A New Stochastic/Perturbation Method for Large-Scale Global Optimization and Its Application to Water Cluster Problems
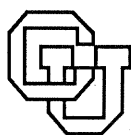
Richard H. Byrd, Thomas Derby, Elizabeth Eskow
Klaas P.B. Oldenkamp, Robert B. Schnabel

University of Colorado at Boulder
DEPARTMENT OF COMPUTER SCIENCE

A New Stochastic/Perturbation Method for Large-Scale Global Optimization and its Application to Water Cluster Problems[1]

Richard H. Byrd
Thomas Derby
Elizabeth Eskow
Klaas P.B. Oldenkamp
Robert B. Schnabel

CU-CS-652-93   May 1993

University of Colorado at Boulder
Department of Computer Science
Campus Box 430
Boulder, Colorado 80309-0430 USA

Abstract

We describe a class of new global optimization methods that has been designed to solve large, partially separable problems. The methods have been motivated by the consideration of problems from molecular chemistry, but should be applicable to other partially separable problems as well. They combine a first, stochastic phase that identifies an initial set of local minimizers, with a second, more deterministic phase that moves from low to even lower local minimizers and that accounts for most of the computational cost of the methods. Both phases make critical use of portions that vary only a small subset of the variables at once. Another important new feature of the methods is an expansion step that makes it easier to find new and structurally different local minimizers from current low minimizers. We give the results of the initial application of these methods to the problem of finding the minimum energy configuration of clusters of water molecules with up to 21 molecules (189 variables). These runs have led to improved minimizers, and interesting structures from the chemistry perspective.

1 Introduction

This paper describes a new large-scale global optimization method and its application to the problem of finding the minimum potential energy configurations of clusters of water molecules. The purposes for this research are twofold: the development of a fairly general purpose large-scale global optimization method for use in solving many large-scale problems with similar structure, and the solution of useful problems in molecular chemistry.

The unconstrained global optimization problem is the problem of finding the lowest minimizer of a nonlinear function f in a domain $D \in R^n$, where D is defined by upper and lower bounds on each variable. It is assumed that the global minimizer lies in the interior of D, and that the function f is twice continuously differentiable. Typically large-scale problems refer to those with hundreds, thousands, or more variables, but because global optimization problems are so computationally expensive to solve, even problems at the low end of this range are considered to be large-scale problems. In addition, the number of local minimizers affects the difficulty of solving global optimization problems, and problems with hundreds or thousands of minimizers are also considered to be large-scale.

The difficulties in solving large-scale global optimization problems arise from the challenges of effectively searching a potentially enormous parameter space, and locating the basin of attraction of the global minimizer, when there are huge numbers of local minimizers. Generally, small-scale global optimization techniques such as the stochastic methods of [8], and the tunneling methods of [4], are ineffective on large-scale problems because the complexity of working on the entire parameter space simultaneously requires too much computation. Extensions of methods for discrete optimization, such as simulated annealing and genetic algorithms, are generally slow and not necessarily effective on large-scale problems either. By using methods that contain stochastic and deterministic features, and that incorporate the partially separable structure of the problem by working on small subsets of the parameter space as well as in the full dimensional space, we have begun to make exciting progress in solving some large-scale global optimization problems.

We began research in large-scale global optimization with the development of a large-scale method and its application to solving Lennard-Jones problems with 15 to 165 parameters [2]. The Lennard-Jones problem is that of determining the structure of clusters of atoms whose potential energy is given by the sum of the pairwise interactions between atoms using the Lennard-

Jones potential energy function. These problems use a simple energy model, but are somewhat representative of other large-scale global optimization problems in molecular chemistry in the number of parameters and local minimizers they contain and in some of the characteristics of the energy surface. The method we developed is based on a combination of stochastic and deterministic techniques, and combines phases that concentrate on one-atom subproblems with phases that work in the full-dimensional parameter space. The techniques used are basically applicable to any partially separable function, which is a function that is the sum of many terms, each of which involves only a small subset of the parameters. The method finds the presumptive global minimizer of all Lennard-Jones problems with up to 90 variables (30 atoms) as well as some larger cases. (Experiments on some larger Lennard-Jones problems using the additional techniques that are introduced in this paper indicate that larger Lennard-Jones problems can be solved as well with the new techniques.) The results of our original method on the Lennard-Jones problems indicate that the method works better on these problems than other general purpose global optimization methods that have been tried, but that for larger problems it is not as effective as special purpose methods that use information regarding the solution structure of the problem (see e.g. [6, 7]).

We now proceed to the more difficult problem of finding the minimum potential energy of clusters of water molecules, as a stepping stone to even more challenging problems such as minimizing the configurations of polymers or proteins. The structure of the water problem is more complex than Lennard-Jones problems because there are two levels of interaction, one between neighboring water molecules and the other within each individual water molecule. This requires a more complicated potential function to express the energy of the system, and more sophisticated global optimization techniques to enable the method to move from already low energy configurations to new, even lower energy configurations. Another motivation for attempting to solve the water cluster problem is that this problem is of interest to the chemistry community because of the possibility of finding new, previously unpredicted low energy structures using global optimization methods. Most of the new algorithmic techniques described in this paper are applicable to other molecular chemistry problems, and may also be appropriate for use in solving other partially separable large-scale problems.

The remainder of the paper describes our new global optimization method and its application to the water cluster problem. Section 2 describes the problem of finding the minimum

2

energy configurations of clusters of water molecules. Section 3 describes the global optimization method used to solve the water cluster problems, highlighting the new features that were necessary to enable the method to work on these problems. Section 4 presents some early test results for the water cluster problems, and conclusions and future research are discussed in section 5.

2   The Water Cluster Problem

Water clustering behavior continues to be a topic of interest in various areas of scientific research, including theoretical chemistry and atmospheric physics. It is also an important component towards understanding the structure of substances such as proteins in a water solution. The problem of finding the structure that a cluster of water molecules assumes at equilibrium can be expressed as the problem of minimizing the potential energy function of the water cluster. Since the potential energy function of even moderate size clusters (e.g. 20 water molecules) has huge numbers of local minima, each with a relatively small basin of attractions and many with energy values that are relatively close to the global minimum, the problem of finding the lowest energy structure is a challenging global optimization problem.

According to [5], the energy and structural data produced by the empirical water dimer potential energy surface function (RWK2-M) described in [3] is consistent with experimental results and hence can be considered an accurate approximation to the potential energy surface of the cluster. We have used this empirical potential function in our experiments and give a general description of it here. See [5] or [3] for the complete details of the function. The function takes the form $\sum_{i \neq j} u(x_i, x_j) + \sum_i v(x_i)$ where each vector x_i gives the nine Euclidean coordinates of the i-th water molecule, and u and v are fairly complex algebraic functions giving the interaction energy between pairs of molecules, and the internal energy of the molecule, respectively. The intermolecular function u is the sum of the Coulomb or electrostatic interactions, the exponential repulsions for oxygen-oxygen and hydrogen-hydrogen interactions, the Morse potential for oxygen-hydrogen interactions, and a dispersion term for oxygen-oxygen distances, while the intramolecular function v sums the nuclei-nuclei Morse oscillator potentials with an additional coupling term [5]. The code for the potential function, in units that have been reduced to atomic units (a.u.) from the original units of kcal/mol and angstroms, was provided to us by Xiping Long.

3

The following are expressions that give the basic mathematical forms of the interaction terms of the potential energy surface mentioned above. We use the symbol r_{AB} to indicate $|R_B - R_A|$ where R_X is the position of the charge carried by atom X. The constants A_{OH}, A_{OO}, A_{HH}, α_{OH}, α_{OO}, α_{HH}, R_M, C_6, C_8, C_{10}, β, a, b, c, α_1, α_2, α_3, f_{12}, D_1, D_2, D_3, are empirical parameters to the RWK2-M potential that have been fitted to the model using experimental data. The values for these parameters are given in [5] and [3].

- The Coulomb interactions between two molecules have the form $\sum_{i,j}^3 \frac{q_i q_j}{r_{ij}}$, where q_i is the charge associated with the O, H^1 or H^2 atom of the first molecule and likewise, q_j for the second molecule.

- The exponential repulsion terms, where XY are OO or HH atoms, are
$$A_{XY} \exp(-\alpha_{XY} r_{XY}).$$

- The Morse potential for oxygen-hydrogen interactions is
$$A_{OH} \exp(-\alpha_{OH}(r_{OH} - R_M))(\exp(-\alpha_{OH}(r_{OH} - R_M)) - 2)$$

- The dispersion term as a function of the oxygen-oxygen distances has the form
$$-f\,[C_6(\frac{g_6}{r_{OO}})C_8(\frac{g_8}{r_{OO}})C_{10}(\frac{g_{10}}{r_{OO}})],$$
where $f = 1 - (cr_{OO})^\beta \exp(-cr_{OO})$ and $g_n = 1 - \exp(-(a\frac{r_{OO}}{N} + b\frac{r_{OO}^2}{\sqrt{N}}))$.

- The Morse oscillator potentials for the intramolecular energy have the form
$$\sum_{i=1}^3 D_i(1 - \exp(-\alpha_i s_i))^2 + f_{12} s_1 s_2,$$
where $s_i = r_{OH^i} \cos(\frac{\theta - \theta_0}{2}) - R_O, i = 1, 2$ and $s_3 = \frac{r_{OH^1} + r_{OH^2}}{R_O} sin(\frac{\theta - \theta_0}{2})$, with R_O = optimal OH bond length, θ = bond angle and θ_O = optimal bond angle.

The structures of neutral water clusters with N, the number of water molecules, equal to 20 and 21, have been studied extensively. Therefore, our initial experiments have been mainly with clusters of molecules of these two sizes. A discussion of the structures expected for these clusters and our experimental results for them is presented in section 4.

It should be mentioned that while the natural parameterization of the water cluster problem, in Euclidean coordinates, has $9N$ parameters for N water molecules, there are only $9N - 6$ degrees of freedom due to the invariance of the problem to translation and rotation. It is possible to reparameterize the model to eliminate this invariance and contain only $9N - 6$

variables. Our experiments have shown, however, that there is no loss in computational efficiency in solving the problem in the full $9N$ parameters; the main issue is that even though the Hessian of the model is singular everywhere, the convergence of local optimization methods is not impaired by using the full parameterization. Therefore the methods described in Sections 3 and 4 solve the water cluster problem using the $9N$ Euclidean coordinates, since this model is easier to supply, differentiate, and understand than a parameterization with $9N - 6$ variables, and its use does not impair the efficiency of the method.

3   The Global Optimization Method

Now we describe the large-scale global optimization method we have developed for solving the water cluster energy minimization problem. This algorithm is closely related to the algorithm we developed for solving the Lennard-Jones energy minimization problem, which is described in [2], but has several important new features that are crucial to its success. This section describes the entire algorithm, with emphasis on the new features. The parallelization of the method also is described. Although, we developed the new method while working on the water potential problem, it is applicable to a wide class of molecular configuration problems, including the simpler Lennard-Jones problem. Indeed, our approach should apply to any large-scale global optimization problem with sufficient partial separability and symmetry among its variables.

A basic idea of our methods is to do some work in the full dimensional parameter space, and some work using only a small subset of parameters. This approach is feasible because of the partially separable property of the objective functions. The choice of the subset is problem dependent, and depends on the natural "unit" for the problem. In the Lennard-Jones application, the subset of variables consists of a single atom, while in the water cluster problem, this subset is a water molecule.

The new method has two phases. The first, sample generation phase uses random sampling and local minimization to identify some promising configurations, that is, local minimizers with reasonably low energies. The second, local minimization improvement phase repeatedly uses special perturbation techniques to improve some local minimizers found in either the first or second phase.

During the first phase, a full dimensional random sample is generated over the domain

space by randomly and independently placing each molecule. The sample points with the highest function values are discarded, and the remaining sample points are improved, by using a subset of the variables technique that randomly samples on and moves one molecule at a time, until the function value for the configuration falls below a specified threshold level. From these improved sample configurations, a subset is selected and used as start points for a local minimization algorithm. Some of the local minimizers found in this phase are passed on to the second phase for improvement.

The second, local minimizer improvement phase successively selects a full dimensional configuration for improvement. The molecule that contributes the most (or second most) to the function value of the selected configuration is chosen, and a global optimization algorithm is applied to the configuration with only this molecule as a variable and the remainder of the configuration fixed. Next, full dimensional local minimizations are performed from the lowest configurations resulting from the single molecule global optimization step. The lowest new configurations that are generated by this process are then merged into the list of local minimizers, and this phase is iterated a fixed number of times.

Algorithm 3.1 below outlines the framework of the global optimization algorithm for the water cluster application.

Algorithm 3.1 - Outline of the Global Optimization Algorithm

Given feasible region $D \in R^n$, $f : D \to R$

1. **Sample Generation Phase**

    (a) **Sampling in Full Domain :** Randomly generate the coordinates of sample configurations in the feasible region D, and evaluate the potential energy at each sample configuration. Discard all sample configurations whose function values are above a global cutoff level.

    (b) **One Molecule Sampling Improvement :** For each remaining sample configuration: While the energy of the sample configuration is above the threshold value, Repeat:

      - Select the molecule that contributes most to the energy function value
      - Randomly sample on new locations of the selected molecule

- Replace this molecule in the sample configuration with the newly sampled location that gives the lowest energy value for the configuration.

  (c) **Start Point Selection :** Select a subset of the improved sample configurations to be start points for local minimizations.

  (d) **Full-Dimensional Local Minimizations :** Perform a local minimization from each start point selected. Collect some number of the best of these local minimizers for improvement in Phase 2.

2. **Local Minimizer Improvement Phase:** For some number of iterations:

  (a) **Select a Configuration :** From the list of full-dimensional local minimizers, select a local minimizer using a heuristic to determine the most promising candidate.

  (b) **Expansion :** Transform the configuration by multiplying the position of each molecule relative to the center of mass of the configuration by a constant factor. (Leave the internal structure of each molecule unchanged.)

  (c) **One Molecule Global Optimization :** Select the molecule whose partial energy, before expansion, has the worst (or second worst) value. Apply a global optimization algorithm to the expanded configuration with only this molecule as a variable.

  (d) **Full-Dimensional Local Minimization :** Apply a local minimization procedure, using all the molecules as variables, to the lowest configurations that resulted from the one-molecule global optimization.

  (e) **Merge the New Local Minimizers :** Merge the new lowest configurations into the existing list of local minimizers.

The two phase structure and small subproblem steps of Algorithm 3.1 are similar to the global optimization method that we used for Lennard-Jones problems, but the algorithm has several important new features that are expected to be applicable to other molecular chemistry problems. The most important is step 2b, which expands the molecular cluster around its center of mass before the one molecule global optimization. This step, which is somewhat analogous to heating in nature and in simulating annealing, creates more openings in the configuration and enables the one molecule global optimization step to find more possible locations for the molecule

being moved. The incorporation of expansion has enabled the algorithm to improve the function values of many local minimizers that it otherwise was unable to improve. We have experimented with different expansion factors, and have found factors from 1.25 to 2 to be the most effective. We have also made preliminary experiments with applying this technique to Lennard-Jones problems and it appears to be effective in that context as well. It is likely that an analogous technique will be applicable to other molecular chemistry problems, although the details of its application may be problem dependent.

Another important new feature for the water cluster problem is the use of heuristics in step 2a to choose which configuration and molecule within that configuration to work on next. For Lennard-Jones problems, the heuristic was simple: select the configuration with the lowest function value that hadn't been previously used, and select the atom that contributed the most (or second most) to the function value of that configuration. For the water problem, we have found it useful to consider more complicated heuristics. The obvious technique is to again choose the configuration with the lowest energy that hasn't been used yet, and the molecule that contributes the most to this energy. It has been found that one should avoid choosing the same molecule many times in a row in a series of consecutive modifications stemming from one configuration. A second technique we have used is to choose the configuration with the lowest function value for $N - 1$ molecules, with the molecule that contributes the most to the function value omitted, and then choose this omitted molecule for the one-molecule global optimization. This has the effect of choosing a configuration and molecule where there is much room for improvement. A third technique is to use a measure for the structural similarity of configurations, and to choose configurations that are structural dissimilar to those that have been modified previously. This requires an effective metric for measuring structural diversity among configurations; experiments for developing and assessing such a metric are in progress. Initially we have used combinations of the first two heuristics, and the results reported in this paper are based on these. We are currently experimenting with all the approaches mentioned above, and with ways to combine them in an integrated manner. We expect that the issue of choosing which configurations to improve will be very important for other molecular chemistry problems as well, and that the approach we are developing will have applicability to other problems.

We have also incorporated one relatively minor feature in our algorithm that utilizes the particular structure of the water cluster problem. In sampling steps 1a and 1b, and the sampling

within step 2c of Algorithm 3.1, the bond lengths and bond angle within each water molecule are kept fixed while the other parameters are sampled freely. This insures that the individual water molecules that are generated are reasonable, rather than allowing random placements of the hydrogens relative to the oxygens. Viewed more generally, this feature turns the variables that are known to have very little variation (here, bond length and bond angle) into constraints for the sampling phase. This technique has very general applicability although its implementation is clearly problem specific.

Finally we briefly discuss the parallel implementation of Algorithm 3.1. All of our implementations of Algorithm 3.1 have been parallel ones; the initial implementation and most of our experiments so far have been on a network of Sun workstations, and recently we have ported the code to an Intel iPSC/2 hypercube. Much of Algorithm 3.1 parallelizes readily at a coarse grain level, although due to the coarse granularity and variable length of tasks such as local minimizations, dynamic scheduling support for tasks is necessary. During the first phase of the algorithm, the sample generation, sample point improvement and start point selection steps (1a-1c) are parallelized by dividing the full dimensional domain space among the processors. The full dimensional local optimizations of step 1d are dealt to processors as they become available by a centralized scheduler. In the local minimizer improvement phase, the scheduler performs steps 2a and 2b, and then the one molecule global optimization of step 2c is performed in parallel using the adaptive, small dimensional parallel global optimization method of [9]. This method allocates sampling and local minimization work to the processors using a centralized scheduler. The implementation on the Sun workstations (for which we present results in section 5) does not parallelize the full dimensional local minimizations of step 2d, but this step has been parallelized in the iPSC implementation. The iPSC implementation also includes another level of parallelization: multiple configurations are selected at once in step 2a and improved concurrently, each using a subset of the processors. Typically this subset is greater than one processor, so that each one molecule global optimization and full dimensional local minimization step is parallelized as well, resulting in two nested levels of parallelism.

4 Experimental Results

We have run our algorithm for several water cluster sizes. Let N denote the number of water molecules in the cluster (meaning that the total number of optimization variables is $9N$). The smaller cases of $N = 2, 6$ and 8 were run to verify the correctness of the algorithm and potential energy function implementations, whereas most of our effort was spent exploring the solutions generated for $N = 20$ and 21 because these two cases are of interest to chemists and are very challenging to solve. The experimental results of Algorithm 3.1 are presented in Table 4.1, along with the results for these same problems from [5]. The energies are given in atomic units (a.u.), i.e. 1 a.u. of energy $= 627.51$ kcal/mol, or 27.21 eV, or 1 hartree. All these results were obtained on a network of five Sun workstations. Continuing work on a more powerful parallel machine and with improved heuristics is in progress and is generating improved results; these are mentioned briefly in Section 5.

Table 4.1: Experimental Results for Water Cluster Problems

N	Energy in a.u. of Best Min found Using Alg. 3.1	Energy in a.u. of Best Min from [5]	Energy in a.u. of Best Min found Using Phase 2 from "special" configurations
2	-.00982	-0.00982	
6	-0.0756	-0.0756	
8	-0.1199	-0.1199	
20	-0.3447	-0.3429	-.3477
21	-0.3608	-0.3579	-.3679

These results show that Algorithm 3.1 was able to find the same minimizer as the dynamic simulated annealing procedure described in [5] for the small cases ($N = 2, 6, 8$), and even lower minimizers for the difficult cases of $N = 20$ and $N = 21$. Assuming the energy of the system is equal to the temperature times Boltzman's constant, the differences of .0018 and .00329 a.u. between the best result of Algorithm 3.1 and the best result from [5] for the 20 and 21 molecule problems, respectively, represent differences in temperature of 568.3°K and 915.6°K, whereas the difference in temperature from absolute zero to room temperature is about 300°K. Thus these differences are quite significant.

By using the best 20 water molecule solution from [5] in Step 2a, however, we located

three even better 20 molecule structures, with the lowest having energy -0.3477 a.u.. By using these three 20 molecule structures, augmented by a randomly sampled 21st molecule, in Step 2a for the 21 molecule problem, we obtained many significantly better 21 molecule solutions, with the lowest having energy -0.3679 a.u.. The differences in temperature from the best minimizers found by running Algorithm 3.1 entirely versus the best minimizers found by using these special configurations just described, for 20 and 21 molecules, are 947.2°K and 2241.6°K, respectively. These additional results demonstrate the strength of phase 2 of the algorithm, but also that with the current amount of computational effort, Algorithm 3.1 obviously does not find the global minimum for difficult problems. In [5], the starting configuration was found to be essential to the success of the simulation. By running our algorithm on parallel computers with considerably greater computational power, we will be able to determine if our current global optimization approach is an effective way to produce minimizers with energies as low as column 3 of Table 4.1, or if more sophisticated techniques are required in phase 1 to produce better starting configurations for phase 2. As mentioned in Section 5, initial indications from runs on more powerful machines are very encouraging.

An indication of the cost of our algorithm is given in Table 4.2 below, which gives the costs for the 8 and 20 molecule problems on a network of five dedicated Sun4 workstations of varying processor speeds. The average time for the total algorithm was computed from the other two average times by assuming that phase 2 was run for 10 iterations, which is typical in the Sun workstation runs. (The cost of a 1-molecule function evaluation is about $\frac{2}{N}$ times the cost of an N molecule function evaluation since $N - 1$ of the $N^2 - N$ intermolecular energies need to be recomputed, and these are the dominant costs.) Table 4.2 indicates that a typical run of Algorithm 3.1 on 5 Sun4 workstations take approximately seven hours for the 20 molecule problem. In addition, the number of function evaluations performed during these runs is quite small compared to the experiments in [5]. This indicates that in order to experiment with larger problems, or compute more intensively on problems of the current size, more powerful machines are required.

These experiments also appear to be producing interesting chemistry issues. Most notably, some of the best 21 molecule structures found so far are dodecahedron-like objects with one molecule in the center, which is the structure expected by the chemists, but others have much more irregular shapes. Several contain cycles of water molecules with 3, 4, 6, or 7 bonds, whereas

Table 4.2: Average Costs for 8 and 20 molecule water cluster problems using Algorithm 3.1

N	N-molecule function evaluations for Phase 1	1-Molecule function evaluations for Phase 1	N-molecule function evaluations for 1 iteration of Phase 2	1-Molecule function evaluations for 1 iteration of Phase 2	Time in minutes for Phase 1	Time in minutes for 1 iteration of Phase 2	Time in minutes for Total Alg
8	2860	47600	1880	25702	22.27	5.02	77.47
20	5551	197700	3248	21861	74.36	35.14	425.76

in a dodecahedron all the bonded water molecules form 5-cycles. In addition, one low energy 21 molecule structure has 35 bonds between water molecules, rather than the 34 bonds found in dodecahedron-like structures, although its energy is not as low as the best dodecehedral-like structures. It is not known yet whether we are close enough to the global minimum for this problem that these irregular shapes are truly near optimal. If they are, or if similar shapes are found near the true global minimum, this would raise interesting issues regarding either the shapes of water clusters, or the accuracy of the empirical energy function which heretofore has been thought to be quite accurate.

5 Summary and Future Research

We have described a new approach to large-scale global optimization that is applicable to a broad class of partially separable problems, including many problems from molecular chemistry. Two of the most important features of the new approach are several portions of the algorithm that concentrate on a small subset of the variables within the larger problem, and a new step that "expands" the current configuration before attempting to modify and improve it. The implementation of both of these portions may be in part problem dependent, but the techniques should be applicable to many partially separable problems. Computational results for the water cluster problem on a network of Sun workstations show that the new techniques are very helpful, and that the algorithm does a good job of finding low local minimizers. The minimizers found are considerably lower than those previously found for this problem.

Very recently, we have ported this algorithm to an Intel iPSC/860 hypercube and have begun running it on the same water cluster problems. Using this computer has enabled us to make runs that are approximately ten times as long (in terms of function evaluations, or local

minimizer improvement steps) as previously, and to try new heuristics for determining what parts of the search space to explore. The preliminary results from these experiments are far superior to those obtained previously on the Sun workstations with shorter runs. On the 21 water molecule problem, the main one we have tried so far on the iPSC/860, we have found many local minimizers with considerably lower energy values than we had found previously. The lowest minimizer found so far has energy value -0.3670, which is -.0062 atomic units lower than the best value found on the Sun workstations and within -.0009 atomic units (284.1K) of the lowest energy found by using "special" configurations in stage 2 of the algorithm (see Section 4). These preliminary results indicate that our global optimization algorithm may be quite effective on the water problem, given reasonable computing power. Indeed, it appears that with the current computing power and heuristics, we may soon find better minimizers from scratch than we have found by applying our techniques to special configurations as described in Section 4. When an extensive computational and algorithmic study is concluded, the results will be reported in a future paper.

Finally, it should be mentioned that an important topic that has not been addressed in this paper is the theoretical properties of our method. Part of the basic approach that we have taken (although not for the large-scale aspects) is motivated by the stochastic methods of [8] and [1]. One of the noteworthy aspects of these methods is that they have strong theoretical guarantees: under reasonable assumptions on the problem, the method will find the global minimizer with probability one while doing only a finite number of local minimizations. Unfortunately, this mathematical property does not correspond to efficient computational performance on large scale problems, and it appears that special techniques for dealing with the large problem size, such as those described in this paper, are necessary to solve large problems efficiently. On the other hand, it would be nice for an efficient large scale global optimization method to also have strong theoretical guarantees. Algorithm 3.1 does not have such guarantees, basically because the sampling in the full dimensional space (Phase 1) is performed only once. By repeating Phase 1 periodically, it seems clear that we could gain the theoretical properties of [8], but it is not clear that this modification would be desirable computationally. We are currently investigating ways to modify Algorithm 3.1 that are both computationally helpful and that lead to strong convergence properties, and will report the results of this work in a future paper.

References

[1] Byrd, R.H., Dert, C.L., Rinnooy Kan, A.H.G., and Schnabel, R.B. (1990). Concurrent stochastic methods for global optimization. *Mathematical Programming*, 46, 1-29.

[2] Byrd, R.H., Eskow, E., and Schnabel, R.B. (1992). A new large-scale global optimizaiton method and its application to Lennard-Jones problems, Technical Report CS-CS-630-92, Department of Computer Science, University of Colorado, Boulder.

[3] Coker, D. F. and Watts, R. O. (1987). Structure and vibrational spectroscopy of the water dimer using quantum simulation, *J. Phys. Chem*,91, 2513-2518.

[4] Levy, A.V. and Montalvo, A. (1985). The tunneling algorithms for the global minimizer of functions. *SIAM Journal on Scientific and Statistical Computing*, 6, 15-29.

[5] Long, X. (1992). Molecular dynamics simulations of clusters – impure van der Waals and $e^- - (H_2O)_n$ systems, Ph.D. Dissertation, Department of Chemistry, University of California, San Diego.

[6] Maier, R.S., Rosen, J.B., Xue, G.L. (1992). A discrete-continuous algorithm for molecular energy minimization. AHPCRC Preprint 92-031, University of Minnesota.

[7] Northby, J.A. (1987). Structure and binding of Lennard-Jones clusters:$13 \leq N \leq 147$ *J. Chem. Phys.*, **87**, 6166-6177.

[8] Rinnooy Kan, A.H.G. and Timmer, G.T. (1984). A stochastic approach to global optimization. In *Numerical Optimization*, P. Boggs, R. Byrd, and R.B. Schnabel, eds., SIAM, Philadelphia, 245-262.

[9] Smith, S.L., Eskow, E., and Schnabel, R.B. (1989). 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.