Parametrization of the Reactive MD Force Field for Zn-O-H Systems

K. S. Shefov[a], M. M. Stepanova[b], S. N. Kotriakhova[c]

St. Petersburg State Institute University, 7-9 University emb, St. Petersburg, 199034, Russia

E-mail: [a] k.s.shefov@gmail.com, [b] mstep@mms.nw.ru, [c] dangemy@gmail.com

We describe a procedure of optimizing the molecular dynamic force field for Zn-O-H chemical systems by means of a new parallel algorithm of a multifactorial search for the global minimum. This algorithm allows one to obtain numerous parameters of the ReaxFF classical force field based on quantum chemical computations of various characteristics of simple compounds. The force field may be then used for simulating of large-scale chemical systems consisting of the same elements by means of classical molecular dynamics. Our current implementation of the algorithm is done in C++ using MPI. We compare characteristics of simple compounds, obtained by 1) quantum chemical techniques, 2) molecular-dynamic methods using reference parameters of the force field, and 3) MD methods using optimized parameters of the force field. With the optimized parameter set we perform MD simulations (using LAMMPS package) of crystals of zinc and zinc oxide of various modifications at the room temperature. Finally, we compare results of the parameter optimization procedure by means of the algorithm described above and results of a parallel implementation of an evolutionary approach to minimum search using dynamic models of Zn and ZnO crystals. Also we discuss advantages and disadvantages of the both methods and their efficiency for extremal problems. All computations are performed with machines of the distributed scientific complex of the Faculty of Physics of St Petersburg State University.

Keywords: numerical simulation, chemically reactive systems, reactive force field, molecular dynamics, parameter optimization, parallel algorithm, absolute extremum search

Introduction

Reactive molecular-dynamic (MD) force field ReaxFF is a function of a large number of parameters. Searching optimal values of these parameters is a complicated problem. To solve it one may use different approaches, each of them having its own advantages and disadvantages. One of the approaches is the multifactorial global search algorithm (MGSA). The aim of this work is to estimate efficiency of MGSA technique being applied to parameter search. We compare quality of parameter sets for Zn-O-H systems obtained by different methods: MGSA, evolutionary algorithm and one-parameter search technique [Raymand, Duin, …, 2010]. Quality of a parameter set is estimated 1) by comparing of physical characteristics obtained by MD-methods and quantum chemistry methods, 2) by stability of crystal lattices of Zn and ZnO during their simulation under room temperature, and 3) by simulation of ZnO crystal growth. For MD-simulations we use package LAMMPS.

Global Search Algorithm

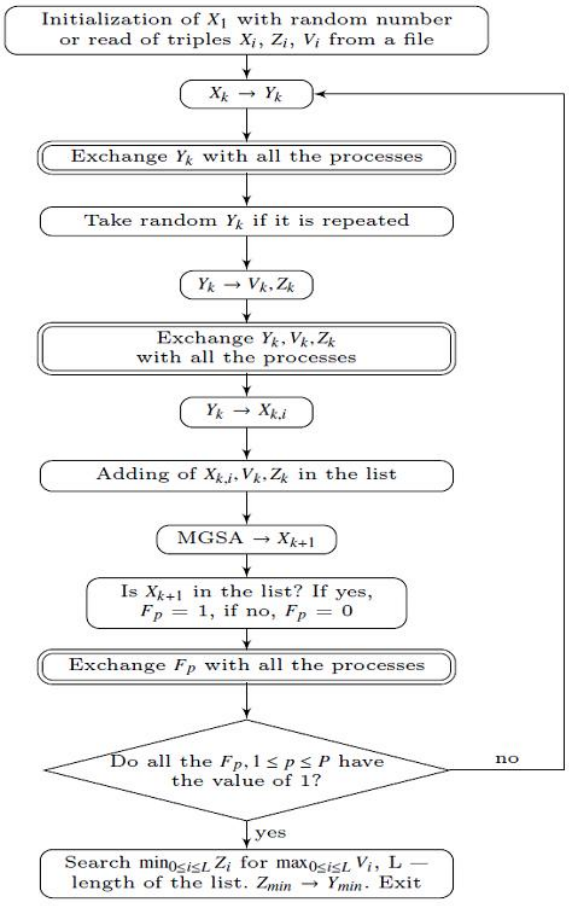


Fig. 1. MGSA parallel implementation flowchart

The global search algorithm (GSA) [Strongin, 1978] allows one to obtain an absolute minimum of a target function on a segment. It is based on probability approach. Assume we already have a set of known target function values in several points on a segment. We use these known values to find an interval between neighboring points on the segment, where absolute minimum location is the most probable. In this interval we take a point corresponding to a mathematical expectation of position of the minimum and compute target function value in it. The point is added to the list of known values and we pass to the next iteration. The algorithm terminates, when distance between points of the segment of two consequent iterations becomes less than a given criteria. In multidimensional variant of the algorithm a function of many variables is mapped on a function of a single variable with a help of Peano scans. Multidimensional definition domain (hypercube) is mapped on a segment of the real axis. To make the global search algorithm parallel we use a technique of rotating scans [Strongin, Gergel, …, 2009]. Each parallel process operates with its own scan rotated by angles $\pm\pi/2$ with respect to the basic scan in particular pair of dimensions. In total one can perform $N\cdot(N - 1)$ of such rotations for N-dimensional domain of definition. Thus, in total the program uses $N\cdot(N - 1) + 1$ processes, each of them performs GSA and at each iteration sends its result to all other processes. Parallel algorithm speeds up convergence and compensates points proximity information loss caused by use of a scan.

As a target function we use function (1).

$$Error = \sum_{k=1}^{L} \sigma_k \left| U_k^{\mathrm{QC}} - U_k^{\mathrm{ReaxFF}} \right| + \sum_{k=1}^{L} \sigma_{L+k} \sqrt{\sum_{\alpha=1}^{A_k} \sum_{i=1}^{3} \left(F_{k\alpha i}^{\mathrm{QC}} - F_{k\alpha i}^{\mathrm{ReaxFF}} \right)^2} \qquad (1)$$

Here U_k are potential energies of training set models (simple chemical compounds, which optimization is based on), $F_{k\alpha i}$ are components of forces acting on atom α of model k, L is number of models in the set, A_k are numbers of atoms in model k, and σ_k are weight factors. Indices QC and ReaxFF mean that corresponding characteristics are obtained either by quantum chemistry methods of by molecular dynamics, respectively. U_k^{ReaxFF} and $F_{k\alpha i}^{\mathrm{ReaxFF}}$ depend on parameters $p_1, p_2, \ldots, p_N$ of the force field. In our case we use multifactorial algorithm, in which in addition to the basic target function (1) we also use several functions-limitations on particular groups of addends in (1). This allows us to set all the weights σ_k equal to 1 and not to solve a complicated problem of weights choice.

MGSA parallel implementation flowchart is in Fig. 1. Here points of a function definition domain (a grid on a hypercube) are named Y_i, and values of a target function (or functions-limitations) in them are named Z_i. Values of index (number of satisfied limitations in Y_i) are named V_i. Points of a grid on the segment [0; 1], on which the hypercube is mapped, are named X_i. Double frames in the flowchart show that on a particular step data exchange between parallel processes takes place. Number of parallel processes is named P. For more details of the algorithm see paper [Stepanova, Shefov, …, 2016].

Quantum chemistry computations

Basic compounds that are included in the training set are presented in Table 1. QC computations are performed with the same methods that are used in the work [Raymand, Duin, …, 2010] that is used for comparison.

Table 1. Compounds computed by quantum chemistry methods

Compound	Modifications	Comp. method
ZnO lattices	P63mc, F-43m, Fm-3m, Pm-3m	CRYSTAL09, DFT/B3LYP Gauss basis, k-points 10x10x10
Zn lattices	hcp, fcc, bcc, sc	ABINIT, DFT/PAW/PBE planewave basis k-points 8x8x8, E_{cutoff}=30 Ha
ZnO thin films (10 layers)	(1 0 -1 0) P63mc, (1 1 -2 0) P63mc, (1 1 0) Fm-3m, (1 0 0) Fm-3m	CRYSTAL09, DFT/B3LYP Gauss basis, k-points 10x10
Molecule ZnOH2	Different bond lengths Zn-OH, different angles OH-Zn-OH and Zn-O-H	GAUSSIAN09 DFT/B3LYP basis 6-311+G
Molecule OHZnOZnOH	Different angles OHZn-O-ZnOH	GAUSSIAN09, DFT/B3LYP basis 6-311+G
8 ZnO cells in vacuum	P63mc, F-43m, Fm-3m	GAUSSIAN09, DFT/B3LYP Gauss basis

In Table 2 and Table 3 there are shown results of computations of ZnO and Zn crystal lattice constants and bulk moduli (BM) compared to computations from the work [Raymand, Duin, …, 2010] and physical experiment data. Sign "+" marks those crystals that are observed in nature.

Parameters optimization

Parameters of the ReaxFF force filed are optimized with the multifactorial global search algorithm [Stepanova, Shefov, …, 2016]. We used function (1) as the target one. The optimization is done under condition that a sum of every group of addends in (1) that corresponds to each particular molecule or crystal does not exceed 1/8 of the estimate of the mean value of this sum. The mean value is

estimated by computing every sum value in 65000 points uniformly distributed over the search domain.

Table 2. Lattice constants and bulk moduli (BM) of zinc oxide crystals

Cryst.	Charact.	QC comp.	QC [Raymand, Duin, …, 2010]	ReaxFF opt.	ReaxFF [Raymand, Duin, …, 2010]	Exper.
P63mc+	a,A	3.28	3.28	3.27	3.29	3.25
P63mc+	c, A	5.28	5.28	5.24	5.28	5.21
P63mc+	BM, GPa	143	136	138	144	141, 143
F-43m	a, A	4.62	4.60	4.62	4.62	N/A
F-43m	BM, GPa	144	162	133	130	N/A
Fm-3m+	a, A	4.34	4.30	4.35	4.29	4.27
Fm-3m+	BM, GPa	183	202	242	283	203, 228
Pm-3m	a, A	2.68	2.68	2.77	2.61	N/A
Pm-3m	BM, GPa	178	183	211	407	N/A

Table 3. Lattice constants and bulk moduli (BM) of zinc metal crystals

Cryst.	Charact.	QC comp.	QC [Raymand, Duin, …, 2010]	ReaxFF opt.	ReaxFF [Raymand, Duin, …, 2010]	Exper.
hcp+	a, A	2.65	2.63	2.74	2.73	2.67
hcp+	c, A	4.98	5.06	4.47	4.46	4.95
hcp+	BM, GPa	74.1	66.7	81.3	87.7	64.5 - 75.1
fcc	a, A	3.90	3.86	3.86	3.86	N/A
fcc	BM, GPa	70.7	81.8	101.2	112.4	N/A
bcc	a, A	3.12	3.06	3.08	3.06	N/A
bcc	BM, GPa	67.8	84.6	73.8	73.5	N/A
sc	a, A	2.64	2.71	2.76	2.71	N/A
sc	BM, GPa	47.4	64.2	34.5	30.2	N/A

The full set of parameters has been preliminarily sorted by descending of their influence on the target function [Stepanova, Shefov, …, 2016]. From that list first 24 parameters are optimized. They are chosen by the largest correlation with the target function addends. The search for minimum is performed on a grid of 4097 points per each parameter. Four parameters are optimized simultaneously. The algorithm has performed 6 passes one time for each 4 parameters. Target function value was reduced 1.5 times during six passes. Each run takes from 36 to 48 hours to finish. Target function value is computed in 1 million grid points in average during each run.

Fig. 2 shows relative values of all 24 parameters that take part in the optimization. The values are measured in per cent of control values [Raymand, Duin, …, 2010] ("Control") taken for 100 %. "MGSA" - parameters optimized by multifactorial algorithm, "EA" - ones optimized by evolutionary algorithm. One can see that the values of parameters differ from method to method, and in some cases they differ relatively a lot. Each method has its own target function, and it results in its own minimum. Below we check how well each of these optimized parameter sets can simulate a real physical system.

Dynamic simulation of crystals

For three sets of parameters (Control, MGSA, EA) we performed simulation of crystals (600 atoms) of zinc and zinc oxide in vacuum at temperature of 300 K. The simulations were performed for all crystal lattices that appear in Tab. 2 and Tab. 3. Those modifications of crystals that occur in nature are adequately simulated with both the parameter set obtained by MGSA and the set obtained by the evolutionary algorithm. This means their lattices retain their density and structure. The rest of crystal modifications either partly deform or completely lose their form. In Tab. 4 as an example we show

snapshots of ZnO (P63mc) crystal in the end of simulation for three different parameter sets. One can see stable enough crystal structure in all the cases.

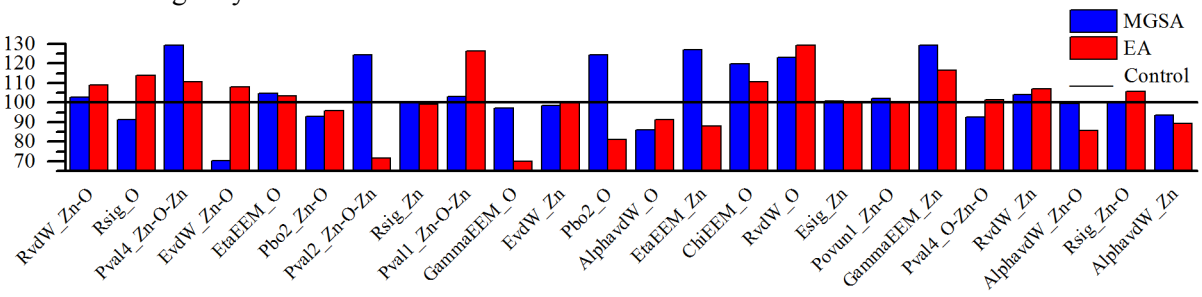
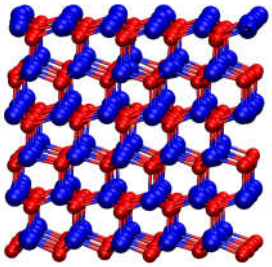
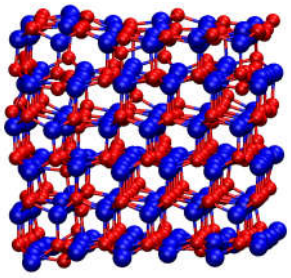
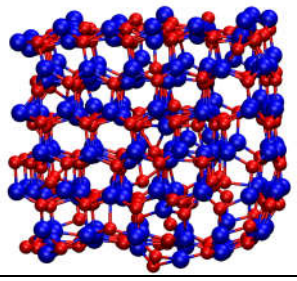


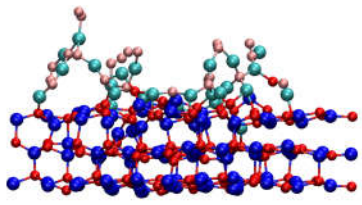
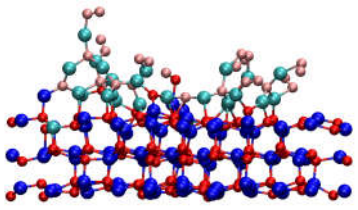
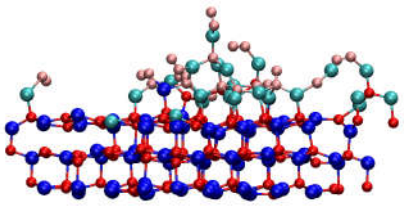
Fig. 2. Comparison of ReaxFF parameters, optimized by different techniques

Table 4. Simulation of ZnO (P63mc) crystal at 300 K

Crystal	Control params	MGSA params	EA params
ZnO (P63mc)			

Using parameters obtained by MGSA we simulate growth of ZnO crystal layer on a (0 0 1) slab of wurtzite ZnO. In Tab. 5 there are illustrated system states with the simulation time of 10, 100 and 300 ps. When the time is 300 ps we observe formation of the first layer of crystal and partly of the second layer. Quality of the particular simulation is not worse than the one in paper [Raymand, Duin, …, 2010].

Table 5. Simulation of ZnO (P63mc) crystal growth

		
t=10 ps	t=100 ps	t=300ps

Conclusion

Both MGSA and the evolutionary algorithm allow one to obtain parameters that give adequate results when simulating particular Zn-O-H systems. Absolute values of parameters differ which is caused by use of different target functions. MGSA converges relatively slow but achieves absolute minimum with a grid precision. Evolutionary approach allow one to quickly localize those domains where minima take place but searching of precise minimum location may turn out to be a very compli-

cated task depending on the form of multidimensional surface. Experience shows that in the given particular problem an absolute minimum of a target function does not always give the best solution since the training set is inevitably limited. For this problem the most effective step would be introduction of additional search criteria. This is implemented for MGSA [Stepanova, Shefov, …, 2016] but is also reasonable for the evolutionary algorithm. Both approaches discussed here are applicable for ReaxFF parameters search for various chemical systems.

References

Raymand D., van Duin A. C. T, Spangberg D., Goddard W. A. III, Hermansson K. Water adsorption on stepped ZnO surfaces from MD simulation // Surface Science. — 2010. — Vol. 604. — P. 741–752.

Стронгин Р.Г. Численные методы в многоэкстремальных задачах. — М.: «Наука». 1978.

Strongin R.G. Chislennye metody v mnogoekstremalnykh zadachakh [Numerical Methods in Multi-Extremal Problems]. — M.: «Nauka». 1978. — 240 p. (in Russian).

Стронгин Р.Г., Гергель В.П., Баркалов К.А. Паралельные методы решения задач глобальной оптимизации // Известия ВУЗов. Приборостроение. — 2009. — Т. 52. — С. 25-33.

Strongin R.G., Gergel V.P., Barkalov K.A. Parallelnye metody resheniya zadach globalnoy optimizatsii [Parallel methods of global optimization problems solution] // Izvestiya Vuzov. Priborostroenie. — 2009. — Vol. 52. — P. 25–33 (in Russian).

Stepanova M.M., Shefov K.S., Slavyanov S.Yu. Multifactorial global search algorithm in the problem of optimizing a reactive force field // Theoretical and Mathematical Physics. — 2016. — Vol. 187, Issue 1. — P. 603–617.

Шефов К.С., Степанова М.М. Реализация и применение параллельного алгоритма глобального поиска минимума к задаче оптимизации параметров молекулярно-динамического потенциала ReaxFF // Компьютерные исследования и моделирование. — 2015. — Т. 7, № 3. — С. 745–752.

Shefov K.S., Stepanova M.M. Realizatsiya i primenenie parallelnogo algoritma globalnogo poiska minimuma k zadache optimizatsii parametrov molekulyarno-dinamicheskogo potentsiala ReaxFF [An implementation of a parallel global minimum search algorithm with an application to the ReaxFF molecular dynamic force field parameters optimization] // Kompyuternye issledovaniya i modelirovaniye. — 2015. — Vol. 7, No 3. — P. 745–752 (in Russian).