Global Optimization For Molecular Clusters Using A New Smoothing Approach *

CHUNG-SHANG SHAO, RICHARD BYRD, BETTY ESKOW and ROBERT B. SCHNABEL
Dept of Computer Science, University of Colorado, Boulder, CO 80309-0430, U.S.A.

Abstract. Strategies involving smoothing of the objective function have been used to help solve difficult global optimization problems arising in molecular chemistry. This paper proposes a new smoothing approach and examines some basic issues in smoothing for molecular configuration problems. We first propose a new, simple algebraic way of smoothing the Lennard-Jones energy function, which is an important component of the energy in many molecular models. This simple smoothing technique is shown to have close similarities to previously-proposed, spatial averaging smoothing techniques. We also present some experimental studies of the behavior of local and global minimizers under smoothing of the potential energy in Lennard-Jones problems. An examination of minimizer trajectories from these smoothed problems shows significant limitations in the use of smoothing to directly solve global optimization methods.

Keywords: Global Optimization, Molecular Chemistry, Lennard-Jones Problems, Smoothing Techniques

1. Introduction

The topic of this research is the development of new smoothing methods for large-scale global optimization problems arising in molecular chemistry applications. Molecular conformation problems give rise to very difficult global optimization problems, and smoothing is a technique that has been used in the chemistry and optimization communities to aid in their solution. Our overall approach is to develop new smoothing techniques that are both effective and inexpensive, and also to consider the integration of smoothing with sophisticated global optimization algorithms.

This paper describes the first stage of this research, the development of new smoothing methods to solve the Lennard-Jones problem. The Lennard-Jones problem is an important molecular conformation test problem for two reasons. First, the problem is a very difficult global optimization problem, since it is believed that the number of its minima

* Research supported by AFOSR Grants No. AFOSR-90-0109 and F49620-94-1-0101, ARO Grants No. DAAL03-91-G-0151 and DAAH04-94-G-0228, and NSF Grant No. CCR-9101795.

grows exponentially as $O(e^{N^2})$ [12], and many minima have energy values near the global minimum. Secondly, its objective function, the Lennard-Jones potential energy function, commonly exists within other molecular conformation problems, such as protein folding problems. In part for this reason, the techniques developed in this paper can also be used in smoothing a wide range of molecular conformation problems including protein folding problems.

The Lennard-Jones problem is to find the minimum energy structure of a cluster of N identical atoms using the Lennard-Jones potential energy. That is, the problem assumes that the potential energy of the cluster is given by the sum of the pairwise interactions between atoms, with these interactions being Van der Waals forces given by the Lennard-Jones 6-12 potential

$$p(r) = \frac{1}{r^{12}} - 2 * \frac{1}{r^6} \tag{1}$$

where r is the distance between two atoms. This potential represents a repulsive-attractive force that is very repulsive at very short distances, most attractive at an intermediate distance, and a very weak attractive force at longer distances. In this formulation, the pairwise equilibrium distance (distance of greatest attraction) is scaled to 1, and the pairwise minimum energy is scaled to -1. If we define the position of the cluster by

$$x = (x_1, x_2, ..., x_N)$$

where x_i is a three dimensional vector denoting the coordinates of the i^{th} atom, then the overall potential energy function is

$$f(x) = \sum_{i=2}^{N}\sum_{j=1}^{i-1} p(d_{ij}) = \sum_{i=2}^{N}\sum_{j=1}^{i-1}\left[\frac{1}{{d_{ij}}^{12}} - 2 * \frac{1}{{d_{ij}}^{6}}\right] \tag{2}$$

where d_{ij} is the Euclidean distance between x_i and x_j. We can now denote the problem as

$$\mathbf{LJ} : \min_{x \in D} f(x) \tag{3}$$

where $f(x)$ is function (2) and D is some closed region in R^n, $n =$ 3N. This problem has been the target of many different computational approaches, such as [2] [7] [9] [10] [11] [12] [13] [15] [16] [20] [23] [25], that make varying amounts of use of the solution structure of Lennard Jones clusters. A large scale global optimization algorithm that does not utilize the solution structure of the clusters has been developed in past few years by [2], and has successfully solved all Lennard-Jones problems with up to 76 atoms.

In practice, interesting molecular conformation problems contain at least 1,000 to 10,000 atoms, and huge numbers of local minimizers. It is expected that it will generally be too expensive to solve problems of such size by using global optimization algorithms directly on the objective function. This realization motivates approaches that seek to improve the effectiveness of global optimization algorithms via transformations of the objective (energy) function.

One transformation method is to use a parameterized set of *smoothed* objective functions. The smoothed functions are intended to retain the coarse structure of the original objective function, but have fewer local minimizers. By selecting different smoothing parameters, objective functions with different degrees of smoothness can be derived. The intent is to first solve the global optimization problem on very smooth problems, and then use this solution to gradually solve the problem on less smooth problems and ultimately the original, unsmoothed problem. The connections between the smoothed problems and the original problems are characterized by *trajectories* of minima in the space of the smoothing parameter(s). Along the trajectories, each point is a minimizer of a smoothed problem with some specific smoothing parameter setting. Usually, one end of the trajectory is a minimizer of the original problem. That is, once a minimizer of a smoothed problem is found, its trajectory usually leads to a corresponding minimizer of the original problem. If a problem is *strongly scaled*, as defined in [17], the trajectory from the global minimizer of a very smoothed problem leads directly to the global minimizer of the original problem. It is also quite possible, however, that the trajectory from the global minimizer of a very smoothed problem will not lead to the global minimizer of the original problem. Indeed, this paper will show that the latter situation appears to be common in Lennard-Jones problems.

At least three distinct smoothing methods have been proposed and applied to Lennard-Jones problems recently. These are the diffusion equation method [14], the effective energy method [6] [7] [20], and the effective energy transformation scheme [8] [24]. All of the above methods transform the original function into a family of smoothed functions via integration of the original objective function. Such integrations are too expensive to compute at run time. For use in global optimization algorithms, either some approximations must be employed, or look-up tables have to be computed in advance. These smoothing functions are described in Section 2 of this paper.

In Section 3 of this paper, a new family of smoothing functions is introduced. As opposed to previous methods, the new functions do not use integrations. Instead, the transformation is performed directly in an algebraic form. Another important difference between the new method and previous methods is that our transformation is specially designed to be applied on Lennard-Jones-like functions. We will show, however, that the techniques that are introduced also apply to other important functions, and that they allow our techniques to be applied to many important empirical energy functions, such as the common empirical energy functions for proteins.

In order to employ the smoothing techniques to solve global optimization problems, it is crucial to understand the behaviors of minimizer trajectories in smoothing parameter space. Section 4 of this paper summarizes some basic properties of minimizer trajectories. In Section 5, we discuss the design and results of experiments applied to both the integral form smoothing functions and the new algebraic smoothing functions. The results of experiments show that tested smoothing techniques have several common behaviors. Importantly, these behaviors imply that there are significant limitations to global optimization algorithms that use simple trajectory tracking approaches.

2. Spatial Averaging Smoothing Techniques

The basic idea of smoothing is to soften the original function by reducing abrupt function value changes while retaining the coarser structure of the original function. In other words, smoothing dampens high gradient values and fine grain fluctuations in the original function. As a result, nearby minimizers will merge after sufficient smoothing is applied to remove the barriers between them. Therefore, smoothing reduces the total number of minima in the problem.

The Lennard-Jones pairwise potential (1) has a pole at distance zero, and thus very large derivative values for distances near zero. The pole and large gradient values create huge barriers that separate similarly structured minimizers in the Lennard-Jones problems. This is the fundamental reason why the Lennard-Jones problem, as well as more complex problems that include the Lennard-Jones potential or similar ones, have so many minima. A Lennard-Jones smoothing technique should remove these barriers in some effective way.

In general, a smoothing technique is based upon a family of functions that is parameterized over a smoothing parameter or a set of smoothing parameters to generate different smoothnesses. Such a family can be represented as

$$\tilde{f}_s : D \to R, \quad s \in S \tag{4}$$

where D is some closed region in R^n, m is the number of smoothing parameters, and S is some subregion of R^m. By varying the smoothing parameter set s, one can create a series of functions that gradually smoothes the original function. For the Lennard-Jones problem, a family of smoothed Lennard-Jones problems can be constructed,

$$\mathbf{LJ}_s : \min_{x \in D} \tilde{f}_s(x) \tag{5}$$

where s is some smoothing parameter set and $\tilde{f}_s$ is a smoothed Lennard-Jones potential function. The intent is that the number of minima is reduced gradually as the objective functions become smoother. It is hoped that often, the smoothed problem will have only one minimizer once the degree of smoothing is sufficiently large.

A general smoothing technique, called spatial averaging, has been studied in various ways in [6] [7] [8] [14] [20] [24]. The fundamental idea of this technique is that the smoothed function value at each point is given by a weighted average of the energy function in a neighborhood of this point determined by a distribution function centered at this point. The Gaussian distribution function is commonly used. In this case, the smoothing transformation is

$$\tilde{f}_s(x) = \int H(f(x'), s') \cdot \exp\left(\frac{-\|x - x'\|^2}{\lambda^2}\right) dx' \tag{6}$$

where λ and s' together constitute the smoothing parameter set, s. The parameter λ determines the scale of the Gaussian distribution, while the parameter s' is used within the function H to transform the original function $f(x)$ into a function that has no poles. The transformation $H(f, s')$ is necessary to make the function integrable. Without this transformation, the Lennard-Jones potential is not integrable due to the infinite value at zero distance. The cost of performing the multidimensional integral (6) is reasonable due to the partially separable form of the objective (2).

Two families of smoothing functions using the spatial averaging technique, the diffusion equation method [14] and the effective energy transformation scheme [24], are experimented within this paper to compare them with the new function family. Both families are based on function (6), however, they differ in how the transformation $H(f, s')$ is performed. The rest of the section briefly discusses these two families of smoothing functions.

Diffusion Equation Method. The diffusion equation method [14] transforms a function to be minimized according to a partial differential equation. In this paper, the function to be transformed is the Lennard-Jones potential $f(x)$ in (2). The method is applied by solving the diffusion equation

$$\Delta F = \partial F / \partial t \qquad t > 0 \tag{7}$$

for $F(x,t)$, where Δ is the Laplacian operator with

$$F(x,0) = f(x). \tag{8}$$

The parameter t takes on the meaning of time.

If the Lennard-Jones potential $f(x)$ is viewed as an initial temperature field in (8), the solution of equation (7), $F(x,t)$, can be seen as expressing how the temperature field reaches an equilibrium state through time. This process is equivalent to a spatial averaging process, and $F(x,t)$ can be regarded as a family of functions of x parameterized by t. As t increases, smoother functions are obtained. Eventually, beyond some time t^*, the temperature field becomes uniform. The parameter t is a smoothing parameter for our purpose, where $t = 0$ represents the original Lennard-Jones potential and $t = t^*$ gives the smoothest function.

Equation (7) has an analytical solution

$$F(x,t) = \left[2(\pi t)^{1/2}\right]^{-n} \int_{R^n} f(y) \cdot \exp\left(-\frac{1}{4t}\|x-y\|^2\right) dy \tag{9}$$

where $\|x-y\|$ is the length of the vector of $x-y$. There are two difficulties in applying the above integration to the Lennard-Jones potential (2). First, the Lennard-Jones potential contains infinite poles which make (9) non-integratable. Secondly, it would be costly to evaluate the integration numerically if an analytical formula were not available.

To overcome these difficulties, the diffusion equation method uses a sum of Gaussians to approximate the Lennard-Jones pairwise potential (1) throughout the region of physical interest,

$$p(r) \approx \sum_{k=1}^{L} a_k \cdot \exp\left(-b_k r^2\right) \tag{10}$$

where L is large enough to describe the potential accurately. Reference [14] shows that two terms are sufficient to approximate the Lennard-Jones pairwise potential within the interesting physical range. However, it is necessary to add a third term to ensure that the energy surface will only contain one minimum at a highly smoothed level before it becomes completely uniform. Therefore, $L = 3$ is used for experiments in both [14] and this paper. Using (10), the full Lennard-Jones potential is approximated by

$$f(x) \approx \sum_{i=2}^{N} \sum_{j=1}^{i-1} \sum_{k=1}^{L} a_k \cdot \exp(-b_k \|x_i - x_j\|^2) \tag{11}$$

which no longer contains poles. This approximation is the transformation H in equation (6).

By substituting (11) into (9), it is straightforward to solve the integration. The solution of (9) is a family of smoothing functions for the Lennard-Jones potential,

$$\tilde{f}_t(x) = \sum_{i=2}^{i} \sum_{j=1}^{i-1} U_{ij}(d_{ij}, t) \tag{12}$$

where

$$U_{ij}(d_{ij}, t) = \sum_{k=1}^{L} a_k (1 + 8b_k t)^{-\frac{3}{2}} \cdot \exp\left(\frac{-b_k d_{ij}^2}{1 + 8b_k t} \right) \tag{13}$$

and $d_{ij} = \|x_i - x_j\|$. Now, function (13) can be practically evaluated in global optimization computations.

Effective Energy Transformation Scheme. Unlike the diffusion equation method, the effective energy transformation scheme [24] directly smoothes a function by using spatial averaging. The transformation is similar to (6),

$$\tilde{f}_\lambda(x) = C_\lambda \int f(x') \cdot \exp\left(\frac{-\|x - x'\|^2}{\lambda^2} \right) \, dx', \tag{14}$$

where λ is a positive number and C_λ is a normalization constant such that

$$C_\lambda \int \exp\left(\frac{-\|x\|^2}{\lambda^2} \right) \, dx \;= 1. \tag{15}$$

Due to the integration in (14), the effective energy transformation has the same difficulties as the diffusion equation method. However, the effective energy transformation scheme chooses another approach to solve the difficulties. In [24], it shows that if the objective function is partially separable, (14) is also partially separable and the computation will be much simpler. Since the Lennard-Jones potential (2) is a partially separable function, the transformation (14) can be applied on the Lennard-Jones potential:

$$\tilde{f}_\lambda(x) = C_\lambda \int \left(\sum_{i=2}^{N} \sum_{j=1}^{i-1} p(\|r_{ij}\|) \right) \cdot \exp\left(\frac{-\|x - x'\|^2}{\lambda^2} \right) \, dx' \tag{16}$$

$$= \sum_{i=2}^{N} \sum_{j=1}^{i-1} < p >_{\sqrt{2}\lambda} (\|r_{ij}\|) \tag{17}$$

where $r_{ij} = x_i - x_j$ and

$$< p >_{\sqrt{2}\lambda} (\|r_{ij}\|) = C_{\sqrt{2}\lambda} \int p(\|r'_{ij}\|) \cdot \exp\left(-\frac{\|r_{ij} - r'_{ij}\|^2}{2\lambda^2}\right) dr'_{ij}. \quad (18)$$

From (16) to (17), a multi-dimension integration becomes a summation of one-dimensional integrations.

In (18), the pairwise potential, p, still contains a pole at $r = 0$. One has to remove this pole from function p for (18) to be integratable. To solve this difficulty, the effective energy transformation scheme replaces the pairwise potential p with

$$p'(r) = \begin{cases} p(r), & r \geq r_{min} \\ h_a + h_b \cdot \sqrt{1 - r^2/r_a^2}, & r < r_{min} \end{cases} \quad (19)$$

where r_{min} is a cutoff distance determined by a chosen cutoff value on $p(x)$, and h_a, h_b and r_a are constants determined such that the two functions in (19) connect smoothly at $r = r_{min}$. By replacing p with p', (18) becomes integratable with function p'.

In contrast to the diffusion equation method, the integral in (18), with $p'(r)$ replacing $p(r)$, is still not analytically integrable. To avoid the expense of evaluating this integral numerically at runtime for each value of r, the effective energy transformation scheme, therefore, evaluates the integral at many (200 per λ value) points in advance and creates a two-dimensional lookup table, with parameters r and λ. In global optimization computations, all function evaluations are computed through interpolating the data in the lookup table with cubic splines.

3. A New Family of Smoothing Functions

In this section we introduce a new, algebraic method of smoothing the Lennard-Jones potential energy function and compare the characteristics of this method with the spatial averaging approach.

The new family of smoothing functions for the Lennard-Jones pairwise potential is

$$\tilde{p}(r, P, \gamma) = (\frac{1 + \gamma}{r^P + \gamma})^2 - 2 * (\frac{1 + \gamma}{r^P + \gamma}) \quad (20)$$

where r is the distance between atoms, and γ and P are the two smoothing parameters. This equation is equal to the original Lennard-Jones pairwise potential (1) when $\gamma = 0$ and $P = 6$. For any values of $\gamma >= 0$ and $P > 0$, it attains its minimum value of -1 at $r = 1$, as does the Lennard-Jones pairwise potential. Furthermore, function (20) only contains simple algebraic computations, and is nearly as inexpensive to evaluate as the original Lennard-Jones pairwise potential (1).

From (20), we can define the family of smoothed Lennard-Jones potentials as

$$\tilde{f}_{<P,\gamma>}(x) = \sum_{i=1}^{N} \sum_{j=2}^{i-1} \tilde{p}_{<P,\gamma>}(d_{ij}) \tag{21}$$

and the smoothed Lennard-Jones problem by

$$\mathbf{LJ}_{<P,\gamma>} : \min_{x \in D} \tilde{f}_{<P,\gamma>}(x). \tag{22}$$

The two smoothing parameters in (20), P and γ, serve two different roles. As P becomes smaller than six, the function becomes smoother and the basin around its minimizer becomes broader. Fig. 1.1 illustrates this, showing the potential between two atoms for different values of P while γ is fixed at zero. The values of the function near the equilibrium distance of one clearly are reduced as P decreases. However, the function value still goes to infinity as the distance approaches zero.

The smoothing parameter γ has the effect of removing the pole from the Lennard-Jones potential. Fig. 1.2 shows the smoothed function for various values of γ, with $P = 6$ in all cases. This figure illustrates that the y-intercept of the smoothing function, $\tilde{p}(0, P, \gamma)$, decreases as γ increases. In addition, increasing γ reduces the values of the function for $r < 1$, but has minimal effect for $r > 1$. The relation between γ and $\tilde{p}(0, P, \gamma)$ can be easily obtained from equation (20) and is given by

$$\tilde{p}(0, P, \gamma) = \frac{1}{\gamma^2} - 1. \tag{23}$$

It should be noted that the relation (23) is not a function of P and that $\tilde{p}(0, P, \gamma)$ approaches -1 as γ approaches infinity. In addition, when $\gamma > 0$ and $P > 1$, equation (20) has a zero derivative at $r = 0$. This implies that the complete smoothed function given by (21) is continuously differentiable. Fig. 1.3 clearly illustrates the independence of the y-intercept upon the parameter P, as well as the zero derivative at $r = 0$.

There are four significant differences between the proposed smoothing functions and the spatial averaging smoothing functions. First, the proposed formula (20) is cheaper to evaluate. This is important for molecular configuration global optimization problems, since energy functions may be evaluated hundreds of thousands or millions of times. Compared to the diffusion equation method (13), the proposed formula only contains one power function but (13) has three exponential

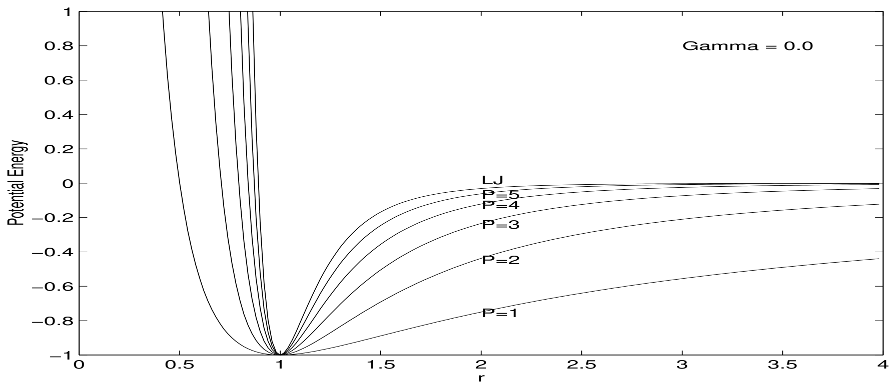


Figure 1.1. Potential Energy Curves Between Two Atoms on Different Settings of P at $\gamma = 0.0$

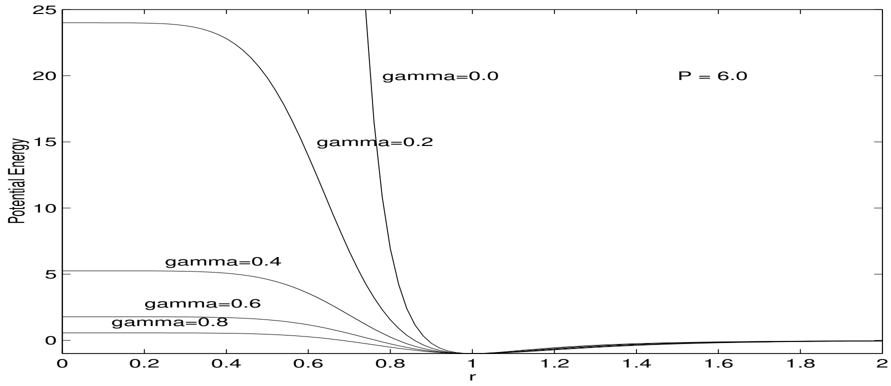


Figure 1.2. Potential Energy Curves Between Two Atoms on Different Settings of γ at P = 6.0

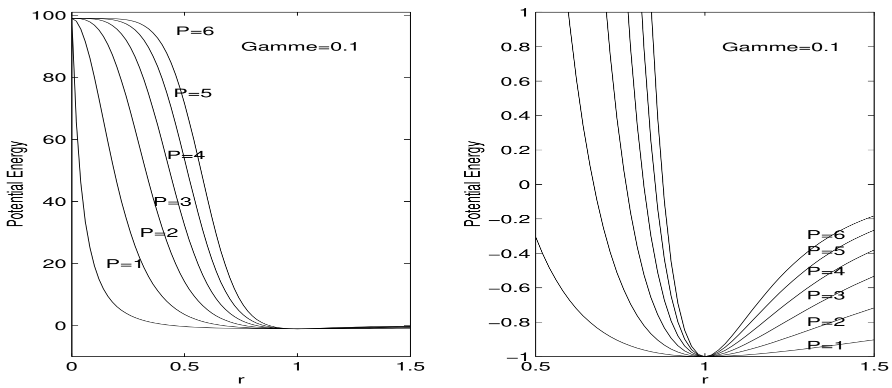
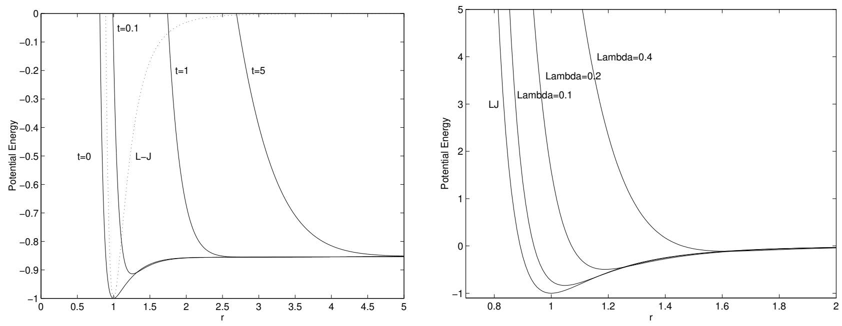


Figure 1.3. Potential Energy Curves Between Two Atoms on Different Settings of P at $\gamma = 0.1$



(a) (b)

Figure 2. Pairwise Potential Energy Curves on Different Settings for Spatial Averaging Smoothing Function Families. (a) Diffusion Equation Method, (b) Effective Energy Transformation Scheme.

function evaluations. On the other hand, the effective energy transformation scheme evaluates (18) through interpolating a pre-computed table which involves considerably more computation per value of r than (20). In addition, for large global optimization problems, the precomputed table must be expanded to cover larger distance, r, and more smoothing levels, λ. The table then consumes more memory space as the problem size grows larger.

The second difference is that the proposed pairwise formula's minimum does not change location as its smoothing parameter varies. It is fixed at $r = 1$ and the value is -1 in (20). Due to the integration, the spatial averaging technique shifts the location and the value of the minima in the pairwise potential. The curves of two spatial averaging function families are shown in Fig. 2. This figure implies that the distance between atoms of optimal configuration of the spatial averaging function at different smoothing levels will grow larger. This characteristic will affect the use of spatial averaging smoothing schemes, particularly by making trajectory tracking require smaller steps in the smoothing parameter space. In addition, for the effective energy transformation scheme, one must be sure that the pre-computed table covers sufficient distance for all the smoothing levels that are used.

Thirdly, the proposed smoothing technique transforms the Lennard-Jones potential without special treatment of the pole. In contrast, the diffusion equation method must use a Gaussian series approximation and the effective energy transformation scheme has to truncate the Lennard-Jones potential more directly to perform the integration. Note that Fig. 2b shows a large difference between the potential of diffusion equation method at $t = 0$ and the Lennard-Jones potential. The difference is mainly from the third approximation term in function (10).

Finally, the proposed method as described so far is restructed to the Lennard-Jones potential, while the diffusion equation method and effective energy transformation scheme are quite general. However, the new approach taken in equation (20) can readily be extended to other functions of $(1/r)$, such as the electrostatic term in empirical potentials for proteins. As we mention briefly in Section 6, this property has enabled us to apply our new analytic smoothing approach to general protein conformation problems (see [1]).

In [21], we have investigated how the shapes of our new family of smoothing functions compare with the class of functions in Fig. 2b from the effective energy transformation scheme. If we allow a domain scaling to make the locations of the minimizers the same, we show that there exist parameter choices in (20) for which (20) and the functions from Fig. 2b (using values of λ considered in papers on this approach) are nearly indistinguishable over the range of values of r for which the functions are evaluated in practice. The functions do differ for values of r very close to 0. We would not expect our class of functions to be as similar in a graphical sense to the set of diffusion equation approximations shown in Fig. 2a, since the latter differs so significantly from the true Lennard-Jones potential as $t \to 0$. Still, in the Section 5 we will see that our new family and the diffusion equation family behave quite similarly in trajectory-tracking experiments on a Lennard-Jones cluster.

While making the present study, we discovered a recent paper [18] that proposes a smoothing function quite similar to (20). They propose a smoothing function where r in (1) is replaced with $(r+\alpha)/(1+\alpha)$ for some positive constant α. Note that our smoothing function (20) may be viewed as replacing r^6 in (1) by $(r^P+\gamma)/(1+\gamma)$.

4. Behaviors of Minimizer Trajectories in Smoothed Space

It is important to understand the effects of the smoothing techniques on the objective functions, since the surface of smoothed functions can be dramatically different from the original functions. To explore the behaviors of smoothing techniques, minimizer trajectory tracking is a useful tool, since it is the behaviors of these minimizers in smoothing space which are important in global optimization. The behaviors of minimizer trajectories can directly affect the approaches and performances of global optimization algorithms that utilize smoothing techniques.

As mentioned in Section 1, varying the smoothing parameters gives rise to trajectories in the space of the smoothing parameters that link minimizers of functions with different amounts of smoothing. Each point on a given trajectory is a minimizer of some smoothed problem. These trajectories can be denoted as

$$traj_i : S \to D. \tag{24}$$

Since each point on a given trajectory is a minimizer, the points on the trajectory satisfy

$$\nabla_x \tilde{f}_s(traj_i(s)) = 0. \tag{25}$$

In practice, once one has found a local minimizer for a given value of the smoothing parameter(s), one can trace its trajectory by slowly changing the smoothing parameter and conducting a series of local minimizations. Each local minimization uses the minimizer from the previous value of the smoothing parameter(s) as the starting point for minimizing the function with the new smoothing parameter value.

In the best case, the trajectory containing the global minimizer of highly smoothed problems will also contain the global minimizer of the original problem. In this case, by finding the global minimizer of a very smooth problem (which hopefully is fairly easy to do), and then tracking this trajectory through a careful sequence of local minimizations on less and less smoothed functions, one can locate the global minimizer of the original problem. However, this relationship between the global minimizer of very smoothed problems and the global minimizer of the original problem does not always exist. In order to understand the incorporation of smoothing in global optimization, one needs to be aware of the possible pitfalls in trajectory tracking.

The global minimizer of the original problem can fail to lie on the same trajectory as the global minimizer of very smoothed problems for one of at least four reasons. The most important of these is that the order of minimizers may change as the function becomes smoother, due to the function value surface being transformed by different amounts in different regions. For example, as the function is smoothed by spatial averaging smoothing techniques, a low minimizer that is surrounded by very high barriers may become a higher minimizer than an initially somewhat-higher minimizer that is surrounded by much lower barriers. The trajectories containing these minimizers will *flip orders* at some point. If the trajectory containing the global minimizer of the original problem flips its order with another trajectory, then tracking the global minimizer of a very smooth problem back will not lead to the global minimizer of the original problem. The experiments in this paper indicate that this situation appears to be common.

Secondly, any trajectory may terminate beyond a given set of smoothing parameter values. This occures when a minimizer is smoothed away beyond these smoothing parameter values. For the trajectory of the global minimizer of the original problem, it is easy to see that this condition can only present if the trajectory has first flipped its order as described previously. When such termination occurs, the process of tracking to smoother objectives will generally jump to a different trajectory and continue. If the trajectory from the global minimizer of the original objective terminates in this way, then the global minimizer cannot be found by continuous tracking from any smoother minimizer.

The third possibility is the analogous termination of the trajectory from a smoothed minimizer as the parameters make the function less smooth. This behavior, which we believe to be less common, occurs when smoothing introduces a new minimizer at some point. The experiments in this paper verify that this may occur. If this happens for a trajectory coming from the global minimizer of a highly smoothed problem, no trajectory can lead continuously to the global minimizer of the original problem. Finally, two trajectories may merge as the function becomes more smoothed, if two minimizers merge. This behavior appears to be much less common, and can be treated as a variant of trajectory termination.

Related behaviors to those just described have been reported in the paper [17] on state trajectories. State trajectories are intended to capture the transitions between *valleys* of minimizers in the smoothing hyperspace. Three types of patterns are described in [17]: *strong scaling*, *weak scaling*, and *poor scaling*. These correspond to the abilities of trajectories of smoothed problems in the formulation of [17] to lead to the original global minimizer, with strong scaling corresponding to the "best" case mentioned above.

In the next section, a series of trajectory tracking experiments is presented. These experiments are based on Lennard-Jones cluster problems. The results show the behaviors discussed in this section commonly appear in different sizes of the Lennard-Jones problems. For example, Fig. 5 will illustrate that each of the first three types of behaviors mentioned above occurs in a 30 atom Lennard-Jones cluster problem.

5. Trajectory Tracking Experiments

We now present results of trajectory tracking experiments with Lennard-Jones clusters of 9, 30 and 34 atoms, that are designed to explore the behaviors of the proposed new smoothing techniques. From the range of experiments in this section we are able to make meaningful observations about smoothing techniques for molecular configuration problems in general, and about our new family of smoothing functions.

We also present results of trajectory tracking experiments on the 9 atom problem using the spatial averaging smoothing techniques discussed in Section 2, which are intended to compare the spatial averaging techniques with the proposed new method. Among these experiments, the results from the 9 atom problem are most important. The problem size, nine, is small enough to find and track all of its minimizers and is complex enough, at the same time, to yield all the interesting behaviors discussed in Section 4; and allows for meaningful comparisons between smoothing techniques.

First we describe our experimental methodology. Each of our trajectory tracking experiments starts by using a global optimization program to find a set of minimizers with the original Lennard-Jones function. This global algorithm, which is described in [2], has successfully found the best-known minimizers for all Lennard-Jones problems with up to 76 atoms. It is believed that this program finds all the minimizers of the 9 atom problem, as discussed later, and all 20 minimizers found are used in the trajectory tracking experiments. The 30 and 34 atom problems each have a huge number of minimizers and only a tiny portion of the minimizers in each problem are used in the trajectory tracking experiments. Due to the characteristics of the global optimization algorithm, however, most minimizers used in both experiments are low energy minima. And, in both problems the best known minimizers are found and used.

After obtaining the minimizers, a tracking procedure is performed on each of these minimizers. Each minimizer is tracked carefully through different values of the smoothing parameters P and γ for the proposed method (or t for diffusion equation method and λ for the effective energy transformation scheme). The tracking procedure for our new function is a series of local minimizations with objective function (21) through a sequence of smoothing parameter sets,

$$\{(P_0, \gamma_0), (P_1, \gamma_1), (P_2, \gamma_2), ...\};$$

where $P_0 = 6, \gamma_0 = 0$ and (P_{i+1}, γ_{i+1}) is smoother than (P_i, γ_i), *ie* $P_{i+1} < P_i, \gamma_{i+1} > \gamma_i$. Analogous procedures are used for the spatial averaging smoothing families. The local minimizations were performed using the BFGS method of the UNCMIN package [19]. Great care was taken to assure that the step sizes in the smoothing parameter space are small enough that the tracking procedure is stable. If the step size is too large, the tracking process may "jump" off of the current trajectory to another trajectory due to the large change in the potential energy landscape. In this procedure, trajectory terminations can be identified by seeing that two or more trajectories merge as the parameters become smoother.

Table I. Number of trajectories with different values of smoothing parameters. A number with bold type denotes that the trajectory of original global minima is no longer global under the smooth setting.

γ	f(0)	P = 6.0	P = 5.0	P = 4.0	P = 3.0	P = 2.0
0.0	inf	20	16	11	**6**	**4**
0.1	99.0	16	15	9	**6**	**1**
0.2	24.0	16	12	**7**	**5**	**1**
0.3	10.1	16	10	**6**	**4**	**1**
0.4	5.25	14	8	**5**	**1**	**1**
0.5	3.0	11	6	**5**	**1**	**1**
0.6	1.78	10	**5**	**2**	**1**	**1**
0.7	1.04	8	**5**	**1**	**1**	**1**
0.8	0.56	6	**3**	**1**	**1**	**1**
0.9	0.23	**5**	**1**	**1**	**1**	**1**

Finally, a reverse tracking procedure is performed on the minizers which survive through the previous (forward) tracking procedure. This reverse tracking procedure, identifies which original minimizers *survive* through the smoothing process. Additionally, new trajectories that emerge in the forward smoothing process can be identified in the reverse process, by locating trajectories that merge with others as the parameters move to the less smooth direction. However, it is important to note that this procedure only finds new trajectories when original trajectories terminate and merge into them. That is, the experiments are not designed to explore the entire smoothed space to find all trajectories of minimizers. This implies that there might be more trajectories created than the numbers of new trajectories shown in the experiments. Still, experiments on 30 and 34 atoms do reveal this type of behavior.

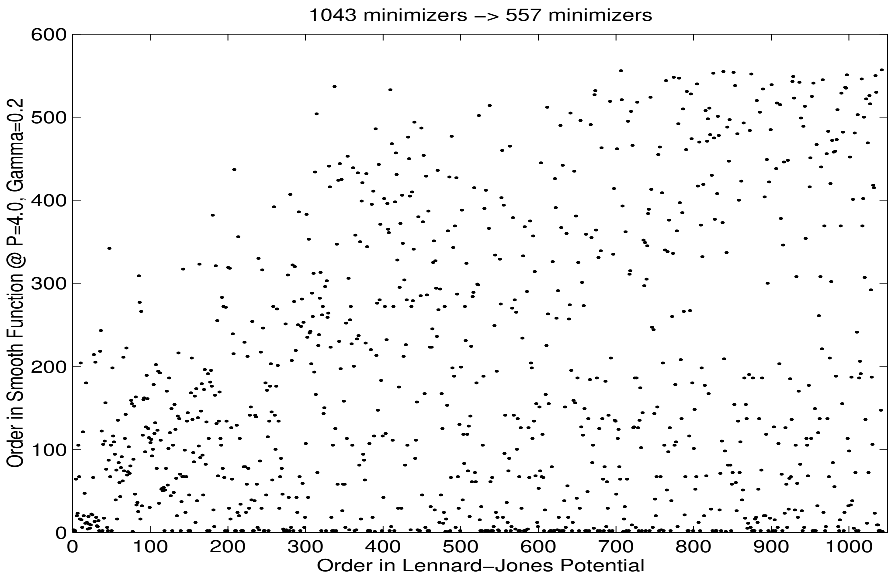
5.1. Experiments on the 9 atom Lennard-Jones problem

We first tried to understand the smoothing ability of the new function family. Our first experiment tracked all minimizers of a 9 atom Lennard-Jones problem to various smoothing parameter combinations. For the original Lennard-Jones problem, we were surprised to find twenty distinct minimizers, instead of the eighteen minimizers that are reported in [11] and have often been cited in the literature. This demonstrates the difficulty of solving the Lennard-Jones problem comprehensively. The reason the 9 atom cluster was chosen is that it is the smallest cluster that has enough minimizers to lead to a rich set of experiments.

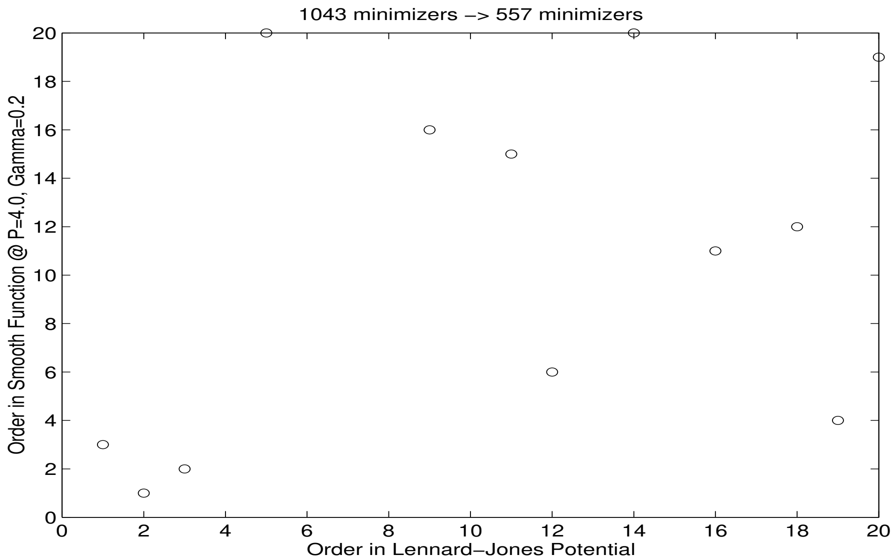
For different values of the smoothing parameters, the number of the original trajectories of the 9 atom problem that still exist at specific smoothing level is recorded in Table I. The steady reduction in the number of minimizers shown in Table I clearly shows that the new smoothing function (21) effectively smoothes the Lennard-Jones problem. The experiment also demonstrates the transitions of trajectories through smoothing parameter space. For example, four trajectories have either terminated or merged with other trajectories between ($P = 6, \gamma = 0.0$) and ($P = 5, \gamma = 0.0$), or between ($P = 6, \gamma = 0.0$) and ($P = 6, \gamma = 0.1$). Once a certain level of smoothing is reached or exceeded, only one minimizer is left, and all remaining trajectories have terminated.

Most importantly, however, Table I clearly illustrates one of the biggest limitations of the smoothing approach, *order flips*. This means that the orders of trajectories can change through smoothing parameter space. That is, at a given smoothing level, a trajectory A can have a higher smoothed minimizer than a trajectory B, even though the initial unsmoothed minimizer that started the trajectory A had a lower function value than the unsmoothed minimizer that started the trajectory B. In the worst case, the global minimizer's trajectory can drop its order dramatically as it is smoothed. If this happens, this means the global minimum can not be found by simply tracking the lowest minimizers of very smoothed problems backwards. In Table I, the smoothing levels where the global minimizer's trajectory is no longer the lowest have been denoted in boldface. It is seen that for this problem, this phenomenon *always* occurs by the time there are five or fewer trajectories remaining.

A more detailed analysis of the results of the 9 atom trajectory tracking experiment reveals that the trajectory of the unsmoothed global minimizer first drops in order to second place once a certain level of smoothing is reached. Then, at some greater value of smoothing, this trajectory terminates and local minimizations from it jump into the current smoothed global minimizer's trajectory. This trajectory is the one that originates from the second lowest minimizer of the original problem. A reverse tracking procedure has been performed on the only minimizer remaining at ($P = 2, \gamma = 0.9$). It is not surprising that the reverse tracking procedure leads to the second lowest minimizer of the original problem, instead of the global minimizer. This simple problem illustrates that if the global minimizer flips its order, the global minimum can not be found by only reverse tracking the best smoothed minimizer, and that if the global minimizer's trajectory terminates, the global minimum can not be found by reverse tracking from any minimizer of a very smoothed problem.

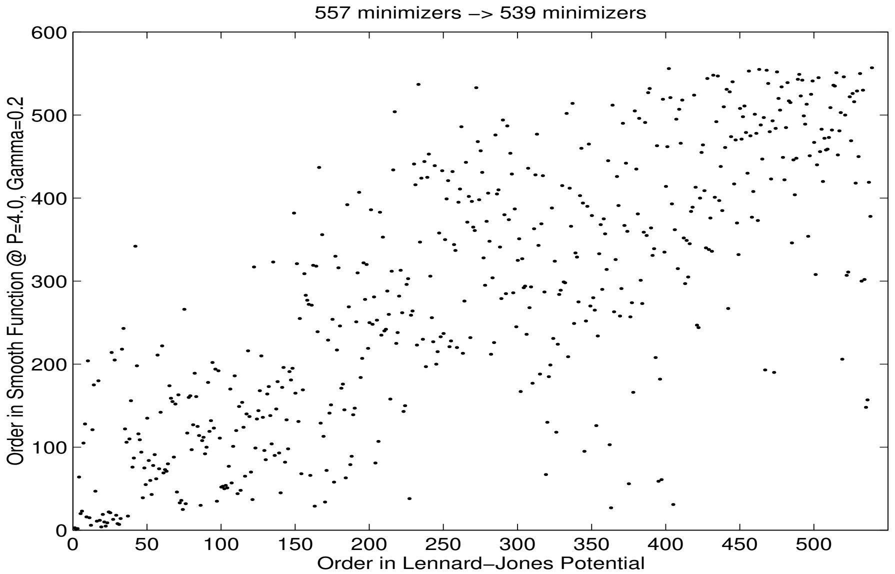


(a)

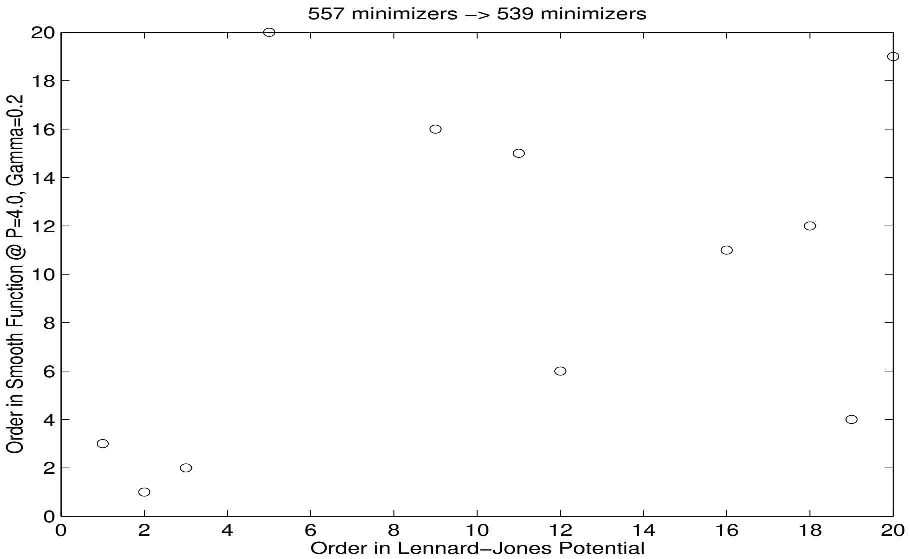


(b)

Figure 3. Order flipping experiment on the 30-atom problem: (a) 1,043 the Lennard-Jones minimizers are tracked to smoothing level $P = 4, \gamma = 0.2$. The horizontal axis represents the trajectory's order in the Lennard-Jones potential, and the vertical axis represents the trajectory's order in the smoothing function, $\tilde{f}_{<4.0,0.2>}$. (b) The first twenty minimizers of Fig. 3a.



(a)



(b)

Figure 4. Order flipping experiment with reverse tracking on the 30-atom problem: (a) 557 minimizers of smoothing level $P = 4, \gamma = 0.2$ are tracked back to the Lennard-Jones potential. The horizontal axis represents the trajectory's order in the Lennard-Jones potential, and the vertical axis represents the trajectory's order in the smoothing function, $\tilde{f}_{<4.0,0.2>}$. (b) The first twenty minimizers of Fig. 4a.

5.2. Experiments on large problems

Next, we examined the order flip phenomenon on a larger problem size. First, the Lennard-Jones problem with 30 atoms was tested. This size was chosen somewhat arbitrarily as a moderate sized cluster. We began by locating 1,043 of the lowest minimizers for this problem, including the global minimizer, using our global optimization algorithm discussed in the next section. Unlike the 9 atom problem, it is not feasible to find all the minima for the 30 atom problem. Indeed, according to [11], we have found only a tiny portion of all the minimizers. All 1,043 minimizers were tracked to the smoothing level (4.0,0.2) by the same procedure discussed earlier in this section. The order flip graphs are shown in Fig. 3a and 3b.

The order flips in the 30 atom problem are much more extreme than the flips in the 9 atom problem. This experiment clearly demonstrates that order flips among trajectories is a significant issue when using smoothing. Fig. 3b shows that the trajectory of the original global minimizer has dropped down to third place by smoothing level (4.0,0.2), More dramatically, the minimizers in fourth and sixth to eighth positions in the original function are all below twentieth place by this smoothing level. Similarly, the trajectories which are in seventh to tenth positions for smoothing level (4.0,0.2) were all below twentieth place for the original function. This means that if we run a global optimization algorithm at this smoothing level and then simply track the trajectories of the best smoothed minimizers back to the original function, we will miss some of the best Lennard-Jones minimizers and will find several poor Lennard-Jones minimizers. Furthermore, since the original global minimizer has already dropped in order, it is likely that its trajectory will terminate at some further smoothing level and that if we start reverse tracking from that level we could miss the original global minimizer entirely.

In Fig. 3, almost half of the minimizers' trajectories have terminated and collapsed into other minimizers' trajectories. From the data in these figures, we cannot tell which original trajectories have terminated and which have led directly to the minimizers of the smoothed problem, since there is no guarantee that the trajectory with the lower original minimizer is the one that exists longest. In order to tell exactly which original minimizers survived through the tracking process, we have conducted a reverse tracking procedure on each of the 557 distinct smoothed minimizers remaining in the 30 atom problem at smoothing level (4.0,0.2). The result is shown in Fig. 4. By removing the minimizers whose trajectories terminate during forward tracking, Fig. 4a clearly confirms the phenomenon of order flipping in the smoothed 30

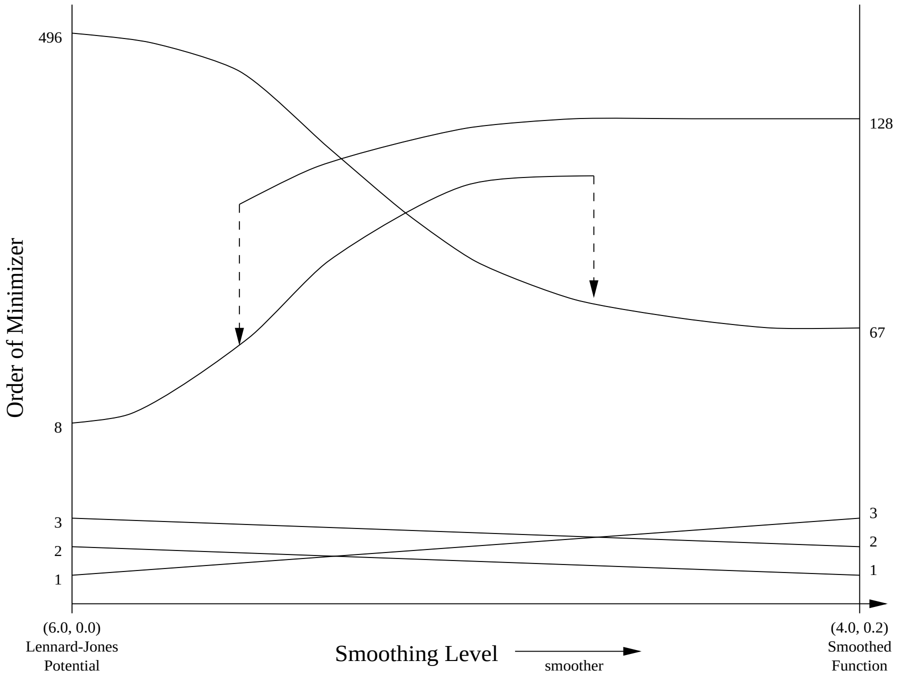


Figure 5. Trajectories of selected minimizers for 30 Atom problem.

Atom Lennard-Jones problem. That is, the reverse tracking experiment also produces a multitude of large order flips. In addition, Fig. 4b shows that the same pattern of flips is observed for the lowest minimizers in the reverse tracking experiment as in the forward tracking experiment. Taken together, these experiments indicate that order flips are likely to be an important issue to consider in using smoothing in global optimization algorithms.

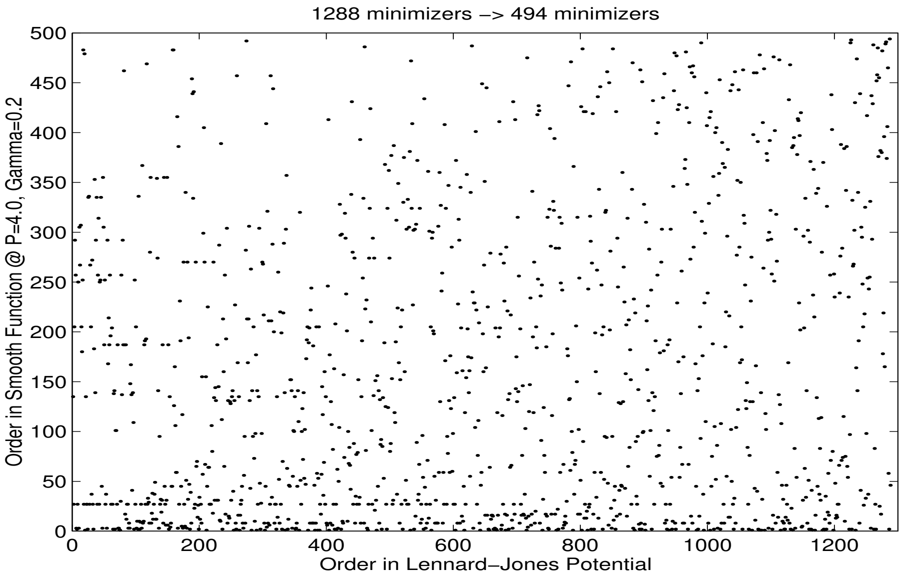
Another interesting observation from the reverse tracking experiment is that some trajectories terminate during the reverse tracking process. In this experiment, 18 of the 557 trajectories terminate before the original problem is reached. This termination as the problem becomes less smooth confirms one of the types of behavior of smoothing trajectories that was mentioned in Section 4, and implies the emergence of new trajectories as the problem becomes smoother. After such a trajectory terminates during reverse tracking, the tracking procedure will fall to another, lower trajectory. This behavior is far less common in the experiments than trajectories terminating during forward tracking, since it occurs only when a new minimizer *emerges* as the function becomes smoother, whereas smoothing tends to reduce the number of minimizers overall. A sampling of trajectories that shows all three behaviors - order flip, trajectory termination in the forward tracking,

Table II. Results of tracking best 20 Lennard-Jones minimizers for 34 atom problem to minimizers of smoothing level ($P = 4.0, \gamma = 0.2$), and results of tracking smooth minimizers that lead back to best 20 Lennard-Jones minimizers.
Trajectory tracking status: "Jump" - trajectory terminates and falls into other trajectory. "Cont" - trajectory is tracked to smoothing level (4.0,0.2) without termination. "None" - did not find trajectory leading to specified Lennard-Jones minimizers.

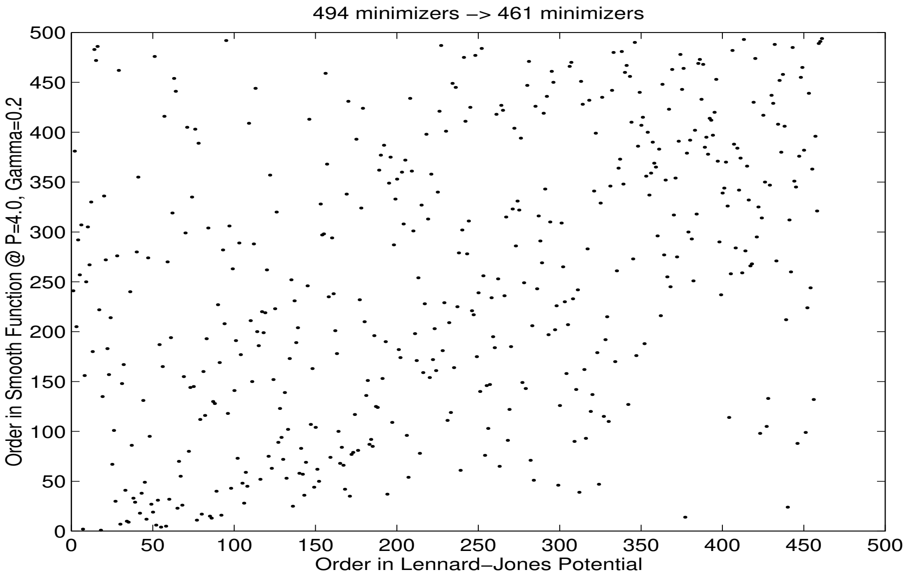
Best 20 of 1288 Original Lennard-Jones Minimizers		Best 20 LJ Min. → Smooth Min. @ (4.0,0.2)			Smooth Min. @ (4.0,0.2) → Best 20 LJ Min.	
		Trajectory Tracking Status	Final Smoothed Minimizer	Order in Smoothed Potential	Trajectory Tracking Status	Order in Smoothed Minimizers
1	-150.044528	Jump	-215.667165	135	Jump	241
2	-149.997007	Jump	-216.732431	27	Jump	381
3	-149.921622	Cont	-215.300955	205	Cont	205
4	-149.886921	Cont	-214.933377	292	Cont	292
5	-149.795387	Cont	-215.088910	257	Cont	257
6	-149.780940	Jump	-218.359634	3	None	
7	-149.672121	Cont	-218.390471	2	Cont	2
8	-149.494653	Jump	-216.732431	27	None	
9	-149.482330	Cont	-215.103792	250	Cont	250
10	-149.470798	Jump	-218.359634	3	None	
11	-149.434941	Cont	-214.873379	305	Cont	305
12	-149.223885	Cont	-215.060548	267	Cont	267
13	-149.162201	Jump	-214.844140	307	None	
14	-149.095462	Jump	-215.300955	205	None	
15	-149.093101	Cont	-215.427124	180	Cont	180
16	-149.091680	Jump	-215.101311	252	None	
17	-149.082624	Cont	-213.391740	483	Cont	483
18	-149.058111	Jump	-216.732431	27	None	
19	-149.054405	Cont	-213.646593	479	Cont	479
20	-149.010289	Cont	-219.480360	1	Cont	1

and trajectory termination in the reverse tracking - for the 30 atom problem is given in Fig. 5.

Next, we conducted the same experiments for the 34 atom Lennard-Jones problem as for the 30 atom problem. The 34 atom problem appears to be one of the most difficult Lennard-Jones problems among the first 75; the global minimizer, which has an energy value of -150.045, appears to be in a much narrower and more isolated basin of attraction than the second lowest minimizer with energy value -



(a)



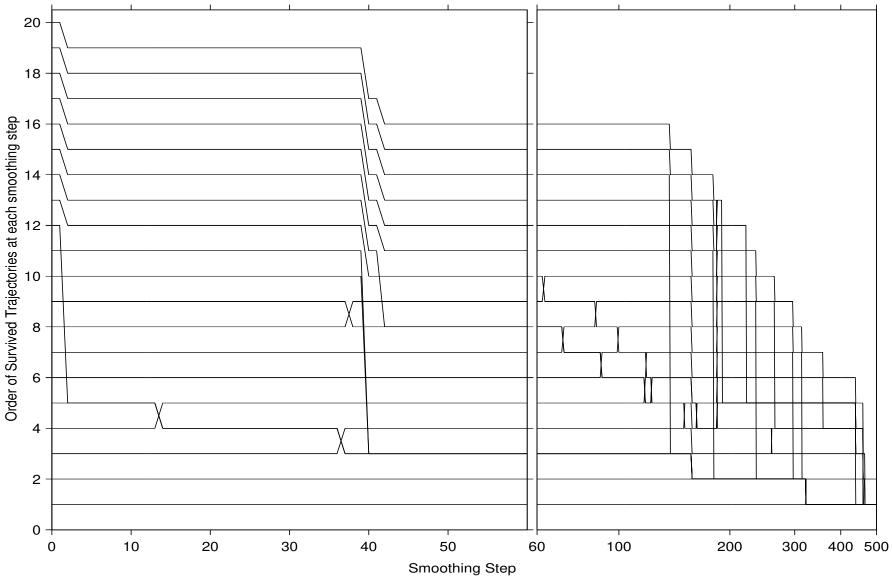
(b)

Figure 6. Order flipping experiment on the 34-atom problem: The horizontal axis represents the trajectory's order in the Lennard-Jones potential, and the vertical axis represents the trajectory's order in the smoothing function, $\tilde{f}_{<4.0,0.2>}$. (a) 1288 Lennard-Jones minimizers are tracked to smoothing level ($P = 4, \gamma = 0.2$). (b) 494 minimizers of smoothing level ($P = 4, \gamma = 0.2$) are tracked back to corresponding Lennard-Jones potential minimizers.

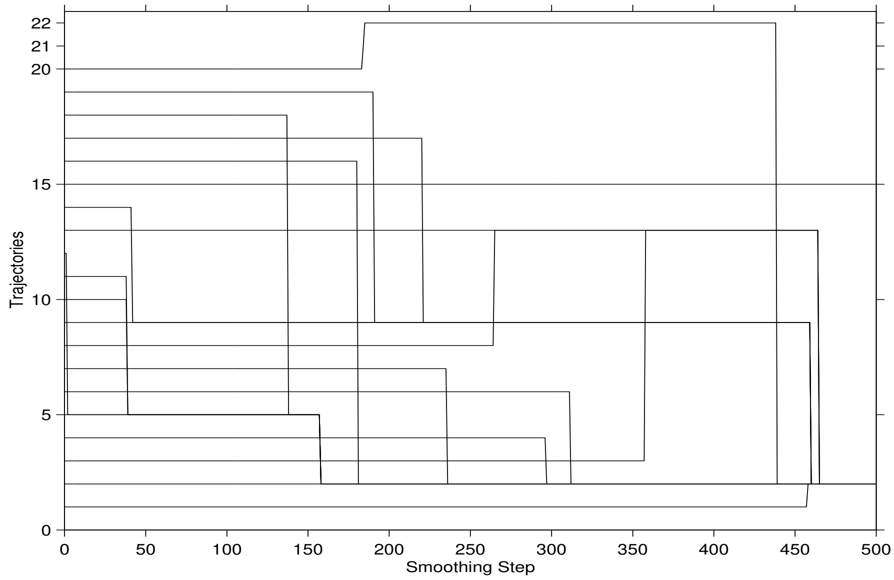
149.997. For this problem we began by locating 1288 of the lowest minimizers and tracking them to smoothing level (4.0, 0.2), resulting in 494 distinct minimizers at this smoothing level. Then we tracked these 494 minimizers in reverse back to the original Lennard-Jones function, resulting in 461 minimizers. The correspondence between the unsmoothed and smoothed minimizers when tracked in these two directions is shown in Figures 6a and 6b. These figures show that the order flip phenomenon is even more pronounced for this problem than for the 30 atom problem, and is really quite extreme. Table II summarizes the results of these tracking experiments for the best minimizers of the unsmoothed problem. For each of the twenty lowest Lennard-Jones minimizers, after tracking to the smoother function, the objective value of the smoothed minimizer reached is shown in column 4 and its order (among the 494 found)in column 5. In column 3, the designation *Jump* indicates the trajectory of the unsmoothed Lennard-Jones minimizer terminated, and the tracking procedure jumped to another trajectory, while *Cont* indicates a single continuous trajectory was followed. It is seen that almost all of the 20 lowest Lennard-Jones minimizers, including the global, tracked to relatively poor smoothed minimizers. In the reverse tracking from the 494 smoothed minimizers, we determined which of these minimizers led to one of the best twenty Lennard-Jones minimizers. For each of these best 20, column 7 gives the order of the lowest smoothed minimizer from which reverse tracking led to it. If the trajectory from the smoothed minimizer terminated during reverse tracking, and the tracking procedure jumped to another trajectory, column 6 says *Jump*; if the trajectory went all the way, the entry is *Cont*. The entry *None* means that no smoothed minimizer was found that would track to that unsmoothed minimizer. It is seen from column 7 that only 2 of the 20 lowest smoothed minimizers lead to any of the best 20 unsmoothed minimizers via reverse tracking, and that the 241st lowest smoothed minimizer is the best one that leads back to the global minimizer, though with a trajectory jump. Thus, reverse tracking 20 lowest smoothed minimizers we found would not find the global minimizer, and would find only two of the best 20.

5.3. Comparison with spatial averaging methods

Finally, we tried to compare the proposed smoothing techniques with the two spatial averaging smoothing techniques, the diffusion equation method and the effective energy transformation scheme, which were discussed in Section 2. The 9 atom problem is again used in this experiment. With only handful of minizers, it is possible to perform detail comparisons. As in the 30 and 34 atom experiments, we tracked each of

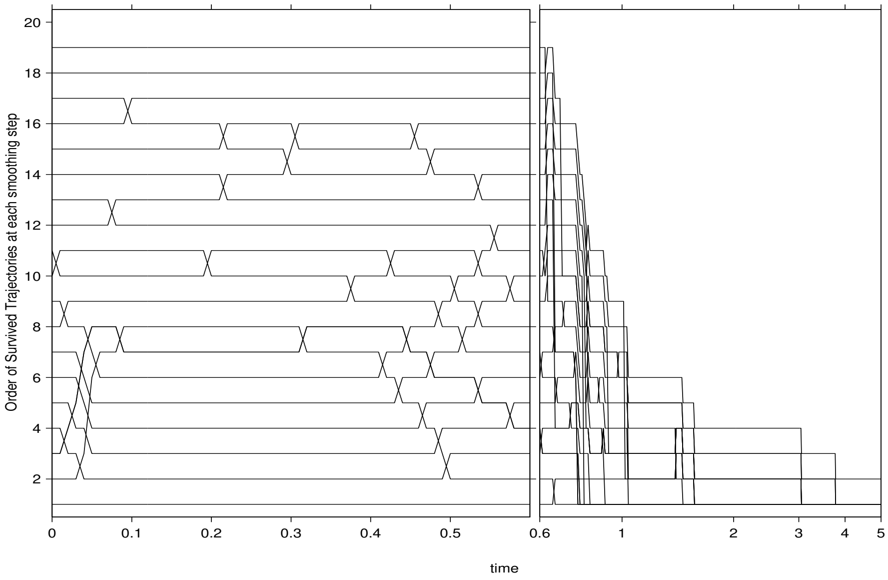


(a)

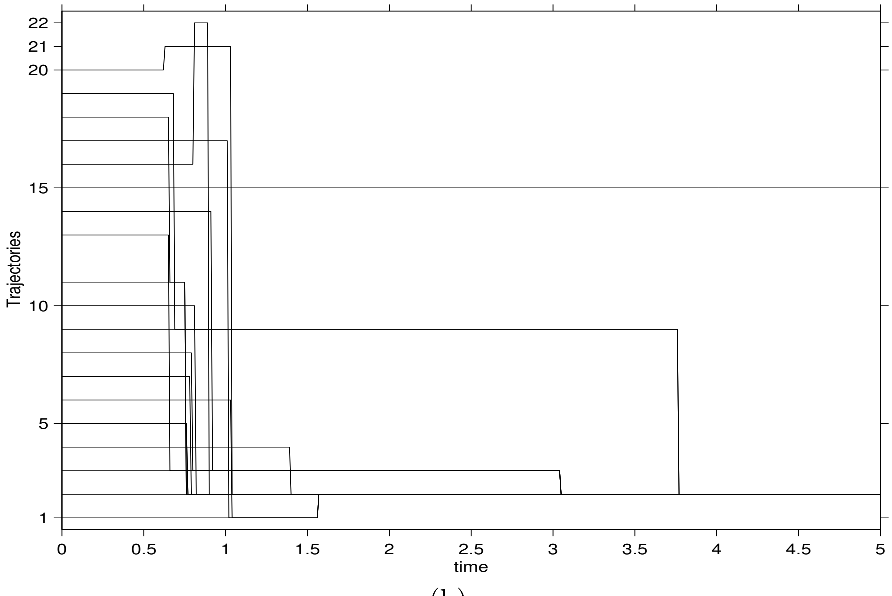


(b)

Figure 7. The 9-atom trajectory tracking from (6.0,0.0) to (4.0,0.7) of the proposed method. (a) Order Flipping Diagram. At each smoothing step, the trajectories are reordered by their function values at that smoothing step. The order flipping is shown as crossover on trajectories. (b) Trajectory Termination Diagram. The trajectories are arranged in their original order with new trajectories numbered above 20.



(a)



(b)

Figure 8. The 9-atom trajectory tracking from t=0 to t=5 of the diffusion equation method. (a) Order Flipping Diagram. At each smoothing step, the trajectories are reordered by their function values at that smoothing step. The order flipping is shown as crossover on trajectories. Only 19 trajectories in the diagram, since one is already merged at t=0. (b) Trajectory Termination Diagram. The trajectories are arranged in their original order with new trajectories numbered above 20.

the original twenty minimizers to some smoothing level for each of the three smoothing methods. However, the tracking procedure is slightly different from the previous experiments. After every forward tracking step, a reverse tracking back to the previous smoothing parameter value is taken immediately. The result is compared to the minimizer in previous step. With this procedure, forward jumps between trajectories can be immediately identified when the reverse step produces a different minimizer from that in the previous step. If the minimizer identified in the reverse step is different from all originally tracked minimizers in previous smooth parameter value, we can conclude that the original trajectory terminates and falls into a newly emerged trajectory.

There is no clear way to define equivalent smoothing levels among different smoothing methods. In this experiment, for each smoothing method, we merely tracked the trajectories to a smoothing level where only two trajectories are left. For this problem, this level is $(P = 4.0, \gamma = 0.7)$ for the proposed new method, and $t = 5$ for the diffusion equation method. For the effective energy transformation scheme, we tracked minimizers up to $\lambda = 4$ which is considered a very smooth level in the context of [6] and [7], but only four trajectories terminated. Graphical analysis of the pairwise potential function shows that this is because even for high λ values, the effective energy transformation scheme with the parameters and software supplied by [24] does not smooth the Lennard-Jones function much. In fact, the well of the function is slightly narrower for $\lambda = 4$ then for $\lambda = 0$ once a rescaling is applied so that the minimizers of the pairwise potentials are the same. The reason for this behavior seems to be the high cutoff value used in [6] [7] and [24] for the effective energy transformation scheme in the experiment; see function (19). The large positive values near the zero distance have a strong effect on the spatial averaging process and prevent significant smoothing. This seems to mean that the cutoff value would need to be lowered to obtain the desired smoothing effect. However, we did not attempt to repeat the experiment with different parameters than those supplied by the authors.

Fig. 7 shows results of the experiment for the new smoothing function, and Fig. 8 contains diagrams for the diffusion equation method. Each figure contains two diagrams. The first diagram presents how trajectories change their orders at each smoothing level. The second diagram, on the other hand, depicts how trajectories terminate and fall into other trajectories. In addition, the second diagram shows some new trajectories which the original trajectories fall into.

Figures 7 and 8 show qualitatively very similar behaviors between the proposed method and the diffusion equation method. First, at the end of tracking procedure, the 2nd and 15th trajectories survived in

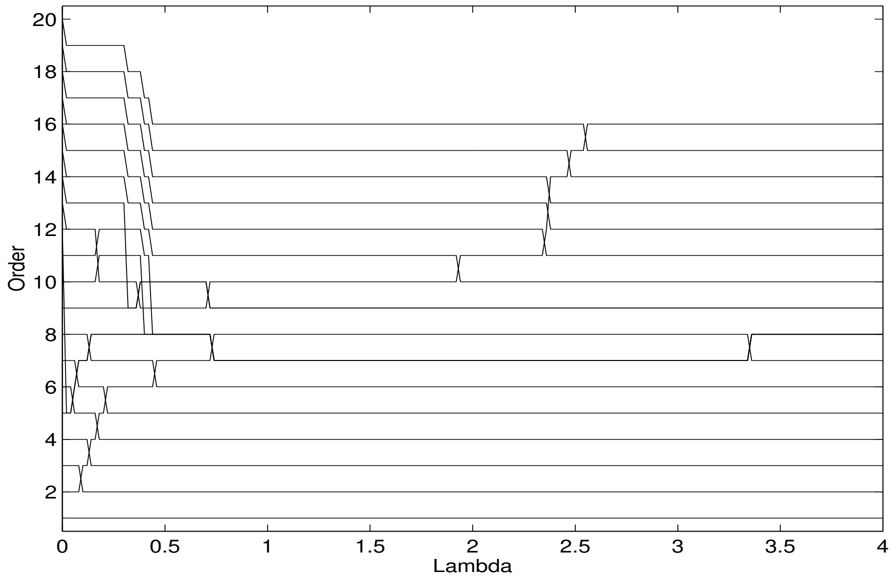


Figure 9. The 9-atom trajectory tracking from $\lambda = 0$ to $\lambda = 4$ of the effective energy transformation scheme. Order Flipping Diagram. At each smoothing step, the trajectories are reordered by their function values at the smoothing step.

both methods, shown in Fig. 7b and Fig. 8b. Secondly, in both methods, the global minimizer's trajectory first flips its order, then terminates and jumps to the second best minimizer's trajectory. Furthermore, both methods introduce some significant number of order flips as shown in Fig. 7a and Fig. 8a. Finally, both methods create some new trajectories along the smoothing process, which are labeled as 21 and 22 in trajectory termination diagrams. In Fig. 7b, the 20th trajectory first jumps to a new trajectory numbered as 21 in the figure. It then immediately jumps to yet another new trajectory numbered as 22. Fig. 8b shows both the 16th and 20th trajectories terminate and jump to new trajectories. The many similarities between the two smoothing approaches in the 9 atom cluster lead us to believe that the spatial averaging techniques would give similar results to the new smoothing method on other problems. Most importantly, the spatial averaging techniques also introduce some significant order flips in the course of the smoothing procedure.

By observing Fig. 8a carefully, one will notice that there are only 19 minimizers at $t = 0$. That is because the 12th minimizers has already merged with the 5th minimizer. This is caused by the approximation (10) used in the diffusion equation method. As shown in Fig. 2a, the function is quite smooth even at $t = 0$. This means there is a significant *gap* between the true Lennard-Jones potential and the smoothed function. In Fig. 7a, it is seen that trajectory 12 terminates and merges into trajectory 5 almost immediately, far before any other merges. This is another similarity between the behaviors of the two methods.

Fig. 9 shows the results for the effective energy transformation scheme. Only four trajectories terminate and merge into other trajectories. All four terminations happen before $\lambda = 1$. However, in the experiment, the scheme also introduces some order flips. Notably, the 12th minimizer's trajectory which also quickly terminates and merges into the 5th minimizer's trajectory as the other two methods.

From these results, we expect that behaviors like order flips and trajectory termination are shared by different smoothing techniques, and that they may be common in different problems. An important lesson from these experiments is that one should not optimistically assume that the global minimizer of a smoothed function will directly lead to the global minimizer in the original function, or even that reverse tracking a significant number of the lowest minimizers of a smoothed function will lead to the global minimizer of the original function. That is, smoothing followed by the trajectory tracking is not adequate as a general method for solving global optimization problems arising from molecular configuration.

6. Summary and Future Research

In this paper, a new family of functions for smoothing the Lennard-Jones potential energy function is introduced. The smoothing functions naturally remove the pole from the Lennard-Jones potential and widen the basin of attraction of its minimizer, without changing the location or energy value of the minimizer. Also, they are easy to evaluate.

The experiments in this paper carefully examine the properties of the new smoothing method and general behaviors of smoothing techniques. They also compare the new function family with two spatial averaging smoothing techniques, the diffusion equation method and the effective energy transformation scheme. They demonstrate that the proposed smoothing function family effectively reduces the number of minima in the Lennard-Jones problems with different sizes. They also indicate that the proposed method has similar smoothing behaviors to the spatial averaging approach. Thus it may be advantageous to use due to its lower cost and ease of implementation.

Our experiments also show that smoothing techniques can introduce some potentially undesirable behaviors into smoothed functions. The undesired behaviors, which appear to be are shared among different families of smoothing functions, are order flips, termination of trajectories, and emergence of new trajectories. These behaviors show that an algorithm that simply finds the global minimizer of a very smoothed problem and tracks it (or possibly the lowest few minimizers) back to the unsmoothed problem cannot be expected to locate the global minimizer of the unsmoothed problem in general. Some more sophisticated and robust algorithms need to be developed in order to take advantage of smoothing techniques in global optimization. Such algorithms should able to accommodate the ill behaviors of smoothing techniques discussed in this paper.

The larger goal of this research is to apply the types of techniques described in this paper to more complex molecular conformation problems, including the protein folding problem. The algebraic smoothing technique described in this paper can be generalized to other functions (usually with poles) that are likely to benefit from smoothing, such as electrostatic forces or attractive/repulsive forces given by different formulas than the Lennard-Jones potential. Currently, as discussed in [1], we are exploring another very similar formula which can smooth both the Lennard-Jones potential and electrostatic potential simultaneously.

$$\tilde{f}_{<P,\gamma>}(x) = \begin{cases} \left(\frac{1+\gamma}{r_{ij}^2+\gamma}\right)^P - 2\left(\frac{1+\gamma}{r_{ij}^2+\gamma}\right)^{P/2} & \text{for Lennard-Jones potential} \\ \frac{q_{ij}}{\sigma_{ij}}\sqrt{\frac{1+\gamma}{r_{ij}^2+\gamma}} & \text{for electrostatic potential} \end{cases} \tag{26}$$

In addition, we have developed an approach, based on [2], which utilizes the smoothing technique to dramatically reduce the number of minima at early stages, but does not rely solely on trajectories to find the global minima at final stages. Instead, minimizers from each smoothed stage serve as starting configurations for a global minimization algorithm at each successive, less smoothed stage. This approach has been recently applied to protein polyalanine problems with the new smoothing functions (26). Some encouraging results have been obtained and show the advantages of using the smoothing technique in global optimization. These results will be reported in a forthcoming paper [1].

References

1. Azmi, A., Byrd, R.H., Eskow, E., and Schnabel, R.B., *The New Protein Smoothing Paper* in preparation, 1997.
2. Byrd, R.H., Eskow, E., and Schnabel, R.B., *A New Large-Scale Global Optimization Method and its Application to Lennard-Jones Problems* Technical Report CU-CS-630-92, Dept. of Computer Science, University of Colorado, revised, 1995.

3. Byrd, R.H., Derby, T., Eskow, E., Oldenkamp, K., and Schnabel, R.B. *A New Stochastic/Perturbation Method for Large-Scale Global Optimization and its Application to Water Cluster Problems*, In *Large-Scale Optimization: State of the Art*, W. Hager, D. Hearn, and P. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 71–84, 1994.
4. Byrd, R.H., Dert, C.L., Rinnooy Kan, A.H.G., and Schnabel, R.B *Concurrent stochastic methods for global optimization*, *Mathematical Programming*, **46**, 1-29, 1990.
5. Byrd, R.H., Eskow, E., van der Hoek, A.,Schnabel, R.B, Shao, C-S., and Zou, Z. *Global optimization methods for protein folding problems*, In *Proceedings of the DIMACS Workshop on Global Minimization of Nonconvex Energy Functions: Molecular Conformation and Protein Folding*, P. Pardalos, D. Shalloway, and G. Xue, eds., American Mathematical Society, 29-39, 1995.
6. Coleman, T., Shalloway, D. and Wu, Z. *Isotropic Effective Energy Simulated Annealing Searches for Low Energy Molecular Cluster States*, Technical Report CTC-92-TR113, Center for Theory and Simulation in Science and Engineering, Cornell University, 1992.
7. Coleman, T., Shalloway, D. and Wu, Z. *A Parallel Build-up Algorithm for Global Energy Minimizations of Molecular Clusters Using Effective Energy Simulated Annealing*, Technical Report CTC-93-TR130, Center for Theory and Simulation in Science and Engineering, Cornell University, 1993.
8. Coleman, T., and Wu, Z. *Parallel Continuation-Based Global Optimization for Molecular Conformation and Protein Folding*, Technical Report CTC-94-TR175, Center for Theory and Simulation in Science and Engineering, Cornell University, 1994.
9. Farges, J., DeFeraudy, M.F., Raoult, B., and Torchet, G. *Cluster models made of double icosahedron units*, *Surface Sci.*, **156**, 370-378, 1985.
10. Freeman, D.L. and Doll, J.D. *Quantum Monte Carlo study of the thermodynamic properties of argon clusters: the homogeneous nucleation of argon in argon vapor and "magic number" distributions in argon vapor*, *J. Chem. Phys.*, **82**, 462-471, 1985.
11. Hoare, M.R. *Structure and dynamics of simple microclusters*, *Adv. Chem. Phys.*, **40**, 49-135, 1979.
12. Hoare, M.R. and Pal, P. *Physical cluster mechanics: statics and energy surfaces for monatomic systems*, *Adv. Phys.*, **20**, 161-196, 1971.
13. Judson, R.S., Colvin, M.E.i, Meza, J.C., Huffer, A. and Gutierrez, D. *Do Intelligent Configuration Search Techniques Outperform Random Search for Large Molecules?*, SAND91-8740, 1991.
14. Kostrowicki, J., Piela, L., Cherayil, B.J., and Scheraga, A. *Performance of the Diffusion Equation Method in Searches for Optimum Structures of Clusters of Lennard-Jones Atoms*, *J. Phys. Chem.*, **95**, 4113-4119, 1991.
15. Maranas, C. D. and Floudas C. A. *A Global Optimization Approach for Lennard-Jones Microclusters*, Preprint, Department of Chemical Engineering, Princeton University, June 1992.
16. Northby, J.A. *Structure and binding of Lennard-Jones clusters:*$13 \leq N \leq 147$, *J. Chem. Phys.*, **87**, 6166-6177, 1987.
17. Oresic, M., and Shalloway, D. *Hierarchical Characterization of Energy Landscapes using Gaussian Packet States*, *J. Chem. Phys.*, **101**, 9844-9857, 1994.
18. Pillardy, J., and Piela, L. *Molecular Dynamics on Deformed Potential Energy Hypersurfaces*, *J. Phys. Chem.*, **99**, 11805-11812, 1995.

19. SCHNABEL, R.B., KOONTZ, J.E., AND WEISS, B.E. *A modular system of algorithms of unconstrained minimization*, *ACM Transactions on Mathematical Software*, **11**, 419-440, 1985.
20. SHALLOWAY, D. *Packet annealing : a deterministic method for global minimization, with application to molecular conformation. Preprint, Section of Biochemistry, Molecular and Cell Biology*, *Global Optimization*, C. Floudas and P. Pardalos, eds., Princeton University Press, 1991.
21. SHAO, C-S., BYRD, R.H., ESKOW, E., AND SCHNABEL, R.B. *Global optimization for molecular clusters using a new smoothing approach*, In *Large-Scale Optimization with Applications, Part III: Molecular Structure and Optimization*, L.T. Biegler, T.F. Coleman, A.R. Conn, and F.N. Santosa, eds., The IMA Volumes in Mathematics and Its Applications, **94**, Springer, New York, 163-199, 1997.
22. SMITH, S.L., ESKOW, E., AND SCHNABEL, R.B. *Adaptive, asynchronous stochastic global optimization algorithms for sequential and parallel computation*, In *Proceedings of the Workshop on Large-Scale Numerical Optimization*, T.F. Coleman and Y. Li, eds. SIAM, Philadelphia, 207-227, 1989.
23. WILLE, L.T. *Minimum-energy configurations of atomic clusters: new results obtained by simulated annealing*, *Chem. Phys. Lett.*, **133**, 405-410, 1975.
24. WU, Z. *The Effective Energy Transformation Scheme as a Special Continuation Approach to Global Optimization with Application to Molecular Conformation*, Technical Report CTC-93-TR143, Center for Theory and Simulation in Science and Engineering, Cornell University, 1993.
25. XUE, G. *Improvement of the Northby Algorithm for Molecular Conformation: Better Solutions,* , *J. Global Optimization*, **4**, 425. 1994.