DIMACS Technical Report 93-89
December 1993

Optimization Methods for Computing Global Minima of Nonconvex Potential Energy Functions

by

Panos M. Pardalos[1]
Department of Industrial and Systems Engineering
University of Florida, Gainesville, FL 32611
Email: `pardalos@math.ufl.edu`

David Shalloway
Section of Biochemistry, Molecular and Cell Biology
Cornell University, Ithaca, NY 14853
Email: `disj@cornell.cit.cornell.edu`

Guoliang Xue[2]
Department of Computer Science and Electrical Engineering
University of Vermont, Burlington, VT 05405.
Email: `xue@cs.uvm.edu`

[1]Research supported in part by DIMACS.
[2]Also a researcher at the Army High Performance Computing Research Center, University of Minnesota, Minneapolis, MN 55415.

DIMACS is a cooperative project of Rutgers University, Princeton University, AT&T Bell Laboratories and Bellcore.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

The minimization of potential energy functions plays an important role in the determination of ground states or stable states of certain classes of molecular clusters and proteins. In this paper we introduce some of the most commonly used potential energy functions and discuss different optimization methods used in the minimization of nonconvex potential energy functions. A very complete bibliography is also given.

1 Introduction

The minimization of potential energy functions plays an important role in the determination of ground states or stable states of certain classes of molecular clusters and proteins. Since, in almost all the cases, the potential energy function is nonconvex and therefore has many local minimizers, the minimization of the potential energy function is a very hard problem.

For many years, this has been a field of Chemists and Physicists. In recent years, however, many researchers from Optimization and Computer Science have paid close attention to this problem. Their interest in this problem is two-fold. Firstly, they find that most of these potential energy minimization problems are ideal as test problems for global minimization algorithms. Second, with the advanced techniques in optimization algorithms and the use of massively parallel computers, they can, for certain instances, find lower energy states for some of the well studied problems. The minimization of nonconvex potential energy functions arising from molecular conformation or protein folding problems has become one of the most important problems for interdisciplinary research [8, 57, 24]. It should be noted that this is a very hard problem. Current methods are limited by the problem size since, when the cluster size grows, the number of local minimizers often grows exponentially. The problem requires intensive computing time, ranging from weeks to months of CPU time on a Cray supercomputer even for a middle sized problem.

In this paper, we will introduce some of the most commonly used potential energy functions and describe most of the approaches taken to tackle the minimization problem. A very complete reference list is also provided at the end of the paper.

2 Potential Energy Functions

In this section, we describe some of the most commonly used potential energy functions and address some of their properties.

Of the many cases of interest, the simplest is that of two-body central forces between the component atoms. Although this potential energy function does not actually model the real system in many cases – ignoring angle dependent conditions, bond constraints, and dihedral angle conditions – it provides a useful prototype that enormously simplifies the design of computer programs.

Given a cluster of N atoms in $3-$dimensional space, the potential energy function of the cluster is defined as the summation (over all of the pairs) of the two-body interatomic pair potentials. Let the center of the N atoms be $a_1, \cdots, a_N$. The potential energy function is defined as follows.

$$V_N(a_1, \cdots, a_N) = \sum_{1 \le i < j \le N} v(||a_i - a_j||), \qquad (1)$$

where $|| \bullet ||$ is the Euclidean norm and $v(r)$ is the interatomic pair potential.

Although many types of function $v(r)$ may be used in physical models, it is necessary to apply some restrictions if we want $v(r)$ to have satisfactory behavior. In general, we would like $v(r)$ to be continuous at least down to a hard core r_{min} and to possess derivatives up to the second order over the interval (r_{min}, ∞). More specifically, we are only interested in cases where $v(r)$ is a *well potential* satisfying the following conditions:

c1. $v(r) \to 0^-$ as $r \to \infty$;
c2. $v(r) \to \infty$ as $r \to r_{min}$ and $r_{min} \geq 0$;
c3. $v'(r_0) = 0$ for a unique r_0 with $r_{min} < r_0 < \infty$;
c4. $v''(r_0) > 0$ and $v(r_0) < 0$.

Pair potentials of interest in clusters include the following [19, 26]:

p1. $v(r) = (n-m)^{-1}[mr^{-n} - nr^{-m}], (n > m)$ (Mie)
p2. $v(r) = 2r^{-12} - r^{-6}$ (Lennard-Jones)
p3. $v(r) = [1 - e^{\alpha(1-r)}]^2 - 1$ (Morse)
p4. $v^{\alpha\beta}(r) = z^\alpha z^\beta / r + Ae^{-r/\rho}$ (Born-Meyer)

From the above conditions, we can easily see that the potential energy function is continuous and always has a global minimizer with a minimum energy function value less than zero.

From condition c1, it is also clear that the interatomic force is very small when two atoms are far away from each other. Therefore short range potential energy functions can be introduced where the interatomic potential is defined to be zero if the distance between two atoms is greater than r_{max}. Although simple as it sounds, the idea of a short range potential is very important, for in most cases, when the cluster size is very large, an atom only interacts with a very small portion of the whole cluster. Therefore a lot of time is wasted in computing the full range potential. For this reason, there have been many techniques developed for the fast computation of short range potential energy functions.

3 Optimization Methods

In this section, we will discuss many of the optimization methods used in previous studies to minimize the potential energy functions. These methods are grouped in several subsections in the sequel. For general global optimization algorithms, we refer readers to the two classical references [29] and [51].

3.1 Lattice Based Search

The difficulty of the Lennard-Jones cluster problem arises from the fact that it is a global optimization problem with an exponential number of local minima [26]. It is axiomatic that the solution of a global optimization problem, or multiple-minima problem, will be improved by any physical understanding of the problem which serves to restrict the search space, or domain, which contains the global solution. In the case of the pure Lennard-Jones cluster, the critical assumption is that *a well-defined set of lattice structures contains at* *least one initial cluster configuration which relaxes to the ground state.* Works along this line include [1, 13, 14, 15, 25, 26, 40, 49, 54, 55, 71, 78, 79, 80]. The support for this assumption comes primarily from numerical investigations [26, 49]. For the Lennard-Jones cluster, the lattice structures consist of an icosahedral core and particular combinations of surface lattice points. Given this physical insight into the problem, the approach outlined below begins with a discrete optimization algorithm applied to a relatively small sample of initial states. The resulting lattice minima are then relaxed as a continuous minimization problem. This method can be described at a high level of abstraction:

1. Define S as a set of lattices.
2. Define a potential function for the discrete problem.
3. For each element in S,
   Repeat
      Perform lattice search based discrete optimization
      to identify minimum energy lattice conformations.
   Until stopping criteria satisfied.
4. Define a potential function for continuous problem.
5. Perform relaxation (continuous minimization) for
   each minimum energy lattice conformation.

Suppose we have a cluster of N atoms. We need to choose a base lattice structure which has $M(\geq N)$ lattice points. The lattice search is then equivalent to partition the M lattice to a set of N filled sites and $M-N$ empty sites. For each such partition, there is a corresponding potential energy value. The goal is to find a partition such that the corresponding potential energy is minimized. During the relaxation phase, the atoms (starting at the lattice minimizer) are allowed to go off the lattice sites in order to minimize the potential energy function value.

Lattice based search has lead to great success in the determination of lower-energy structures with Lennard-Jones clusters and other related clusters. Northby has used this approach to report best-known solutions for Lennard-Jones clusters of sizes $13 \leq N \leq 150$. [80] has made improvement on the Northby algorithm and found states with lower potentials for clusters with sizes $65, 66, 75, 76, 77$. However, whether this kind of approach will lead to success in more complicated clusters and in protein folding is still not clear at this time.

3.2 Simulated Annealing and Genetic Algorithms

Simulated annealing (SA) is a general purpose optimization technique that has been proposed by Kirkpatrick et al. [31]. This method is an extension of a Monte Carlo method developed by Metropolis et al. [45], to determine the equilibrium states of a collection of atoms at any given temperature T. Since the method was first proposed in [31, 32], much research has been conducted on its use and properties. Besides its applications to other problems, SA has been widely used in the search of lower energy configurations of molecular clusters. Wille

[74], for example, used SA and found a new minimum for a Lennard-Jones cluster with 24 atoms. [58] and [39] applied Monte Carlo minimization approaches to protein folding. Xue [79] has developed a two-level simulated annealing algorithm that performed very well for large Lennard-Jones clusters.

Genetic algorithms (GA) are search algorithms based on the mechanics of natural selection and natural genetics. They combine survival of the fittest among string structures with a structured yet randomized information exchange to form a search algorithm with some of the innovative flair of human search. The first strategy of GA is the use of a breeding population in which individuals who are more fit in some sense have a higher chance of producing offspring and passing their genetic information onto succeeding generations. The second strategy of GA is the use of crossover in which a child's genetic material is a mixture of his/her parents'. The third strategy of GA is that of mutation, meaning that genetic material is occasionally corrupted, leading to individuals who may or may not be more fit than they would have been otherwise, but always maintaining a certain level of genetic diversity in the population. In reference [30], both SA and GA are implemented to compute the global minimizers of the Lennard-Jones cluster. Results show that both GA and SA perform progressively better relative to random search as the problem size increases, compared to the Nelder-Mead simplex algorithm. It was also found that SA and GA have complimentary strengths which implies that a hybrid GA-SA method would be more efficient than either one alone.

3.3 Spatial Averaging Methods

The NP-complete character of global optimization implies that approximations will be needed to obtain solutions in feasible time. The primary question is,"What approximations?" Particularly for molecular conformation problems, it is natural to explore the utility of spatial averaging. These methods use a spatially-averaged objective function to approximately locate the global minimizer, and then iteratively increase accuracy by sequential minimizations with smoothed objective functions which have been averaged over smaller and smaller regions of space. The success of spatial averaging approaches depends on the validity of what we call the "strong spatial-scaling hypothesis": that the position of the global minimizer varies in a continuous manner as the energy (objective) function is progressively smoothed by averaging over increasingly large regions of space. Even when the strong hypothesis does not hold, the"weak spatial-scaling hypothesis"–that the global minimizer varies in a continuous manner except for a bounded number of discontinuities of bounded extent–may be valid. The weak condition can also be exploited to organize searches although more complicated algorithms will be needed. The validity of these hypotheses for specific problems is not known and is an interesting topic for future examination.

We can characterize the extent of spatial averaging by parameter Λ (which may be a scalar or tensor, depending on whether averaging is isotropic or anisotropic). A natural way to define an averaged energy function $\tilde{H}_{\Lambda,f}$ is by the convolution with a sampling function

S_Λ,

$$\tilde{H}_{\Lambda,f}(R) \equiv f^{-1}\Big\{\int S_\Lambda(R-R')f[H(R')]dR'\Big\} \tag{2}$$

where R is a multidimensional vector representing all the coordinates in the molecule.

One of the simplest and most useful forms for S_Λ is a Gaussian

$$\begin{aligned} S_\Lambda(R) &\equiv C(\Lambda)e^{-R\Lambda^{-2}R} \\ C(\Lambda) &\equiv \pi^{-d/2}Det^{-1}(\Lambda) \end{aligned} \tag{3}$$

where d is the total dimensionality of R. The function f included in (2) allows for non-linear averaging. Two choices motivated by physical considerations are $f(x) = x$ and $f(x) = e^{-x/k_BT}$. These choices correspond respectively to the "diffusion equation" and "effective energy" methods which are described below. Wu [77] has presented a general discussion of transformations of the form of (2).

A highly smoothed $\tilde{H}_{\Lambda,f}$ (from which all high spatial-frequency components have been removed) will in most cases have fewer local minima than the unsmoothed ("bare") function, so it will be much easier to identify its global minimum. If the strong spatial-scaling hypothesis is correct, the position of this minimum can then be iteratively tracked by local-minimization as Λ decreases. As $\Lambda \to 0$, the position will approach the global minimizer of the bare objective function.

In its basic form, this is simply deformation of the energy surface and can be viewed as a homotopy transformation [72]. Unfortunately, it is not very useful in the general case because exponential (in the number of atoms) effort is required to evaluate the multi-dimensional integral in (2). However, most of the terms appearing in macromolecular energy functions appear as sums of two-body potentials like the pair-potential terms in the microcluster potential energy (1). As described below, (2) can be approximated for this class of energy function.

3.3.1 Diffusion Equation Method

Choosing $f(x) = x$ and Λ as a scalar (i.e., a multiple of the identity matrix), $\Lambda^2 \leq 4t$, the averaged energy $\tilde{H}_{\Lambda,f}(R) \equiv H(R,t)$ in (2) is a solution to the diffusion-equation

$$\frac{\partial^2}{\partial R^2}H(R,t) = \frac{\partial}{\partial t}H(R,t) \tag{4}$$

with boundary condition

$$H(R,0) = H(R) \tag{5}$$

$H(R,t)$ can be viewed as the solution for the time-dependent dispersion of heat through a uniformly conducting medium if t is associated with physical time. Although t does not have this meaning in the global minimization context, thinking of the spatial averaging in this manner is a convenient heuristic.

The "diffusion equation method" [53] begins by considering $H(R,t)$ at large "time" t such that

$$\frac{\partial\ H(r,t)}{\partial\ R} = 0 \tag{6}$$

has only one solution. t is then gradually decreased in steps and (6) is solved at each step by local minimization methods to track the position of a minimizer as $t \rightarrow 0$. Hopefully, this will be the global minimizer.

Piela, Scheraga, Kostrowicki and coworkers have applied this method to some standard mathematical test cases and a number of model problems, including microclusters of Lennard-Jones atoms, and small peptides with encouraging results. The method found the global minimum when applied to standard test cases for low dimensionality ($n \leq 10$) (e.g., Goldstein-Price, Griewank, ...) [34]. Performance in medium dimensionality model problems was good, but not flawless. For example, in the microcluster problem, the lowest energy state was obtained in 9 out of 14 cases investigated (where the number of atoms ranged from 5 to 55. Even when the global minimum was not found, relatively low-energy local minima were identified [35]. Performance of the method on small peptides was also tested using a typical protein model potential [36]. The method found the global minimum of a simplified model of dialanine restricted to 2 degrees of freedom. When tested on a model of the pentapeptide Met-enkephalin, with 19 degrees of freedom, the method found the general region of the global minimizer, but not the minimizer itself.

The diffusion-equation is made computationally feasible by approximating the interatomic two-body potential energy functions $u(|r_i - r_j|)$ by a sum of Gaussians. That is,

$$u(r) \cong \sum_{k=1}^{L} a_k e^{-b_k r^2} \tag{7}$$

Since integrals of products of Gaussians can be analytically evaluated, (2) with (3) and (7) can be reduced to an algebraic sum. However, there are some problems to this approach. First, because the Lennard-Jones potential (and a number of other potentials of interest) becomes infinite as $r \rightarrow 0$ ("hard-core" repulsive interaction), the integral in (2) diverges. This problem was addressed simply by replacing the Lennard-Jones potential with a truncated version that goes to a finite value as $r \rightarrow 0$ [35]. However, the effect of this procedure and sensitivity of the obtained solution to the truncation parameter is unknown. It should be possible to overcome these problems by replacing $H(R)$ in boundary conditions (5) with the exponentially transformed form $exp[-\beta H(R)]$. The utility of this procedure is being investigated (H. Scheraga, personal communication).

3.3.2 Effective Energy Methods

This class of methods [61, 62] is motivated by considering the physical process of thermally annealing a molecular structure from high temperature, where it is largely disordered, down to a low temperature T_{lo}, where the conformation fluctuates in a narrow range about the global minimizer. At every temperature T, the probability density that the molecular conformation is near conformation R is given by the Gibbs/Boltzmann distribution

$$p_T(R) = \frac{e^{-H(R)/k_BT}}{Z(T)} \tag{8}$$

where k_B is Boltzmann's constant and $Z(T)$ is the partition function:

$$Z(T) \equiv \int e^{-H(R)/k_BT} dR.$$

At very low temperature $p_{T_{lo}}(R)$ can be approximated as

$$p_{T_{lo}}(R) \cong c(\Lambda[T_{lo}])e^{[-(R-R^0)\Lambda^{-2}(T_{lo})(R-R^0)]} \tag{9}$$

where the symmetric tensor $\Lambda(T_{lo})$ parametrizes the moment of inertia of the probability density and R^o is the global minimizer.

The global minimization problem can be viewed as that of interpolating from (8) to (9). Simulated annealing, in which the stochastic Metropolis algorithm is used to search the space with probability density $p_T(R)$ [31], is one method for doing this; but it is too inefficient for the large dimensionality continuous problems posed by molecular structure prediction. A key to good performance is the identification and utilization of appropriate sizes for the stochastic jumps. The appropriate size scales will depend on T and on the character of the energy landscape in a complicated manner. Intuitively, we guess that these scales corresponds to the magnitudes of the thermal fluctuations experienced by a physical system at temperature T. Different size scales $\Lambda_\alpha(T)$ may be needed in different regions.

3.3.3 Packet Annealing

The packet-annealing method [61] is designed to focus attention on the appropriate temperature-dependent size scales in a systematic manner. It begins by approximating $p_T(R)$ as a sum of Gaussians

$$p_T(R) \approx \tilde{\tilde{p}}_T(R) \equiv \sum_{\{\alpha\}_T} p_\alpha C[\frac{\Lambda_\alpha(T)}{\sqrt{2}}]e^{-[R-R^o_\alpha(T)]\Lambda_\alpha^{-2}(T)[R-R^o_\alpha(T)]} \tag{10}$$

The parameters $\Lambda_\alpha(T)$ which govern the widths of the Gaussian packets centered at positions R^o_α identify the dominant spatial scales in these regions. The convergence of $p_T(R)$ to $p_{T_{lo}}(R)$ is then traced by following the development as $T \rightarrow T_{lo}$ of the parameters p_α, Λ_α, and R^o_α that describe the Gaussian "packets". The number of Gaussians needed will vary with T, so algorithms for determining when to add and eliminate packets are also required.

The equations that govern the behavior of the packets (i.e., the behavior of p_α, Λ_α, and R^o_α) in T depend on the "effective energy" $\tilde{H}_{\Lambda_\alpha,T}(R)$:

$$\tilde{H}_{\Lambda,T}(R) \equiv -k_BT \log\left\{\frac{\int e^{-H(R')/k_BT}e^{-(R-R')\Lambda^{-2}(R-R')}dR'}{Z(T)}\right\} \tag{11}$$

This transformation is of the form (2) with $f = e^{-x/k_BT}$. (The factor $Z(T)$ is included for completeness. However, since it only contributes an R-independent constant to $\tilde{H}_{\Lambda_\alpha,T}(R)$, in practice it can be ignored. Note that this definition of $\tilde{H}_{\Lambda_\alpha,T}(R)$ differs by an additive normalization constant from that used in [61, 62]).

The packet parameters are determined in terms of $\tilde{H}$ by the packet annealing equations:

$$p_\alpha(T) = e^{-\frac{\tilde{H}_{\Lambda_\alpha,T}(R^o_\alpha)}{k_BT}} \tag{12}$$

$$\frac{\partial\ \tilde{H}_{\Lambda_\alpha,T}(R)}{\partial\ R}|_{R=R^o_\alpha} = 0 \tag{13}$$

$$\frac{\partial^2\tilde{H}_{\Lambda_\alpha,T}(R)}{\partial\ R^2}\ |_{R=R^o_\alpha} = \frac{k_BT}{\Lambda^2_\alpha} \tag{14}$$

The packet annealing procedure begins at a high temperature where the Λ_α are large and where there are only a small number of (or one) solutions to (13) and (14). Temperature is then gradually reduced by a stepwise annealing procedure while (13) and (14) are solved iteratively by perturbative methods. (13) is similar to (6) and fixes R^o_α. (14) is unique to this method and identifies the dominant spatial scale in the region near R^o_α. In principle, the bifurcation points of (14) identify points of packet branching (i.e., where packets are added), but practical algorithms to identify these points in high-dimensionality problems have not yet been developed.

As in the diffusion equation method, efficient evaluation of the integral defining $\tilde{H}_{\Lambda_\alpha,T}(R)$ in (11) is a key to practical success. Approximation schemes have been used which exploit the fact that H is the sum of two-body potential terms: this suggests approximating the multidimensional integral in (11) by factorizing it into the product of computationally feasible two-body integrals [62]. This factorization approximation has been used in the practical tests of the method described below, although its accuracy is not known.

A simplified implementation of the packet annealing method has been applied to the Lennard-Jones microcluster problem, and results similar to those obtained with the diffusion-equation method were obtained (i.e., the global minimum was identified in 14 out of 20 tested cases; [61]. Use of the full method, incorporating packet branching, might improve performance, but this has not yet been tested.

3.3.4 Effective Energy Simulated Annealing

$\tilde{H}_{\Lambda_\alpha,T}$ can be used directly in stochastic simulated annealing methods. Since the high frequency spatial fluctuations have already been integrated out in (10), it is to be expected that stochastic Metropolis sampling with the effective energy will converge to the global minimum more rapidly than sampling with the bare energy function. Tests of a simple procedure of this type applied to the microcluster problem showed that this is the case and that order-of-magnitude improvements in efficiency can be obtained on low-dimensionality (number of atoms ≤ 100) problems [10]. Performance degraded with increasing dimensionality, possibly because only isotropic averaging (i.e., Λ a scalar) was used; the use of anisotropic averaging may overcome this difficulty.

Effective energy simulated annealing has also been adapted into a build-up procedure which is more efficient for large dimensionality problems. This method, when applied to the microcluster problem, found a structure for the 72-atom cluster lower than any other found in previous studies [11].

3.3.5 Effective Energy and Thermodynamic Free Energy

The concept of effective energy was previously introduced in the "renormalization group" approach to statistical mechanics [75]. $\tilde{H}_{\Lambda_\alpha T}(R)$ can be viewed as a partially integrated form of the partition function in which the high, but not the low, spatial frequency components have been integrated out by the convolution with the Gaussian. In effective energy methods, each packet represents a metastable physical state of the system; that is, a collection (parametrized by R^o_α and Λ_α) of microscopic conformations (parametrized by R) that equilibrate amongst themselves in times much shorter than those required for transitions between packets. It can be shown that $\tilde{H}_{\Lambda_\alpha,T}(R^o_\alpha)$ is an approximation to the thermodynamic free-energy of the metastable state corresponding to the set of conformations centered around conformation R^o_α. At a finite temperature, $\tilde{H}_{\Lambda_\alpha,T}(R^o_\alpha)$ depends both on the energy of conformation R (the enthalpic contribution) and on the size of the basin in the surrounding energy landscape (the entropic contribution). Narrow minima will have higher effective energies than wide minima of comparable depth and the algorithms will tend to seek out the wider minima. Thus, minimization using the effective energy is closely related to minimization of the free energy. For physical problems, this is superior to standard minimization of the potential energy (i.e., enthalpy) since it is the states of low free energy that are of physical interest.

3.3.6 Comparison of Diffusion Equation and Effective Energy Methods

For bounded $H(R)$, $H(R,t)$ in the diffusion equation method can be viewed as the high-temperature limit of the effective energy $\tilde{H}_{\Lambda_\alpha,T}(R)$, where $\Lambda = 4t$ is a scalar. (This can be seen by expanding $\exp[-H(R')/k_BT]$ in a Taylor's series.) Beyond this, the methods differ in that: 1) t (or, equivalently, Λ) is regarded as a free parameter in the diffusion equation method while it is constrained by (14) in effective energy methods, 2) anisotropic Λ are permitted in effective energy methods, and 3) in principle, the trajectories (in T) of multiple minima (packets) can be tracked in effective energy methods while only one minimum is tracked in the diffusion equation method. Successful identification of the global minimizer in the diffusion equation method requires that the strong spatial scaling hypothesis be valid for the problem under consideration. In principle, tracking multiple minima (packets) may allow packet annealing methods to find the global minimizer when only the weak, but not the strong, spatial scaling hypothesis is valid. Similarly, allowing for anisotropic averaging may be an important advantage in high-dimensionality problems where basins-of-attraction are expected to be highly anisotropic. However, the price for this increased robustness and flexibility is greater algorithmic complexity; a complete packet annealing algorithm which tracks multiple anisotropic packets has yet to be developed.

3.3.7 Schroedinger-Equation Methods

Other approaches using smoothed energy functions have also been suggested. These include the use of solutions of the Schroedinger equation [63] or a mean-field approximation to the Schroedinger equation [50]. The practical difficulty with these methods is the need to identify good trial wavefunctions and an appropriate reduced basis set for problems with large number of degrees of freedom.

3.4 Other Optimization Methods

C.D. Maranas and C.A. Floudas [41, 42, 43] have developed very efficient global optimization methods for molecular conformation. In [4, 5] general purpose global optimization algorithms have been proposed; these algorithms can (without knowledge on the lattice structure) find minimizers as good as the ones reported by Northby for the Lennard-Jones clusters of size in the range $n \leq 147$, with only a few exceptions. In addition, these minimizers are almost as good as the ones reported in [49]. We refer interested readers to the specific references.

3.5 General Purpose Macromolecular Structure Software Packages

One significant software packages for molecular dynamics and molecular conformation is CHARMM, reported by Brooks, Bruccoleri, Olafson, States, Swaminathan and Karplus. CHARMM (Chemistry at HARvard Macromolecular Mechanics) is a highly flexible computer program which uses empirical energy functions to model macromolecular systems. The program can read or model build structures, energy minimization by first- or second-order derivative techniques, perform a normal mode or molecular dynamics simulation, and analyze the structural, equilibrium, and dynamic properties determined in these calculations. CHARMM is available on the Cray supercomputers at the Minnesota Supercomputer Institute and many other sites. Similar programs are available from other sources. These include AMBER [73], CEDAR [7], and GROMOS ([66] and references therein).

3.6 Fast Multipole Algorithm

In one of the 1987 ACM Distinguished Dissertations [21], Leslie Greengard presented an algorithm for the rapid evaluation of the potential and force fields in large-scale systems of particles. This algorithm is now called the fast multipole algorithm.

Greengard observes that an amount of work of the order $O(N^2)$ has traditionally been required to evaluate all pairwise interactions in a system of N particles, unless some approximation or truncation method is used. As a result, large-scale simulations have been extremely expensive in some cases, and prohibitive in others.

Greengard studied the multipole expansion and noticed that in order to obtain a relative precision ϵ (with respect to the total charge), $O(|\log(\epsilon)|n)$ amount of work is sufficient.

For example, a usual algorithm for computing the Lennard-Jones potential energy function of a cluster of n atoms would require $O(n^2)$ time since there are $\frac{n(n-1)}{2}$ pairs in all. However, with the fast multipole algorithm, the same potential energy function can be computed in $O(n)$ time. Computational results as well as parallel algorithms are reported in [23] This is still a fruitful direction of research. Readers are referred to the references [23, 21, 22, 6]

3.7 Parallel Algorithms

Besides conventional sequential algorithms for molecular conformation and protein folding, there have been many parallel algorithms developed for vector computers, SIMD computers and MIMD computers. In [3], an optimal hypercube direct N-body solver was implemented on the Connection Machine-2. In [65], parallel approaches to short range molecular dynamics simulation on the Connection Machine-2 was reported. [44] reported molecular dynamics simulation of liquids on the Connection Machine. [9] presented parallel algorithms on a MIMD architecture. [64] presented molecular dynamics algorithms on vector computers. [40, 78, 79] presented parallel algorithms for molecular conformation on the Connection Machine-5.

4 Conclusions

The minimization of nonconvex potential energy functions is a fascinating and utterly hard problem which is one of the most important problems for interdisciplinary research. We have introduced some of the most commonly used potential energy functions and discussed a variety of methods that have been developed to tackle this minimization problem. It is our belief that important results in this area will be obtained in the near future with groups of researchers from different fields working together. It has been therefore the goal of this paper to provide interested readers a brief introduction to the problems and corresponding literature.

References

[1] L.L. Boyer and J.Q. Broughton, Statics and Dynamics of Icosahedrally Twinned and Single-Crystal fcc Clusters, *Physical Review B*, Vol. 42(90), pp. 11461-11468.

[2] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *Journal of Computational Chemistry*, Vol. 4(1983), pp. 187-217.

[3] J.P. Brunet, A. Edelman, J.P. Mesirov, An Optimal Hypercube Direct N-body Solver on the Connection Machine, in Proceedings of *IEEE/ACM Supercomputing'90*, pp. 748-752, IEEE Computer Society Press 1990.

[4] R.H. Byrd, E. Eskow, R.B. Schnabel, and S.L. Smith, Parallel Global Optimization: Numerical Methods, Dynamic Scheduling Methods, and Application to Molecular Configuration, *Technical Report CU-CS-553-91, University of Colorado at Boulder, Department of Computer Science, Boulder, CO.*, October 1991.

[5] R.H. Byrd, E. Eskow, R.B. Schnabel, Global Optimization Methods for Molecular Configuration Problems, Presented at the *Fourth SIAM Conference on Optimization*, May 11-13, 1992, Chicago, IL.

[6] J. Carrier, L. Greengard and V. Rokhlin, A Fast Adaptive Multipole Algorithm for Particle Simulations, *SIAM Journal on Scientific and Statistic Computing*, Vol. 9(1988), pp. 669-686.

[7] M. Carson and J. Hermans, The Molecular Dynamics Workshop Laboratory, in *Molecular Dynamics and Protein Structure*, J. Hermans (ed.), University of North Carolina: Chapel Hill (1985), pp. 165-166.

[8] H. S. Chan and K. A. Dill, The Protein Folding Problem, *Physics Today*, February 1993, pp. 24-32.

[9] T.W. Clark and J.A. McCammon, Parallelization of a Molecular Dynamics Non-Bonded Force Algorithm for MIMD Architecture, *Computers and Chemistry*, Vol. 14(1990), pp. 219-224.

[10] T. Coleman, D. Shalloway, and Z. Wu, Isotropic Effective Energy Simulated Annealing Searches for Low Energy Molecular Cluster States, *Compututational Optimization and Applications*, Vol. 2(1993) pp. 145-170.

[11] T. Coleman, D. Shalloway, and Z. Wu, A Parallel Build-Up Algorithm for Global Energy Minimizations of Molecular Clusters using Effective Energy Simulated Annealing, *Journal of Global Optimization*, in press.

[12] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag, 1988.

[13] J. Farges, M.F. De Feraudy, B. Raoult and G. Torchet, Noncrystalline Structure of Argon Clusters. I. Polyicosahedral Structure of Ar_N Clusters, $20 < N < 50$, *Journal of Chemical Physics*, Vol. 78(1983), pp. 5067-5080.

[14] J. Farges, M.F. De Feraudy, B. Raoult and G. Torchet, Cluster Models Made of Double Icosahedron Units, *Surface Science*, Vol. 156(1985), pp. 370-378.

[15] J. Farges, M.F. De Feraudy, B. Raoult and G. Torchet, Noncrystalline Structure of Argon Clusters. I. Polyicosahedral Structure of Ar_N Clusters, $50 < N < 750$, *Journal of Chemical Physics*, Vol. 84(1986), pp. 5067-5080.

[16] D.M. Ferguson and D.J. Raber, A New Approach to Probing Conformational Space with Molecular Mechanics: Random Incremental Pulse Search, *Journal of the American Chemical Society*, Vol. 111(1989), pp. 4371-4378.

[17] D.M. Ferguson, W.A. Glauser and D.J. Raber, Molecular Mechanics Conformational Analysis of Cyclononane Using the RIPS Method and Comparison with Quantum-Mechanical Calculations, *Journal of Computational Chemistry*, Vol. 110(1990), pp. 903-910.

[18] D.M. Ferguson and D.J. Raber, Molecular Mechanics Calculations of Several Lanthanide Complexes: An Application of the Random Incremental Pulse Search, *Journal of Computational Chemistry*, Vol. 11(1990), pp. 1061-1071.

[19] I.Z. Fisher, *Statistical Theory of Liquids*, University of Chicago Press, 1964.

[20] D.G. Garrett, K.D. Kastella, D.M. Ferguson, New Results on Protein Folding from Simulated Annealing, submitted to *Journal of the American Chemistry Society*, 1992.

[21] L. Greengard, *The Rapid Evaluation of Potential Fields in Particle Systems*, The MIT Press, Cambridge, Massachusetts, 1988.

[22] L. Greengard and V. Rokhlin, A Fast Algorithm for Particle Simulations, *Journal of Computational Physics*, Vol. 73(1987), pp. 325-348.

[23] L. Greengard and W.D. Gropp, A Parallel Version of the Fast Multipole Method, *Computers and Mathematics with Applications*, Vol. 20(1990), pp. 63-71.

[24] W.F. van Gunsteren and H.J.C. Berendsen, Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry, *Angew. Chem. Int. Ed. Engl.*, Vol. 29(1990), pp. 992-1023.

[25] D.A. Hinds and M. Levitt, A Lattice Model for Protein Structure Prediction at Low Resolution, *Proceedings of the National Academy of Sciences USA*, Vol. 89(1992), pp. 2536-2540.

[26] M.R. Hoare, Structure and Dynamics of Simple Microclusters, *Advances in Chemical Physics*, Vol. 40(1979), pp. 49-135.

[27] D. Hohl, R. Idaszak, and R.O. Jones, Quantum Molecular Modeling with Simulated Annealing – A Distributed Processing and Visualization Application, in Proceedings of *ACM/IEEE Supercomputing'90*, pp. 816-825, IEEE Computer Society Press 1990.

[28] J. Danna Honeycutt and Hans C. Andersen, Molecular Dynamics Study of Melting and Freezing of Small Lennard-Jones Clusters, *Journal of Physical Chemistry*, Vol. 91(1987), pp. 4950-4963.

[29] R. Horst and H. Tuy, *Global Optimization, Deterministic Approaches*, 2nd Edition, Springer-Verlag, 1993.

[30] R.S. Judson, M.E. Colvin, J.C. Meza, A. Huffer, and D. Gutierrez, Do Intelligent Configuration Search Techniques Outperform Random Search for Large Molecules?, *Sandia Report SAND91-8740, Sandia National Laboratories, Center for Computational Engineering, Livermore, CA.*, December 1991.

[31] S. Kirkpatrick, C.D. Gelatt, Jr., M.P. Vecchi, Optimization by Simulated Annealing, *Science*, Vol. 220(1983), pp. 671-680.

[32] S. Kirkpatrick, Optimization by Simulated Annealing: Quantitative Studies, *Journal of Statistical Physics*, Vol. 34(1984), pp. 975-986.

[33] A. Kolinski, M. Milik, and J. Skolnick, Static and Dynamic Properties of a New Lattice Model of Polypeptide Chains, *Journal of Statistical Physics*, Vol. 94(1991), pp. 3978-3985.

[34] J. Kostrowicki and L. Piela, Diffusion Equation Method of Global Minimization: Performance for Standard Test Functions, *Journal of Optimization Theory and Applications*, Vol. 69(1991), pp. 269-284.

[35] J. Kostrowicki, L. Piela, B.J. Cherayil and H.A. Scheraga, Performance of the Diffusion Equation Method in Searches for Optimum Structures of Clusters of Lennard-Jones Atoms, *Journal of Physical Chemistry*, Vol. 95(1991), pp. 4113-4119.

[36] J. Kostrowicki and H.A. Scheraga, Application of the Diffusion Equation Method for Global Optimization to Oligopeptides, *Journal of Physical Chemistry*, Vol. 96(1992), pp. 7442-7449.

[37] S.M. Le Grand and K.M. Merz Jr., The Application of the Genetic Algorithm to the Minimization of potential Energy Functions, *Journal of Global Optimization*, Vol. 3(1993), pp. 49-66.

[38] M. Levitt, Protein Folding by Restained Energy Minimization and Molecular Dynamics, *Journal of Molecular Biology*, Vol. 170(1983), pp. 723-764.

[39] Z. Li and H. Scheraga, Monte Carlo-Minimization Approach to the Multiple-Minima Problem in Protein Folding, *Proceedings of the National Academy of Sciences USA*, Vol. 84(1987), pp. 6611-6615.

[40] R.S. Maier, J.B. Rosen, G.L. Xue, A Discrete-Continuous Algorithm for Molecular Energy Minimization, in Proceedings of *IEEE/ACM Supercomputing'91*, pp. 778-786, IEEE Computer Society Press 1992.

[41] C.D. Maranas and C.A. Floudas, Global Optimization of Lennard-Jones Microclusters, *Journal of Chemical Physics*, 97, 10, 7667 (1992).

[42] C.D. Maranas and C.A. Floudas, Global Optimization for Molecular Conformation Problems, *Annals of Operations Research*, Vol. 42(1993), pp. 85-117.

[43] C.D. Maranas and C.A. Floudas, Global Minimum Potential Energy Conformations of Small Molecules, *Journal of Global Optimization*, in press.

[44] A.I. Melcuk, R.C. Giles and H. Gould, Molecular Dynamics Simulation of Liquids on the Connection Machine, *Computers in Physics*, May/Jun 1991, pp. 311-318.

[45] N. Metropolis, A. Rosenbluth, A. Teller, E. Teller, Equation of Several State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, Vol. 21(1953), pp. 1087-1892.

[46] S.G. Nash, User's Guide for TN/TNBC: Fortran Routines for Nonlinear Optimization, *Report 397, Mathematical Sciences Department*, The Johns Hopkins University, 1984.

[47] S.G. Nash, Preconditioning of Truncated-Newton Methods, *SIAM Journal on Scientific and Statistic Computing*, Vol. 6(1985), pp. 599-616.

[48] D. R. Nelson, T. Piran, and S. Weinberg (Editors), Statistical Mechanics of Membranes and Surfaces, *World Scientific*, 1989.

[49] J.A. Northby, Structure and Binding of Lennard-Jones Clusters: $13 \leq n \leq 147$, *Journal of Chemical Physics*, Vol. 87(1987), pp. 6166-6178.

[50] K.A. Olszewsky, L. Piela and H.A. Scheraga, Mean Field Theory as a Tool for Inter-Molecular Conformational Optimization. I. Tests on Terminally Blocked Alanine and Met-Enkephalin, *Journal of Physical Chemistry*, Vol. 96(1992), pp. 4672-4676.

[51] P.M. Pardalos and J.B. Rosen, *Constrained Global Optimization, Algorithms and Applications*, Springer-Verlag Lecture Notes in Computer Science 268, 1987.

[52] H.G. Petersen and J.W. Perram, Molecular Dynamics on Transputer Arrays. I. Algorithm Design, Programming Issues, Timing Experiments and Scaling Projections, *Molecular Physics*, Vol. 67(1989), pp. 849-860.

[53] L. Piela, J. Kostrowicki and H.A. Scheraga, The Multiple-Minima Problem in the Conformational Analysis of Molecules. Deformation of the Potential Energy Hypersurface by the Diffusion Equation Method, *Journal of Physical Chemistry*, Vol. 93(1989), pp. 3339-3346.

[54] B. Raoult, J. Farges, M.F. De Feraudy and G. Torchet, Stability of Relaxed Lennard-Jones Models Made of 500 to 6000 Atoms, *Z. Phys. D – Atoms, Molecules and Clusters*, Vol. 12(1989), pp. 85-87.

[55] B. Raoult, J. Farges, M.F. De Feraudy and G. Torchet, Comparison between Icosahedral, Decahedral and Crystalline Lennard- Jones Models containing 500 to 6000 Atoms, *Philosophical Magazine B*, Vol. 60(1989), pp. 881-906.

[56] A. Rey and J. Skolnick, Comparison of Lattice Monte Carlo Dynamics and Brownian Dynamics Folding Pathways of α-Helical Hairpins, Journal of Chemical Physics, Vol. 158(1991), pp. 199-219.

[57] F.M. Richards, The Protein Folding Problem, *Scientific American*, January, 1991, pp. 54-63.

[58] D.R. Ripoll, S.J. Thomas, A Parallel Monte Carlo Search Algorithm for the Conformational Analysis of Proteins, in Proceedings of *IEEE/ACM Supercomputing'90*, pp. 94-102. IEEE Computer Society Press 1990.

[59] T. Schlick and M. Overton, A Powerful Truncated Newton Method for Potential Energy Minimization, *Journal of Computational Chemistry*, Vol. 8(1987), pp. 1025-1039.

[60] E. Shakhnovich, G. Farztdnov, A.M. Gutin and M. Karplus, Protein Folding Bottlenecks: A Lattice Monte Carlo Simulation, *Physical Review Letters*, Vol. 67(1991), pp. 1665-1669.

[61] D. Shalloway, Packet Annealing: A Deterministic Method for Global Minimization. Application to Molecular Conformation, In *Recent Advances in Global Optimization*, C. Floudas and P. Pardalos (eds.), Princeton University Press: Princeton, N.J., (1992) pp. 433-477.

[62] D. Shalloway, Application of the Renormalization Group to Deterministic Global Minimization of Molecular Conformation Energy Functions, *Journal of Global Optimization*, Vol. 2(1992), pp. 281-311.

[63] R.I. Somorjai, Novel Approach for Computing the Global Minimum of Proteins. I. General Concepts, Methods and Approximations, *Journal of Physical Chemistry*, Vol. 95(1991), pp. 4141-4146.

[64] F. Sullivan, R.D. Mountain and J. O'Connell, Molecular Dynamics on Vector Computers, *Journal of Computational Physics*, Vol. 61(1985), pp. 138-153.

[65] P. Tamayo, J.P. Mesirov and B.M. Boghsian, Parallel Approaches to Short Range Molecular Dynamics Simulations, in Proceedings of *IEEE/ACM Supercomputing'91*, pp. 462-470, IEEE Computer Society Press 1991.

[66] J.M. Troyer and F.E. Cohen, Simplified Models for Understanding and Predicting Protein Structure, in *Reviews in Computational Chemistry II*, pp. 57-80, K.B. Lipkowitz and D.B. Boyd eds, VCH Publishers, 1991.

[67] S. Vajda and C. Delisi, Determining Minimum Energy Conformations of Polypeptides by Dynamic Programming, *Biopolymers*, Vol. 29(1990), pp. 1755-1772.

[68] W.F. van Gunsteren, H.J.C. Berendsen, J. Hermans, W.G.J. Hol and J.P.M. Postma, Computer Simulation of the Dynamics of Hydrated Protein Crystals and Its Comparison with X-Ray Data, *Proceedings of the National Academy of Sciences of the USA*, Vol. 80(1983), pp. 4315-4319.

[69] D.G. Vlachos, L.D. Schmidt, and R. Aris, Structures of Small Metal Clusters: Phase Transitions and Isomerization, *AHPCRC Preprint 91-69, University of Minnesota, Minneapolis, MN 55415*, 1991.

[70] D.G. Vlachos, L.D. Schmidt, and R. Aris, Structures of Small Metal Clusters: Low Temperature Behavior, *AHPCRC Preprint 91-70, University of Minnesota, Minneapolis, MN 55415*, 1991.

[71] Benjamin W. van de Waal, Stability of Face-Centered Cubic and Icosahedral Lennard-Jones Clusters, *Journal of Chemical Physics*, Vol. 90(1989), pp. 3407-3408.

[72] L. T. Watson, R.T. Haftka, F.H. Lutze, R.H. Plautt, and P.Y. Shin, The Application of Globally Convergent Homotopy Methods to Nonlinear Optimization, In *Advances in Numerical Partial Differential Equations and Optimization*, S. Gomez, J.P. Hennart, and R.A. Tapia (eds.), SIAM, Philadelphia, PA, 1991, pp. 284-298.

[73] P.K. Weiner and P.A. Kollman, AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions, *Journal of Computational Chemistry*, Vol. 2(1981), pp. 287-303.

[74] L.T. Wille, Minimum-Energy Configurations of Atomic Clusters: New Results Obtained by Simulated Annealing, *Chemical Physics Letters*, Vol. 133(1987), pp. 405-410.

[75] K.G. Wilson, The Renormalization Group: Critical Phenomena and the Kondo Problem. *Reviews of Modern Physics.*, Vol. 47(1975), pp. 773-840.

[76] S.R. Wilson and W. Cui, Applications of Simulated Annealing to Peptides, *Biopolymers*, Vol. 29(1990), pp. 225-235.

[77] Z. Wu, The Effective Energy Transformation Scheme as a General Continuation Approach to Global Optimization with Application to Molecular Conformation, *Technical Report CTC93TR143*, Advanced Computing Research Institute, Cornell University, Ithaca, 1993.

[78] G.L. Xue, R.S. Maier, J.B. Rosen, Minimizing the Lennard-Jones Potential Function on a Massively Parallel Computer, in Proceedings of *1992 ACM International Conference on Supercomputing*, pp. 409-416, ACM Press, 1992.

[79] G.L. Xue, Molecular Conformation on the CM-5 by Parallel Two-Level Simulated Annealing, *Journal of Global Optimization*, in press.

[80] G.L. Xue, Improvement of the Northby Algorithm for Molecular Conformation: Better Results, accepted for publication in *Journal of Global Optimization*.