DESIGN OF A NOVEL HYBRID OPTIMIZATION ALGORITHM

Dimitris V. Koulocheris and Vasilis K. Dertimanis
Vehicles Laboratory, National Technical University of Athens, Iroon Politechniou 9, 157 80, Athens, Greece

Keywords: Hybrid optimization, Evolution strategy, Deterministic mutation, Line–search, Trust–region, Vehicles.

Abstract: The interrelation of stochastic and deterministic optimization algorithms, as well as the exploitation of the advantages that each counterpart presents simultaneously, is studied in this paper. To this, a hybrid optimization algorithm is developed, which consists of a conventional Evolution Strategy that maintains its recombination and selection phases unaltered, while its mutation operator is replaced by well–known deterministic methods, such as line–search and/or trust–region. The alteration results in superior performance of the novel algorithm, compared to other instances of Evolutionary Algorithms, as exploited out in tests using Griewangk and Rastrigin functions. The proposed algorithm is further examined through its implementation to the structural optimization problem of a full–car suspension model, with satisfying results.

1 INTRODUCTION

Numerical optimization, either deterministic (Nocedal and Wright, 2006), or stochastic (Schwefel, 1995; Baeck, 1996), has shown to be a very powerful tool in engineering, with implementation in a very wide area of applications, including structural design (Rao, 1996; Alkhatib et al., 2004; Koulocheris et al., 2003a), system identification (Koulocheris et al., 2003c), control (Fleming and Purshouse, 2002) and fault diagnosis (Dertimanis, 2006; Chen and Patton, 1999).

The corresponding schemes that have been for a long time the subject of significant research in the field of numerical optimization, are mostly divided into two main categories, deterministic and stochastic: the former, usually build a local quadratic model of the function of interest and converge rapidly to a local stationary point, given a "*good initial guess*" for the parameter vector, while the latter perform in a wide area of the search space, since, generally, the optimization procedure is conducted in parallel. Yet, both suffer from serious drawbacks, as the deterministic methods depend drastically on the initial parameter vector provided and frequently stuck in local optima, while the stochastic ones present very slow convergence rate (Vrazopoulos, 2003). To this, the idea of combining the diverse characteristics of these two optimization categories into a hybrid algorithmic structure, follows naturally. Surprisingly, at least in the engineering research field, relative works are rather limited (Koulocheris et al., 2004), the almost exclusive use of GA (refer to Appendix A for notation) is utilized (Koh et al., 2003), while applications are scarcely ever reported (Dertimanis et al., 2003). It should be noted though, that the problem of accelerating conventional EA has been faced using different techniques, such as neural networks (Papadrakakis and Lagaros, 2002).

This paper presents a methodology of interconnecting stochastic and deterministic optimization algorithms, in a way that exploits the advantages of both of them and results into a method that shows faster convergence rate, as well as increased reliability in the search for the global optimum. Among EA, the stochastic component has been selected to be the $[\mu/\rho\ (+/,)\ \lambda]$–ES, while the deterministic one belongs to the family of quasi–Newton methods and it is currently implemented using either line–search, trust–region, or a combination of both. To this, the currently proposed version of the algorithm integrates previous ones (Koulocheris et al., 2008; Koulocheris et al., 2004; Vrazopoulos, 2003), so that a more robust and flexible scheme is developed. In order to get insight about the performance of the novel optimization method, it is tested with the Griewangk and Rastrigin functions and compared with the conventional ES (in fact its multi–membered *plus* and *comma* versions), as well as a meta version of EP (Baeck, 1996). Consequently, it is applied to the problem of optimizing the characteristics of a suspension system used in ground vehicles.

V. Koulocheris D. and K. Dertimanis V.
DESIGN OF A NOVEL HYBRID OPTIMIZATION ALGORITHM.
DOI: 10.5220/0002166501290135
In *Proceedings of the 6th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2009)*, page
ISBN: 978-989-8111-99-9

The rest of the paper is organized as follows: in Sec. 2 the novel algorithm is presented and in Sec. 3, indications of its performance are illustrated, through the evaluation by theoretical objective functions, as well as an application example, corresponding to the problem of optimizing the riding comfort of a passenger vehicle. In Sec. 4 some final remarks are given, together with suggestions for further research.

2 THE HYBRID ALGORITHM

2.1 Description

The proposed hybrid algorithm with deterministic mutation aims, as already mentioned, at interconnecting the advantages of both optimization approaches. Deterministic methods are characterized, if the optimization function is regular, by a high convergence rate and accuracy in the search for the optimum. On the other hand, EA show a low convergence rate but they can search on a significantly broader area for the global optimum.

$[\mu/\rho\ (+/,)\ \lambda,\nu]$–hES is based on the distribution of the local and the global search for the optimum and it consists of a super-positioned stochastic global search, followed by a independent deterministic procedure, which is activated under conditions in specific members of the involved population. Thus, every member of the population contributes in the global search, while single individuals perform the local search. Similar algorithmic structures, the theoretical background of which pertains to the simulation of insects societies (Monmarche et al., 2000; Rajesh et al., 2001), have been presented by (Colorni et al., 1996; Dorigo et al., 2000; Jayaraman et al., 2000).

The stochastic platform has been selected to be the ES, while the deterministic counterpart is a quasi–Newton algorithm (see Sec. 2.2). It must be noted that the selection of ES among the other instances of EA is justified via numerical experiments in non–linear parameter estimation problems (Schwefel, 1995; Baeck, 1996), which have provided significant indication that ES perform better than the other two classes of EA, namely GA and EP.

The conventional ES is based on three operators that take on the recombination, the mutation and the selection tasks. In order to maintain an adequate stochastic performance in the new algorithm, the recombination and selection tasks are retained unaltered (refer to (Beyer and Schwefel, 2002) for a brief discussion about the recombination phase), while its strong local topology performance is utilized through the substitution of the original mutation operator by a quasi–Newton one.

A very important matter that affects significantly the performance of the $[\mu/\rho\ (+/,)\ \lambda,\nu]$–ES involves the members of the population that are selected for mutation: there exist indications (Koulocheris et al., 2003b) that the reason for the poor performance of EA in non–linear multimodal functions is the loss of information through the non-privileged individuals of the population. Thus, the new deterministic mutation operator is not applied to all λ recombined individuals but only to the ν worst among the $(\mu\ (+/,)\ \lambda)$, where ν is an additional algorithm parameter. This means that a sorting procedure takes place twice in every iteration step: the first time in order to yield the ν worst individuals and the second to support the selection operator, which succeeds the new deterministic mutation operator. This modification enables the strategy to yield the corresponding local optimum for each of the selected ν worst individuals in every iteration step. The advantage is reflected in terms of increased convergence rate and reliability in the search for the global optimum, while three other alternatives were tested. In these, the deterministic mutation operator was activated by:

- every individual of the involved population,
- a number of privileged individuals, and
- a number of randomly selected individuals.

The above alternatives led to three types of problematic behavior. More specifically, the first increased the computational cost of the algorithm without the desirable effect. The second alternative led to premature convergence of the algorithm to local optima of the objective function, while the third generated unstable behavior that led to statistically low performance.

2.2 The Deterministic Mutation

As noted, quasi–Newton type methods replace the original mutation of ES. Yet, unlike earlier versions (Vrazopoulos, 2003), it is not wise to limit the operator in a line–search framework, since trust–region and mixedcombined methods have also proven to be competitive alternatives, or to enforce the exclusive use of the BFGS Hessian update, as analytical or finite–difference derivative information may, in some cases, be either available, or costless to compute. This fact leads to the optional implementation of full Newton methods, but the term quasi–Newton shall be preserved, in order to cover the majority of the problems faced in practice. Thus, in the following it is assumed that the gradient of the objective function is approximated using finite–differences, while the Hessian is calculated using the powerful BFGS update.

The currently presented version of the $[\mu/\rho\ (+/,)\ \lambda,\nu]$–ES offers three alternatives to be used as mutation operators, which are briefly discussed in the following.

2.2.1 Line–search

A line–search algorithm is build in a simple idea: at iteration k, given a descent direction $\mathbf{p}_k$, take a step in that direction that yields an "acceptable" parameter vector, that is,

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \cdot \mathbf{p}_k \tag{1}$$

for some λ_k that makes $\mathbf{x}_{k+1}$ an acceptable next iterate. Since,

$$\nabla^2 f(\mathbf{x}_k) \cdot \mathbf{p}_k = \nabla f(\mathbf{x}_k) \tag{2}$$

is utilized, the proposed mutation operator implements a cubic polynomial line–search procedure for the determination of λ_k, that satisfies both Wolfe conditions (Nocedal and Wright, 2006). It must be noted that in every iteration the full quasi–Newton step ($\lambda_k = 1$) is always tested first.

2.2.2 Trust–region

If in Eq. 1 the full quasi–Newton step is unsatisfactory, it means that the quadratic model fails to approximate the objective function in this region. Instead of calculating a search direction, trust–region methods calculate a shorter step length by solving the problem,

$$m_c(\mathbf{x}_k + \mathbf{p}) = f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)^T \cdot \mathbf{p} + \frac{1}{2}\mathbf{p}^T \nabla^2 f(\mathbf{x}_k) \cdot \mathbf{p} \tag{3}$$

where m_c is the quadratic model, subject to,

$$\|\mathbf{p}\| \le \delta_k \tag{4}$$

so that,

$$(\nabla^2 f(\mathbf{x}_k) + \xi \cdot I) \cdot \mathbf{p}_k = \nabla f(\mathbf{x}_k) \tag{5}$$

for some $\xi > 0$. Trust–region mutation utilizes two alternatives for the calculation of ξ: the locally constrained optimal ("hook") step and the double dogleg step (Dennis and Schnabel, 1996).

2.2.3 Combined Trust–region / Line–search

The third alternative that the proposed algorithm offers as mutation operator, is a combined trust–region / line–search framework. To this, Eqs. 3-5 are solved approximately for the direction $\mathbf{p}_k$ and if the full quasi–Newton step does not result in a sufficient decrease of the objective function, a line–search is performed, which guarantees, under certain conditions, a lower objective function value. The corresponding algorithm is described in (Nocedal and Yuan, 1998).

2.3 Termination Criteria

The termination criteria are distinguished as local, referring to the deterministic mutation and global, referring to $[\mu/\rho\ (+/,)\ \lambda,\nu]$–ES. For the former, standard tests that are presented in detail in (Nocedal and Wright, 2006) and (Dennis and Schnabel, 1996) are utilized:

- Objective function value smaller than a specified tolerance,
- relative gradient norm less than a specified tolerance,
- relative distance between two successive iterations less than a specified tolerance,
- not a descent current direction, and
- maximum mutation operator iterations exceeded.

In addition, the proposed algorithm terminates if at least one of the following occurs:

- Absolute difference between worse and best objective function less than a specified tolerance,
- maximum function evaluations exceeded, and
- maximum iterations exceeded.

3 NUMERICAL RESULTS

3.1 Performance Evaluation

In order to assess the performance of the proposed algorithm, a number $N = 100$ of independent tests were utilized using the Griewangk

$$f(\mathbf{x}) = 1 + \sum_{i=1}^{n} \frac{x_i^2}{400 \cdot n} - \prod_{i=1}^{n} cos\left(\frac{x_i}{\sqrt{i}}\right) \tag{6}$$

and the Rastrigin

$$f(\mathbf{x}) = 10 \cdot n + \sum_{i=1}^{n} x_i^2 - 10 \cdot cos(2 \cdot \pi \cdot x_i) \tag{7}$$

functions, with $n = 50$ parameters and known minimum at $\mathbf{x}_m = \mathbf{0}$, $f(\mathbf{x}_m) = 0$. For the tests, a version of the algorithm with $\mu = 15$ parents and $\lambda = 100$ offspring was used, while the recombination type was panmictic intermediate with $\rho = 2$ parents for the generation of each offspring. For the mutation, the trust–region approach (using the double dogleg step) of the relative operator was implemented and in every iteration the $\nu = 3$ worse vectors were mutated. The selection was made among all the involved population (that is both parents and offspring, choice that is denoted by the $+$ sign of the full notation).

Table 1: Statistical results of the compared methods: Griewangk's function.

Method	$\overline{P}$	P_{min}	P_{max}	Termination Reason (%)		Mean CPU time (s)
				Convergence	Max. Iterations	
$(15+100,3)$–hES	-9.90	-10.06	-9.52	100	0	10
$(15+100)$–ES	-0.10	-0.14	-0.06	0	100	28
$(15,100)$–ES	-2.07	-2.83	-1.09	0	100	28
meta–EP	-0.12	-0.18	-0.09	0	100	25

Table 2: Statistical results of the compared methods: Rastrigin's function.

Method	$\overline{P}$	P_{min}	P_{max}	Termination Reason (%)		Mean CPU time (s)
				Convergence	Max. Iterations	
$(15+100,3)$–hES	1.14	0.00	2.05	100	0	5
$(15+100)$–ES	2.63	2.55	2.69	0	100	26
$(15,100)$–ES	2.56	2.35	2.66	0	100	26
meta–EP	2.59	2.52	2.62	0	100	24

Regarding the comparisons, two similar instances of the conventional ES were used, that is the $(15+100)$–ES and the $(15,100)$–ES with panmictic recombination, while a version of the meta–EP with 100 population members and 10 random members for comparison was activated. Taking under consideration the possibility of a large spectrum of orders in the final objective function value, the following quantity was formulated,

$$P_j = log_{10}(f_{final}), \quad j = 1, \ldots, 100 \tag{8}$$

and three statistics qualified the results, that is the mean, the minimum and the maximum values of the P_j's, out of the set of all the independent tests. I must be noted in every iteration, that prior to the execution of every corresponding code, the random number generator was reset, in order to initialize all the compared algorithms from the same population. As far as the termination criteria are concerned, the tolerance for the convergence of the population and the number of iterations were set equal to $macheps^{1/3}$, where *macheps* the computer precision, and 100, respectively.

The results are illustrated in Tabs.1–2, where it is clear that the hybrid algorithm has outperformed all other EA. Indeed, the $(15+100,3)$–hES with a trust–region mutation returned the best statistics among the four, while it converged in all the independent tests. On the contrary, the EA didn't managed to converge within the specified number of iterations and required $2.5-5$ times more CPU time in order to execute. Yet, in the Rastrigin function the hybrid algorithm showed premature convergence, an issue that requires further investigation. In any case, the above resulted provide significant indication about the performance of the novel algorithm and enforce its application to engineering structural problems, as the one presented next.

3.2 Application

The hybrid algorithm algorithm was subsequently applied to the problem of optimizing the performance of a passenger vehicle, in order to improve the ride comfort, under a vibration environment that generated vehicle–road interaction forces with certain spectral characteristics, corresponding to the Draft–ISO formulation (Cebon, 2000). To this, an equivalent linear full–car model with seven degrees of freedom was utilized, which is presented in Figs. 1(a)–1(b). The objective was the optimization of the suspension system under explicit structural and geometric constraints. Since a vibration environment was of interest, the root–mean–square value of the vertical acceleration,

$$f(\mathbf{x}) = \frac{1}{T}\int_0^T \ddot{x}_M^2(t)dt \tag{9}$$

was selected as objective function, subject to the following constraints:

1. Parameter Bounds:

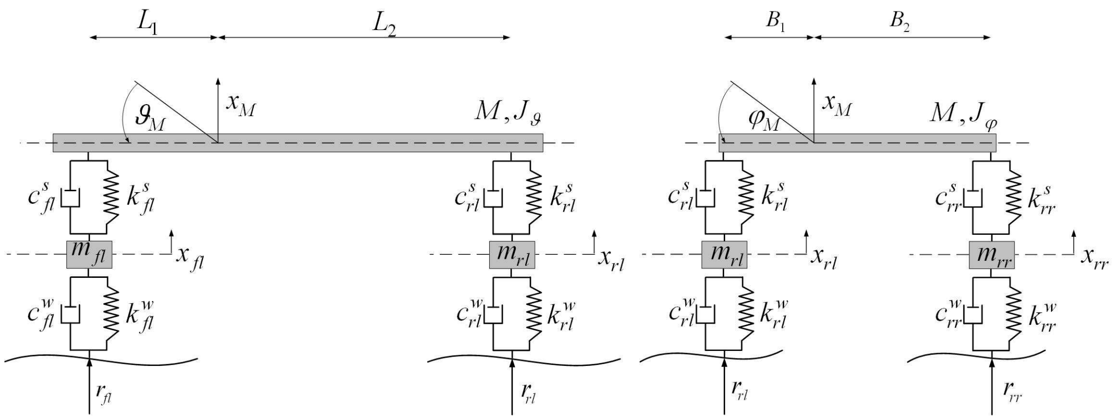
$$1000 \leq k_{ij}^s \leq 50000 \quad (N/m) \tag{10}$$

$$100 \leq c_{ij}^s \leq 5000 \quad (N{\cdot}s/m) \tag{11}$$

for $i = f, r$ and $j = l, r$.

2. Geometry:

$$|x_M(t) + L_k{\cdot}\theta_M(t) + B_k{\cdot}\phi_M(t) - x_{ij}(t)| \leq 0.100m \tag{12}$$



(a) (b)

Figure 1: Structural model of a passenger vehicle: (a) pitch–bounce view and (b) roll–bounce view.

$$|x_{ij}(t)| - r_{ij}(t)| \leq 0.075m \tag{13}$$

for $k = 1, 2$, $i = f, r$ and $j = l, r$.

It can be proved (see (Rao, 1996) for details) that a constrained optimization problem with low and high bounds for the involved parameters can be transformed into an unconstrained one, by applying a simple change of variable, a procedure that followed here, so that the penalty functions that were added in Eq. 9, concerned only the second type of constraints.

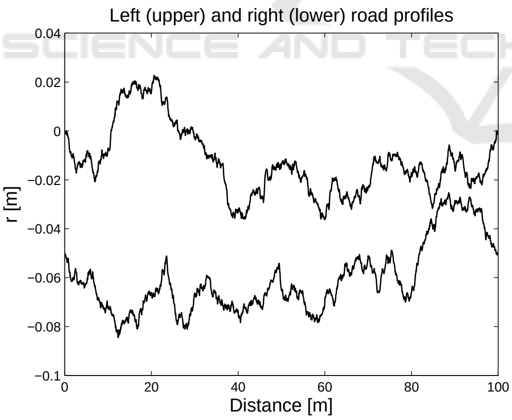


Figure 2: Two tracks across an "average" isotropic surface.

Algorithms's performance characteristics were examined via 50 Monte–Carlo experiments, each one consisting of a certain profile realization (see Fig. 2 for a single realization of the road surface topography) and 20 independent tests, for the same version of the algorithm as before, that is the $(15 + 100, 3)$–hES.

The results are displayed in Figs. 3–4. Figure 3 displays the performance of the objective function in every Monte Carlo experiment. The horizontal line refers to the mean value of the 20 independent tests, while the vertical lines to the standard deviation of the 20 values of every Monte–Carlo experiment. It appears that the hybrid algorithm presented high statistical consistency, fact that is further supported by the suspension results that are illustrated in Fig. 4, from which clear suggestions about the front/rear suspension set up can be made. Yet, the relatively high standard deviations of the suspensions' stiffness indicate that more intuition is required about the role of these structural parameters to the root–mean–square acceleration, with respect to the mathematical model.

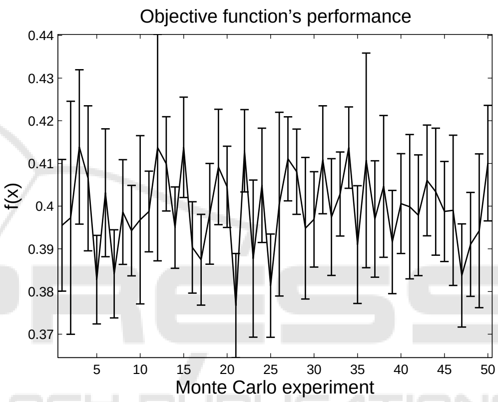


Figure 3: Mean value and dispersions of the objective function, with respect to the Monte Carlo experiments.

4 CONCLUSIONS

A novel hybrid optimization method was presented in this paper, which attempts to combine the diverse characteristics of deterministic and stochastic opti-

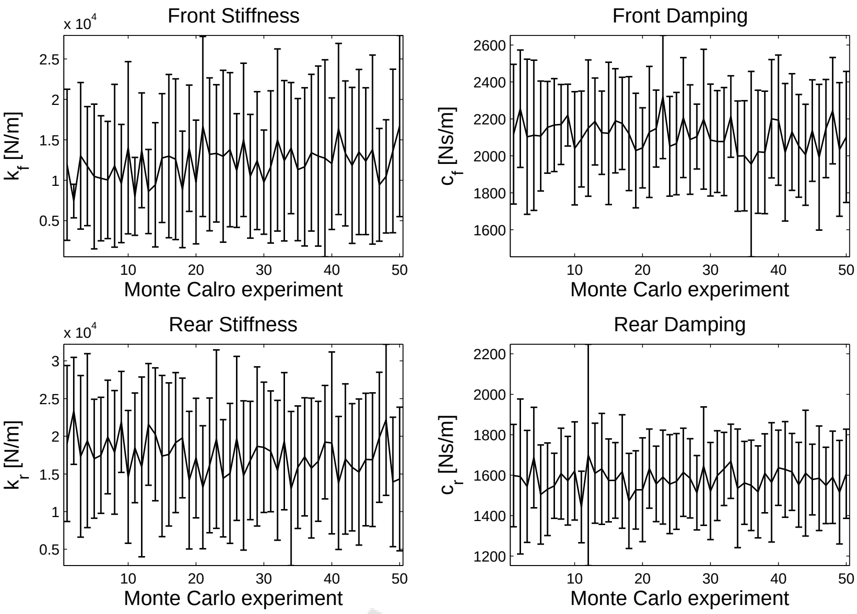


Figure 4: Mean value and dispersions of the parameter vector, with respect to the Monte Carlo experiments.

mization algorithms. That is, to interconnect fast local convergence and increased reliability in the search of the global optimum, without depending on initial values, or suffer from low convergence rate. To this, the corresponding scheme that was developed maintains the stochastic kernel of ES and replaces the original mutation operator by relative methods that utilize derivative information and act on the non-privileged population members, resulting in a more efficient performance.

The proposed algorithm was compared to conventional instances of EA using standard test functions, such as the Griewangk and Rastrigin ones, showing significant evidence about its performance, and subsequently was applied to the problem of optimizing the performance of a passenger vehicle with satisfying results that suggest, not only its use in other engineering problems, but also further investigation about its design parameters, as well as the user–supplied controls.

REFERENCES

Alkhatib, R., Nakhaie Jazarb, G., and Golnaraghi, M. (2004). Optimal design of passive linear suspension using genetic algorithm. *Journal of Sound and Vibration*, 275(3–5):665–691.

Baeck, T. (1996). *Evolutionary* A*lgorithms in* T*heory and* P*ractice*. Oxford University Press, New York.

Beyer, H. and Schwefel, H. (2002). Evolution Strategies: a comprehensive introduction. *Natural Computing*, 1(1):3–52.

Cebon, D. (2000). *Handbook of* V*ehicle*–R*oad* I*nteraction*. Swets & Zeitlinger, Lisse.

Chen, J. and Patton, R. (1999). *Robust model–based fault diagnosis of dynamic systems*. Kluwer Academic Publishers, Massachusetts.

Colorni, A., Dorigo, M., Maffioli, F., Maniezzo, V., Righini, G., and Trubian, M. (1996). Heuristics from nature for hard combinatorial optimization problems. *International Transactions on Operational Research*, 3(1):1–21.

Dennis, J. and Schnabel, R. (1996). *Numerical* M*ethods for* U*nconstrained* O*ptimization and* N*onlinear* E*quations*. Society for Industrial and Applied Mathematics, Philadelphia.

Dertimanis, V. (2006). *Fault* M*odeling and* I*dentification in* M*echanical* S*ystems*. PhD thesis, School of Mechanical Engineering, GR 157 80 Zografou, Athens, Greece. In greek.

Dertimanis, V., Koulocheris, D., Vrazopoulos, H., and Kanarachos, A. (2003). Time–series parametric modeling using Evolution Strategy with deterministic mutation operators. In *Proceedings of the International Conference on Intelligent Control Systems and Signal Processing*, pages 328–333, Faro, Portugal. IFAC.

Dorigo, M., Bonabeau, E., and Theraulaz, G. (2000). Ant algorithms and stigmergy. *Future Generation Computer Systems*, 16(8):851–871.

Fleming, P. and Purshouse, R. (2002). Evolutionary algorithms in control systems engineering: a survey. *Control Engineering Practice*, 10(11):1223–1241.

Jayaraman, V., Kulkarni, B., Karale, S., and Shelokar, P. (2000). Ant colony framework for optimal design and scheduling of batch plants. *Computers and Chemical Engineering*, 24(8):1901–1912.

Koh, C., Chen, Y., and Liaw, C.-Y. (2003). A hybrid computational strategy for identification of structural parameters. *Computers & Structures*, 81(2):107–117.

Koulocheris, D., Dertimanis, V., and Spentzas, C. (2008). Parametric identification of vehicle structural characteristics. *Forschung im Ingenieurwesen*, 72(1):39–51.

Koulocheris, D., Vrazopoulos, H., and Dertimanis, V. (2003a). Design of an I-Beam using a deterministic Multiobjective optimization algorithm. *WSEAS Transactions On Circuits & Systems*, 2(4):794–799.

Koulocheris, D., Vrazopoulos, H., and Dertimanis, V. (2003b). Hybrid Evolution Strategy for the design of welded beams. In *Proceedings of EUROGEN*, Barcelona, Spain.

Koulocheris, D., Vrazopoulos, H., and Dertimanis, V. (2003c). Vehicle Suspension system identification using Evolutionary Algorithms. In *Proceedings of EUROGEN*, Barcelona, Spain.

Koulocheris, D., Vrazopoulos, H., and Dertimanis, V. (2004). A Hybrid Evolution Strategy for vehicle suspension optimization. *WSEAS Transactions On Systems*, 3(1):90–95.

Monmarche, N., Venturini, G., and Slimane, M. (2000). On how Pachycondyla apicalis ants suggest a new search algorithm. *Future Generation Computer Systems*, 16(8):937–946.

Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer–Verlag, New York, 2^{nd} edition.

Nocedal, J. and Yuan, Y. (1998). Combining trust region and line search techniques. In Yan, Y., editor, *Advances in Nonlinear Programming*, pages 153–175. Kluwer Academic Publishers.

Papadrakakis, M. and Lagaros, N. (2002). Reliability–based structural optimization using neural networks and Monte Carlo simulation. *Computer methods in applied mechanics and engineering*, 191(32):3491–3507.

Rajesh, J., Gupta, S., Rangaiah, G., and Ray, A. (2001). Multi–objective optimization of industrial hydrogen plants. *Chemical Engineering Science*, 56(3):999–1010.

Rao, S. (1996). *Engineering Optimization: Theory and Practice*. John Wiley & Sons Ltd., New York, 3^{rd} edition.

Schwefel, H. (1995). *Evolution & Optimum Seeking*. John Wiley & Sons Ltd., New York.

Vrazopoulos, H. (2003). *New Optimization Methods of Mechanical Systems*. PhD thesis, School of Mechanical Engineering, GR 157 80 Zografou, Athens, Greece. in greek.

APPENDIX A: NOTATION

μ	number of parent population
ρ	number of recombination population
λ	number of offspring
ν	number of mutation population
$(+/,)$	plus / comma version of ES
EA	Evolutionary Algorithms
ES	Evolution Strategy
EP	Evolutionary Pogramming
GA	Genetic Algorithms
hES	hybrid Evolution Strategy