Chapter 1
A Parallel Global Optimization Method for Solving Molecular Cluster and Polymer Conformation Problems*

Richard H. Byrd, Elizabeth Eskow, André van der Hoek, Robert B. Schnabel †
Klaas P. B. Oldenkamp‡

Abstract

Identifying the conformations that a molecular cluster or polymer assumes in nature is an important problem with many practical applications in biology and medicine. It is believed that the naturally occurring molecular conformations minimize or nearly minimize the potential energy of the molecular cluster or polymer. The problem of finding the molecular configuration(s) with the lowest potential energy is a challenging global optimization problem with a potentially huge number of local solutions over a very large parameter space. We have developed a method to solve these types of problems, and have experimented with both molecular cluster and polymer applications, with very promising results. We have implemented the method on powerful massively parallel computers because of the enormous computational requirements of solving these types of problems. The parallel algorithms are interesting asynchronous, multi-level parallel algorithms.

1 Introduction

The general problem of finding the naturally occurring three dimensional structure of a polymer given its sequence of monomers includes the problem of finding the native state of a protein given its sequence of amino acids, which is known as the protein folding problem and is one of the fundamental problems in biophysical science [5]. Finding the three dimensional structure of small clusters of water molecules continues to be a topic of interest in atmospheric physics, and has been studied since the late 1800's [2]. The naturally occurring structure for both of these types of problems is believed to be the structure that minimizes or nearly minimizes the potential energy of the chain or cluster of molecules in question. The problem of finding this minimum energy configuration is a difficult and expensive global optimization problem, whose solution requires the construction of new algorithms and the use of powerful computers. The problems are difficult because they have huge numbers of local minimizers, all with small basins of attraction, and many with energy values close to the global minimizer. In this paper, we summarize our algorithmic and parallel computation research on these problems.

We have worked on two different cluster problems, Lennard-Jones [4] and water [3], but will only discuss the water cluster problem in this paper. We use the Coker/Watts

*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 Grants No. CDA-8922510 and CCR-9101795.

†Department of Computer Science, University of Colorado at Boulder, Boulder, CO

‡Department of Operations Research, Econometric Institute, Erasmus University Rotterdam, The Netherlands

energy function [6] to give the potential energy of water clusters. This has the form $\sum_{i \neq j} u(x_i, x_j) + \sum_i v(x_i)$ where each vector x_i gives the coordinates of the i-th molecule, and the functions u and v give the interaction energy between pairs of molecules and the internal energy of the molecule, respectively. The polymer problem we are working on initially is the protein polyalanine, using the CHARMM energy function to compute the potential energy [1]. This function is the sum of internal energy terms for bonded atoms plus nonbonded interactions. The internal energy terms include the bond potential, bond angle potential and dihedral angle potential for both proper and improper dihedral angles. Letting a_i,a_j,a_k,a_l represent atoms with i bonded to j, j bonded to k and k or j bonded to l, these terms are $\sum_{bond}(a_i, a_j) + \sum_{angle}(a_i, a_j, a_k) + \sum_{dihedral}(a_i, a_j, a_k, a_l)$. The nonbonded terms include the Van der Waals interactions and the electrostatic potential and are $\sum_{VdW}(a_m, a_n) + \sum_{elec}(a_m, a_n)$ where atoms a_m and a_n are not bonded. The CHARMM function uses cartesian coordinates for the atoms, whereas the parameters for the global optimization are based on an internal set of parameters given in terms of bond lengths, bond angles and dihedral angles. This parameterization is more natural because the dihedral angles are the crucial parameters to be varied in the optimization and can be easily isolated in the internal parameterization. In addition, if only the dihedral angles are used as optimization parameters, the global optimization is much more efficient because the number of variables is greatly reduced. Using the internal parameterization leads to some interesting algorithmic challenges, however. In the cluster problem, the cartesian parameterization is used throughout.

We have developed a fairly general large scale global optimization method for these types of problems, including some techniques that are specialized for molecular cluster or chain type problems. A key feature of the method is working in smaller dimensional subspaces in certain portions of the algorithm. We review this method briefly in Section 2. The methods are implemented on parallel computers because they require large amounts of computation, and the parallel algorithms are interesting because they are asynchronous and use several levels of parallelism. We summarize the parallel implementation and experimental results on its efficiency in Section 3, and the overall test results in Section 4.

2 The Global Optimization Method

The framework for our current global optimization algorithm is given in Algorithm 1. The method combines an initial phase that locates a beginning set of low minimizers, with a second, more deterministic phase for progressing from low to even lower local minimizers.

The initial phase is derived in part from previous stochastic methods based on the work of [11]. For the cluster problem, random clusters are generated by sampling molecule locations independently over a 3 dimensional region. Then the best sample clusters are randomly perturbed, one molecule at a time, until the energy value reaches a desired threshold level. This perturbation step is needed to find even reasonably low energy configurations. Due to the chain nature of polymers, a considerably different sampling approach is needed for this problem. A buildup procedure is used which samples on each dihedral angle some fixed number of times, and selects the dihedral angle which produces the best partial energy for the part of the chain built so far, before proceeding to the next dihedral angle. For both problem types, a subset of the best sample points are then selected as start points for full dimensional local minimizations, and some number of the best minimizers generated are passed to the second phase of the algorithm.

The second phase accounts for most of the computational effort and success of the method. The basic idea of this phase is to select a configuration from the list of local minimizers, and then select a small subset of its variables for improvement. In the water problem, this small subset is one water molecule; the heuristics for choosing which molecule to optimize are described in [10]. For the polymer problem, the subset is a small number of dihedral angles; the techniques for making that choice are explained in [12]. Once the subset has been determined, a global optimization procedure is executed to find the best new positions for the chosen parameters while holding the remaining parameters fixed. The resulting configurations are then "polished" by applying a full dimensional local minimization using all the variables. The new full-dimensional local minimizers are then merged with those found previously and the entire phase is iterated a fixed number of times.

The method has essentially the same framework for both polymer and cluster type problems with the following exceptions. 1.) The initial sampling approach is quite different for the two problems, using standard sampling with small subset sample improvement for clusters, and a buildup procedure for chains. In general, the initial sampling phase appears to be application dependent, but all that is needed from this phase is some method of locating some reasonably low energy configurations. 2.) For clusters, in phase 2, the cluster is expanded before the small dimensional global optimization step. The expansion permits the method to move to significantly different structures, and has proven to be crucial to the method's success. There is no analogous step in polymer problems due to the nearly fixed nature of bond lengths and angles. 3.) The small dimensional global optimization(2b) is done for a single water molecule for the water cluster problem, and a subset of dihedral angles in the polymer case. 4.) In the cluster algorithm, the phases that work on small subsets of the parameters, sample improvement(1a) and small dimensional global optimization(2b), use an efficient one-molecule energy evaluation that is far cheaper than a full evaluation. The polymer algorithm realizes a similar cost savings in the sampling buildup procedure(1a), but, due to the chain structure, does not currently realize a cost savings in energy evaluations in the small dimensional global optimization(2b). A technique for reducing the costs of these evaluations is currently under investigation, and is mentioned briefly in the section on future research.

3 Parallel Implementation

We have implemented this method to run in parallel on the Intel Delta at Caltech and a Kendall Square KSR1. Parallelizing phase 1 is relatively straightforward; step 1a is distributed among the available processors, and the minimizations(1b) are dealt to processors as the processors become available. However in phase 2, the main part of the algorithm, several levels of parallelism are exploited, and asynchronous scheduling techniques are used, making this an interesting parallel application. In the parallel implementation of phase 2, multiple configurations are worked on simultaneously, each by a group of processors. (If there are, say, 64 processors as is common in our experiments, it is not effective to work on 64 configurations at once, from the viewpoint of processing configurations that are likely to lead to the global minimizer. Rather, we have found it effective to work on 4-16 configurations at once, each with a group of processors.) Each group of processors works asynchronously from the other groups on steps 2b and 2c, and uses a parallel small-scale global optimization algorithm for its group. A centralized scheduler performs step 2a for all of the groups and coordinates the algorithm. This multi-level parallel algorithm appears to make effective utilization of both parallel machines.

Algorithm 1 - Outline of the Global Optimization Algorithm

1. **Coarse Identification of Configurations**
   (a) **Sampling in Full Domain with Sample Point Improvement :** Randomly generate the coordinates of sample configurations in the specified sampling domain, evaluate the energy at each new sample point, and improve energy values by resampling some or all of the coordinates.
   - **Cluster Sampling Improvement:** After initial sampling, for some (low energy) sample points: Select the molecule that contributes the most to the energy function value, randomly sample on new locations for this molecule, replace this molecule in the sample configuration with the new location that gives the lowest energy value, and repeat until energy is below a threshold value.
   - **Polymer Sampling Improvement:** Build up the sample configuration by sequentially generating each dihedral angle in the polymer: randomly sample the current dihedral angle a fixed number of times and select the dihedral angle which gives the lowest energy function value for the partial polymer generated so far, backtracking if no sufficiently low energy value is found.

   (b) **Full-Dimensional Local Minimizations :** Perform a local minimization from a subset of the improved sample points. Save these local minimizers for Step 2a.

2. **Improvement of Local Minimizers :** For some number of iterations:
   (a) **Select a Configuration :** From the list of full-dimensional local minimizers, select a local minimizer to improve, and a subset of its variables to improve upon. This selection uses a heuristic based on the total energy function value, the partial energy function values of the subsets of variables, and the history of selections already made in this phase.
   - **Cluster Configuration Expansion :** Expand the configuration around its center of mass by a fixed factor, generally 1.25-2. (There is no analogous step for polymers.)

   (b) **Global Optimization on a small subset of variables :** Apply a fairly exhaustive small-scale global optimization algorithm to the energy of the selected configuration using the selected small subset of the parameters as variables.

   (c) **Full-Dimensional Local Minimization :** Apply a local minimization procedure, with all parameters as variables, to the lowest configurations that resulted from the global optimization of the previous step, and merge the new local minimizers into the list of local minimizers.

Table 1 gives indicative timing results for a small number of iterations of the water cluster problem executed on the Caltech Delta. The total number of configurations chosen during step 2a of Algorithm 1 for each timing experiment was only 64, whereas a typical production run to find the best water configurations might explore thousands of configurations, but the speedup results should not be different. A minimum of 3 processors are required to execute the second phase of Algorithm 1 on the Delta; due to the mesh architecture of the machine it was convenient to use 4 processors for the initial measurement and double the number of processors for each consecutive one. The results show that as

Table 1: Timing Results on the Caltech Delta
for Phase 2 of the Water Cluster Problem

Number of Processors	Number of Iterations	Number of Simult Configs	Time in Minutes for Asynch Method	Time in Minutes for Synch Method
4	64	1	386.35	382.61
8	32	2	194.55	209.84
16	16	4	97.66	108.42
32	8	6	49.96	58.31
64	4	16	25.77	30.48
128	2	32	14.26	16.36

soon as there is more than one configuration worked on simultaneously, it is more effective to use the asynchronous method described above, rather than to synchronize the work of each group of processors at each iteration. The results show that the asynchronous method is also more efficient in terms of processor utilization: i.e. the efficiency of using up to 64 processors asynchronously is $\geq$ 93.7%, and falls to 84.6% at 128 processors, whereas the efficiencies of the synchronous method fall to 78.5% for 64 processors and 73% for 128 processors.

4 Results

For the water problem, we have mainly run our algorithm on clusters of 20, 21 and 32 water molecules, because results of minimizing these same clusters and energy function, using a dynamic simulated annealing procedure, have been obtained by Long [9]. We have obtained many configurations with significantly lower energies than those obtained in [9]. The best solutions obtained by running our algorithm have energies of -0.348183, -0.369011 and -0.585062 atomic units (a.u.) for 20, 21 and 32 molecules respectively. These energy values are approximately 0.005, 0.01 and 0.02 a.u. lower than the best structures found in [9], whereas at room temperature, only vibrational states with energies about 0.001 a.u. above the ground state are possible. The structures for 20 and 21 molecules have collapsed dodecahedral (for 20) and dodecahedral (for 21) shapes, which are in agreement with the shapes of the lowest energy clusters found in [9].

For the polymer problem, so far we have experimented with 20, 30, 40 and 58 residue polyalanine. The structures of the results for 20, 30 and 40 are very regular alpha-helices, as expected, but we have no comparative energy values from other methods using the same input data and the CHARMM energy function. Work on the 58 residue problem is still in progress. There are conflicting results in the literature regarding what shape to expect for this problem. Wilson and Cui [13] found alpha-helical structures for polyalanine with up to 80 residues using a simulated annealing method; however, Head-Gordon and Stillinger [8] show that for 58 residue polyalanine, a lower minimizer exists whose structure is two parallel alpha-helices connected by a beta bend. Interestingly, so far our algorithm has found two very low energy minimizers for the 58 problem with almost identical energies, each with one of the two different shapes just mentioned. Both conformations require more work to make the alpha-helix(ces) completely regular at the ends, and it will be exciting to compare the energies of the two final, completely regular structures and perhaps gain insight into the true global minimum conformation for this problem.

5 Conclusions and Future Research

The problem of finding the lowest energy configurations for clusters and polymers is an important and difficult problem with many practical applications. The method discussed in this paper has begun to successfully solve problems with magnitudes of hundreds of variables on massively parallel computers in ways that make efficient use of the machines. Practical problems may have thousands or tens of thousands of variables, and these still pose a great challenge.

Future research will include efforts to make our algorithm more efficient in several ways. The full dimensional local minimizations of steps 1b and 2c might be done inexactly at early stages. In the polymer application, a second area for cost improvement would be to decrease the cost of evaluating the potential energy function in the small-dimensional global optimizations (step 2b) by optimizing over consecutive groups of dihedral angles and ignoring the interaction energies of certain portions of the polymer chain. A third area for possibly decreasing costs is to make use of a "smoothed" potential energy function that should have fewer local minimizers than the original problem [7]. The goal would be to use the global minimizer(s) of the smoothed function as starting points for consecutively less smoothed problems.

Finally, our future work will continue to attempt to solve larger and more difficult problems within the application area of polymers and proteins.

References

[1] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *A program for Macromolecular Energy, Minimization, and Dynamics Calculations*, J. Comp. Chem., 4 (1983) pp. 187–217.

[2] I. P. Buffey and W. Byers Brown, *Structure of water clusters with the aid of molecular graphics*, Chem. Phys. Let., 109 (1984) pp. 59–65.

[3] R. H. Byrd, T. Derby, E. Eskow, K. Oldenkamp and R. B. Schnabel, *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, 1994, pp. 71-84.

[4] R. H. Byrd, E. Eskow and R. B. Schnabel, *A New Large-Scale Global Optimization Method and its Application to Lennard-Jones Problems*, University of Colorado Technical Report CU-CS-630-92.

[5] H. S. Chan and K. A. Dill, *The Protein Folding Problem*, Phys. Today, (1993), pp. 24–32.

[6] D. F. Coker and R. O. Watts, *Structure and Vibrational Spectroscopy of the Water Dimer Using Quantum Simulation*, J. Phys. Chem., 91, (1987), pp. 2513–2518.

[7] T. Coleman, D. Shalloway and Z. Wu, *Isotropic Effective Energy Simulated Annealing Searches for Low Energy Molecular Cluster States*, Technical Report CTC92TR113, Cornell University, Advanced Computing Research Institute, (1991).

[8] T. Head-Gordon and F. H. Stillinger, *Enthalpy of Knotted Polypeptides*, J. Phys. Chem., 96, (1992), pp. 7792-7796.

[9] X. Long, *Molecular Dynamics Simulations of Clusters – Impure van der Waals and* $e^- - (H_2O)_n$ *Systems*, Ph.D. Thesis, Department of Chemistry, University of California, San Diego, 1992.

[10] B. Oldenkamp, *Parallel Global Optimization: Applications to Molecular Configuration Problems*, Masters Thesis, Erasmus Universiteit Rotterdam, 1993.

[11] A. H. G. Rinnoy Kan and G. T. Timmer, *A Stochastic Approach to Global Optimization*, in Numerical Optimization, SIAM, Philadelphia, 1984, pp. 245–262.

[12] A. van der Hoek, *Parallel Global Optimization of Proteins*, Masters Thesis, Erasmus Universiteit Rotterdam, 1994.

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