International Congress on Evolutionary Methods for Design,
Optimization and Control with Applications to Industrial Problems
EUROGEN 2003
G. Bugeda, J. A- Désidéri, J. Periaux, M. Schoenauer and G. Winter (Eds)
© CIMNE, Barcelona, 2003

GLOBAL SEARCH METHODS FOR NONLINEAR OPTIMISATION: A NEW PROBABILISTIC-STOCHASTIC APPROACH

G. Bugeda*, E. Balsa-Canto*, C. Thamotheram†, E. Oñate* and F. Zarate*

* International Center for Numerical Methods in Engineering (CIMNE)
Universidad Politécnica de Cataluña
Campus Norte UPC, 08034 Barcelona, Spain
e-mail: cimne@cimne.upc.es, web page: http://cimne.upc.es/

† Dept. of Aeronautics, Imperial College of Science, Technology, and Medicine,
London, UK.
Web page: http://www.ic.ac.uk

Key words: global optimisation, stochastic-probabilistic approaches, Monte-Carlo, optimal design, shape optimisation.

Abstract. *In this work the problem of overcoming local minima in the solution of nonlinear optimisation problems is addressed. As a first step, the existing nonlinear local and global optimisation methods are reviewed so as to identify their advantages and disadvantages. Then, the major capabilities of a number of successful methods such as genetic, deterministic global optimisation methods and simmulated annealing, are combined to develop an alternative global optimisation approach based on a Stochastic-Probabilistic heuristic.*

The capabilities, in terms of robustness and efficiency, of this new approach are validated through the solution of a number of nonlinear optimisation problems. A well know evolutionary technique (Differential Evolution) is also considered for the solution of these case studies offering a better insight of the possibilities of the method proposed here.

1 INTRODUCTION

Many everyday problems in engineering, decision sciences and operations research may be formulated as optimisation problems. Three are the key components of an optimisation problem: a mathematical model which describes the phenomena as a set of *constraints* on a number of unknown *decision variables*, and a measure of quality, the *objective function*, whose maximum (or the minimum of its negative or inverse) corresponds to the optimal solution. Some examples include the calculation of a set of unknowns to obtain an optimal design, maximum production, minimum environmental impact, maximum quality etc…

The nature of the decision variables, the objective function and the constraints determine the type of optimisation problem and the different levels of complexity. The cases to be considered in this work are of continuous and non-linear nature. For this type of problems, if the objective function turns out to be smooth and differentiable, analytical methods will produce the exact solution, if not, for example in the design of aerodynamic components and systems, the use of suitable techniques to explore the entire search space are then required. Moreover, many real world optimisation problems, are not only non-linear, but constrained and large scale which makes their solution really complex.

Except for very simple cases, non-linear optimisation problems can not be solved analytically. Therefore during the last decades a number of different methods have been proposed for the their numerical solution.

Local optimisation methods

Can be applied to the solution of linear and non-linear optimisation problems. Some of these techniques, such as Downhill Simplex and Powell's method do not require the derivatives of the objective function. Although very popular, due to its easy applicability, it has been demonstrated that the use of extra information regarding the search space in terms of, at least, the gradient, as in the steepest-descent or quasi-Newton methods, can considerably speed up the optimisation process. Newton strategies additionally require the second partial derivatives, thus building a quadratic internal model, and achieving up to quadratic convergence properties[i-ii].

Although very efficient in the solution of uni-modal and/or large scale optimisation problems, in a multi-modal environment these algorithms move "downhill" from their respective starting points, hence, converging to the closest local optimum.

Global optimisation methods.

In many engineering optimisation problems, it will be rarely possible to write down the objective function in a closed form, and a simulation model is often required so as to reproduce reality [iii]. In general, these models will not behave smoothly and the objective function may contain numerous local optima with corresponding function values varying significantly.

These difficulties motivated the development of the so called global optimisation methods. Among the different possibilities, *stochastic optimisation methods* will be considered in this work as they are quite easy to use and implement and present the ability to escape from local solutions. However, global optimality can not be guaranteed except in an asymptotic probabilistic sense as the number of iterations tends to infinite.

These strategies can be classified in two main groups: *adaptive sequential methods*, that generate a new point in each iteration, as for example the Simulated Annealing [iv], which takes analogies form physical systems. And *population based methods*, which generate a set of points in each iteration. Many of the methods in this last group are somehow inspired in biological processes. For example, the evolutionary algorithms (EAs) and genetic algorithms (GAs) introduced in the early 70s by Holland and applied for the first time to practical problems by Goldberg[v], mimic the concept of the natural selection and survival of the fittest.

Although adaptive sequential methods have very good properties, population based strategies are becoming more a more popular as they are able to build up an overall picture of the search space. In any case, the successful methodologies combine effective mechanisms of exploration of the search space and exploitation of the previous knowledge obtained by the search.

2 NEW PROBABILISTIC-STOCHASTIC OPTIMISATION METHOD

This work presents the recent experiences on the development of a new alternative for the solution of general multi-modal problems which is based on a probabilistic-stochastic approach and which takes advantage of different elements from other successful methods, both local and global, reviewed above.

The main characteristics of this approach are the following:

▸ *Population based*, as in Genetic Algorithms: the well known Monte Carlo method is used to generate the populations and extract statistical information of the search space. It also takes advantage of the typical GAs operators:

- Regarding the *selection process*, the probabilistic stochastic method proposed successively choose the search area for the next generation using a probabilistic greedy criterion. Thus, giving a higher probability of exploration to the areas where the current best information is being extracted from.
- Regarding the *re–combination*, in evolutionary algorithms components of the design variable vector are randomly exchanged between parents to try and maintain the diversity within the population. This is being done implicitly in the stochastic method by taking random values of the design variables in relation to their probability density functions.
- Finally, *mutation* in evolutionary algorithms is used to generate individuals outside of the range of the existing population to explore new areas of the search space. This happens in the new approach method when the initial uniform distribution is changed to a normal distribution. Thus, values of the design variable are allowed to be generated outside the range of the initial distribution with a correspondingly lower probability.

▸ *Gradient-like information*, as in local and also global deterministic methods. It has been largely demonstrated that the use of gradient information clearly enhances the convergence properties of minimisation methods. Considering that many of the problems to be solved might not be differentiable the use of gradient-like information which provides a measure of the closeness to a solution, is proposed in this method.

- *Uphill movements*, as in simulated annealing: in order to avoid getting trapped in local solutions.

Probabilistic-stochastic method main elements are described in the following sections.

3 MONTE CARLO SIMULATION (MCS)

Monte Carlo (MC) methods are stochastic techniques, based on the use of random numbers and probability statistics, which allow the simulation of processes under a large number of different conditions so as to obtain a statistical measure of the outputs.

If a variable X can take a real value in the interval a ≤ X ≤ b, and if the upper limit b, is a deterministic real value, x. Then the probability density function will describe the expected concentration and spread of the random variable X within the range [a, b].

$$P[x \le X \le x + \Delta x] = F(x + \Delta x) - F(x) = p(x)dx \qquad (1)$$

where: $p(x) = \frac{F(x+\Delta x) - F(x)}{dx}$ and F(x) is the probability that X is less than or equal to x.

As Δx approaches zero, p(x) approaches dF(x)/dx, the first derivative of the cumulative distribution function usually called the probability density function, hereafter referred to as PDF. The cumulative distribution frequency is therefore recoverable from the PDF as:

$$F(x) = \int_a^x p(x)dx \qquad (2)$$

The MCS works by producing a random uniform number, U(0,1), then inverting the PDF describing the design variable parameters to obtain the corresponding random deviate[vi-vii].

This process is carried out for every design variable until all design variables have been assigned a random deviate in accordance with their PDF. The input variables are then entered into the mathematical model and the values of the output variables are recorded. The process is known as taking a shot and the amount of shots has to be sufficient to ensure the Monte Carlo simulation has converged from a statistical point of view. This is done by considering the confidence intervals on the mean as described below.

With a given set of observation of values the true mean and a range that includes the mean true value with a high level of certainty are estimated.

4 PROBABILISTIC - STOCHASTIC HEURISTIC (PSH)

The PDF of the decision variables, as described above, is the main guide for the optimisation process, as it dictates the area of the search space to be explored and with what frequency. Different probabilistic transition conditions were established to alter the probability density functions so as to examine the search space.

The population is initialized so that the entire search space is explored. This is done by defining the input design vectors, **x**, PDF as uniform within the feasible design space.

$$p(x) = \frac{1}{b-a} \qquad a \le x \le b \tag{3}$$

The following iterations were defined by sampling the fittest P% of the current generation, G, and taking the mean, μ_{G+1}, and the standard deviation, σ_{G+1}, of each design variable to define a normal distribution as the input PDF for the next generation, G+1:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-1/2\left(\frac{x-\mu}{\sigma}\right)^2\right) \qquad -\infty \le x_i \le +\infty \tag{4}$$

This basic process is illustrated in Figure 1 for a very simple unidimensional function:

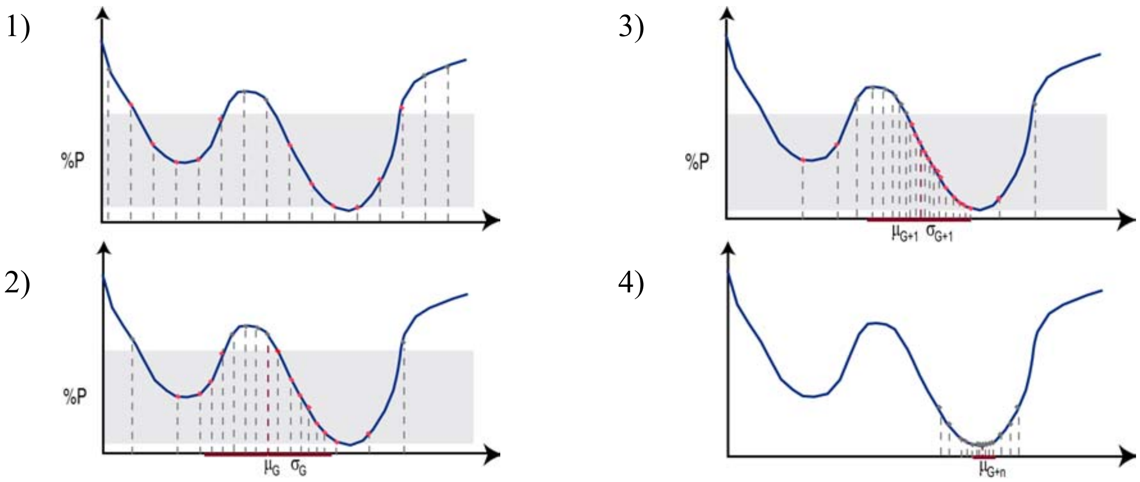

Figure 1: Schematic behavior of the PSh method for an unidimensional multimodal function.

The standard deviation is a measure by how much values within a sample differ from the mean and therefore implicitly information of the gradient For one generation there are NP population members, the standard deviation is defined as:

$$\sigma = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{NP-1} \tag{5}$$

If values for the design variables are produced outside the feasible space, they are penalised creating the effect of truncated normal distributions.

This process is repeated until convergence occurs, that is, when two consecutive iterations show less than one percent improvement in the objective function. This approach will be called from now PSh*.

Note that as the process follows the standard deviation is being reduced behaving as the norm of the gradient vector in a gradient based optimization method providing extra information on the closeness of the solution.

Additional conditions are also imposed when defining the normal distributions, in order to benefit from knowledge on where the top P% of the population is located in relation to the time spent searching in that area. This was done by considering the probability of where the mean μ_{G+1} lies in relation to μ_G. The so called *standard mean displacement* is calculated as follows:

$$Standard\ Mean\ Displacement = \frac{\mu_{G+1} - \mu_G}{\sigma_G} \tag{6}$$

As the magnitude of the standard mean displacement increases the possibility of exploring a new area of the search space also increases, this is equivalent to a lower probability of re-exploring already visited space with the corresponding computational cost reduction (this is illustrated in Figure 2 for one decision variable).

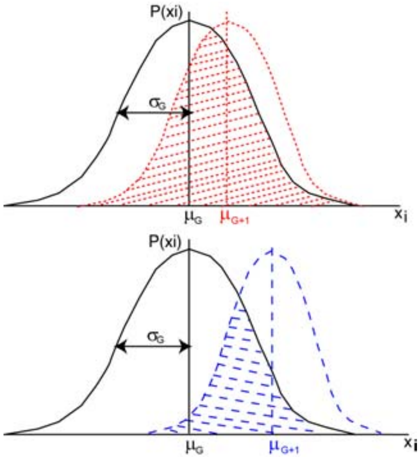

Figure 2: Illustration of the standard mean displacement effect.

In order to avoid unnecessary waste of search time the magnitude of the standard mean movement my be modified by a parameter G, $\mu_{G+1} = G(\mu_{G+1} - \mu_G) + \mu_G$, so as to reduce the probability of re-exploring already visited space.

When the search is shifted in a given direction these conditions have the effect of giving the search a slight push in that direction. The magnitude of the push (G) indicates the greediness of the search. By increasing the search space we are taking into the account that the global minimum is not in the immediate region currently being searched and allowing the algorithm to take uphill steps. The method PSh* with these new conditions on the standard mean displacement will be called PSh.

5 APPLICATION TO AERODYNAMIC SHAPE OPTIMISATION

Numerical optimisation methods have been largely applied in the aerodynamic sector[viii-x].. All solution approaches are based on some structural or fluid dynamic simulation method and are coupled with appropriate optimisation algorithms to adjust a first guess performance to shapes that perform according to specified targets.

. The problems to be considered here are related to airfoil design. Typical targets include prescribed pressure or velocity distributions, lift range, maximum lift, minimal drag, etc.., under geometrical constraints that may include one or more of the following: thickness ratio, maximum slope, etc...

This section summarizes the results obtained in the solution of an inverse problem, which consists on recovering an aerofoil shape from a target coefficient of pressure distribution. The candidate aerfoil shapes were parameterized using two Bézier curves[xi], one for the mean camber line and one for the half thickness distribution (t) of the airfoil. By superimposing the half thickness normal to the mean camber line a wide range of feasible airfoils can be produced. Regarding the numerical simulation of the aerodynamic problem the panel method[xii] was considered.

In order to evaluate PSh* and PSh posibilities, a well know evolutionary method, Differential evolution (De[xiii-xiv]), was selected as a basis for comparison. DE makes use of the evolution scheme and has demonstrated to be very robust and efficient for the solution of multimodal optimisation problems.

Figures 3 and 4 show the convergence curves for DE, and the methods PSh* and PSh proposed here, for two different selected targets:

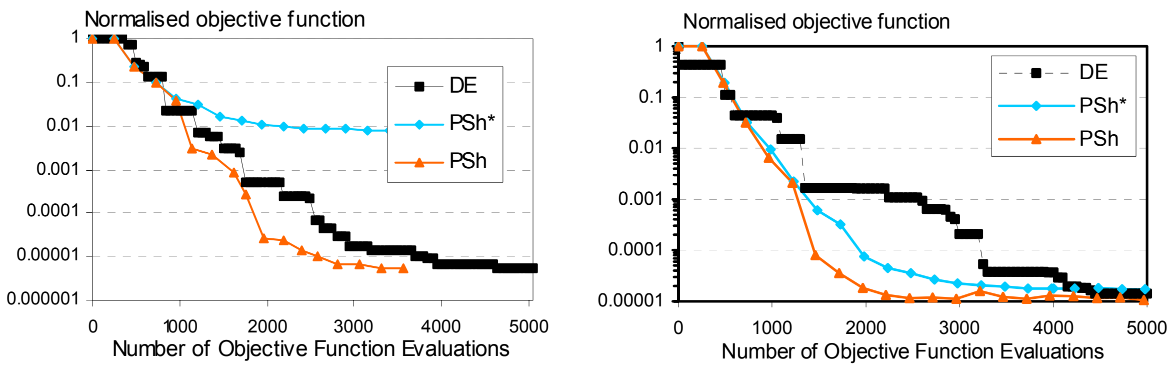

Figure 3-4: Comparison of convergence curves.

As it may be seen from the figures, all methods present the typical behaviour of stochastic approaches, fast convergence to the vicinity of the solution but very slow refinement till arriving to the global. Note that for the first case PSh* converges to a local solution whereas PSh with further conditions on the standard mean deviation is able to escape and finally arrive quite rapidly to the same solution as DE. For the second case, although all methods have converged to almost the same solution, velocities of convergence were quite different. Particularly, both PSh* and PSh were faster than DE, arriving to the region of the optimal solution in around a half to three quarters the time required by DE. Remark that imposing conditions on the standard deviation displacement makes PSh to converge at a quicker rate than PSh*.

6 CONCLUSIONS AND FUTURE WORK

The new probabilistic stochastic methods developed in this work, particularly PSh, presented good performance, showing that is both robust and efficient, outperforming DE for the cases considered. One of the main characteristics of this method is its ability to arrive to the vicinity of the solution very rapidly, which qualifies it to implement possible hybrid methods which combine global optimisation techniques with local ones to improve convergence properties.

The authors feel that it is a combination of the properties described in section 3 that give the probabilistic–stochastic PSh a wide scope for future development. It is expected that large improvements could be made by directly linking the magnitude of the greediness to the standard mean movement. This would give a dynamic link between where the best information is being obtained in the search and the amount of time spent searching in that area, providing a more adaptive probabilistic self adaptation strategy parameter as opposed to the static one used in this work.

REFERENCES

[i] Bazaraa, M.S., H.D. Sherali and C.M. Shetty. “Nonlinear Programming: Theory and Applications”. John Wiley & Sons, New York. 1993.
[ii] Nash, S.G. and A. Sofer. “Linear and Nonlinear Programming”. McGraw- Hill, 1996.
[iii] Floudas, C.A. “Deterministic Global Optimization: Theory, Methods and Applications”. Kluwer Acad., Holland, 2000.
[iv] Kirkpatrick, S., C. D. Gelatt Jr., M. P. Vecchi, "Optimization by Simulated Annealing", Science, 220, 4598, 671-680, 1983.
[v] Goldberg, D.E. “Genetic Algorithms in Search, Optimization and Machine Learning”. Addison Wesley Longman, 1989.
[vi] Hurtado, J. E. , F. Zárate, E. Oñate. “The Monte Carlo method. Application to the stochastic analysis of sheet stamping process”, Publication CIMNE Nº 184, Barcelona, February 2000.
[vii] Zárate, F., J. Hurtado, E. Oñate, J.A. Rodríguez: “Stac - MCS software”, CIMNE, International Centre for Numerical Methods in Engineering, 2001.
[viii] Hämäläimen, J.P., Mäkinen, R.A.E., Tarvainen, P. and Toivanen, J. “Evolutionary Shape Optimisation in CFD With Industrial Applications”, ECCOMAS, Spain, 2000.
[ix] Obayashi, S., Nakahashi, K., Oyama, A. and Yoshino, N. “Design optimization of supersonic wings using evolutionary algorithms ”. Presented at ECCOMAS98, John Wiley & Sons, ltd. 1998.
[x] Sasaki, D., Obayashi, S., Sawada, K., and Himeno, R., "Multiobjective Aerodynamic Optimization of Supersonic Wings Using Navier-Stokes Equations," ECCOMAS, Spain, 2000.
[xi] Bézier, P. “Numerical control. In: Mathematics & Applications”, Wiley, 1972.
[xii] Eppler, R. “Airfoil design and data”, Springer-Verlag, 1989.

[xiii] Storn, R. and K. Price. “DE – A simple and efficient adaptive scheme for global optimisation over continuous spaces”, TR-95-012, March 1995.
[xiv] Storn, R. and Price, K., "Differential Evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces", J. of Global Optimization, Kluwer Acad. Pub., 1997, Vol. 11, pp. 341 - 359.