Stochastic parallel algorithm of structural evolution of large Morse clusters

A N Kovartsev1, D A Popova-Kovartseva1

1Samara National Research University, Moskovskoe Shosse, 34A, Samara, Russia, 443086

Abstract. The formation of large Morse atomic clusters is a complex and time-consuming task, especially when the cluster stiffness equals 14. Currently, optimal cluster conformations are known for sizes up to 240 atoms. The article proposes a parallel stochastic algorithm for the structural evolution of Morse clusters, which makes it possible to find the optimal conformations of large atomic clusters. A significant number of problems associated with the complexity of the task are resolved using the proposed algorithm for optimizing the cluster structure. The results of optimization of Morse clusters for \( N > 240 \) are given.

1. Introduction
The Morse cluster is a quite precise mathematical model of a real cluster, usually understood as a pool of elements of the same type (atoms/ions/molecules) each connected with all others. The cluster model is by far not the only one. For instance, the Lennard-Jones potential is often used in literature. However, unlike the Lennard-Jones cluster, the Morse cluster has a more complex potential and a more diverse set of structural conformations. Let us cite the Lennard-Jones potential (1) and the Morse potential (2) for comparison [1]:

\[
v(r) = LJ(r) = \varepsilon \left( \frac{r_0}{r} \right)^6 \left( \left( \frac{r_0}{r} \right)^6 - 2 \right), \tag{1}
\]

\[
v(r_j) = M(r, \rho) = e^{(1-r)} \left( e^{\rho(1-r_j)} - 1 \right), \tag{2}
\]

where \( \rho \) is the rigidity ratio of the cluster (in this article, only “rigid” clusters with the rigidity ratio of \( \rho = 14 \) are reviewed); \( r_j \) is the distance between atoms; and \( \varepsilon = r_0 = 1 \).

The interaction energy of all atoms in the cluster can be calculated as the sum of pairwise interactions’ energies:

\[
v(X) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} v(r_{ij}), \tag{3}
\]

where \( X = (x_1, ..., x_N)' \), \( x_i \in R^3 \) are the coordinates of the atoms’ centers.

In this case, the optimal cluster is a conformation of atoms that ensures the global minimum of its potential energy (3). For the general case, A. Nemirovsky and D. Yudin proved in 1979 [2] that the global optimization (GO) problem belongs to the class of \( NP \)-complete problems. Considering high
dimensionality of atom conformation optimization problems, the GO problems are difficult to solve. This is why the problem of minimization of potential energy of the Morse cluster has been often used for development of global optimization methods.

It is worth noting that the above cluster models are so popular due to the global optimization problems aimed at forecasting the structure of biological molecules, protein folding and docking. Such problems kicked off research aimed at forecasting atomic cluster structures as a test ground for forecasting the structure of biological molecules. Both are often viewed within the framework of the conformation analysis performed to find optimal conformations of molecules and research their physical and chemical properties.

From the thermodynamic point of view, protein folding is a transition of a protein molecule to the most statistically probable conformation of the lowest potential energy. One of the features of the folding kinetics is the Levinthal’s paradox [3]. According to it, if a small protein molecule (about 100 of amino-acid residues) sequentially sampled all possible conformations, it would require a time longer than the age of the universe to arrive at its native form. Yet it is empirically known that the normal time scale for protein folding is milliseconds. This is explained by the fact the protein molecule does not sample all theoretically possible conformations; instead, the molecule finds the shortest way possible over the hyperplane of potential energy to the point corresponding to the native conformation of the protein.

Research of clusters is of utmost importance for various areas of human life. The data obtained during such research is applied in the medicine (including in development of new pharmaceuticals), in metallurgy (for metal modelling) as well as in development of new materials. The above data are also essential to understand processes of catalysis, condensation of water vapors in clouds, calculation of electronic and dynamic characteristics of nanomaterials, creation of new sources of light, etc. A key problem in this area of research is finding the geometrical structure of an atomic cluster (conformation) that corresponds to the minimum of the interaction energy of the atoms in that cluster [4,5] (i.e., the minimum amount of potential energy in the cluster).

It is important that the “curse of dimensionality” of global optimization of a potential function defined by many variables poses the main problem in formation of optimal conformations of atomic clusters. Theoretical research of GO methods allow concluding that no feasible direct methods of GO exist for functions with more than 30 variables [6]. This means that the existing practice of finding optimal conformations is based upon using stochastic algorithms of GO.

The success of their application depends on using a large number of heuristics on the properties of the subject of research. Currently, globally optimal conformations for Morse clusters are built for clusters formed by up to 240 atoms [7]. However, building even larger Morse atomic clusters (N>240) still poses a great problem due to high computational loads.

The article reviews problems of formation of globally optimal conformations of large (N>240) Morse atomic clusters. The proposed approach is based upon finding rational heuristics for forming initial approximations of Morse atomic clusters and using local methods of optimization of functions with a large number of variables. In fact, this article is a continuation of article [8] that dealt with the rules of forming initial approximations of conformations of Morse atomic clusters based upon a geometrically sound method.

2. Task formulation

Nowadays, the task of finding conformations that would ensure the global minimum of potential energy of a cluster is solved by applying optimization methods that use additional information on cluster properties. An example of such additional information is the fact that atomic lattices in the Morse cluster conformation gravitate toward a symmetrical spherical shape [9]. This, in its turn, has caused development of geometrically sound GO methods for Morse clusters [4].

It is worth noting that the idea of taking cluster conformation geometry into account for minimization of its potential energy has often been viewed with distrust; yet researchers have returned to it time and again [10,11,12]. This study only reviews the problem of finding optimal configurations of the Morse cluster in its most complex description, when $\rho =14$, and atoms can only touch each
other. In this case, the problem of finding optimal configurations (i.e., conformations with the lowest potential energy) of Morse atomic clusters is directly related to the problem of close packing of spheres [13]. By default, it is assumed that the most critical factors for clusters are:

- **sphericity of conformation forms** (the closer to the sphere, the better);
- **achievability of close packing of atoms.**

For rigid spheres, the closest packing is achieved in a face-centered cubic lattice. However, as many experiments prove, the most popular conformation is a decahedral lattice (figure 1). In this case, the spheres populate slightly more than 74% of the space [14]. At the same time, for “soft” atoms of the Morse cluster, an even closer packing may be achieved using icosahedral and dodecahedral meshes [11, 13, 15, 16].

Finding the optimal cluster conformation means finding such an allocation of cluster atoms that would ensure optimal spatial allocation of all elements, considering the interactions of each single atom with all others and ignoring the chemical properties of the cluster. Therefore, to find the optimal conformation of a cluster consisting of 30 atoms, one has to solve a GO problem with 90 variables, which, as was mentioned above, is simply impossible with modern “direct” GO algorithms.

An analysis [4] of globally optimal Morse cluster conformations for $p = 14$, registered in the Cambridge database [7], shows that it is always possible to find such a position for one of the coordinate axes (e.g., axis $z$) at which cluster atoms will be positioned in layers, forming an indistinct layered structure (figure 1). An indistinct layer is understood here to be an arrangement of a group of atoms (spheres) whereby the atom centers lie not on the plane but in a sufficiently “thin” layer—$z_j \pm \Delta z_j$. At the same time, it appears that cluster conformations of the next generation with a larger number of atoms “inherit” the geometry of their predecessors: from “parent” to “child” clusters.

![Figure 1. Layered formation of a cluster structure.](image)

However, at certain moments, the geometrical structure of globally optimal conformations undergoes qualitative changes. From figure 2, it is evident that for $N=201$, the offset icosahedral structure becomes optimal. In [7], such a structure is referred to as “close-packed.” Yet in this case inheritance of parent properties from preceding clusters is also evident. At the same time, cluster conformations for $N=192, 196, 200$ belong to the type of icosahedral structures.

An analysis of registered globally optimal conformations of Morse atomic clusters [7] allowed discovering that the selection of optimal location of cluster atoms originates from a limited area of close allocation of atoms of the icosahedral lattice. A part of atoms that form the cluster core is contained in all globally optimal clusters of the corresponding structural type [8].

### 3. The algorithm of structural evolution of Morse clusters

The idea of the algorithm proposed in this article is quite simple and based upon using close-packed lattices—a priori known locations of atoms of the cluster to be formed. Figure 3 shows an asymmetrical “inclined” icosahedral lattice that contains 14 layers and may be used for finding optimal conformations of close-packed clusters up to 270 atoms in size. Obviously, optimal
conformation of Morse clusters will be closely fitted inside the used lattice as long as there is a vacant space in it. We only have to “correctly” select the necessary atoms forming the configuration of minimal potential energy. This heuristic allows reducing the problem of parameter global optimization to a much simpler problem of discrete structural optimization.

Figure 2. Layers of globally optimal Morse clusters $\rho = 14$.

The proposed Monte-Carlo method of finding the globally optimal cluster conformation is based upon the algorithm of its dynamic expansion. The main principle of this algorithm is a rationally organized expansion of the cluster core with atoms up to the required size. The cluster core may be formed using the statistics observed for smaller optimal clusters.

Figure 3. An “inclined” close-packed lattice.

The cluster dynamic expansion algorithm consists of the following steps:
**Step 1.** An adjacency matrix of atoms rationally allocated in nodes of the searched lattice is formed. Two atoms are considered to be connected with a strong potential bond if the distance between the atoms is \( p < 1.08 \). The bond graph is described in the system typologically, as description of bonds of all graph nodes. For this purpose, the following structure array is used:

\[
Gr = \begin{cases}
  \text{from} & \text{number of the current node}; \\
  \text{to} & \text{list of adjacent nodes}; \\
  \text{deg} & \text{node degree}.
\end{cases}
\]

In fact, the description of every position of the lattice includes a reference to all adjacent nodes of the lattice.

**Step 2.** Study [5] shows that large clusters inherit typical configurations of interior layers of smaller clusters. Let us call the inheritable part of the atoms the cluster core. The cluster core conditionally folds into one node 0. To build a cluster of the required size, we need to know the bonds of the cluster core (node 0) with free (vacant) nodes of the searched lattice. Next, we form the adjacency matrix of vacant nodes of \( IG_{0} \) graph and the list of nodes of the cluster core \( Top0 \). Then, the cluster core potential is calculated.

**Step 3.** The main idea of the proposed algorithm may be reduced to dynamic stepwise expansion of a cluster by connecting the most promising lattice nodes (cluster atoms) from the number of vacant nodes.

**Step 3.1.** Let us call the list of cluster core nodes as \( Top0 \) and the list of nodes connected at this step of the algorithm as \( Top1 \). The rest are considered vacant. Then the set of current cluster nodes is \( Top11 = Top0 \cup Top1 \), and the incomplete cluster looks like \( XS = \bigcup_{i=Top1} X_{i} \).

Its potential energy may be calculated using the \( v(11) \) formula.

**Step 3.2.** Next, we calculate the possible contribution to the cluster energy of all vacant lattice nodes \( XS \), if this \( X_{j} \not\in Top1 \) is connected to the cluster. Using this calculation, we may build a matrix of effectiveness of vacant nodes

\[
TopG = \begin{pmatrix}
  T_{1} & T_{2} & \cdots & T_{n_{1}} \\
  v(X_{1}) & v(X_{2}) & \cdots & v(X_{n_{1}}) \\
  v(XS \cup X_{1}) & v(XS \cup X_{2}) & \cdots & v(XS \cup X_{n_{1}})
\end{pmatrix},
\]

where \( T_{j} \) is the vacant node number; \( v(X_{j}) \) is the potential contribution of this vacant node to the total energy of the cluster; \( v(XS \cup X_{j}) \) is the total energy of the cluster with the node \( T_{j} \) included.

**Step 3.3.** The \( TopG \) matrix is sorted by the second indicator, the result of which is a new list of vacant nodes ordered by the criterion of contribution of potential energy.

\[
TopS = \begin{pmatrix}
  \tilde{T}_{1} & \tilde{T}_{2} & \cdots & \tilde{T}_{n_{1}} \\
  \tilde{v}(X_{1}) & \tilde{v}(X_{2}) & \cdots & \tilde{v}(X_{n_{1}}) \\
  \tilde{v}(XS \cup X_{1}) & \tilde{v}(XS \cup X_{2}) & \cdots & \tilde{v}(XS \cup X_{n_{1}})
\end{pmatrix}
\]

The \( TopS \) matrix is then reduced to \( m \) of its members satisfying the following condition:

\[
|\tilde{v}(X_{j}) - \tilde{v}(X_{i})| \leq 0.02
\]

As a result, we have the \( TopMax \) matrix that consists of the best candidates for adding to the cluster.

**Step 3.4.** Next, two options are possible.

1. If \( \|TopMax\| = 1 \), then the first node of the \( TopS \) list is selected.
2. If \( \|TopMax\| = n_{i} > 1 \), then we choose one of equiprobable nodes of the \( TopMax \) set.

Then, the “winning” node \( TP_{j} \) is added to the cluster and deleted from the set of vacant nodes.
Step 3.5. This iterative process continues until the cluster grows to the required size.

**Step 4.** The process of dynamic stepwise cluster expansion (steps 1–3) is repeated many times. The number of attempts is preset. The starting situation is equal for all attempts. We remember the conformation that sets the record for the criterion (2).

4. **The parallel algorithm of structural evolution of Morse clusters**

Traditionally, multi-processor computers and parallel algorithms have been used for conformation analysis of molecular clusters. The use of parallel algorithms is due to application of global optimization methods in this area. Direct GO methods [6] that are able to find the solution with the preset precision require about $2.4 \cdot 10^{12}$ objective function calls, which is simply not feasible even for high-performance computers. This is why heuristic optimization algorithms, such as the local-stochastic “basin-hopping” (BH) method and many its modifications are primarily used as GO algorithms [15].

The BH method uses original parallel algorithms related to organization of random perturbations of the atomic cluster structure. The parallel algorithm itself often uses the “manager-worker” scheme. This approach, however, only works with clusters up to $N=80$, because clusters of $N>80$ require significantly more random perturbations.

Recently, there has been an increase in popularity of various population algorithms [10]: genetic algorithms, the beehive method, the ant colony method etc. However, they too demand significant time to execute. To speed up the GO more significantly, additional knowledge of chemical bonds of atoms or molecules of the cluster to be optimized may be useful.

In our case, exceptionally high performance of the proposed algorithm of Morse cluster structural evolution was achieved by means of the developed sequential dynamic expansion algorithm (SDEA) for the cluster. The algorithm is based upon using a large number of useful heuristics described in [4, 5, 8]. As a result, creation of one cluster isomer conformation using the stochastic SDEA algorithm takes only about 0.6 seconds on a normal computer. All SDEA calculations are performed independently from one another, and one of them potentially finds the optimal cluster configuration. For this reason, the calculations may be grouped in units and allocated to processors of a high-performance computer. Obviously, the acceleration of calculations is directly proportional to the number of processors.

The basis of the proposed stochastic parallel algorithm is the idea of stochastic decimation of the cluster core. One of the found optimal clusters of a smaller size is selected as the core (originally, the core is selected from the clusters of the Cambridge database [7]). Then a certain number of atoms is randomly registered in the core. They are used as basis for finding lacking atoms with the SDEA dynamic expansion algorithm. This approach expands the variability of direct search of cluster conformations by the SDEA algorithm and increases the possibility to find the globally optimal conformation.

![Figure 4. The parallel algorithm of cluster structural evolution.](image)

Computational experiments performed on globally optimal clusters known from the DB [7] show that the possibility $p$ to find the globally optimal conformation varies from 0.1 to 0.001. If we
generously assume that $p=0.0001$, then, with the level of confidence probability of $d=0.98$, we will need about 40,000 tests to reach the globally optimal conformation.

The proposed parallel algorithm (see the model on figure 4) was implemented in a program using a graphical-symbolical programming process [17]. The MPI process was used as a communication medium. The program was compiled and run on the Sergei Korolev high-performance computational cluster of the Samara University.

Identical program modules are run on parallel processors $p_1$, $p_2$, …$p_n$. All modules use the same input data: the cluster core, the search lattice. The parallel modules have no common data. Each of the processors searches for its own record-setting solutions, out of which the globally optimal conformation is chosen. Figure 5 shows accelerations of a parallel program consisting of four processes. These accelerations were reached during a search of optimal clusters of $N= 234, 235, 236, 237, 238$.

![Figure 5](image)

**Figure 5.** Acceleration of the parallel algorithm (S) achieved by optimizing clusters of sizes N.

| Number of atoms in the cluster $N$ | Atom interaction energy $\nu$ (ρ = 14) | |
|-----------------------------------|---------------------------------------|---|
|                                  | Close-packed structure | Icosahedral structure |
| 239                              | -1144.0616                  | -1144.5482               |
| 241                              | -1155.08652                 | -1154.113               |
| 242                              | -1160.10568                 | -1160.57356             |
| 243                              | -1166.11859                 | -1165.114               |
| 244                              | -1172.1300                  | -1171.11                |
| 245                              | -1176.1500                  | -1175.113               |
| 246                              | -1181.14472                 | -1179.96957             |
| 247                              | -1186.1930                  | -1184.3192              |
| 248                              | -1192.1703                  | -1190.7000              |
| 249                              | -1197.1831                  | -1195.1001              |
| 250                              | -1203.19575                 | -1199.999               |
| 251                              | -1207.2207                  | -1205.9481              |
| 252                              | -1212.2456                  | -1211.11133             |
| 253                              | -1217.26437                 | -1216.14                |
| 254                              | -1223.23465                 | -1222.126               |
| 255                              | -1228.2897                  | -1227.14                |
| …                                | …                              | …                        |
| 260                              | -1254.299                   | -                      |
| 266                              | -1284.4123                  | -1283.9000              |
| …                                | …                              | …                        |
| 318                              | -1560.11539                 | -                      |
| …                                | …                              | …                        |
| 389                              | -1935.87887                 | -                      |
5. Results of computational experiments

Adequacy of the proposed method is proved by results shown in table 1. The proposed algorithm was successfully used to find optimal Morse cluster conformations for N>240 and ρ = 14. As evident from the table, the authors used the SDEA algorithm to calculate potential cluster energies with dimensionality higher than in those registered in the Cambridge databases [7]. It should be noted that optimal clusters were created more often on an asymmetrical “inclined” lattice, which correspond to the tendencies observed for the clusters included in the database [7]. It is interesting that for the cluster with N=239, an optimal conformation was found with the potential energy lower than in the database [7] (ν_{239} = −1144.5328).

6. Conclusion

This study presents an algorithm of parallel stochastic structural evolution of Morse clusters that allows finding optimal initial approximations of conformations of large Morse atomic clusters. The proposed method allows organizing quite simple parallel calculation processes which significantly expands possibilities of search of large Morse cluster conformations.

In the future, the authors are looking forward to study possibilities to further refine the proposed algorithm.

7. References

[1] Morse P 1929 Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels Physics Reviews 34 57-64
[2] Nemirovskij A S and YUdin D B 1979 Slozhnost' zadach i effektivnost' metodov optimizacii (M.: Nauka) p 384 (in Russian)
[3] Levinthal C 1969 How to Fold Graciously Mossbauer Spectroscopy in Biological System: Proceedings of a meeting held at Allerton House 22-24
[4] Kovartsev A N 2017 Geometric method for large morse clusters formation Computer Optics 41(1) 118-125 DOI: 10.18287/2412-6179-2017-41-1-118-125
[5] Kovartsev A N and Klimashova G P 2018 The use of information technology in clarifying the rules for the formation of initial approximate conformal Morse clusters based on a geometrically-based method Perspektivnye informacionnye texnologii 87-90 (in Russian)
[6] Kovartsev A N and Popova-Kovartseva D A 2011 On efficiency of parallel algorithms for global optimization of functions of several variables Computer Optics 35(2) 256-261
[7] Table of the global minima of Morse clusters for N 81-240 URL: http://staff.ustc.edu.cn/~clj/morse/table.html (accessed October 7, 2016)
[8] Kovartsev A N and Klimashova G P 2018 Rules for the formation of initial approximations of the conformation of Morse atomic clusters on the basis of a geometrically grounded method Information Technology and Nanotechnology (ITNT-2018) 1635-1644
[9] Ren L and Cheng L 2012 Structural prediction of (Au_{20})_n (N=2-40) clusters and their building-up principle Computational and Theoretical Chemistry 984 142-147
[10] Locatelli M and Schoen F 2013 Global Optimization: Theory, Algorithms and Applications (SIAM) p 439
[11] Cheng L, Feng Y, Yang Jie and Yang Jinlong 2009 Funnel hopping: searching the cluster potential energy surface over the funnels Journal of Chemical Physics 130(21) 214112
[12] Shao N, Huang W, Gao Y, Wang L M, Li X, Wang L S and Zeng X C 2010 Probing the structural evolution of medium-sized gold clusters: Au \textit{n} (n = 27–35) Journal of the American Chemical Society 132(18) 6596-6605
[13] Berri R S, Smirnov B M and Strizhev A Y 1997 Competitors of cubic face-centered and icosahedral cluster structure Journal of Experimental and Theoretical Physics 112(3) 1082-1090 (in Russian)
[14] Hales T C 2005 A proof of the Kepler conjecture Annals of Mathematics 162(3) 1065-1185
[15] Doye J P K and Wales D J 1997 Structural consequences of the range of the interatomic potential: a menagerie of clusters Journal of the Chemical Society 93 4233-4244
[16] Pullan W 2010 Unbiased Geometry Optimisation of Morse Atomic Clusters. *WCCI IEEE World Congress on Computational Intelligence* (Barcelona, Spain) 4496-4502

[17] Kovartsev A N, Zhidchenko V V, Popova-Kovartseva D A and Danilenko A N 2019 A Technique for the Visual Modeling of Parallel Algorithms: The PGRAPH Language. *Izvestiya Samarskogo nauchnogo tsentra RAN* 21(1) 145-149 (in Russian)