Parameter Optimization of Tersoff Interatomic Potentials Using a Genetic Algorithm*

Yoko SAITO**, Naoya SASAKI**, Hiroshi MORIYA**, Akiko KAGATSUME** and Shingo NORO**

A method that gives the parameters of advanced Tersoff interatomic potentials for describing nonequilibrium atomic structures has been developed. This method uses a genetic algorithm to optimize the Tersoff potential parameters fitted to first-principles-calculated cohesive energies of various carbon systems, including bulk systems with atomic defects and amorphous, surface, or cluster systems under stress. These optimized parameters converge towards a set of Tersoff potential parameters that well describes not only crystals but also amorphous systems.

Key Words: Computational Mechanics, Molecular Dynamics, Numerical Analysis, Tribology, Interatomic Potential, First Principle, Genetic Algorithm, Optimization, Carbon

1. Introduction

Rapid developments in computer processing capability have pushed the development of computer simulation technology for atomic-scale phenomena. In industry, atomic-scale analysis is becoming necessary as semiconductors are scaled down and recording density of hard-disk drives is increased.

To simulate atomic motion precisely, first-principles molecular dynamics calculations are performed. However, first-principles molecular dynamics cannot deal with a large number of atoms and cannot be applied to phenomena ruled by the dynamic motion of atoms in the bulk or the surface of materials. To avoid such problems, classical molecular dynamics, which assumes interatomic potentials and solves each motion of atoms classically, is used. The problem of classical molecular dynamics is that we must assume interatomic potentials; it is almost impossible to find a perfect potentials that can describe all the diverse bonding geometries.

However, some empirical potential functions and their parameters for a specific system have been proposed. One of the most well-known is the Lennard-Jones potential\(^1\). It is a pair potential that is a function of the relative distance between two elements only, and it can precisely describe the configuration of rare-gas atoms. On the other hand, many-body potentials, which include additional terms for many-body interaction, have been proposed and applied to materials such as semiconductors and polymers\(^2\).

To apply classical molecular dynamics to semiconductors or hard-disk-drive media, interatomic potentials of covalent systems such as silicon or carbon films are required. The Tersoff potential is one of the most well-known potentials and provides a realistic description of bond order depending on the local environment (number of neighbors)\(^3\). However, the parameters that define the Tersoff potential have been proposed for only a few systems, e.g., carbon, silicon and germanium\(^4\), and those for other systems are rarely given. This is mainly because the atomic
configurations for fitting parameters have not been finalized and because there are too many parameters
to be determined. Moreover, potentials given by
previous works were fitted to the first-principles calcu-
lated energies around the equilibrium state. To
analyze systems under friction and wear (including
atomic peeling), more precise potentials are required.
Though previous works have reported fitting
methods\(^{69-77}\), there have been no works on a set of
parameters fitted to the first-principles cohesive
energies of crystal, amorphous (including bulk and
surface systems), and cluster systems under stress.

The purpose of this study is to establish a method
to determine optimal parameters of Tersoff inter-
atomic potentials for describing nonequilibrium states
from a large amount of data produced by first-prin-
ciples calculations. We optimize Tersoff potential
parameters for carbon and evaluate new potentials.
In the optimization procedure, we use a genetic algo-
rithm that is effective for optimizing problems in com-
plex systems.

**Nomenclature**

\[ r_{ij} \]: length of the \( ij \) bond

\[ N \]: number of atoms

\[ f_s(r_{ij}) \]: cutoff function between atoms \( i \) and \( j \)

\[ b_{ij} \]: bond-order parameter between atoms \( i \) and \( j \)

\[ \theta_{ij} \]: bond angle between \( ij \) and \( ik \)

\[ R, D \]: cutoff parameters

\[ A, B, \alpha, \beta, \gamma, \delta, \rho, \sigma, \alpha, \beta, n, c, d, h \]: fitting parameters of
Tersoff potential

\[ s \]: set of parameters to be optimized

\[ f(s) \]: fitness of the set of parameter \( s \)

\[ m \]: number of atomic configurations

\[ E_k(s) \]: cohesive energy per atom in configuration \( k \)

\[ E^* \]: reference cohesive energy per atom calculated
from first principles.

**2. Tersoff Potential**

Tersoff interatomic potential is written as follows:

\[ \Phi = \sum_{i=1}^{N} \sum_{j=1}^{N} f_s(r_{ij})[a_{ij} f_b(r_{ij}) + b_{ij} f_a(r_{ij})], \]  \hspace{1cm} (1)

\[ f_s(r) = A \exp(-\lambda r), \]

\[ f_b(r) = -B \exp(-\lambda r), \]  \hspace{1cm} (2)

\[ f_a(r) = \begin{cases} \frac{1}{2} \sin \left( \frac{\pi}{2} r_{ij} - R \right) / D, & 1 < r_{ij} < R - D \\ R - D < r_{ij} < R + D, & 0 \end{cases}, \]  \hspace{1cm} (3)

\[ b_{ij} = (1 + \beta r_{ij}/\delta_{ij})^{-1/2}, \]

\[ \delta_{ij} = \sum_{k=0}^{\infty} f_k(r_{ik}) g(\theta_{ik}) \exp[r_{ij} - r_{ik}], \]  \hspace{1cm} (4)

\[ a_{ij} = \sum_{k=0}^{\infty} f_k(r_{ik}) \exp[r_{ij} - r_{ik}], \]  \hspace{1cm} (5)

Here, \( r_{ij} \) is the length of the \( ij \) bond, \( N \) is the number of atoms, \( f_s(r_{ij}) \) is the cutoff function, and \( b_{ij} \) is the bond-order parameter describing many-body interaction.

And \( \theta_{ij} \) is the bond angle between \( ij \) and \( ik \), \( R \) and \( D \) are the cutoff parameters, and \( A, B, \alpha, \beta, \gamma, \delta, \rho, \sigma, \alpha, \beta, n, c, d, h \) are the other parameters. Moreover, to simplify the form of the potential, we set \( \alpha = 0 \) and \( \delta = \gamma \). By using a genetic algorithm (GA),\(^{69}\) we optimize nine parameters, \( A, B, \lambda, \alpha, \beta, n, c, d, h \). On the other hand, \( R \) and \( D \) are determined from first principles.

**3. Method to Determine Interatomic Potentials**

**3.1 Genetic algorithm**

The GA optimizes a set of parameters to satisfy
an evaluation function and performs a global search
of parameter space. Many sets of parameters are
crossed and reproduced. If we can express the degree
of fitness as an evaluation function, a set of parameters
with different dimensions can be optimized
systematically\(^{69}\). Here, each parameter corresponds
to a genotype and a set of parameters to be optimized
corresponds to an individual. In a simple GA, the best
individual is found by the following procedure:

1. Production of initial population
2. Calculation of the fitness of each individual
3. Determination of surviving probabilities corresponding to the fitness
4. Crossing over of two individuals
5. Mutation of some genotypes randomly

By repeating steps (2) to (5), the entire population
of individuals is evolved. Fitness \( f(s) \) is taken to have the following form:

\[ f(s) = \frac{1}{m} \sum_{k=1}^{m} (E_k^* - E_k(s))^2. \]  \hspace{1cm} (6)

Here, \( s \) is the set of parameters to be optimized, \( m \) is the number of atomic configurations, \( E_k(s) \) is the cohesive energy per atom in configuration \( k \), and \( E_k^* \) is the reference cohesive energy per atom calculated from first principles. Therefore, as \( E_k(s) \) approaches \( E_k^* \), \( f(s) \) decreases. We terminate the cycle of steps (2) to (5) when the individual having the minimum value of \( f(s) \) becomes stationary. If we give the nine parameters to be optimized randomly in steps (1) and (5), it takes too much time to obtain the best set of parameters. We therefore restrict the values within a range including the parameters reported by Tersoff\(^{69}\). The range is chosen as follows. First, by using the Newton–Raphson method, we optimize one parameter to describe the cohesive energies of crystal, amorphous, and cluster systems, respectively. The
optimized parameter varies depending on the atomic system under consideration. Next, in the same way, we optimize other parameters and obtain their dispersions. Parameter $A$ varies by 40%, $B$ by 20%, $\lambda_1$ by 10%, $\lambda_2$ by 10%, $\beta$ by 100%, $n$ by 10%, $c_1$ by 10%, $d^2$ by 10%, and $h$ by 70%. The crossing-over and mutation probabilities are determined by trial and error. In the following calculation, we set the crossing-over probability to 40% and mutation probability to 10%.

3.2 First-principles calculation

Cohesive energies, which are needed for optimization of potential parameters, are calculated from first principles. We use the code based on density functional theory (DFT), CASTEP\textsuperscript{(10)}, made by Molecular Simulation Inc. CASTEP is a plane-wave pseudopotential code appropriate for solids and surfaces. We adopt a generalized gradient approximation (GGA)\textsuperscript{(11)} that takes account of electron density gradient.

3.3 Calculation model

Table 1 lists the ten models used for fitting. Each model includes five to seven atomic configurations under stress. Typical configurations of models 3, 6, 8, and 10 are shown in Fig. 1. The stress range in all models is from $-100$ to $100$ GPa. When atoms are scattered by tensile stress, we use a configuration expanded isotropically just before scattering. High-symmetry polytypes include dimer, graphite, diamond, simple cubic, bcc, and fcc structures. Models 3, 4, and 5, including defects, are configurations equilibrated after 8, 16, and 24 atoms are randomly deleted from eight lattices of diamond. Radial distribution functions of models 3, 5, and 6 are shown in Fig. 2. This figure shows that model 5, including many defects, is similar to amorphous model 6. However, model 3, including a few defects, is similar to model 1 (diamond).

In the calculation, we use periodic boundary conditions for all models. Models 1, 2, and 7, high-symmetry crystals, are calculated as a unit cell. On the other hand, models 3 to 6, and 8, low-symmetry configurations, are calculated as a cell composing 40 to 56 atoms in order to produce a large calculation space.

<table>
<thead>
<tr>
<th>Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Diamond (bulk)</td>
</tr>
<tr>
<td>Model 2</td>
<td>High-symmetry polytypes (bulk)</td>
</tr>
<tr>
<td>Model 3</td>
<td>Type-1 defects</td>
</tr>
<tr>
<td>Model 4</td>
<td>Type-2 defects</td>
</tr>
<tr>
<td>Model 5</td>
<td>Type-3 defects</td>
</tr>
<tr>
<td>Model 6</td>
<td>Amorphous (bulk)</td>
</tr>
<tr>
<td>Model 7</td>
<td>Diamond (surface)</td>
</tr>
<tr>
<td>Model 8</td>
<td>Amorphous (surface)</td>
</tr>
<tr>
<td>Model 9</td>
<td>Cluster with 5 atoms</td>
</tr>
<tr>
<td>Model 10</td>
<td>Cluster with 10 atoms</td>
</tr>
</tbody>
</table>

Table 1 Fitting models

Fig. 1 Calculation models

Fig. 2 Radial distribution function, $g(r)$
4. Optimization of Potential Parameters

4.1 Determination of cutoff radius

The optimized parameters of the potential, $A$, $B$, $\lambda_1$, $\lambda_2$, $\beta$, $n$, $c^2$, $d^2$, and $h$, depend on cutoff parameters $R$ and $D$, which were determined from the model of diamond. Figure 3 shows cohesive energy per atom versus bond length of diamond. The solid diamonds express the data calculated from first principles and solid squares express the data obtained from the parameters reported by Tersoff[14]. The label "exp." in Fig. 3 indicates the measured datum listed in Table 3[12]. The cohesive energy given by Tersoff fits well to that calculated by first-principles calculations when energy reaches a minimum. However, the trough of the line obtained from first-principles calculations is wider than that of the line obtained from the parameters given by Tersoff when bond length is above 2. This difference of energies influences classical molecular dynamics calculations applied to systems describing phenomena such as atomic peeling. In the following, we optimize the parameters of potentials with larger values of cutoff, $R=3.7$ and $D=0.3$, than those given by Tersoff, $R=1.95$ and $D=0.15$. (For reference, we also optimize parameters of potentials with cutoffs given by Tersoff.)

4.2 Parameter optimization for each model

The cohesive energies calculated from first principles and from the parameters given by Tersoff for models 2 to 10 are compared in Fig. 4. The energies for bulk systems are plotted as a function of volume per atom (a) to (e) in Fig. 4. On the other hand, the energies for surface and cluster systems are plotted in order of data number (f) to (i) in Fig. 4. Model 2, of high-symmetry polytypes, is only represented by graphite atomic data. Figures 4(a) and (f), for crystal systems, shows that the energies given by Tersoff fit well to those calculated from first principles. However, Figs. 4(d) and (e), for amorphous systems, shows that the energies given by Tersoff are higher than those calculated from first principles. Similarly, Figs. 4(h) and (i), corresponding to cluster systems, show that the energies given by Tersoff are a little higher than those calculated from first principles.

We individually optimized parameters of potential for each model. In model 2, the six polytypes are used at the same time during optimizing. Three hundred sets of parameters were produced and optimized by the GA. Optimized potentials are illustrated by open circles in Figs. 3 and 4 (open circles hidden by other data are almost equivalent to ones calculated from first principles). The graph shows that optimized potentials for each model are extremely close to the first-principles calculations. On the other hand, the GA does not evaluate each parameter individually but evaluates a set of parameters. Optimized parameters therefore tend to disperse. The dispersions of the parameters are listed in Table 2. They are given as the difference between the maximum and minimum parameters divided by the original Tersoff parameter. The dispersions of parameters $\lambda_1$, $\lambda_2$, $n$, and $d$ are low, but that of $\beta$, $c$, and $h$ are high.

4.3 Parameter optimization for all models

Taking all the first-principles data into account, we must find the best potential that will satisfy the diverse bonding geometries. However, computing power is limited, so we must reject some data used in the GA. Since the final purpose of this work is to analyze friction and wear on a solid surface, cluster models were excluded. Moreover, surface systems were also excluded because the original Tersoff potential describing bulk systems (Fig. 3) can also describe a surface system (Fig. 4 (f)). We, therefore, selected 16 atomic configurations, of which four represent diamond (model 1), two represent dimer (model 2), two represent graphite (model 2), two represent simple cubic (model 2), two represent type-2 defects (model 4), four represent amorphous (model 6) (selected data are indicated by arrows in Figs. 3 and 4). Since many data that fit might result in a decrease in fitness, we increased the number of individuals to 3000.

Each optimized parameter is listed in Table 2 and optimized potential are illustrated in Figs. 3 and 4 by mark X. Compared to the potentials optimized individually, the potentials optimized for all models are a bit different from those calculated from first principles. However, they are in agreement with those obtained by first principles in amorphous systems, especially when atomic volume is large. Fitness of optimized parameters is 0.3, which is more precise than 4.2, the original Tersoff parameters. Furthermore,
Table 2  Tersoff potential parameters

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>( \lambda_1 )</th>
<th>( \lambda_2 )</th>
<th>( \beta )</th>
<th>( n )</th>
<th>c</th>
<th>d</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tersoff</td>
<td>1.3936E3</td>
<td>3.487E2</td>
<td>3.4879</td>
<td>2.2119</td>
<td>1.5724E-7</td>
<td>0.72751</td>
<td>3.8046E4</td>
<td>4.384</td>
<td>-0.57058</td>
</tr>
<tr>
<td>(original)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispersion</td>
<td>39.3</td>
<td>32.7</td>
<td>12.5</td>
<td>14.2</td>
<td>126.9</td>
<td>12.8</td>
<td>118.8</td>
<td>7.7</td>
<td>113.8</td>
</tr>
<tr>
<td>GA</td>
<td>1.1619E3</td>
<td>3.105E2</td>
<td>3.4325</td>
<td>2.1203</td>
<td>9.3716E-9</td>
<td>0.78267</td>
<td>5.2026E4</td>
<td>4.3668</td>
<td>-0.71977</td>
</tr>
<tr>
<td>(R=1.95)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>8.8546E2</td>
<td>3.379E2</td>
<td>3.3637</td>
<td>2.2607</td>
<td>2.3221E-7</td>
<td>0.72335</td>
<td>3.0339E4</td>
<td>4.5570</td>
<td>-0.64932</td>
</tr>
<tr>
<td>(R=3.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4  Cohesive energy for models 2 - 10: model 2 of graphite (bulk); model 3 of type-1 defects (bulk); model 4 of type-2 defects (bulk); model 5 of type-3 defects (bulk); model 6 of amorphous (bulk); model 7 of diamond (surface); model 8 of amorphous (surface); model 9 of a cluster with 5 atoms; model 10 of a cluster with 10 atoms; arrows indicate data used in GA (R=1.95) and GA (R=3.7)

the optimized potential fitted to the energies only for bulk systems can also describe models 7 to 10, surface and cluster systems, well.

Each optimized parameter (when we use \( R=1.95 \) and \( D=0.15 \) given by Tersoff), is listed in Table 2 and optimized potential is illustrated in Figs. 3 and 4 by solid triangles. While the optimized potential is close to that obtained by using \( R=3.7 \), it is not fitted when atomic volume is large. This is because \( R \), cutoff parameter, is small. Fitness of the optimized parameters is 3.1, which is a better fit than that of the original Tersoff parameters.
Table 3 Experimentally measured and calculated values

<table>
<thead>
<tr>
<th></th>
<th>Experimental</th>
<th>First principles</th>
<th>Optimized potential (R=1.95)</th>
<th>Optimized Potential (R=3.7)</th>
<th>Tersoff (original)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohesive energy (a)</td>
<td>1.177(^{(12)})</td>
<td>1.223</td>
<td>1.177</td>
<td>1.128</td>
<td>1.189</td>
</tr>
<tr>
<td>(\times 10^{-19}J/atom)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lattice constant (b)</td>
<td>3.568(^{(12)})</td>
<td>3.556</td>
<td>3.484</td>
<td>3.526</td>
<td>3.568</td>
</tr>
<tr>
<td>(\times 10^{-10}m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulk modulus (c)</td>
<td>517(^{(13)})</td>
<td>526</td>
<td>503</td>
<td>529</td>
<td>547</td>
</tr>
<tr>
<td>(GPa)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defect energy (d)</td>
<td></td>
<td>0.890</td>
<td>0.887</td>
<td>0.871</td>
<td>0.972</td>
</tr>
<tr>
<td>(\times 10^{-19}U/atom)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Cohesive energy of diamond  
(b) Lattice constant of diamond  
(c) Bulk modulus of diamond  
(d) Defect energy for model 5, (includes relaxation effect)

5. Discussion

5.1 Optimized parameters

Table 2 shows that parameters \( A \) and \( h \), which describe repulsive force and the cosine of stable bond angle, respectively, tend to be small. Since cohesive energies calculated from the original parameters given by Tersoff are high in amorphous systems, especially when atomic volume is large, parameter \( A \) becomes a little lower. A small value of \( h \) implies that a stable bond angle between three atoms is bigger than that given by Tersoff. This is because we used many atomic configurations, including amorphous systems, for fitting. In these systems, coordination number is lower than that of crystal systems. Since we set parameter \( a \) to zero, \( \eta_0 \) in Eq. (5) is arbitrary and the coefficient of the repulsive term in Eq. (1), \( a_o \), is one. If we optimize \( a \) and \( \eta_0 \) in Eq. (5), the potential can be obtained more precisely.

If we consider the long-range attraction force in the calculation of dynamic motion of atoms, we should take \( R=3.7 \) and \( D=0.3 \) as cutoff parameters. However, in case of a many-body potential, if the cutoff radius is expanded \( x \) times, computation time for calculating molecular dynamics takes more than three powers of \( x \). Therefore, we have to take into account some short-range cutoff parameters. As mentioned before, potential obtained by short-range cutoff parameters can accurately describe amorphous systems around a stable point. However, note that cutoff parameters should be evaluated before fitting because optimized potential strongly depends on them.

5.2 Optimization method

It was found that the GA is well suited to handle complex systems. However, it is under development and there is room for evaluating the probability of crossing-over and mutation, which were adjusted by trial and error. By changing the probabilities, it is possible to obtain a better-fitted potential. On the other hand, the GA uses an evaluation function defined by a set of parameters, so each optimized parameter is not assured to take the best value. To complement the optimization method, we should control the dispersions of parameters to be determined by decreasing the number of them and keeping parameters \( \lambda_1, \lambda_2, n, \) and \( d \) fixed.

Since the structure of the amorphous systems in this study is just one example, we can expect to obtain better potentials by using some configurations with different radial distribution functions. Moreover, the energies of stable structures of diamond and graphite are slightly less accurate because we also used unstable structures for fitting. We thus must improve the accuracy of the calculated potential at the minimum point by using a weight function.

5.3 Evaluation of the optimized potentials

Cohesive energy, lattice constant, bulk modulus, and defect energy are compared in Table 3. Cohesive energy, lattice constant, and bulk modulus are close to those values obtained from experiment. Defect energy obtained from optimized parameters is closer to that calculated from first principles than that of original. In future, better optimized potentials for precisely describing an unstable structure will be developed and we must then develop a way to evaluate these new potentials.

6. Summary

Tersoff interatomic potentials were optimized by using a genetic algorithm (GA) to fit parameters to first-principles-calculated cohesive energies of various atomic systems, including nonequilibrium and amorphous structures. The GA is useful for optimizing the potential parameters by taking account of diverse
configurations. Optimized parameter $A$, which describes repulsive force, and parameter $h$, which describes the cosine of the stable bond angle, tend to be lower than those given by Tersoff. The parameters optimized by the GA, especially $\beta$, $c$, and $h$, strongly depend on the fitting models. Moreover, optimized potential only fitted to that for bulk systems can also describe surface and cluster systems.

Acknowledgments

This study was performed by the Science and Technology Agency of the Japanese Government through the Special Coordination Fund for Promoting Science and Technology.

References


