On a local search based method for a class of global optimization problems

Thesis Summary

K. M. Rajesh

Indian Institute of Management Calcutta, 2006

Supervisor: Asim Kumar Pal

This thesis is in the area of unconstrained Global Optimization (GO). Specifically, the problem is to find a global minimum for an unconstrained continuous function (Dixon & Szegö, 1978; Törn & Žilinskas, 1989; Kan & Timmer, 1989; Horst & Tuy, 1990). Such problems are ubiquitous in many fields of science, engineering and management domains. Data fitting problems across disciplines, including regression, multidimensional scaling, clustering, etc. (Goffe, Ferrier, & Rogers, 1994; Groenen & Heiser, 1996; Goffe, 1997; Jerrell & Campione, 2001), neural network learning (Hassoun, 1995), molecular cluster problems in computational chemistry (Wales, Doye, Miller, Mortenson, & Walsh, 2000), and finding the equilibrium point in an economy (Wu & Wang, 1998) are some of the examples of GO problems. Since the GO problem is NP-Hard (Vavasis, 1995), most prevalent methods involve the use of stochastic techniques[1].

[1] Hence the phrase "Stochastic Global Optimization" (SGO) is popular in this field.

Funnel functions[2] are an important class of such unconstrained GO problems. Of late, there is growing interest on funnel functions among the optimization community, especially in fields such as computational biology and chemistry. These functions were conjectured while studying protein folding (Leopold, Montal, & Onuchic, 1992; Bryngelson, Onuchic, Socci, & Wolynes, 1995). It is understood now that the objective function in many optimization problems are funnel-like in nature (Wales, 2003). Molecular clustering problems like Lennard-Jones (LJ) and Morse (Locatelli & Schoen, 2002; Doye, Leary, Locatelli, & Schoen, 2004; Wales et al., 2006), densely packing geometrical objects (Addis, Locatelli, & Schoen, 2005b; Specht, 2006) and virus capsids (Wales, 2005) are few examples. Similar functions are found in discrete domain too, where they are called "globally convex" functions (Hu, Klee, & Larman, 1989; Boese, Kahng, & Muddu, 1994).

Multistart local search (Dixon & Szegö, 1978), a theoretically sound global optimization methodology which is also found to work well for various kinds of GO problems, fails to detect the global minimum, even for benchmark funnel functions (Locatelli, 2005). This has led to development of many heuristics in the recent times. Some of the prominent heuristics here work with the local minimum mapping of the objective function instead of the actual objective function. Basin Hopping (BH) algorithm (Wales & Doye, 1997), Monotonic Basin Hopping (MBH) algorithm (Leary, 2000), Local Optimum Smoothing (LOS) algorithm (Addis, 2004; Addis, Locatelli, & Schoen, 2005a), Two-Phase MBH (Locatelli & Schoen, 2003), and Population set based Basin Hopping (Grosso, Locatelli, & Schoen, 2005) are some of the important heuristics under this category. A number of variants of the BH algorithm have also appeared in recent literature (Goedecker,

[2] also called "funnel-like" functions

2004; Iwamatsu & Okabe, 2004; Cheng, Cai, & Shao, 2005).

This thesis proposes efficient metaheuristics for stochastic global optimization of unconstrained funnel functions. We explore the possibility of using simplex based search procedure for funnel functions. Specifically, we use the Nelder-Mead (NM) algorithm (Nelder & Mead, 1965; Olsson & Nelson, 1975), which is still a popular direct search method. It was invented for function minimization as an alternative to descent-based local minimization. In this thesis, we use NM on the local minima mapping of the given objective function, and we call this hybrid Nelder-Mead Local Search (NMLS) algorithm. Towards this end, each vertex of NM would be replaced by its nearest local minimum. It is shown that NMLS does not always converge to the funnel bottom of single funnel functions in one dimension. To overcome this, we propose NMLS-S, which is a simpler variant of NMLS. In NMLS-S outside and inside contraction steps of NMLS are replaced by a shrink step. Convergence to the funnel bottom in the case of one-dimensional single funnel functions is obtained for NMLS-S. Preliminary observations about the case with dimensions more than one are made. It is to be mentioned here that the literature so far has a very poor record of theoretical proofs: be it for NM, MBH or LOS - old and modern algorithms alike (Lagarias, Reeds, Wright, & Wright, 1998; McKinnon, 1998; Kelley, 1999; Leary, 2000; Addis, 2004; Addis et al., 2005a). Neither does this minimize the practical utility of these algorithms nor does it minimize the importance of theoretical results. This is similar to the case with other popular class of algorithms like Genetic Algorithms and Tabu Search. Obtaining theoretical results on the convergence of NMLS and NMLS-S, for both single and multiple funnels, for dimensions more than one, should therefore be an important area of future research.

<table>
<tr><td>Algorithms (Ω)</td><td>NMLS (N1), NMLS-S (N2), MBH (MB) and LOS (LO)</td></tr>
<tr><td>Objective function (f)</td><td>Rastrigin (8 variants), Ackley, and Levy</td></tr>
<tr><td>Complexity of f</td><td>Dimensions ($n = 20, 50$ and 100)
Amplitude (A) (three levels)*
Perturbation (P) (four leves)*
Asymmetry in size of level sets*
Asymmetry in function values of local minima*</td></tr>
<tr><td>Unit of experiment</td><td>$\langle \Omega, f, n, \Delta \rangle$**
1000 runs per experiment (some cases 100)</td></tr>
<tr><td>Primary metrics for $\langle \Omega, f, n, \Delta \rangle$</td><td>Success rate, $Success$
Local searches per success with and without stopping criteria, $LSWOS/S$ and $LSWS/S$
Smoothing per success for LOS with and without stopping criteria, $Sm/S - wos$ and $Sm/S - ws$
Remediation per success for NMLS and NMLS-S with and without stopping criteria, $R/S - wos$ and $R/S - ws$</td></tr>
<tr><td>Selection of Δ</td><td>Manual (depends on nonlinear behaviour of Ω)</td></tr>
<tr><td>Secondary metrics for $\langle \Omega, f, n \rangle$</td><td>Expected succes rate, $E(Success)$
Expected number of local searches per success, with and without stopping, $E(LSWOS/S)$ and $E(LSWS/S)$
Range of 95% confidence interval, $I-95$, (normalized with 1 for MB and LO and with $\sqrt{n}$ for N1 and N2)
Optimistic (O) and Pessimistic (P) estimates for each</td></tr>
<tr><td>Performance profile</td><td>Each Ω Vs the best Ω (for each secondary metric)</td></tr>
<tr><td>Ratios (R1 and R2) of secondary metrics***</td><td>Lower Estimate (LE) and Upper Estimate (UE)</td></tr>
</table>

* Variants for Rastrigin function only

** Δ denotes Radius of the neighbourhood hypersphere for MBH and LOS, and side-length of the simplex for NMLS and NMLS-S, normalized w.r.t. the side-length of the searching hypercube

*** R1 for NMLS Vs MBH & LOS, and R2 for NMLS-S Vs MBH & LOS

Table 1: Schema of Experiments on the Systematic Set

Extensive empirical investigations were carried out on the four algorithms: NMLS, NMLS-S, MBH and LOS. The experiments were categorized as *systematic* and *exploratory* sets. Detailed experiments on popular benchmark functions are presented under the systematic set[3]. Scoping experiments on problems of special interest are categorized under the exploratory set.

Results of the systematic set have clearly demonstrated the superiority of the NMLS and NMLS-S algorithms over the competing (or existing) algorithms MBH and LOS for low level functions. The strengths of the proposed algorithms have been observed in their robustness in finding the global minima as well as in the efficiency of the search. NMLS and NMLS-S uniformly outperformed MBH and LOS with respect to all performance metrics, across various complexities of each benchmark function tested. All these algorithms deteriorate in their performance with increasing complexity of the objective function[4]. However, the deterioration is slower for NMLS and NMLS-S compared to MBH and LOS, as asymmetries are introduced in the size of the level sets. Further, NMLS and NMLS-S are found to be much less sensitive to the choice of tunable parameters[5]. Between NMLS and NMLS-S, the latter has been found to work better.

The exploratory set of experiments are still in a preliminary stage to make any conclusive observations. We have tried to explore in a few directions and documented the initial findings. Cases of mixing of complexities of the objective function, multifunnel functions[6] and Lennard-Jones clusters were the problems tried. The algorithms proposed in the thesis were able to find the putative global

[3]Refer Table 1 for the experimental setup of systematic experiments.

[4]Complexities considered are dimensions, amplitude, perturbation, asymmetry in level sets and asymmetry in the function values of local minima.

[5]Two key parameters were studied: Neighbourhood size and stopping number of iterations

[6]A two level nested algorithm was tried.

optimum for many configurations of LJ clusters. This fact is encouraging, but the probe need to be made much deeper.

Efforts to make NMLS and NMLS-S more efficient on multifunnel functions, performance of these algorithms on real life problems, finding new applications of funnel functions, especially in the broad discipline of management, work on the theoretical characterizations of funnel functions and study of the relation of the funnel functions to the "globally convex" functions are some of the natural extensions of the thesis work.

Contributions

Literature survey: Funnel functions and multilevel funnels in unconstrained global optimization have been discussed with the help of a directed graph representation since no clear definition of funnel or funnel-like functions has been found in the literature which is widely accepted. However, one definition of funnel found in the literature has been used to argue that there could be multiple categories (some not so trivial) of single funnel functions in dimensions more than one.

Nelder Mead based hybrids: Two simple hybrids based on Nelder Mead algorithm and local optimization called the Nelder Mead Local Search (NMLS) and NMLS-Shrink variant (NMLS-S) are proposed.

Convergence issues: Non-convergence of NMLS and convergence of NMLS-S to the funnel bottom in one dimension are demonstrated. Convergence issues of NMLS and NMLS-S *vis-à-vis* MBH for multiple dimensions are discussed. It is argued that NMLS and NMLS-S would perform better than MBH and

LOS for funnel functions (which is subsequently substantiated by empirical study).

Empirical analysis on systematic set: Based on the results on a systematic set of experiments the following observations were made on the performance of NMLS and NMLS-S with respect to MBH and LOS. Note that observations 2 and 3 involve the following complexities of objective function − amplitude, perturbation, asymmetry in local searches and function values of local minima[7]. **For studying these complexities, only Rastrigin function was used with different variants in our experiments.**

Observation 1: Performance of NMLS & NMLS-S are uniformly better than that of MBH & LOS.

Observation 2: Performance of each of the four algorithms generally deteriorates with increasing complexity of the objective function.

Observation 3: NMLS & NMLS-S are less sensitive to asymmetry in the size of the local minimum basins compared to MBH & LOS. For other complexities of the objective function the relative behaviour is in favour of NMLS & NMLS-S.

Observation 4: NMLS & NMLS-S are robust to side-length (of the simplex) whereas MBH & LOS are too sensitive to radius (of the neighbourhood hypersphere). Both NMLS and NMLS-S work for any side-length above a minimum threshold.

Observation 5: NMLS & NMLS-S are robust to the number of iterations in the stopping phase, whereas MBH & LOS are not so.

Observation 6: Performance of NMLS-S is better than that of NMLS.

[7] Dimension is another

Empirical analysis on exploratory set: An exploratory set of experiments were conducted to study the performance of the algorithms on problems of special interest. These experiments are still in a preliminary stage to make any conclusive observations. We have tried to explore in the following directions and documented the initial findings:

- Mixing of complexities of the objective function.
- Multifunnel functions
- LJ cluster

Keywords: Unconstrained function minimization, Funnel functions, Local minimum mapping, Monotonic Basin Hopping, Local Optima Smoothing, Simplex Based Search, Nelder Mead Local Search, NMLS-Shrink variant

References

Addis, B. (2004). *Global optimization using local searches*. Unpublished doctoral dissertation, Global Optimization Laboratory, University of Florence.

Addis, B., Locatelli, M., & Schoen, F. (2005a). Local optima smoothing for global optimization. *Optimization Methods and Software*, *20*(4, 5), 1-21.

Addis, B., Locatelli, M., & Schoen, F. (2005b). *Packing circles in a square: New putative optima obtained via global optimization* (Technical Report No. DSI 1-2005). Italy: Global Optimization Laboratory, University of Florence. (`http://www.optimization-online.org/DB_HTML/2005/03/1088.html`; last accessed on 21 November 2006)

Boese, K. D., Kahng, A. B., & Muddu, S. (1994). A new adaptive multi-start technique for combinatorial global optimizations. *Operations Research Letters*, *16*, 101-113. (The correspoding technical report of UCLA-CS Department can be found at `ftp://ftp.cs.ucla.edu/tech-report/1993-reports/930015.pdf`; last accessed on 20 October 2005)

Bryngelson, J. D., Onuchic, J. N., Socci, N. D., & Wolynes, P. G. (1995). Funnels, pathways and the energy landscape of protein folding: A synthesis. *Proteins: Structure, Function and Genetics*, *21*(3), 167-195.

Cheng, L., Cai, W., & Shao, X. (2005). An energy-based perturbation and a taboo strategy for improving the searching ability of stochastic structural optimization methods. *Chemical Physics Letters*, *404*, 182-186.

Dixon, L. C. W., & Szegö, G. P. (1978). *Towards global optimization 2.*

Amsterdam: North-Holland.

Doye, J. P. K., Leary, R. H., Locatelli, M., & Schoen, F. (2004, Fall). The global optimization of Morse clusters by potential energy transformations. *INFORMS Journal of Computing*, *16*(4), 371-379. (The original Technical Note available in `http://www.optimization-online.org/DB_HTML/2003/07/689.html`; last accessed on 23 October 2005.)

Goedecker, S. (2004, June). Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *Journal of Chemical Physics*, *21*, 9911-9917.

Goffe, W. L. (1997). *A toolkit for optimizing functions in Economics* (Technical Report). Economics Working Paper Archive at WUSTL. (`http://ideas.repec.org/e/pgo13.html`; last accessed on 4 October 2005)

Goffe, W. L., Ferrier, G. D., & Rogers, J. (1994). Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, *60*(1-2), 65-99. (`http://cook.rfe.org/anneal-preprint.pdf`; last accessed on 20 October 2005)

Groenen, P. J. F., & Heiser, W. J. (1996). Tunneling method for global optimization in multi-dimensional scaling. *Psychometrika*, *61*, 529-550.

Grosso, A., Locatelli, M., & Schoen, F. (2005). *A population based approach for hard global optimization problems based on dissimilarity measures* (Technical Report). Italy: Global Optimization Laboratory, University of Florence. (`http://http://www.optimization-online.org/DB_HTML/2005/02/105%6.html`; last accessed on 2 October 2005)

Hassoun, M. H. (1995). *Fundamentals of artificial neural networks.* New Delhi: Prentice Hall of India.

Horst, R., & Tuy, H. (1990). *Global optimizaton: Deterministic approaches.* Berlin:

Springer-Verlag.

Hu, T. C., Klee, V., & Larman, D. (1989, September). Optimization of globally convex functions. *SIAM Journal of Control and Optimization*, *27*(5), 1026-1047.

Iwamatsu, M., & Okabe, Y. (2004, October). Basin hopping with occasional jumping. *Chemical Physics Letters*, *399*, 396-400.

Jerrell, M. E., & Campione, W. A. (2001). Global optimization of econometric functions. *Journal of Global Optimization*, *20*, 273-295.

Kan, A. H. G. R., & Timmer, G. T. (1989). Global optimization. In G. L. Nemhauser, M. J. Todd, & A. H. G. R. Kan (Eds.), *Handbook in operations research and management science, volume i: Optimization* (p. 631-662). Amsterdam: Elsevier Science B. V.

Kelley, C. T. (1999). Detection and remediation of stagnation in the Nelder-Mead algorithm using a sufficient decrease condition. *SIAM Journal on Optimization*, *10*(1), 43-55.

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal on Optimization*, *9*(1), 112-147.

Leary, R. H. (2000, December). Global optimization on funneling landscapes. *Jounral of Global Optimization*, *18*(4), 367-383.

Leopold, P. E., Montal, M., & Onuchic, J. N. (1992, September). Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(18), 8721-8725. (`http://www.pnas.org/cgi/reprint/89/18/8721`; last accessed on May 3 2006)

Locatelli, M. (2005). On the multilevel structure of global optimization problems.

Computational Optimization and Applications, *30*, 5-22.

Locatelli, M., & Schoen, F. (2002). Fast global optimization of difficult Lennard-Jones clusters. *Computational Optimization and Applications*, *21*(1), 55-70.

Locatelli, M., & Schoen, F. (2003). Efficient algorithms for large scale global optimization: Lennard-Jones clusters. *Computational Optimization and Applications*, *26*, 173-190.

McKinnon, K. I. M. (1998). Convergence of the Nelder-Mead simplex method to a non-stationary point. *SIAM Journal on Optimization*, *9*(1), 148-158.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, *7*, 308-313.

Olsson, D. M., & Nelson, L. S. (1975, February). The Nelder-Mead simplex procedure for function minimization. *Technometrics*, *17*(1), 45-51.

Specht, E. (2006). *Packomania.* (`http://www.packomania.com`; last accessed on 30 July 2006)

Törn, A., & Žilinskas, A. (1989). *Lecture Notes in Computer Science: Global Optimization* (Vol. 350). Heidelberg: Springer-Verlag.

Vavasis, S. A. (1995). Complexity issues in global optimization: a survey. In R. Horst & P. Pardalos (Eds.), *Handbook of global optimization* (p. 27-41). Dordrecht: Kluwer Publications.

Wales, D. J. (2003). *Energy landscapes.* Cambridge: Cambridge University Press.

Wales, D. J. (2005). The energy landscape as a unifying theme in molecular science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *363*(1827), 357-377.

Wales, D. J., & Doye, J. P. K. (1997). Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing upto 110 atoms. *Journal of Physical Chemistry A*, *101*, 5111-5116.

Wales, D. J., Doye, J. P. K., Dullweber, A., Hodges, M. P., Naumkin, F. Y., Calvo, F., et al. (2006). *The Cambridge Cluster Database.* (http://brian.ch.cam.ac.uk/CCD.html; last accessed on 4 October 2005)

Wales, D. J., Doye, J. P. K., Miller, M. A., Mortenson, P. N., & Walsh, T. R. (2000). Energy landscapes: From clusters to biomolecules. *Advances in Chemical Physics*, *115*, 1-111.

Wu, L., & Wang, Y. (1998). An introduction to SA algorithms for computation of economic equilibrium. *Computational Economics*, *12*, 151-169.