Analele Universităţii din Timişoara
Vol. XXXIX, Fasc. 2, 2001
Seria Matematică–Informatică

On the Explorative Power of Differential Evolution Algorithms

Presented at 3rd Int. Workshop "Symbolic and Numeric Algorithms on Scientific Computing" - SYNASC 2001, October 2-4, 2001, Timişoara, Romania

Daniela ZAHARIE

Abstract. The aim of this paper is to analyze some theoretical properties of Differential Evolution algorithms (DE). The main result is the analytical expression which describes the evolution of the expected population variance, a measure of the explorative power of population-based optimization methods. The theoretical result agrees with some empirical ones obtained by experimental studies and allows to explain some of the behavior of DE.

Key words: differential evolution, evolutionary strategies, global optimization, explorative power

AMS Subject Classification: 90C56, 68W20

1. Introduction The Differential Evolution algorithm (DE) is a recently introduced method (Storn and Price, 1997) for finding the global minimum of nonlinear and nondifferentiable functions defined on continuous domains. DE algorithms are based (as all evolutionary algorithms) on the idea of evolving a population of explorers of the solutions space in order to find the optimum. Besides its good convergence properties and suitability for parallelization, DE's main assets are its conceptual simplicity and ease of use.

DE is conceptually close to Evolution Strategies (ES) (Bäck and Schwefel, 1995) which belong to the class of Evolutionary Algorithms (EA). But DEs are easier to be implemented than the ES. For instance, the ES mutation is implemented by adding a random variable (which must be simulated) while the DE mutation is implemented through a simple arithmetical operation. However DE has proven to be very efficient in solving different optimization problems. Some of the applications of DEs are (Lampinen, 1999): fusions of information from multiple sensors used in planning and control of robotic systems, hybrid learning algorithms in regression neural networks, optimizing scheduling problems with nonlinear objective functions and multiple dependent restrictions, design of hierarchical fuzzy logic controllers, solving highly nonlinear constraint satisfaction problems.

Since its invention, DE has been tested extensively against artificial and real world optimization problems. The experimental studies allowed to state some heuristic rules of choosing the parameters. Despite the huge amount of experiments made on DE, without theoretical results the picture of its behavior can not be complete. However a theory of differential evolution is still missing.

The aim of this paper is to make a step into the theoretical analysis of DE. We tried to approach the DE behavior from the viewpoint of ESs. We started from the following statement (Beyer, 1998): "the ability of an EA to find a global optimal solution depends on its ability to find the right relation between exploitation of the elements found so far and exploration of the search space". More specifically, exploitation means the ability of an algorithm to use the information collected so far such that the search is oriented to the goal while the exploration is the process that allow to introduce new information into population.

Regarding the role of each evolutionary operator on the exploitation/exploration process it is accepted (Beyer and Deb, 1999) that in ESs the purpose of selection (based on the values of the objective function) is the exploitation of the search space by emphasizing good solutions and the purpose of variation operators (mutation and recombination) is the exploration of the search space.

In ESs theory, the explorative power is considered to be influenced by the population variability. A measure of the population variability is its variance. To achieve their goal of driving the exploration process, the variation operators should increase the population variance. Results regarding the explorative power of ESs have been reported in (Beyer, 1998), (Beyer, 1999), (Beyer and Deb, 1999).

In this paper we analyze the explorative power of the DE by computing the evolution of the population variance during the algorithm's iterations. The paper is organized as follows. In Section 2 is presented the overall structure of DE and some parallels to classical ESs are made. The main result (an analytical expression of the dependence between the population variance and the generation number) is presented in Section 3 and the explorative power of DE is compared with that of an ES analyzed in (Beyer and Deb, 1999). Section 4 presents some statistical estimates of the population variance. The experimental results agree with the theoretical result.

2. The Differential Evolution algorithm We consider the problem of finding the global minimum of a function $f : D \subset \mathbb{R}^n \to \mathbb{R}$. Both ESs and DEs are based on the evolution of a (usually) fixed size population $P = \{x_1, x_2, \ldots, x_m\}$, $x_l \in D$, $l = \overline{1, m}$.

The population is initialized with elements randomly selected from D (frequently an uniform distribution is used). After initialization an iterative process is started and at each iteration (called generation) a new population is produced. The iterative process cycles until a stopping criterion is satisfied (e.g. a maximum number of generations have been passed or no significant improvements are made).

The construction of a new population is based on three main steps: *mutation*, *recombination* (called also *crossover*) and *selection*. Usually, the first two (also called variation operators) do not use information on the objective function, having the main aim of introducing variation in the population (exploration of the search space). The selection uses objective function information to favor the perpetuation of the best elements.

The structure of a classical $(m + m')$-ES (Bäck and Schwefel, 1995) is:

`Initialize` $P(0) = \{x_1(0), \ldots, x_m(0)\}$, $g = 0$
`Repeat`
  Recombination step:
    `For` $l = \overline{1, m'}$ `generate` y_l (by mixing r elements of $P(g)$)
  Mutation step:
    `For` $l = \overline{1, m'}$ `generate` $z_l = y_l + \xi$
  Selection step:
    `From` $\{x_1(g), \ldots, x_m(g), z_1, \ldots, z_{m'}\}$ `select the best` m
    `elements and form the new population` $P(g+1)$
  $g = g + 1$ (increment the generation counter)
`Until a stopping criterion is satisfied.`

The selection of the parents used in the recombination step has a random character and the mutation step is based on the addition of a random vector ξ (usually, the components of ξ are mutually independent normally distributed random variables).

The structure of DE (Storn and Price, 1997) is:

`Initialize` $P(0) = \{x_1(0), \ldots, x_m(0)\}$, $g = 0$
`Repeat`
Mutation step:
`For` $l = \overline{1,m}$ and $i = \overline{1,n}$ `generate`
$y_l^i = x^i_{\alpha_1^l}(g) + F \cdot (x^i_{\alpha_2^l}(g) - x^i_{\alpha_3^l}(g))$
Crossover step:
`For` $l = \overline{1,m}$ and $i = \overline{1,n}$ `generate`
$z_l^i = \begin{cases} y_l^i & \text{with probability } p_c \\ x_l^i(g) & \text{with probability } 1 - p_c \end{cases}$
Selection step:
`For` $l = \overline{1,m}$
`If` $f(z_l) < f(x_l(g))$ `then` $x_l(g+1) = z_l$ `else` $x_l(g+1) = x_l(g)$
$g = g + 1$ (increment the generation counter)
`Until a stopping criterion is satisfied.`

In the above algorithm $\alpha_1^l, \alpha_2^l, \alpha_3^l$ are distinct elements of $\{1, 2, \ldots, m\}$ randomly selected for each l, F is a parameter which weights the "differential perturbation" and p_c is a parameter which controls the proportion of perturbed elements in the new population. In the original DE (Storn and Price, 1997) some restrictions are imposed: $\alpha_i^l \neq l$ for all i and at least one component of z_l is perturbed. These restrictions are ignored in our considerations. On the other hand the terminology on DE (Storn and Price, 1997) does not contain terms like mutation or crossover. We used them here only to identify the main steps of the algorithm and to facilitate the comparison with ESs.

The main differences between the $(m + m')$-ES and DE are:
(a) Unlike the ES mutation, based on addition of a random vector, the mutation step in DE is based on adding the difference of two vectors to a third one. As it is argued in (Beyer and Deb, 1999), this type of perturbation exhibits some sort of self adaptation.
(b) The crossover used in DE can be viewed as an operator which mixes the elements of two populations (the current one and that obtained after mutation). On the other hand, the recombination in ES mixes the elements of the current population, the mutation being applied after recombination.

(c) The selection used in DE is, for each element of the population, as the selection used in $(1+1)$-ESs. So, from the selection viewpoint, DE can be viewed as a set of $(1+1)$-ESs which work simultaneously. These separate algorithms change information during the mutation and crossover steps.

3. Computation of the expected population variance The explorative power of an EA expresses the ability of the algorithm to explore the search space. A measure of the explorative power of an EA is the population variance (Beyer and Deb, 1999). The selection step may increase or decrease the population variance. To avoid any premature convergence or stagnation, the variation operators must adjust the population variance such that it has a reasonable value from one generation to another. Thus if the selection decreases the population variance then it is necessary that the variation operators increase it.

Since the aim of this section is to compute the population variance let us define it for a population $x = \{x_1, \ldots, x_m\}$ of scalar elements ($x_l \in R$). The variance of population x is:

$$\mathrm{Var}(x) = \frac{1}{m}\sum_{l=1}^{m}(x_l - \overline{x})^2 = \overline{x^2} - \overline{x}^2$$

where $\overline{x} = \dfrac{1}{m}\sum_{l=1}^{m} x_l$ is the population mean and $\overline{x^2} = \dfrac{1}{m}\sum_{l=1}^{m} x_l^2$ is the quadratic population mean.

If the elements of the population are affected by some random elements, then its variance, $\mathrm{Var}(X)$, is a random variable and the measure of the explorative power should be its expected value, $E(\mathrm{Var}(X))$, where the mean E is taken with respect to all random elements which affect X. In the following the random populations and its elements will be denoted with capital letters while the deterministic populations will be denoted with small letters.

In DE, as in others evolutionary algorithms, the population, usually, consists not of scalars but of n-dimensional vectors. Since evolutionary transformations are made similarly and independently to all components it suffices to restrict the analysis to one-component elements.

The main result of this paper is the following theorem.

Theorem 1. *Let $x = \{x_1, \ldots, x_m\}$ be the current population, $Y = \{Y_1, \ldots, Y_m\}$ the intermediate population obtained after applying the mutation and $Z = \{Z_1, \ldots, Z_m\}$ the population obtained by crossing over the* *populations X and Y. If F is the parameter of the mutation step and p_c is the parameter of the crossover step then*

$$E(Var(Z)) = (2F^2 p_c + 1 - \frac{2p_c}{m} + \frac{p_c^2}{m})\, Var(X). \tag{1}$$

Proof. The proof contains two steps. In the first step is determined the influence of the mutation on the expected population variance while in the second step is computed the variance after crossover.
Step 1. Each element, Y_l of the population Y is obtained as $Y_l = x_{\alpha_1^l} + F\cdot(x_{\alpha_2^l} - x_{\alpha_3^l})$ where α_1^l, α_2^l and α_3^l are three elements of $\{1,2,\ldots,m\}$ selected through a random sampling without replacement. This type of sampling assures that the selected indices are different one to each other. Since sampling properties do not depend on the index l, to simplify the notations we shall drop the index l. It will be specified only if it is necessary. The indices α_1, α_2 and α_3 can be viewed as random variables with values in $\{1,2,\ldots,m\}$ and they have the property: $P(\alpha_i = k) = 1/m$ for each $i \in \{1,2,3\}$ and $k \in \{1,\ldots,m\}$. Thus x_{α_i} is also a random variable and

$$E(x_{\alpha_i}) = \sum_{k=1}^{m} P(\alpha_i = k)x_k = \frac{1}{m}\sum_{k=1}^{m} x_k = \overline{x},$$

$$E(x_{\alpha_i}^2) = \sum_{k=1}^{m} P(\alpha_i = k)x_k^2 = \frac{1}{m}\sum_{k=1}^{m} x_k^2 = \overline{x^2}.$$

The variables α_1, α_2, α_3 are not mutually independent, since their values must satisfy the condition of being distinct. Thus for $i \neq j$ the mean $E(x_{\alpha_i}x_{\alpha_j})$ is not necessarily equal to $E(x_{\alpha_i})E(x_{\alpha_j})$.

Since $P(\alpha_i = k, \alpha_j = l) = \dfrac{1}{m}\cdot\dfrac{1}{m-1} = \dfrac{1}{m^2-m}$ it follows that

$$E(x_{\alpha_i}x_{\alpha_j}) = \sum_{k\neq l} P(\alpha_i = k, \alpha_j = l)x_k x_l = \frac{1}{m^2-m}\sum_{k\neq l} x_k x_l =$$

$$= \frac{1}{m^2-m}\left[\left(\sum_{k=1}^{m} x_k\right)^2 - \sum_{k=1}^{m} x_k^2\right] = \frac{1}{m^2-m}(m^2\overline{x}^2 - m\overline{x^2}) =$$

$$= \frac{m}{m-1}\overline{x}^2 - \frac{1}{m-1}\overline{x^2}.$$

We compute now $E(\overline{Y^2})$.

$$E(\overline{Y^2}) = \frac{1}{m} E\left[\sum_{l=1}^{m} (x_{\alpha_1^l} + F \cdot (x_{\alpha_2^l} - x_{\alpha_3^l}))^2\right] =$$

$$= \frac{1}{m} E\left[\sum_{l=1}^{m} ((x_{\alpha_1^l})^2 + 2F x_{\alpha_1^l} x_{\alpha_2^l} - 2F x_{\alpha_1^l} x_{\alpha_3^l} + F^2 (x_{\alpha_2^l} - x_{\alpha_3^l})^2)\right] =$$

$$= \overline{x^2} + F^2 (\overline{x^2} - 2E(x_{\alpha_2^l} x_{\alpha_3^l}) + \overline{x^2}) = (2F^2 \frac{m}{m-1} + 1)\overline{x^2} - 2F^2 \frac{m}{m-1} \overline{x}^2. \quad (2)$$

To obtain the second equality we used the property $E(x_{\alpha_1} x_{\alpha_2}) = E(x_{\alpha_1} x_{\alpha_3})$. The second term which we have to compute is $E(\overline{Y}^2)$. Using the properties of x_{α_1}, x_{α_2}, x_{α_3} we obtain:

$$E(\overline{Y}^2) = \frac{1}{m^2} E\left[\sum_{l=1}^{m} (x_{\alpha_1^l} + F(x_{\alpha_2^l} - x_{\alpha_3^l}))\right]^2 =$$

$$= \frac{1}{m^2} E\left[\sum_{l=1}^{m} (x_{\alpha_1^l} + F(x_{\alpha_2^l} - x_{\alpha_3^l}))^2\right] +$$

$$+ \frac{1}{m^2} E\left[\sum_{k \neq l} (x_{\alpha_1^l} + F(x_{\alpha_2^l} - x_{\alpha_3^l}))(x_{\alpha_1^k} + F(x_{\alpha_2^k} - x_{\alpha_3^k}))\right].$$

The first term of $E(\overline{Y}^2)$ is $\frac{1}{m} E(\overline{Y^2})$. Since for each element of Y the indices α_1, α_2 and α_3 are obtained by a new sampling, independent from the other ones, it follows that for $k \neq l$ the variables $x_{\alpha_1^l} + F(x_{\alpha_2^l} - x_{\alpha_3^l})$ and $x_{\alpha_1^k} + F(x_{\alpha_2^k} - x_{\alpha_3^k})$ are mutually independent. Thus

$$E[(x_{\alpha_1^l} + F(x_{\alpha_2^l} - x_{\alpha_3^l}))(x_{\alpha_1^k} + F(x_{\alpha_2^k} - x_{\alpha_3^k}))] =$$

$$= E[(x_{\alpha_1^l} + F(x_{\alpha_2^l} - x_{\alpha_3^l}))] \cdot E[(x_{\alpha_1^k} + F(x_{\alpha_2^k} - x_{\alpha_3^k}))] =$$

$$= E(x_{\alpha_1^l}) \cdot E(x_{\alpha_1^k}) = \overline{x}^2.$$

Hence we obtained that

$$E(\overline{Y}^2) = \frac{1}{m} E(\overline{Y^2}) + \frac{m-1}{m} \overline{x}^2.$$

Thus

$$E(\mathrm{Var}(Y)) = E(\overline{Y^2}) - E(\overline{Y}^2) = \left(2F^2 + \frac{m-1}{m}\right)\mathrm{Var}(x). \qquad (3)$$

Step 2. Each element of the population Z is a random variable with the structure

$$Z_l = \begin{cases} Y_l & \text{with probability } p_c \\ x_l & \text{with probability } 1-p_c \end{cases} \qquad l = \overline{1,m}.$$

To determine $E(\mathrm{Var}(Z))$ we have to compute $E(\overline{Z^2})$ and $E(\overline{Z}^2)$. Using Lemma 4.1 from (Zaharie, 2001) to compute the mean of Z_l^2 one obtains:

$$E(\overline{Z^2}) = \frac{1}{m}\sum_{l=1}^{m} E(Z_l^2) = \frac{1}{m}\sum_{l=1}^{m}((1-p_c)x_l^2 + p_c E(Y_l^2)) =$$

$$= (1-p_c)\overline{x^2} + p_c E(\overline{Y^2}).$$

On the other hand,

$$E(\overline{Z}^2) = \frac{1}{m^2}E\left[\left(\sum_{l=1}^{m} Z_l\right)^2\right] = \frac{1}{m^2}\left[E\left(\sum_{l=1}^{m} Z_l^2\right) + E\left(\sum_{k\neq l} Z_k Z_l\right)\right] =$$

$$= \frac{1}{m}E(\overline{Z^2}) + \frac{1}{m^2}\sum_{k\neq l} E(Z_k)E(Z_l) \qquad (4)$$

since Z_k and Z_l are mutually independent.

Using again Lemma 4.1 (Zaharie, 2001) the second term in the last relation can be expanded as follows:

$$\sum_{k\neq l} E(Z_k)E(Z_l) = \sum_{k\neq l}((1-p_c)x_k + p_c E(Y_k))((1-p_c)x_l + p_c E(Y_l)).$$

But $E(Y_l) = E[x_{\alpha_1^l} + F\cdot(x_{\alpha_2^l} - x_{\alpha_3^l})] = E(x_{\alpha_1^l}) = \overline{x}$. Thus

$$\sum_{k\neq l} E(Z_k)E(Z_l) = \left[\sum_{l=1}^{m}((1-p_c)x_l + p_c\overline{x})\right]^2 - \sum_{l=1}^{m}((1-p_c)x_l + p_c\overline{x})^2 =$$

$$= (m(1-p_c)\overline{x} + mp_c\overline{x})^2 - \sum_{l=1}^{m} \left[(1-p_c)^2 x_l^2 + 2p_c(1-p_c)\overline{x}x_l + p_c^2\overline{x}^2\right] =$$

$$= m^2\overline{x}^2 - m(1-p_c)^2\overline{x^2} - 2p_c(1-p_c)m\overline{x}^2 - mp_c^2\overline{x}^2 =$$

$$= (m^2 - 2mp_c + mp_c^2)\overline{x}^2 - m(1-p_c)^2\overline{x^2}.$$

Replacing the last relation in (4) one obtains

$$E(\overline{Z}^2) = \frac{1}{m}E(\overline{Z^2}) + \frac{1}{m^2}(m^2 - 2mp_c + mp_c^2)\overline{x}^2 - \frac{1}{m}(1-p_c)^2\overline{x^2}.$$

Thus

$$E(Var(Z)) = \frac{m-1}{m}\Big((1-p_c)\overline{x^2} +$$

$$+p_c\left(1 + 2F^2\frac{m}{m-1}\right)\overline{x^2} - 2p_cF^2\frac{m}{m-1}\overline{x}^2\Big) -$$

$$\left(1 - \frac{2p_c}{m} + \frac{p_c^2}{m}\right)\overline{x}^2 + \frac{(1-p_c)^2}{m}\overline{x^2} =$$

$$= \left(1 + 2F^2p_c - \frac{2p_c}{m} + \frac{p_c^2}{m}\right)\mathrm{Var}(x). \tag{5}$$

and the theorem is proved.

If $X^{(0)}$ is the initial population and $X^{(g)}$ is the population obtained after g generations then:

$$E(\mathrm{Var}(X^{(g)})) = \left(1 + 2F^2p_c - \frac{2p_c}{m} + \frac{p_c^2}{m}\right)^g \mathrm{Var}(X^{(0)}). \tag{6}$$

The expected population variance computed for an ES without selection, with a mutation based on addition of a random vector (having mutually independent normally distributed components with zero mean and a variance σ^2) and dominant recombination (global discrete recombination) is (for details see (Beyer, 1999)):

$$E(\mathrm{Var}(X^{(g)})) = \left(1 - \frac{1}{m}\right)^g \mathrm{Var}(X^{(0)}) + m\left[1 - \left(1 - \frac{1}{m}\right)^g\right]\sigma^2. \tag{7}$$

For great values of m, (6) becomes

$$E(\mathrm{Var}(X^{(g)})) = (1 + 2F^2p_c)^g\mathrm{Var}(X^{(0)}). \tag{8}$$

and (7) becomes

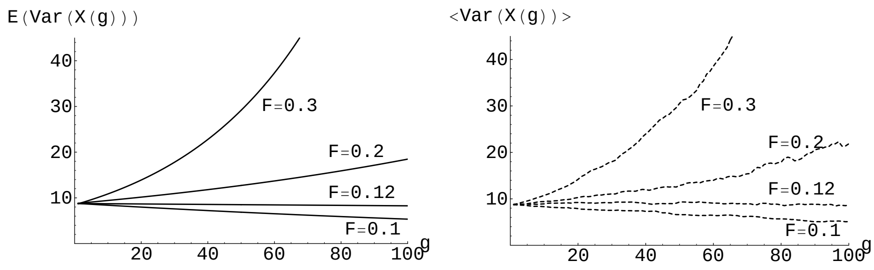
$$E(\mathrm{Var}(X^{(g)})) = \mathrm{Var}(X^{(0)}) + g\sigma^2. \quad (9)$$

Since $2F^2p_c + 1 > 1$ it follows that the expected population variance (after applying mutation and recombination) is greater in the case of DE algorithm than in the case of the ES algorithm analyzed in (Beyer, 1999). Thus the DE algorithm has an explorative power greater than that of some classical ES algorithms. This could explain the good performance of the DE algorithm.

4. Experimental results The aim of the experimental study was not to emphasize the efficiency of DE with respect to other optimization methods but to analyze the evolution of the population variance. The empirical expected population variance ($< Var(X(g)) >$) has been obtained by applying the DE algorithm to some classical test functions and by collecting the results from 100 independent runs. The test functions which we used are:
1. (Rastrighin):
$f_R(x_1,\ldots,x_n) = 0.2\sum_{k=1}^{n} x_k^2 - 2(\cos(10\sum_{k=1}^{n} x_k) - 1)$, $x_k \in [-5.12, 5.12]$, $n = 20$.
2. (Ackley):
$f_A(x_1,\ldots,x_n) = -20\exp(-0.2\sqrt{(\sum_{k=1}^{n} x_k^2)/n} - exp(\sum_{k=1}^{n} cos(2\pi x_k)) + 20 + e$, $x_k \in [-32, 32]$, $n = 20$.

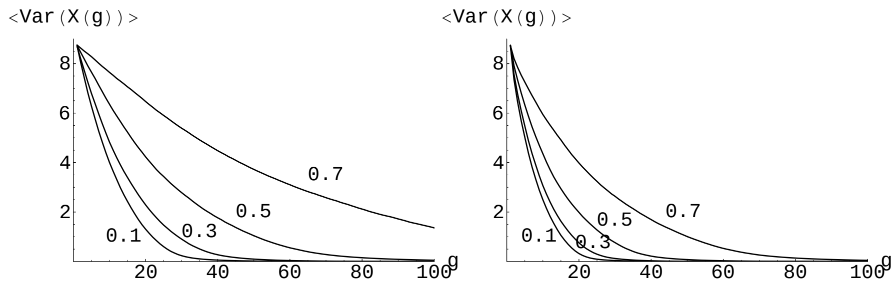
In all experimental studies we used a population of size $m = 50$ and a fixed value of $p_c = 0.5$. Fig. 1 illustrates, for different values of parameter F, that the obtained theoretical result agrees with the empirical results obtained after applying DE mutation and crossover. At each generation, the empirical variance has been computed by averaging the sample variance for all n components and for all 100 independent runs. The population variance behavior suggest that there exists a critical value F_c under which the variance decreases and over which it increases. F_c is the positive solution of the equation $2F^2 + p_c/m - 2/m = 0$ and for $m = 50$, $p_c = 0.5$ it is $F_c = 0.1224$.

The evolution of the population variance after the selection is also applied is illustrated in Fig.2 for two types of selections (DE vs. ES) and for different values of F, while the mutation and crossover are both of DE type. We remark that the ES-selection induces a greater decrease in the population variance than the DE-selection.



(a) (b)

Fig. 1. Evolution of the expected population variance after mutation and crossover ((a) theoretical dependence; (b) empirical results for Rastrigin function)



(a) (b)

Fig. 2. Evolution of the empirical population variance after mutation, crossover and selection for different values of F (Rastrigin function) ((a) DE-type selection; (b) ES-type selection)

The ability of the DE algorithm to find the global minimum of Rastrigin's and Ackley's function is illustrated by the results in Table 1. As stopping criterion for the DE algorithm we used: "$f_* < 10^{-3}$ or $g > 1000$" (f_* is the best value of the objective function found into the population and g is the number of generations). The algorithm is successful if the first part of the condition is true (in this case g^* denote the number of generation until the optimum is reached). In Table 1, $\bar{\cdot}$ denotes the average over the 100 independent runs. Results in Table 1 suggest that values of F lower or near the critical value, F_c, could induce premature convergence (the population variance vanishes before the optimum is reached) which agrees with the theoretical results.

	Rastrigin				Ackley			
F	Succ.	$\overline{g^*}$	$\overline{f_*}$	$< V >$	Succ.	$\overline{g^*}$	$\overline{f_*}$	$\overline{V}$
0.1	3%	122	0.1897	0	0%	–	1.9548	0
0.12*	22%	134	0.0468	10^{-6}	1%	194	1.5043	0
0.15	57%	129	0.0256	$4 \cdot 10^{-6}$	18%	211	0.5165	0
0.2	97%	126	0.0009	$11 \cdot 10^{-6}$	64%	213	0.1162	0
0.3	100%	147	0.0008	$13 \cdot 10^{-6}$	100%	245	0.0009	0
0.5	100%	255	0.0008	$14 \cdot 10^{-6}$	100%	425	0.0009	0
0.7	100%	549	0.0008	$14 \cdot 10^{-6}$	100%	926	0.0009	0

Table 1. Behavior of DE algorithm for Rastrigin and Ackley functions ($m = 50$, $p_c = 0.5$)

REFERENCES

Bäck,T. and H.P. Schwefel (1995). Evolution Strategies I. Variants and their Computational Implementation, in G.Winter, J. Periaux, M.Galan, P.Cuesta (eds), *Genetic Algorithms in Engineering and Computer Science*, Wiley Chichester, pp. 111-126.

Beyer,H.G.(1998). On the Explorative Power of ES/EP-like Algorithms, *Technical Report*, University of Dortmund.

Beyer,H.G.(1999). On the Dynamics of EAs without selection, in W. Banzaf, C. Reeves (eds.), *Foundations of Genetic Algorithms*, Morgan Kaufmann, pp.5-26.

Beyer,H.G. and K. Deb(1999). On the Analysis of Self Adaptive Evolutionary Algorithms, *Technical Report*, CI-69/99, University of Dortmund.

Lampinen,J.(1999). A Bibliography of Differential Evolution Algorithm, Technical Report, Lappeenranta University of Technology, Department of Information Technology, Laboratory of Information Processing.

Storn,R. and K. Price(1997). Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, *Journal of Global Optimization*, **11**(4), pp.341-359.

Zaharie,D.(2001). Recombination operators for evolutionary algorithms, *Proc. of the 27th Summer School "Applications of Mathematics in Engineering and Economics"*, eds. B.I. Cheshankov, M.D.Todorov, Heron Press, Sofia, 2001.

University of the West,
Faculty of Mathematics
Department of Computer Science
B-dul V. Pârvan, 4
Timişoara, 1900, ROMANIA
email:dzaharie@info.uvt.ro

Received January 2002