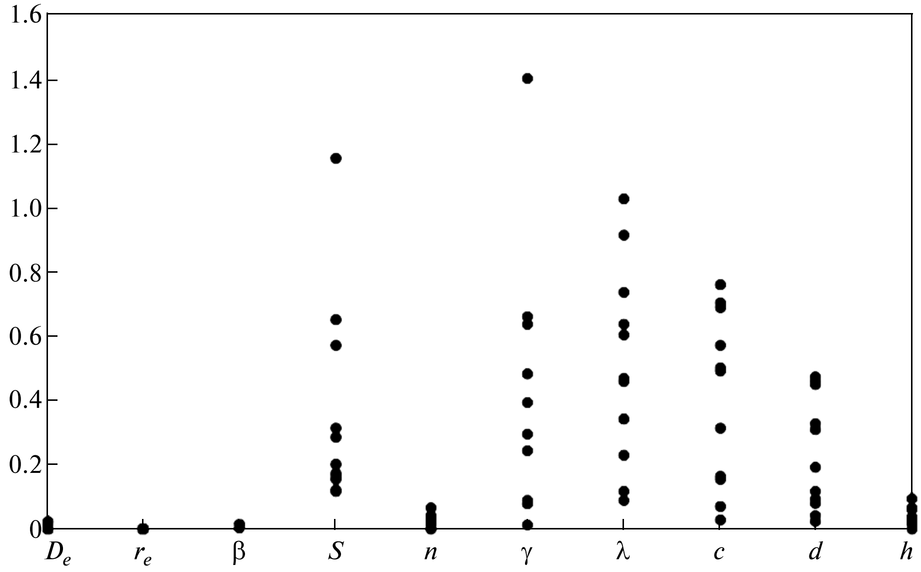

Fig. 2. Deviations of ten parameters from their reference values for the local minima found.

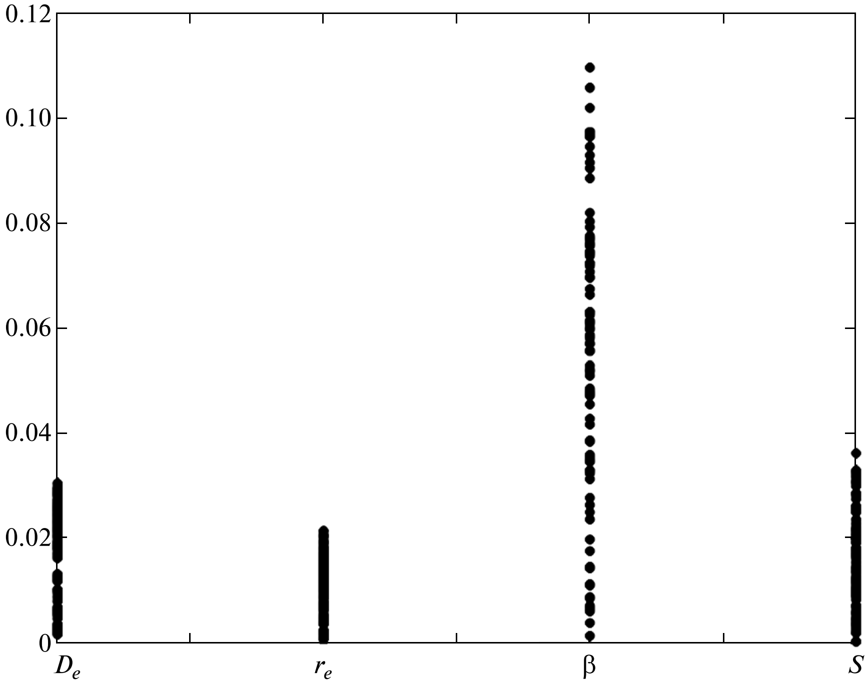

Fig. 3. Deviations of four parameters from their reference values for the local minima found.

$$\hat{\xi} = (2.36,\ 2.34,\ 1.46,\ 1.48,\ 0.94,\ 1.25\times 10^{-6},\ 1.46,\ 113031,\ 14.25,\ -0.42).$$

The basic components of the objective function are

$$E(\xi^0) = -4.63,\quad a(\xi^0) = 5.42,\quad B(\xi^0) = 0.98,\quad C'(\xi^0) = 0.51,\quad C_{44}(\xi^0) = 0.80,\quad \zeta(\xi^0) = 0.52.$$

The parameter values found coincide with their reference values up to the second digit.

These results suggest that, for the given numerical accuracy, some of the parameters are not uniquely determined. Whether or not this accuracy of the identified parameters is acceptable will be determined in a molecular-dynamical simulation, which is intended to perform in the future.

The following approaches can be proposed for improving parameter identification:

1. Additional terms can be introduced into the objective function.
2. The "troublesome" potential parameters can be fixed at physically justified a priori values.
3. The numerical accuracy can be improved substantially by tuning the parameters of the methods used or by applying other computational methods. Specifically, global Lipschitz optimization methods are intended to be considered (see [12, 13]).

CONCLUSIONS

We addressed the problem of identifying the parameters of an interatomic potential, which is a stage in the mathematical simulation of materials. The problem was given in the general formulation and was specified for the Tersoff potential. Numerical experiments for silicon were performed. The results of the experiments were used to compare the effectiveness of the optimization algorithms used and to estimate the possibility of identifying various parameters. Approaches were proposed for overcoming the ambiguities in the identified parameters.

The material properties in the above procedures were calculated at fixed relative coordinates of the atoms (at fixed relative positions of the atoms in an elementary cell). However, this approach does not guarantee that the given atom positions correspond to the minimum of energy for the resulting potential. Accordingly, the next stage was the relaxation of the structure. For this purpose, the forces acting on each of the atoms in an elementary cell (basis atoms) were calculated and the atoms were displaced according to the driving forces so that their positions corresponded to the minimum of energy. The directions of the forces were computed relying on the potential energy gradient.

If the atoms with the potential parameters found are shifted significantly from the required equilibrium positions or the material properties differ substantially from the required ones, a secondary fitting procedure correcting some of the resulting parameters has to be performed so that the coordinates of those atoms whose deviations are beyond the accepted range are added to the set of reference values. This study is planed to be conducted in the future.

ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation, project no. 14-11-00782.

REFERENCES

1. K. K. Abgaryan, "Application of optimization methods for simulation of multilayer semiconductor nanostructures," Tr. Inst. Sist. Anal. Ross. Akad. Nauk Din. Neodnor. Sist. **53** (3), 6–9 (2010).
2. K. Abgaryan, "Multiscale computational model of multilayer semiconductor nanostructures," *Fourth International Conference on Optimization Methods and Applications (OPTIMA-2013)* (Petrovac, Montenegro, 2013), pp. 13–14.
3. P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," Phys. Rev. B **136**, 864–871 (1964).
4. W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," Phys. Rev. A **140**, 1133–1138 (1965).
5. K. Abgaryan, D. Bazanov, and I. Mutigillin, "The simulation of the nitridation of the sapphire surface (001)," *Proceedings of the 21st International Conference on Interaction of Ions with the Surface* VIP **2**, 99–103 (2013).
6. K. K. Abgaryan, N. V. Bykov, and I. V. Mutigullin, "Application of parallel computations and molecular dynamics method for modeling the growth of multilayer semiconductor nanostructures," *Proceedings of the 9th International Conference on Nonequilibrium Processes in Nozzles and Jets (NPNJ 2012), Alushta, May 25–31, 2012* (Mosk. Aviats. Inst., Moscow, 2012), pp. 457–459.
7. J. Tersoff, "Empirical interatomic potential for silicon with improved elastic properties," Phys. Rev. B **38**, 9902–9905 (1988).
8. D. Powell, *Elasticity, Lattice Dynamics, and Parameterization Techniques for the Tersoff Potential Applied to Elemental and Type III–V Semiconductors* (University of Sheffield, distributed, 2006).
9. C. Kittel, *Introduction to Solid State Physics* (Gostekhteorizdat, Moscow, 1957; Wiley, New York, 1976).
10. S. Juzhong, T. Stirner, W. E. Hagston, A. Leyland, and A. Matthews, "A simple transferable interatomic potential model for binary oxides applied to bulk α-AL$_2$O$_3$ and the (0001) α-AL$_2$O$_3$ surface," J. Crystal Growth **290**, 235–240 (2006).
11. M. A. Posypkin, "Solution of global optimization problems with the use of distributed systems," Program. Prod. Sist., No. 1, 23–29 (2010).
12. Yu. G. Evtushenko, "Numerical methods for finding global extrema (case of a nonuniform mesh)," USSR Comput. Math. Math. Phys. **11** (6), 38–54 (1971).
13. Ya. D. Sergeev and D. E. Kvasov, *Diagonal Global Optimization Methods* (Fizmatlit, Moscow, 2008) [in Russian].

Translated by I. Ruzanova