A Neural Network Based Generalized Response Surface Multiobjective Evolutionary Algorithm

M. Farina

Abstract - The practical use of multiobjective optimization tools in industry is still an open issue. A strategy for reduction of objective function calls is often essential, at a fixed degree of Pareto Optimal Front (POF) approximation accuracy . To this aim an extension of single-objective NN-based GRS methods to Pareto Optimal Front (POF) approximation is proposed. Such an extension is not at all straightforward due to the complex relation between the POF and Pareto Optimal Set (POS). As a consequence of such a complexity, it is extremely difficult to identify a multi-objective analogue of the single-objective current optimum region; consequently the design domain search space zooming strategy, which is the core of a GRS method, is to be carefully reconsidered when POF approximation is concerned.

Keywords— **Evolutionary multiobjective optimization, NN interpolation, response surface methods.**

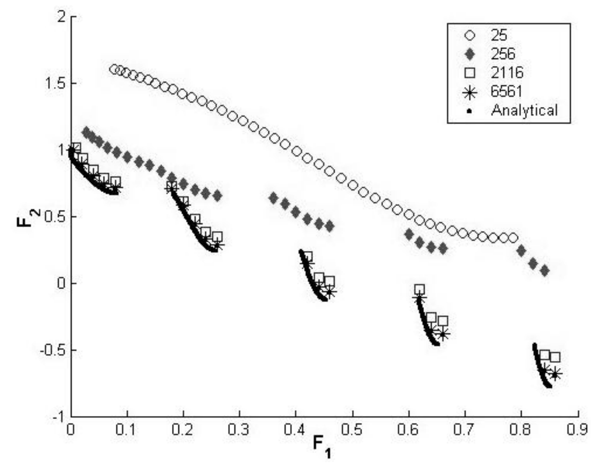

Fig. 1. Different POF approximation with NSESA on DTZ3 with increasing NN training cost (shown in the square box).

I. Introduction

DESPITE evolutionary multiobjective optimization has now come to a full maturity, both in terms of methodologies [1], [2] and algorithms development, [3], [4], [5], [6] [7], [8], [9], the application of the wide variety of available multiobjective optimization methods to problems arising from industrial design (we refer to electromagnetic devices shape design [10], [11], [12], [13], [14], [15] and fuzzy controllers design), is still not fully straightforward due to the computational cost of objective function evaluation (being often non-linear or coupled FEM (Finite Element Method) in the first case and a long time-domain full-system simulation in the second case).

Three alternative (to available MOEAs) approaches have been considered, being specifically devoted to the reduction of objective function calls. They are useful and meaningful when, on one hand, the number of objective function calls that can be afforded from the point of view of an industrially practical computational cost is much smaller than the threshold number required for convergence of available powerful MOEAs, but on the other hand the designer nevertheless wants the few affordable solutions to be convergent towards the POF and to be diverse each other. This means that solutions are to be few but distributed all along the POF. The three approaches can be summarized as follows:

- Build specific MOEAs for tiny populations.
- Adapt Generalised Response Surfaces Methods (GRSM), a well established technique for single-objective optimization, to POF approximation.
- Reconsider a particular preference function method with hybrid global-evolutionary and local-deterministic search in an innovative way.

Regarding the first approach, a non-dominated sorting based MOEA for tiny population (NSESA) was developed and applied to several electromagnetic shape design optimization problems [11], [12]. The third approach is described in [16] . The paper is devoted to the second approach. Generalized Response Surfaces (GRS) methods are a well established technique for single-objective optimization [17], [18] in case of time consuming objective function evaluation. The essential idea of such methodologies is to consider, throughout the optimization, two different objective functions; the first one is the true function which is to be evaluated in as few cases as possible,

M. Farina is with STMicroelectronics, Soft Computing and Nano-Organics Operation, Via C. Olivetti, 2 Agrate Brianza (MI), Italy and with the University of Pavia, Department of Electrical Engineering, Pavia IT; E-mail: marco.farina@st.com .

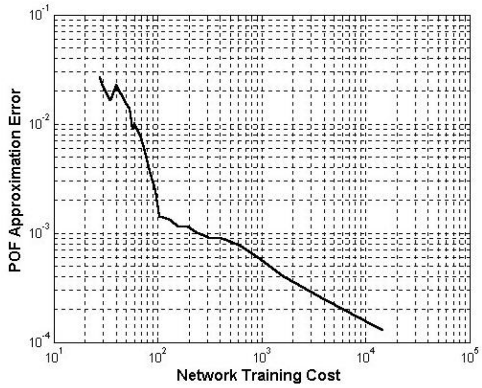

Fig. 2. POF approximation error with NSESA on DTZ3 versus NN training cost.

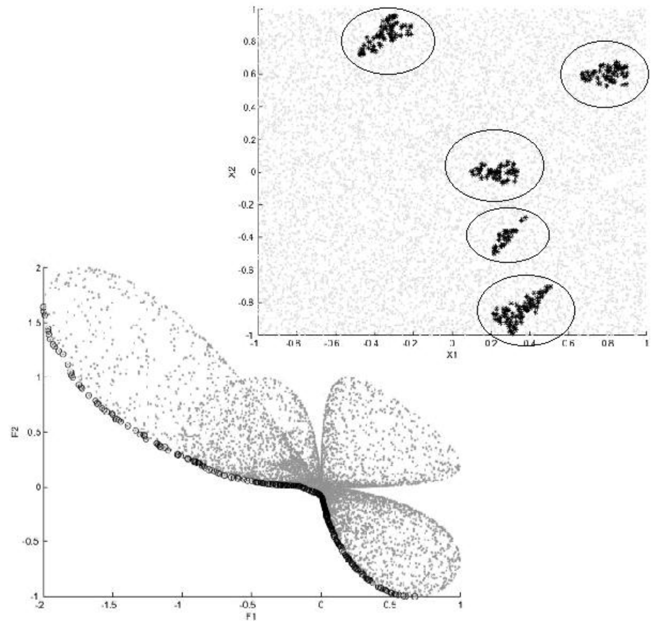

Fig. 3. An analytical problem where a non-trivial zooming strategy is necessary (see the appendix for equations); left: objective domain, right: design domain; gray dots: exhaustive sampling, ∘: POF, ∗: POS.

due to its computational cost, the second one is the interpolation of the true objective function via some interpolation technique (polynomial, multi-quadrics [17], [18], Neural Network [19]) and it can be evaluated as many time as is needed. A powerful stochastic global search algorithm (GA, ES, SA, DE) can be run on the interpolated function up to full convergence in order to be sure of escaping local minima, with no limits in objective function (the almost costless interpolated one) evaluation number. Moreover an iterative strategy alternating search of a new optimum and updating of the interpolation quality in the current optimum region is performed. The aim of single-objective GRS methods is on one hand to increase convergence speed and on the other hand to improve interpolation quality only in the area of current optimum with a dramatic reduction in the interpolation training set size. The extension of such a strategy to POF approximation is considered in the paper.

II. NN training cost and POF approximation accuracy

Before building a GRS-MOEA iterative strategy the dependance of POF approximation error on the accuracy of NN-interpolation of objective functions is considered. This is an essential point that can give several insight and guidelines for the iterative updating of NN training set in the full iterative strategy [19]. Each objective function is approximated by means of radial basis NN of increasing cost, where the cost of a network is meant as the number of elements in the training set (input/target couples). As an example several POFs for problem DTZ3 (see [20] for equations) corresponding to different NN training cost, are shown in figure 1. The training set is always a regular grid in the search space. Once the grid is built and the NN is trained for both objective functions, an NSESA [12] is run on the interpolated function and the POF is obtained. As can be seen from figure 1 the non connected POF for DTZ3 requires a high training cost on objectives NN-interpolation for a satisfactorily approximation. In order to give a quantitative expression of the POF approximation error, the following formula can be considered :

$$e = \frac{1}{m}\sum_{j=1}^{m} \min_{i=1:n} \|\mathrm{pof}_i - f_j^{NN}\|_2 \tag{1}$$

where **pof** is the exact POF when the formula for the exact POF is available or a reference numerical POF when the analytical expression is not available, $\mathbf{f}^{NN}$ is the NSESA m individuals solution on NN-interpolated functions. This error versus NN-interpolation cost is plotted in figure 2; the stopping criterion for all runs with different NN-interpolated functions was set to be the same (a convergence based one). The log-scale decrease is essentially divided into two parts, a faster one, for moderate cost (<100 training nodes) nets and a slower one, for high cost (>100 training nodes) nets.

III. Iterative GRS-MOEA

The general structure of a GRS method for multiobjective optimization is essentially similar to the single-objective counterpart (see the introduction). Three different kind of objective functions (for each criteria) are

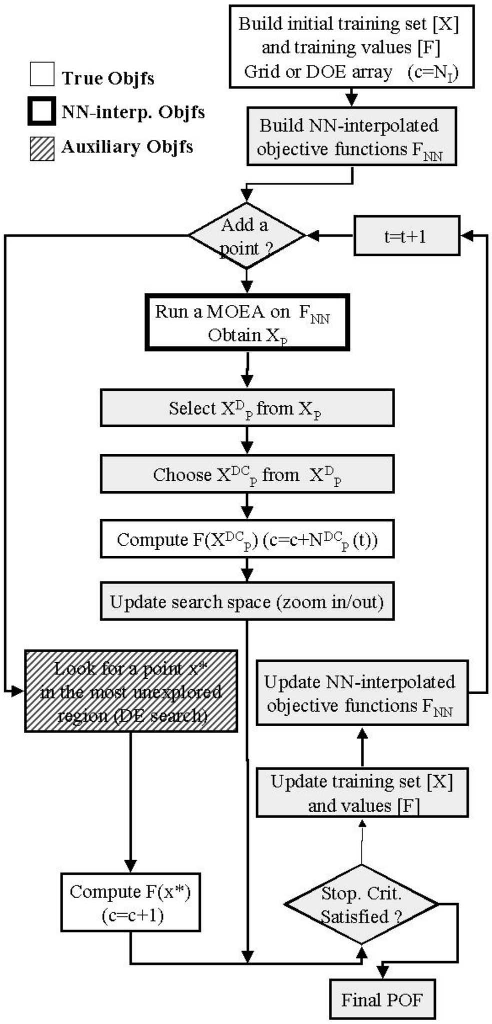

Fig. 4. Principle flowchart of the proposed NN-based GRS-MOEA; the use of three different objective function for each criterion is outlined together with the contribution of each block to the cost c.

considered: the true one, the interpolated one and an auxiliary one, which is used for the addition of a point in the most unexplored area of the search space. As will be shown, both problems - conflicting by their nature - of solution accuracy (in terms of both Pareto-convergence and diversity) and of avoiding local fronts traps, are fully considered; the main steps of the proposed NN-based GRS-MOEA are here described in some details:

Step 1– If the design domain is an n-dimensional box, build an initial training set $[X]$ as an n-dimensional regular p^n points grid, where p is the number of points in each edge of the n-dimensional rectangular search space. If on the other hand the design domain search space is bounded by complex non-linear constraints, a design of experiments strategy for the choice of the first interpolating node set $[X]$ is to be used; the contribution of this step to the cost is N_I .
- Evaluate the true objective functions for all points on the set in order to build the first training values array $[F]$.

Step 2 Build the first NN interpolation $F_{NN}(0)$ using all training nodes and values of step 1.

Step 3– Start an iterative procedure, t being the iteration counter. If $\langle t, k_a \rangle$=0 (k_a being a parameter to be defined $\approx$ 2-10) add a node $\mathbf{x}^*$ (information point) in the interpolation nodes set $[X]$ in the most unexplored area of the whole search space, and compute the true objectives values $F(\mathbf{x}^*)$ for that point. The contribution of this block to the cost is 1. The addition of information points increases the probability of jumping out of eventual local fronts and requires (in the proposed implementation) an optimization itself (see further details below).
- If on the other hand $\langle t, k_a \rangle \neq 0$ run a MOEA on the current NN-interpolated functions $F_{NN}(t)$ and obtain a current solution set $[X_P]$ for the interpolated functions;
- extract non-dominated solutions and obtain the set $[X_P^d]$,
- select some solutions discarding those that are too close in design space and obtain $[X_P^{dc}]$;
- compute true objective functions $F([X_P^{dc}])$ on $[X_P^{dc}]$; the contribution of this block to the cost

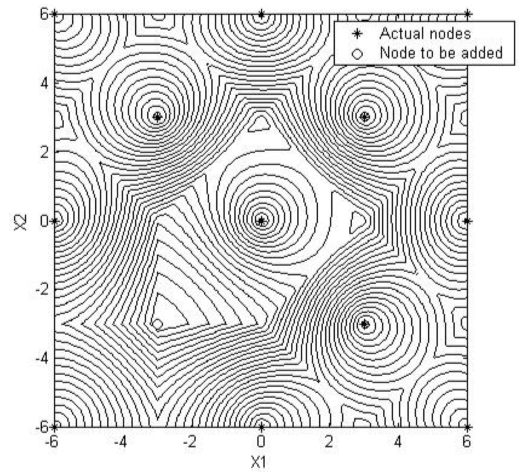

Fig. 5. Example of the addition of an information point ○ to the NN-interpolation node set ● as a minimum point of the additional objective function F_{add} (equivalue lines are shown).

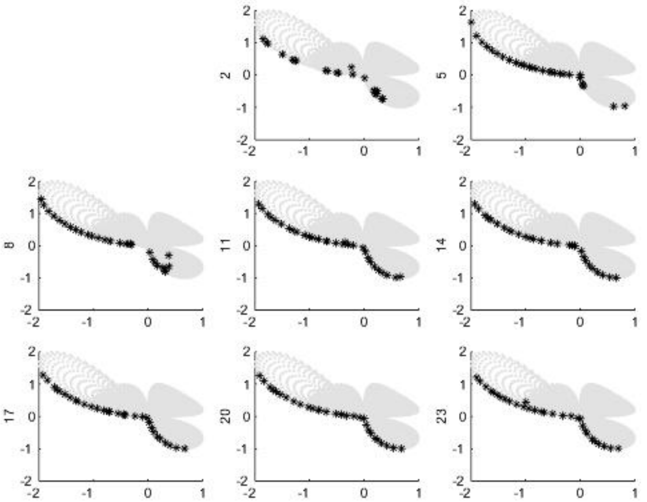

Fig. 6. Behavior of the proposed GRS-MOEA strategy on problem 4 in the objective domain; black $*$: current POF, gray dot: search space sampling.

is $N_P^{DC}(t)$.

– update the search space in the design domain (see further details below)

Step 4 If the termination criterion is satisfied stop the search, otherwise add $[X_P^{dc}]$ and $F([X_P^{dc}])$ to the training set and values and build an updated NN-interpolated functions $F_{NN}(t+1)$; after that go to the next iteration .

The flowchart of the described methodology is shown in figure 5. The true objective function evaluation number (cost of the algorithm) up to iteration $\widetilde{t}$ can be evaluated as follows:

$$c(\widetilde{t}) = N_I + \sum_{\substack{t=1 \\ \langle t, k_a \rangle \neq 0}}^{\widetilde{t}-1} [size(X_p^{dc}(t))] + \frac{\widetilde{t}-1}{k_a} \qquad (2)$$

We point out here that, in order not to loose computational efforts, all points where the true objective functions are computed, are added to the NN training set; $c(t)$ is thus also the current NN training cost at iteration t. At each iteration a new NN interpolation is built on the current training set. This is why in figure 1 and 2 the NN training cost is considered. On the other hand, as can be seen easily from the flowchart, the pseudo-cost (number of NN-interpolated function evaluation) is extremely hight but it is supposed to be negligible with respect to the cost (the case of very time consuming true objective function evaluation is always considered). Moreover, the cost (not the pseudo-cost) does not depend on the number of desired solutions on the POF and on the population size (and in general on the search power) of the MOEA in step 3. The updating of the search space in design domain is non-trivial with respect to the single-objective counterpart. The current POS region may be essentially connected or disconnected; in the first case an n-dimensional box, centered on the center of weights of the current POS can be considered. In the second case the sum of different boxes centered on centers of weight of all parts of the POS is to be used. As an example, a test problem (see the appendix for equations) where this second strategy is to be used, is shown in figure 3 in both design domain and objective domain. The POS is a set of disconnected regions while the POF is non-convex but connected. Some results on this test problem are shown in next section. The choice of a point in the most unexplored area of the search space, to be added to interpolation set, is performed via an optimization procedure. Given a certain set of N_{NN} interpolation nodes X the following auxiliary objective function, based on distances, can be built:

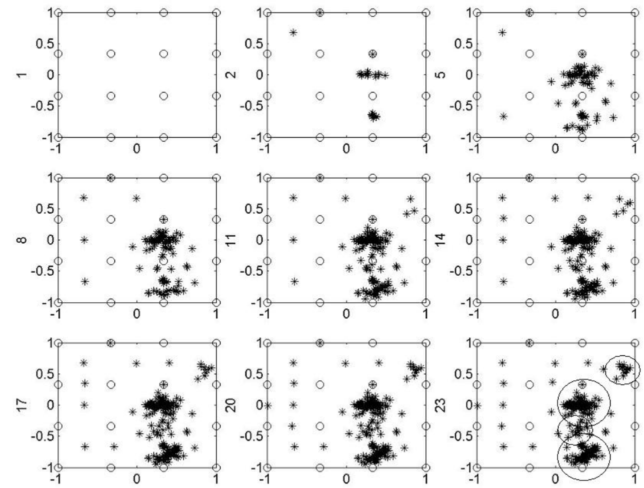

Fig. 7. Behavior of the proposed GRS-MOEA strategy on problem 4 in the design domain, .

$$F_{add}(X) = \frac{\sqrt{\sum_{i=1}^{N}(d_{av}-d_i)^2} - \min_{i=1:N} d_i}{d_{av}}; \qquad (3)$$

where d_i is the Euclidean distance between X and the i-th interpolation node and d_{av} is the average distance. An example of a particularly symmetrical case is shown in figure 5 where the set of current actual nodes ($*$), the new node to be added ($\circ$) and the contour plot of function F_{add} have been plotted. As can be seen the node is added where the F_{add} function has its minimum and the position indeed corresponds to the most unexplored area.

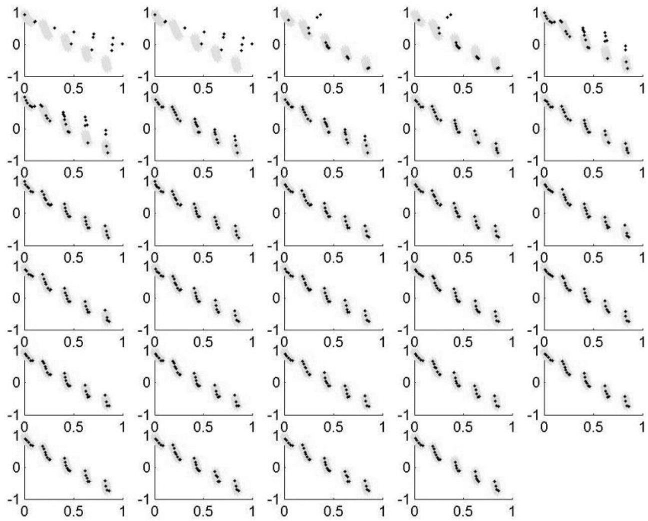

Fig. 8. Behavior of the proposed GRS-MOEA strategy on a DTZ3 problem in the objective domain;black dotted: current POF, gray: exact front.

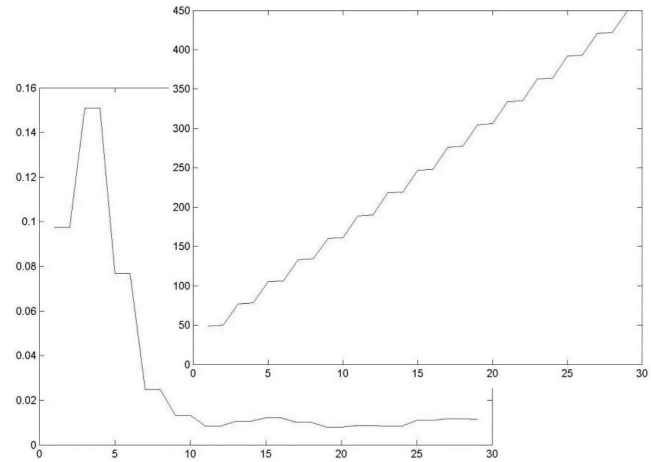

Fig. 9. DTZ3 test case: POF approximation error (left) and true objective function evaluations (right) *versus* GRS-MOEA iterations.

Due to the possible presence of local minima for function F_{add}, when the number of nodes increases, this optimization problem can be solved with a stochastic algorithm (in the test case presented in the following section, a DE algorithm has been used). The use of the aforementioned strategy for the addition of information points is derived from single-objective GRS methods ([17], [18]) and could also be substituted by a suitable DOE strategy for the evaluation of the most unexplored area in the search space. Moreover the choice of NN as interpolation technique is of course arbitrary and other techniques could also be done, such as multiquadrics or polynomial; the power and flexibility of NN with respect to other techniques was outlined in previous studies on GRS methods for single-objective problems [17], [18]. From this point of view the proposed method behaves in a similar way to single-objective GRS methods. On the other hand when the proposed method is considered the quality of approximation is to be evaluated on the resulting POF and not on the single interpolated objective functions.

IV. Test Cases

Two test cases are presented; the first one has been already shown in figure 3 when outlining the difficulties in the zooming strategy and corresponding equations are listed in the appendix. A POF analytical expression is not available but the problem was chosen because the POS is clustered while the POF is connected. As can be seen from figures 7 three of the four cluster are correctly approximated despite the small number of points in he initial grid (16). The current POF is shown in figure 6 for the corresponding iterations. Moreover, the addition of information points corresponding to F_{add} minima can be also seen. As can be seen, despite the initial POF is far from the exact one (the initial training point set was intentionally chosen to be small) the GRS-MOEA is able to converge and to approximate all the branches of the disconnected DTZ3 POF and POS. As can be seen from figure 7 the information points locates in the centers of squares delimited by the initial grid points. This is a correct behavior as shown for the example in figure 5. The second one is the DTZ3 problem (see [2] for equations); it was chosen because of the difficulties in approximating the discontinuity of the POF and because of the known analytical expression for the POF can be used for monitoring the POF approximation error with formula 1. The story of the proposed GRS-MOEA run on DTZ3 is shown in figure 8 in the objective domain. A relevant reduction of true objective function evaluation is obtained (449 for test run in figure 8) at a given number of solutions on the POF. The behavior of the cost function c, depends on k_a and on the number of current Pareto optimal solutions that are considered after each MOEA in step 3. The stair-like plot of c (figure 9 on the right for the run in figure 8.) is due to the addition of one information point every two iteration (flat segments) and the addition of new Pareto Optimal solutions whose number may change at each iteration (sloping segments). When the POF equation is known, as it is the case of DTZ3 problem, at each iteration t of the GRS-MOEA the POF approximation error e can be computed with equation 1. Figure 9 shows on the left the plot of e versus t for the run in figure 8; after an initial increase e quickly decreases towards a flat value showing convergence and stability of the method (several other test cases have been performed on different analytical test problems showing the same behavior).

V. Conclusions

The extension of GRS methods to Pareto Optimal Front approximation requires deep modifications of the classical single-objective strategy. The validity of the proposed NN based GRS-MOEA is shown on test cases in terms of POF approximation errors, convergence and diversity properties and stability. Essential preliminary information on the relationship between POF approximation error and NN training cost are to be considered when building the iterative strategy. With such a strategy a strong reduction of objective function calls is obtained at a given degree of POF approximation accuracy. Further work are in course of development for a better zooming strategy toward the current POS region in order to improve convergence speed. Moreover the application of the proposed strategy to multiobjective optimization of fuzzy controllers for car suspension system is in progress.

Appendix

Equations for test case in figure 3

$$\begin{cases} \min\limits_{-1<x_1,x_2<1}(f_1, f_2) \\ f_1 = -2\exp(15(-(x_1-.1)^2 - x_2^2))+ \\ -\exp(20(-(x_1-.6)^2-(x_2-.6)^2))+ \\ \exp(20(-(x_1+.6)^2-(x_2-.6)^2))+ \\ \exp(20((-x_1-.6)^2-(x_2+.6)^2))+ \\ \exp(20(-(x_1+.6)^2-(x_2+.6)^2)) \\ f_2 = 2\exp(20(-x_1^2-x_2^2))+ \\ \exp(20(-(x_1-.4)^2-(x_2-.6)^2))+ \\ -\exp(20(-(x_1+.5)^2-(x_2-.7)^2))+ \\ -\exp(20(-(x_1-.5)^2-(x_2+.7)^2))+ \\ -\exp(20((x_1+.4)^2-(x_2+.8)^2)) \end{cases} \tag{4}$$

References

[1] K. Miettinen, *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Dordrecht, THE NEDERLANDS, 1999.

[2] K. Deb, *Multi-objective Optimization Using Evolutionary Algorithms*, John Wiley and Sons, ISBN 0471 87339 X, Chichester, UK, 2001.

[3] E. Zitzler and K. Deb et al. editors, *Evolutionary Multi-Criterion Optimization, proceedings of EMO2001, Zurich*, vol. 1993, 2001.

[4] K. Deb, "Non-Linear Goal Programming using Multi-Objective Genetic Algorithms," Tech. Rep. CI-60/98, University of Dortmund, Germany, 1999.

[5] K. Deb and T. Goel, "A hybrid multi-objective evolutionary approach to engineering shape design," *Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization (EMO-2001) Zurich, Switzerland*, pp. 385–399, 2001.

[6] K. Deb, "Nonlinear goal programming using multi-objective genetic algorithms," *PJournal of the Operational Research Society*, vol. 52, no. 3, pp. 291–302, 2001.

[7] M. Eisenring, L. Thiele, and E. Zitzler., "Conflicting criteria in embedded system design," *IEEE Design & Test*, vol. 17, no. 2, pp. 51–59, April-June 2000.

[8] M. Laumanns, E. Zitzler, and L. Thiele., "A unified model for multi-objective evolutionary algorithms with elitism," *Congress on Evolutionary Computation (CEC-2000)*, pp. 46–53, July 2000.

[9] Carlos A. Coello Coello, "An updated survey of evolutionary multiobjective optimization techniques: State of the art and future trends," in *1999 Congress on Evolutionary Computation*, Piscataway, NJ, 1999, pp. 3–13, IEEE Service Center.

[10] M. Farina, A. Bramanti, and P. Di Barba, "Combining global and local search of non-dominated solutions in inverse electromagnetism," *Proc. of EUROGEN2001 Athens, Greece*, 2001.

[11] P. Di Barba, M. Farina, and A. Savini, "An improved technique for enhancing diversity in pareto evolutionary optimization of electromagnetic devices," *COMPEL Intl. Journal for computation and mathematics in Electrical and Electronic Engineering*, vol. 20, no. 2, pp. 482–496, 2001.

[12] P. Di Barba, M. Farina, and A. Savini, "Multiobjective design optimization of real-life devices in electrical engineering: a cost-effective evolutionary approach," *Lecture notes in Computer Science*, vol. 1993, no. 2, pp. 560–573, 2001.

[13] P. Di Barba, M. Farina, and A. Savini, "Multicriteria strategy for the optimization of air-cored solenoid systems," *Studies in Applied Electromagnetics and Mechanics*, vol. 18, pp. 475–478, 1999.

[14] P. Di Barba and M. Farina, "Multiobjective shape optimisation of air cored solenoids," *COMPEL International Journal for computation and mathematics in Electrical and Electronic Engineering*, vol. 21, 2002.

[15] P. Di Barba, M. Farina, and A. Savini, "Optimisation of an electrostatic micromotor using a genetic algorithm," *COMPEL International Journal for computation and mathematics in Electrical and Electronic Engineering*, vol. 19, no. 2, pp. 576–581, 2000.

[16] M. Farina, "A minimal cost hybrid strategy for pareto optimal front approximation," *Evolutionary Optimization, , (2002)*, vol. 3, no. 1, pp. 41–52, 2001.

[17] K. Rashid, M. Farina, J.A. Ramirez, J.K. Sykulski, and E.M. Freeman, "A comparison of two generalized responce surface methods for optimization in electromagnetics," *COMPEL*, vol. 3, no. 20, pp. 741–752, 2001.

[18] M. Farina and J.K. Sykulski, "Comparative study of evolution strategies combined with approximation techniques for practical electromagnetic optimization problems," *to appear on IEEE Trans. Mag.*, 2001.

[19] A. Bramanti, P. Di Barba, M. Farina, and A. Savini, *Combining Response Surfaces and Evolutionary Strategies for Multiobjective Pareto-Optimization in Electromagnetics*, vol. 9 of *Studies in Applied Electromagnetics and Mechanics*, JSAEM, Tokyo, 2001.

[20] Zitzler E. and Thiele L., "Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach," *IEEE Tansactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, 1999.