AD-A285 635

PAGE Dist: A

1. AGENCY		3. REPORT TYPE AND DATES COVERED
	8 July 1994	Final , 1 Dec 1991 – 30 Nov 1993

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
New Methods for Large Scale Local and Global Optimization	AFOSR 90-0109

6. AUTHOR(S)

Richard Byrd and Robert Schnabel

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
University of Colorado at Boulder Boulder, CO 80309	153-7645

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
Air Force Office of Scientific Research/NM 110 Duncan Avenue, Suite 100 Bolling Air Force Base, DC 20332-0001	

11. SUPPLEMENTARY NOTES

DTIC
ELECTE
OCT 2 1994
D

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution unlimited. DTIC QUALITY INSPECTED 2	A

13. ABSTRACT (Maximum 200 words)

We have pursued all three topics described in the proposal during this research period. A large amount of effort has gone into the development of large scale global optimization methods for molecular configuration problems. We have developed new general purpose methods that combine efficient stochastic global optimization techniques with several new, more deterministic techniques that account for most of the computational effort, and the success, of the methods. We have applied our methods to Lennard-Jones problems with up to 75 atoms, to water clusters with up to 31 molecules, and polymers with up to 58 amino acids. The results appear to be the best so far by general purpose optimization methods, and appear to be leading to some interesting chemistry issues. Our research on the second topic, tensor methods, has addressed several areas. We have designed and implemented tensor methods for large sparse systems of nonlinear equations and nonlinear least squares, and have obtained excellent test results on a wide range of problems. We have also developed new tensor methods for nonlinearly constrained optimization problem, and have obtained promising theoretical and preliminary computational results. Finally, on the third topic, limited memory methods for large scale optimization, we have developed and implemented new, extremely efficient limited memory methods for bound constrained problems, and new limited memory trust regions methods, both using our recently developed compact representations for quasi-Newton matrices. Computational test results for both methods are promising.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Global optimization, molecular configuration, parallel computation, nonlinear equations, constrained optimization	3
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

Final Technical Report

U.S. Air Force Grant AFOSR-90-0109
New Methods for Large Scale Local and Global Optimization

R. Byrd, R. Schnabel
University of Colorado at Boulder

1 December 1991 - 30 November 1993

1. Summary

We have made significant progress on all three topics described in the proposal during the grant period. A large amount of effort has gone into the development of large scale global optimization methods for molecular configuration problems. We have developed new general purpose methods that combine efficient stochastic global optimization techniques with several new, more deterministic techniques that account for most of the computational effort, and the success, of the methods. We have applied our methods to Lennard-Jones problems with up to 75 atoms, to water clusters with up to 32 molecules, and to polymer structures of up to 58 amino acids. The results appear to be the best so far by general purpose optimization methods, and appear to be leading to some interesting chemistry issues. Our research on the second topic, tensor methods, has addressed several areas. We have designed and implemented tensor methods for large sparse systems of nonlinear equations and nonlinear least squares, and have obtained excellent test results on a wide range of problems. In conjunction we have developed parallel versions of some of these methods and shown that they obtain very good speedups. We have also developed new tensor methods for nonlinearly constrained optimization problem, and have obtained promising theoretical and preliminary computational results for these methods. Finally, on the third topic, limited memory methods for large scale optimization, we have developed and implemented new, extremely efficient limited memory methods for bound constrained optimization problems, and new limited memory trust region methods, both using our recently developed compact representations for quasi-Newton matrices. We have a well-tested code for bound constrained optimization which we plan to make publicly available in the immediate future.

In the following three sections we describe our research on these topics, and also briefly discuss our continuing research plans in these areas.

2. Large Scale Global Optimization Methods for Molecular Configuration Problems

We have been developing global optimization methods for finding the lowest energy configurations of molecular structures. To do this, one must find the lowest (global) minimum of energy functions that generally have very many parameters and huge numbers of local minimizers. Therefore, these are very difficult global optimization problems. Our approach is to develop fairly general purpose methods that do not utilize any knowledge of the solution structure, and are applicable to a broad class of partially separable large scale global optimization problems. The methods combine efficient stochastic global optimization techniques with several new, more deterministic perturbation techniques. So far we have applied our

methods to Lennard-Jones problems with up to 75 atoms, to water clusters with up to 32 molecules whose energy is given by the Coker/Watts potential, and to polymers with up to 59 amino acids with potential energy given by the CHARMM package. The results appear to be the best so far by general purpose optimization methods, and appear to be producing some interesting chemistry issues.

Our methods combine an initial phase that locates some initial low local minimizers and is derived from previous stochastic methods, and a second, more deterministic phase for progressing from low to even lower local minimizers that is new and accounts for most of the computational effort, and the success, of the methods. Both phases make critical use of new portions that vary only a small subset of variables (an atom for Lennard-Jones, a molecule for water, or a small set of torsion angles for polymers) at once. In the initial phase this is used to improve the sample points, in the second phase, to move one atom or molecule in an existing configuration to better positions via the one atom/molecule global optimizations. These steps are relatively inexpensive due to the small number of variables involved and the separability of the energy function. An expansion of the cluster before the one molecule global optimization was added for the water problem, and is crucial to the success of the method because it permits the method to move to significantly different structures. Heuristics for deciding which configuration and molecule to improve next have also been important to the success of the water method. These methods could be applied to any partially separable function, although the determination of the unit to vary at once would be problem dependent. They could also be applied to ab initio energy functions, but new methods would be required to retain the relative efficiency of the one atom/molecule moves.

We have run a simplified version of our algorithm, with no expansion phase, on Lennard-Jones problems with up to 55 atoms. For up to 30 atoms we find the best known solutions, including a solution for 22 atoms that was unknown prior to Northby's special purpose Lennard-Jones method. For over 30 atoms we usually do not find quite as good a solution as the Northby method, although in preliminary tests on problems with 30-40 atoms, adding the expansion phase enables us to find Northby's solutions. Our solutions are apparently by the far the best that have been produced for these problems by a general purpose global optimization method.

For water, we have mainly run our algorithm on clusters of 20 and 21 water molecules, because results of minimizing these same clusters and energy function, using a dynamic simulated annealing procedure, have been obtained by X. Long at University of California, San Diego. We have obtained many configurations with significantly lower energies than Long's. At present, the best solutions obtained by running our algorithm from scratch have energies of -0.3447 and -0.3608 atomic units (a.u.) for 20 and 21 molecules. These are approximately 0.002 and 0.004 a.u. lower than the best structures found by Long, respectively, whereas at room temperature, only vibrational states with energies about 0.001 a.u. above the ground state are likely. By using Long's best 20 water molecule solution as a starting point of the second phase of our algorithm, 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, as a starting point for the second phase of the 21 molecule problem, we obtained many significantly better 21 molecule solutions, with the lowest having energy -0.3679 a.u.. The computational effort expended by our algorithm is considerably less than Long's. We still do not know, however, whether the structures we have found are global minima. Of the current best structures, some have the expected dodecahedral (for 21) or collapsed dodecahedral (for 20) shapes, but some that are very close to the current minimum have far more irregular shapes. If these are indeed possible vibrational states, this raises interesting questions about either the possible structure of water clusters or the validity of the Coker/Watts energy function.

This work has been presented at the SIAM Conference on Optimization and at workshops at the University of California, San Diego and at Iowa State University. A paper on the initial algorithm and its application to Lennard-Jones problems has been completed and submitted for publication, and several other papers are in preparation. We are actively continuing this work, including the porting of this code (which is already parallel) to Intel distributed memory multiprocessors so that we can make longer runs and run on harder problems, and continued testing and algorithm development in the context of the water problem.

We have also begin working on configuration problems associated with polymers. The polymer problem we have begun working with is the protein polyalanine, using the CHARMM energy function to compute the potential energy. As is common in this area we have been treating the bond lengths and bond angles as fixed, and have been trying to find the optimal values of the dihedral angles. This parameterization is natural because the dihedral angles are the crucial parameters to be varied in the optimization, and more efficient because the number of variables is greatly reduced, but the internal parameterization leads to some interesting algorithmic challenges. We have adapted our general approach to this framework in two ways. In the initial phase, the sampling is done by generating dihedral angles sequentially along the chain, and the angle is resampled if it gives a poor value for the potential up to that point. In the second phase we try to improve local minimizers by selecting a small subset of dihedral angles, and doing global minimization on a the resulting small dimensional problem, followed by full dimensional local minimization as above. So far we have been able to find what appears to be the global solution for problems with 10 and 20 amino acids (20 and 40 dihedral angles). We have just begun working on a problem with 58 amino acids. The results here are incomplete but we are using them to tune the algorithm.

3. Tensor Methods for Large Sparse Nonlinear Problems and for Constrained Optimization

Over the last decade we have developed a new class of methods, called tensor methods, for solving nonlinear equations and unconstrained optimization problems. These methods appear to be considerably more efficient and robust than the best standard algorithms based upon Newton's method. During this research period, we have continued this work in several directions. First, we have completed the development and implementation of an efficient software package for tensor methods, as well as standard Newton or Gauss-Newton methods, for solving large sparse systems of nonlinear equations and nonlinear least squares problems. These methods use efficient, state of the art sparse linear algebra techniques. We conducted substantial tests of this software, that show that the tensor methods have large advantages in robustness and computational cost over standard methods. This research is contained in the Ph.D. thesis of Ali Bouaricha and in upcoming papers, and has been discussed in several talks including our plenary talk on tensor methods at the July 1992 SIAM National Meeting.

Secondly, we developed parallel versions of tensor methods for small to medium size problems, and began the development of Krylov-subspace based tensor methods for very large problems that are amenable to efficient parallelization. The tests of the parallel tensor methods on an Intel hypercube showed that there is no loss in parallel efficiency between tensor methods and standard, linear model based methods. This means that tensor methods have the same advantages in computational costs on parallel as on sequential computers. The preliminary results of the Krylov subspace methods showed tensor methods can still lead to substantial advantages in computational cost over analogous Krylov

subspace based linear model methods, even though there is an extra system that needs to be solved at each iteration. A paper on this work has been accepted for publication at the upcoming SIAM Conference of Parallel Processing for Scientific Computation.

Finally, we have continued to develop tensor methods for nonlinearly constrained optimization problems. Our new methods augment the standard linear model of the constraints by a tensor term. These methods are intended to be especially helpful on problems where the active constraints are (nearly) rank deficient at the solution, an important class of problems that are not solved efficiently by current methods. We developed two types of tensor models for nonlinear constrained problems, and a full global algorithm based upon these models. We have analyzed a simple method from this class and shown that it has fast convergence on problems where the constraint Jacobian at the solution has a null space of dimension one, whereas standard methods are only linearly convergent on such problems. We have made computational tests of one of the new methods, and they appear to be exhibiting substantial gains in efficiency over standard methods. In doing this we have developed a generalization of the standard Kuhn-Tucker conditions for constrained optimization in the case of rank deficiency. This generalization is both mathematically interesting and relevant to computational optimization methods. This work has formed a major part of Dan Feng's Ph.D. thesis, and both the theoretical and computational portions of the work are continuing.

4. Limited Memory Methods for Large Scale Optimization

Our research on limited memory quasi-Newton methods has continued and moved in some new directions. Limited memory methods work by generating a quasi-Newton approximation to the Hessian of the objective function that uses only the most recent updates, resulting in great savings in storage, independent of sparsity of the Hessian. Thus they are an important approach to very large optimization problems where the number of variables is too large to allow a full Hessian approximation to be stored.

Our work on a new compact, closed form representation of limited memory quasi-Newton matrices has been essentially completed. A paper describing this work by the principal investigators together with Jorge Nocedal of Northwestern University was completed during this research period has been accepted for publication by *Mathematical Programming*. This representation is a important part of our extensions of the limited memory approach to constrained optimization and to trust region methods discussed below, and several researchers have shown interest is using it also.

For optimization problems with bound constraints, we have developed an algorithm using this new approach, that is extremely efficient in linear algebra cost. In numerical experiments our method is competitive with the partially separable update method of Conn, Gould and Toint, but it is applicable to a broader class of problems, and in some cases it is easier to implement for a particular application. We have written a paper based on this work with Jorge Nocedal. In addition we have developed and tested software using this method to solve bound constrained optimization problems, which we plan to make publicly available, and which we will describe in a separate paper.

We have also developed, together with doctoral student Xuehua Lu, an efficient implementation of a limited memory symmetric rank one method, using a trust region approach. The trust region allows us to model negative curvature, and the compact representation allows the trust region computations efficiently. Preliminary tests for this approach look promising. A talk on this work was given at the TIMS/ORSA Joint National Meeting in Chicago, May 1993.

5. Publications Resulting from this Grant during this Period

(1) R. H. Byrd, D. Liu, D., and J. Nocedal, "On the behavior of Broyden's class of quasi-Newton methods," *SIAM Journal on Optimization* 2, 1992, pp. 533-557.

(2) R. H. Byrd, R. Tapia, and Y. Zhang, "An SQP augmented Lagrangian BFGS algorithm for constrained optimization," *SIAM Journal on Optimization* 2, 1992, pp. 210-241.

(3) X. Zhang, R. H. Byrd, and R. B. Schnabel, "Parallel methods for solving nonlinear block bordered systems of equations," *SIAM Journal on Scientific and Statistical Computing* 13, 1992, pp. 841-859.

(4) S. Smith and R. B. Schnabel, "Centralized and distributed dynamic scheduling for adaptive parallel algorithms," in *Unstructured Scientific Computation on Scalable Multiprocessors*, P. Mehrotra, J. Saltz, and R. Voigt, eds., MIT Press, Cambridge, Mass., 1992, pp. 301-321.

(5) D. Feng, and R. B. Schnabel, "Globally convergent parallel algorithms for solving block bordered systems of nonlinear equations", *Optimization Methods and Software)* 2, 1993, pp. 269-295.

(6) D. Feng, P. Frank, and R. B. Schnabel, "Local convergence analysis of tensor methods for nonlinear equations", *Mathematical Programming 62*, 1993, pp. 427-459.

(7) R. H. Byrd, J. Nocedal, and R. B. Schnabel, "Representations of quasi-Newton matrices and their use in limited memory methods", *Mathematical Programming 63*, 1994, pp. 129-156.

(8) S. Smith and R. B. Schnabel, "Dynamic scheduling strategies for an adaptive asynchronous parallel global optimization algorithm, University of Colorado Technical Report CU-CS-625-92, submitted to *IEEE Parallel and Distributed Systems*.

(9) 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, submitted *SIAM Journal on Optimization*.

(10) Khalfan, H.F., Byrd, R.H., and Schnabel, R.B., "A Theoretical and Experimental Study of the Symmetric Rank-One Update", *SIAM Journal on Optimization 3*, 1993, pp 1-24.

(11) R.H. Byrd, E. Eskow, R. Schnabel and S. Smith, "Parallel global optimization: numerical methods, dynamic scheduling methods, and application to molecular configuration", in *Parallel Computation*, A.E. Fincham and B. Ford, eds., Clarendon Press, Oxford, 1993

(12) R. H. Byrd, P. Lu, J. Nocedal and C. Zhou, "A limited memory algorithm for bound constrained optimization", Technical Report NAM-07, Dept. of Electrical Engineering and Computer Science, Northwestern University, 1993, submitted to *SIAM Journal on Optimization*.

(13) R.H. Byrd, T. Derby, E. Eskow, K. P. B. 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.W. Hager, D.W. Hearn, and P.M. Pardalos eds., Kluwer Academic Publishers, Dordrecht, 1994.

(14) A. Bouaricha and R.B. Schnabel, "Parallel tensor methods for nonlinear equations and nonlinear least squares", in *Proceedings of Sixth SIAM Conference of Parallel Processing for Scientific*

Computation, SIAM, Philadelphia, 1993, pp. 639-643.

(15) R.H. Byrd, T. Derby, E. Eskow, K. Oldenkamp, R.B. Schnabel and C. Triantafillou, "Parallel global optimization methods for molecular configuration problems", in *Proceedings of Sixth SIAM Conference of Parallel Processing for Scientific Computation*, SIAM, Philadelphia, 1993, pp. 165-169.

6. Professional Personnel Supported by Grant during this Period

R. Byrd, Co-Principal Investigator

R. Schnabel, Principal Investigator

Ali Bouaricha, Research Assistant, 12/91 - 8/92
> Mr. Bouaricha completed his Ph.D in Computer Science in August 1992. His research was on tensor methods and software for large sparse problems, and is described in Section 3.

Dan Feng, Research Assistant, 12/91 - 7/92
> Mr. Feng was a Ph.D student in the Department of Computer Science at the University of Colorado at Boulder. He did his doctoral dissertation on the analysis of tensor methods and their extension to constrained optimization described in Section 3. He completed his Ph.D. in summer 1993. (Since August 1992 he has been supported by a different grant.)

Bart Oldenkamp, Research Assistant, 8/92 - 8/93
> Mr. Oldenkamp is a visiting graduate student from Erasmus University in Rotterdam, Holland. He is working on the development, implementation, and testing of the global optimization algorithms for molecular configuration problems described in Section 2. He will do his M.S. thesis partially under our supervision and then return to Erasmus University for his Ph.D.

Christos Triantafillou, Research Assistant, 9/92 - 12/92
> Mr. Triantafillou is a Ph.D student in the Department of Computer Science at the University of Colorado at Boulder. He worked in fall 1992 on the development of parallel methods for the molecular configuration problems described in Section 2.

Andre van der Hoek, Research Assistant, 8/93 - 11/93
> Mr. van der Hoek is a visiting graduate student from Erasmus University in Rotterdam, Holland. He is working on the development, implementation, and testing of the global optimization algorithms for the protein configuration problems described in Section 2. He will do his M.S. thesis partially under our supervision.

Chung-Shang Shao, Research Assistant, 9/93 - 11/93
> Mr. Shao is a Ph.D student in the Department of Computer Science at the University of Colorado at Boulder. He worked in Fall 1992 on experiments on parallel implementations of quasi-Newton methods, with emphasis on molecular configuration computations.

7. Presentations about Grant Research at Technical Meetings during this Period

(1) R. H. Byrd, "The symmetric rank-one method in a trust region context", invited talk, Sixth IIMAS Workshop on Numerical Analysis and Optimization, Oaxaca, Mexico, January 1992.

(2) R. H. Byrd, "Constrained optimization algorithms using limited memory methods", invited minisymposium talk, Fourth SIAM Conference on Optimization, Chicago, May 1992.

(3) R.H. Byrd "Secant methods for large scale constrained optimization", invited talk, TIMS/ORSA Joint National Meeting, Chicago, May 1993.

(4) R. B. Schnabel, "Global optimization methods for molecular configuration problems", invited minisymposium talk, Fourth SIAM Conference on Optimization, Chicago, May 1992.

(5) R. B. Schnabel, "Tensor methods for nonlinear equations and optimization", invited plenary talk, SIAM Annual Meeting, Los Angeles, July 1992.

(6) R. B. Schnabel, "Global optimization methods for molecular configuration problems", invited talk, Joint UCSD-SDSC Workshop on Computational Aspects of Material Science, San Diego, September 1992.

(7) R. B. Schnabel, "Tensor methods for nonlinear equations and optimization", invited talk, Workshop on New Optimization Techniques for Large Compounds, Ames, Iowa, Jan. 1993.

(8) R. B. Schnabel, "Parallel optimization: the good, the bad, and the ugly", invited talk, DIMACS Workshop on Future Directions for Parallel Optimization, New Brunswick, NJ, Apr. 1993.

(9) R. B. Schnabel, "Parallel nonlinear optimization", invited talk NATO Advanced Study Institute on Optimization, Il Ciocco, Italy, Sept. 1993.

(10) R. B. Schnabel, "Parallel and global optimization", invited talk, ORSA-TIMS National Meeting, Phoenix, AZ, Nov. 1993.