A Parameter-less Evolution Strategy for Global Optimization

Lino Costa
Departamento de Produção e Sistemas,
Escola de Engenharia, Universidade do Minho,
Campus de Gualtar, 4710-057 Braga, Portugal

Abstract- Several evolutionary approaches have been applied to global optimization problems with significant success. Evolution strategies proved to be efficient global optimizers. However, these algorithms have several parameters which the setting is not simple. Thus, it is crucial to investigate how to set dynamically these parameters during the search. In this paper, a new parameter-less evolution strategy, which has only one single parameter to set, is proposed. This algorithm is compared with the traditional evolution strategies considering a set of difficult test problems. The results obtained indicate a promising performance of the new approach.

Keywords: Global Optimization, Meta-Heuristics, Evolutionary Computation, Evolution Strategies

1 Introduction

Several evolutionary approaches have been applied to global optimization problems with success, namely Genetic Algorithms [1] and Evolution Strategies [2] [3]. In the past, Evolution Strategies (ESs) proved to be powerful global optimizers [4]. They use a real coding of decision variables and the adaptation of step sizes for mutation. However, ESs have multiple parameters that are difficult to set since they are problem dependent. So, these algorithms require some expertise in order to make them work conveniently. On the other hand, in general, some previous experimentation is needed in order to tune the algorithm parameters. Moreover, in general, interaction between parameters exists and must be taken into account. Thus, it is relevant to search for algorithms that have a reduced number of parameters to set. These algorithms must took the task of setting parameters from the user. These parameters must be set by the algorithm itself during the search taking into account the features of the problem being solved.

In the proposed approach, the Parameter-less Evolution Strategy (PLES), an effort was made in order to avoid the difficult task of setting initial values for parameters. So, the setting of almost all parameters required by traditional ESs was avoided. Several adaptation rules that avoid the need for the initial setting of parameters were implemented. On the other hand, new operators were developed in order to improve the performance of the adaptation rules.

In order to evaluate its performance, PLES was compared with a traditional $(\mu/\rho + \lambda)$-ES using a set of difficult test problems. With this preliminary experiments, it was intended to validate the new approach in order to identify future developments of the algorithm. The preliminary results obtained indicate a good performance of the new approach. Thus, future developments of the algorithm should be investigated.

In section 2, a short introduction to ESs is presented. Section 3 describes the Parameter-less Evolution Strategy (PLES) implemented. Next, the results of the application to several test problems are presented, as well as some comparisons with traditional ESs. Finally, some conclusions and future work are addressed.

2 Global Optimization with Evolution Strategies

Evolution Strategies are search procedures that mimic the natural evolution of the species in the natural systems. In the past, ESs were applied to the resolution of nonlinear optimization problems without constraints. These problems can be formulated, mathematically, as follows:

$$\min f(x) \text{ with } x \in \mathrm{R}^n \qquad (1)$$

subject to

$$\alpha \leq x \leq \beta$$

where x is the vector of n real decision variables, $f(x)$ is the objective function to minimize, and α and β are the vectors of the lower and upper bounds of the decision variables.

ESs work directly with the real representation of the parameter set, searching from an initial population (a set of points), requiring only data based on the objective function and constraints, and not derivatives or other auxiliary knowledge. Traditionally, two distinct types of ESs differing basically on the selection procedure are considered: the $(\mu + \lambda)$-ES and the (μ, λ)-ES.

In this nomenclature, μ and λ represent, respectively, the parent and offspring population sizes (for many problems, $\lambda/\mu \approx 7$ is suggested [3]). Each population member consists on a tuple of two vectors: a vector of real values representing the decision variables and a vector of real standard deviations used to adapt step sizes during the search. Thus, each decision variable i has an associated standard deviation σ_i. The search starts from an initial population which individuals are, in general, generated at random. The initial standard deviations σ_i can be set according to equation (2), where Δx is a rough measure of the distance from the optimum and n is the dimension of the problem.

$$\sigma_i^{(0)} = \frac{\Delta x}{\sqrt{n}} \qquad (2)$$

Alternatively, if an approximation to the optimum is not known, the setting given by the following equation can be considered:

$$\sigma_i^{(0)} = \frac{\beta_i - \alpha_i}{\lambda\sqrt{n}}. \quad (3)$$

In spite of, traditionally, the search of new points was based on one single operator, the mutation operator, in general, ESs benefit with the introduction of the recombination operator. Thus, the nomenclature for ESs can now be extended, and ESs with recombination are usually referred as $(\mu/\rho+\lambda)$-ES or $(\mu/\rho,\lambda)$-ES. It should be noted that one of the most promising features of ESs is that they use adaptive step sizes for mutation. So, these parameters of the algorithm are themselves optimized during the search.

Next, the basic algorithm and the main features of ESs are presented.

2.1 Algorithm

The basic $(\mu/\rho\overset{+}{,}\lambda)$-ES can be described as follows:

$(\mu/\rho\overset{+}{,}\lambda)$- **Evolution Strategy**

1. Initialization of the individuals of the parent population and corresponding step sizes

$$(x_p^{(0)};\sigma_p^{(0)}) = (x_{p,1}^{(0)},\ldots,x_{p,n}^{(0)};\sigma_{p,1}^{(0)},\ldots,\sigma_{p,n}^{(0)})$$

where $x_{p,i}^{(0)} \sim U(\alpha_i,\beta_i)$ and $\sigma_{p,i}^{(0)} = \frac{\beta_i-\alpha_i}{\lambda\sqrt{n}}$, for all $p = 1,\ldots,\mu$ and $i = 1,\ldots,n$.

2. Let $k = 0$

3. Recombination of the individuals of the parent population

$$(\widetilde{x}_p^{(k)};\widetilde{\sigma}_p^{(k)}) = (x_{u_1,1}^{(k)},\ldots,x_{u_n,n}^{(k)};\sigma_{u_1,1}^{(k)},\ldots,\sigma_{u_n,n}^{(k)})$$

where $u_i \sim U(1,\rho)$ and u_i integer, for all $p = 1,\ldots,\mu$ and $i = 1,\ldots,n$.

4. Step size adaptation
Non-isotropic adaptation

$$(\widehat{x}_p^{(k)};\widehat{\sigma}_p^{(k)}) = (\widetilde{x}_{p,1}^{(k)},\ldots,\widetilde{x}_{p,n}^{(k)};\widetilde{\sigma}_{p,1}^{(k)}e^{z_1}e^{z},\ldots,\widetilde{\sigma}_{p,n}^{(k)}e^{z_n}e^{z})$$

where $z_i \sim N(0,\Delta\sigma^2)$, $z \sim N(0,\Delta\sigma^{\text{‘}2})$, for all $p = 1,\ldots,\mu$ and $i = 1,\ldots,n$.

5. Mutation of the individuals of the parent population

$$(\overline{x}_d^{(k)};\overline{\sigma}_d^{(k)}) = (\widehat{x}_{u,1}^{(k)}+z_{d,1},\ldots,\widehat{x}_{u,n}^{(k)}+z_{d,n};\widehat{\sigma}_{u,1}^{(k)},\ldots,\widehat{\sigma}_{u,n}^{(k)})$$

where $z_{d,i} \sim N(0,\widehat{\sigma}_{u,i}^{(k)})$,
$u = \begin{cases} \mu & \text{if } d = \mu, 2\mu, \ldots, K\mu \text{ with } K \text{ integer} \\ d//\mu & \text{otherwise} \end{cases}$
($//$ states for the rest of integer division), for all $d = 1,\ldots,\lambda$ and $i = 1,\ldots,n$.

6. Selection
If (,)-selection then

$$(\check{x}_q^{(k)};\check{\sigma}_q^{(k)}) = (\overline{x}_q^{(k)};\overline{\sigma}_q^{(k)})$$

for $q = 1,\ldots,\lambda$
Sort all $(\check{x}_q^{(k)};\check{\sigma}_q^{(k)})$, so that $f(\check{x}_a^{(k)}) \le f(\check{x}_b^{(k)})$ for all $a,b = 1,\ldots,\lambda$.
If (+)-selection then

$$(\check{x}_q^{(k)};\check{\sigma}_q^{(k)}) = \begin{cases} (x_q^{(k)};\sigma_q^{(k)}) & \text{if } \quad 1 \le q \le \mu \\ (\overline{x}_{q-\mu}^{(k)};\overline{\sigma}_{q-\mu}^{(k)}) & \text{if } \quad \mu+1 \le q \le \mu+\lambda \end{cases}$$

for $q = 1,\ldots,\mu+\lambda$
Sort all $(\check{x}_q^{(k)};\check{\sigma}_q^{(k)})$, so that $f(\check{x}_a^{(k)}) \le f(\check{x}_b^{(k)})$ for all $a,b = 1,\ldots,\mu+\lambda$.

7. Replace the individuals of the parent population

$$(x_p^{(k+1)};\sigma_p^{(k+1)}) = (\check{x}_p^{(k)};\check{\sigma}_p^{(k)})$$

for $p = 1,\ldots,\mu$.
$k = k+1$.

8. If stopping criterion is not true then return to step 3. else end.

Thus, in a $(\mu/\rho+\lambda)$-ES, at a given generation, there are μ parents, and λ offspring are generated by recombination and mutation. Basically, the recombination operator consists on, before mutation, to recombine a set of chosen parents to find a new solution. On other hand, mutation creates new points by adding random normal distributed quantities. Next, the $\mu+\lambda$ individuals are sorted according to their objective function values. Finally, the best μ of all the $\mu+\lambda$ members become the parents of the next generation (i.e., the selection takes place between the $\mu+\lambda$ members). The $(\mu/\rho,\lambda)$-ES is similar differing, basically, on the selection procedure that is restricted to the offspring population, i.e., the selection takes place between the λ offspring.

2.2 Recombination

Basically, the recombination operator consists on, before mutation, to recombine a set of chosen parents to find a new solution. A given number ρ ($1 \le \rho \le \mu$) of parents are randomly chosen for recombination. When $\rho = 1$ then there is no recombination. Two types of recombination are, mainly, considered: intermediate and discrete recombination. Since, in this work, the recombination implemented was the discrete recombination, only this recombination will be described in detail. In the discrete recombination, each component of the offspring is chosen from one of the ρ parents at random. Thus, for ρ chosen parents (randomly selected from population), the offspring x_p is given by

$$x_p = (x_{u_1,1},\ldots,x_{u_n,n})$$

with $u_1 \in \{1,\ldots,\rho\},\ldots,u_n \in \{1,\ldots,\rho\}$ and $p = 1,\ldots,\mu$. In discrete recombination, the integer uniform random values u_i, for $i = 1,\ldots,n$, allow the selection of which of the ρ parents will give the value of decision variable i. This procedure allows different combinations of the values of the decision variables from existing solutions in the population. Standard deviations are similarly recombined.

2.3 Step size adaptation

During the search, the step sizes for mutation are adapted. Several self-adaptation schemes are possible. One possibility is to actualize the standard deviations σ_i (for each decision variable) according to the equation [3]:

$$\sigma_i^{(k+1)} = \sigma_i^{(k)}e^{z_i}e^{z} \quad (4)$$

where $z_i \sim N(0,\Delta\sigma^2)$, $z \sim N(0,\Delta\sigma^{\text{‘}2})$ and $\Delta\sigma$ and $\Delta\sigma$‘ are parameters of the algorithm. In the experiments conducted only this non-isotropic adaptation rule was considered, other adaptation rules are described by Bäck [5].

2.4 Mutation

Usually, the random numbers $z^{(k)}$ are generated according to a Gaussian or Normal distribution. Besides, it is convenient that small changes occur frequently, but large ones only rarely. So, two requirements arise together for the generation of the random numbers $z^{(k)}$:

- the expected value of the components $z_i^{(k)}$ of $z^{(k)}$ must be equal to zero, i.e., $E(z_i^{(k)}) = 0$ for $i = 1, \ldots, n$, and
- the variances σ_i^2 must be small, for $i = 1, \ldots, n$.

In this sense, the random numbers $z_i^{(k)}$ can be generated according to a Normal distribution with mean zero and variance σ_i^2:

$$z_i^{(k)} \sim N(0, \sigma_i^2) \tag{5}$$

So, mutation consists on adding random numbers with mean zero and variance σ_i^2 to the vector of decision variables, i.e., $x_d = x_u + z$.

3 The Parameter-less Evolution Strategy

In general, evolutionary algorithms have multiple parameters that are difficult to set since they are problem dependent. So, these algorithms require some expertise in order to make them work conveniently. On the other hand, in general, some previous experimentation is needed in order to tune the algorithm parameters. Although the inclusion in ESs of some adaptation rules for some parameters, there are always some parameters that are required to set. Moreover, in general, interaction between parameters is an important issue that must be taken into account. Thus, it is relevant to search for algorithms that have a reduced number of parameters to set.

In the proposed approach, the so-called Parameter-less Evolution Strategy, an effort was made in order to avoid the difficult task of setting initial values for parameters. So, the setting of almost all parameters required by traditional ESs was avoided. Actually, PLES requires the initial setting of one single parameter, the parent population size (μ). The algorithm includes several adaptation rules that avoid the need for initial values of parameters. On the other, the recombination and mutation operators were also substantially modified in order to allow the implementation of particular adaptation rules.

The adaptation rule for standard deviations is based on the success of the parents to generate better offspring. This rule implies the sampling of the search space taking into account the distances between the parents and the generated offspring. Thus, in PLES, the number of offspring corresponds to the number of possible combinations of two parents, i.e.:

$$\lambda = \binom{\mu}{2} = \frac{\mu!}{2!(\mu-2)!}.$$

Each generation, λ offspring are generated by recombination and mutation of all pairs of two parents. This scheme allows the use of the distance between parents and offspring to estimate or adapt the standard deviations. It should be noted that the recombination and mutation operators were also modified in order to collect this information. On the contrary to traditional ESs in which recombination and mutation are applied at distinct steps of the algorithm, in PLES, these operators are applied conjointly.

3.1 Algorithm

The basic PLES algorithm can be described as follows:

Parameter-less Evolution Strategy

1. Initialization of the individuals of the parent population and corresponding step sizes

   $$(x_p^{(0)}; \sigma_p^{(0)}) = (x_{p,1}^{(0)}, \ldots, x_{p,n}^{(0)}; \sigma_{p,1}^{(0)}, \ldots, \sigma_{p,n}^{(0)})$$

   where $x_{p,i}^{(0)} \sim U(\alpha_i, \beta_i)$ and $\sigma_{p,i}^{(0)} = \beta_i - \alpha_i$, for all $p = 1, \ldots, \mu$ and $i = 1, \ldots, n$.
2. Let $k = 0$
3. Recombination and mutation of the individuals of the parent population

   $$(\widetilde{x}_d^{(k)}; \widetilde{\sigma}_d^{(k)}) = (x_{c,1}^{(k)}, \ldots, x_{c,n}^{(k)}; \sigma_{c,1}^{(k)}, \ldots, \sigma_{c,n}^{(k)})$$

   where $x_{c,i}^{(k)} = (x_{a,i}^{(k)} + x_{b,i}^{(k)})/2 + z_{d,i}$ with $z_{d,i} \sim N(0, \sqrt{(\sigma_{a,i}^{(k)})^2 + (\sigma_{b,i}^{(k)})^2}/2)$ and $\sigma_{c,i}^{(k)} = \sqrt{(\sigma_{a,i}^{(k)})^2 + (\sigma_{b,i}^{(k)})^2}/2$ (a and b correspond to a combination of two progenitors), for all $d = 1, \ldots, \lambda$ and $i = 1, \ldots, n$.
4. Step size adaptation

   Success based adaptation

   Compute $D_{da,i} = \widetilde{x}_{d,i}^{(k)} - x_{a,i}^{(k)}$ and $D_{db,i} = \widetilde{x}_{d,i}^{(k)} - x_{b,i}^{(k)}$ with $i = 1, \ldots, n$

   if $f(\widetilde{x}_d^{(k)}) < f(x_a^{(k)})$ then

   $$(x_a^{(k)}; \sigma_a^{(k)}) = (x_{a,1}^{(k)}, \ldots, x_{a,n}^{(k)}; |D_{da,1}|, \ldots, |D_{da,n}|)$$

   if $f(\widetilde{x}_d^{(k)}) < f(x_b^{(k)})$ then

   $$(x_b^{(k)}; \sigma_b^{(k)}) = (x_{b,1}^{(k)}, \ldots, x_{b,n}^{(k)}; |D_{db,1}|, \ldots, |D_{db,n}|)$$

   for all $d = 1, \ldots, \lambda$ and $i = 1, \ldots, n$.

   Non-isotropic adaptation

   $$(\widehat{x}_p^{(k)}; \widehat{\sigma}_p^{(k)}) = (\widetilde{x}_{p,1}^{(k)}, \ldots, \widetilde{x}_{p,n}^{(k)}; \widetilde{\sigma}_{p,1}^{(k)} e^{z_1} e^{z}, \ldots, \widetilde{\sigma}_{p,n}^{(k)} e^{z_n} e^{z})$$

   where $z_i \sim N(0,1)$, $z \sim N(0,1)$, for all $p = 1, \ldots, \mu$ and $i = 1, \ldots, n$.
5. Selection

   $$(\breve{x}_q^{(k)}; \breve{\sigma}_q^{(k)}) = \begin{cases} (x_q^{(k)}; \sigma_q^{(k)}) & \text{if} \quad 1 \le q \le \mu \\ (\widehat{x}_{q-\mu}^{(k)}; \widehat{\sigma}_{q-\mu}^{(k)}) & \text{if} \quad \mu + 1 \le q \le \mu + \lambda \end{cases}$$

   for $q = 1, \ldots, \mu + \lambda$

   Sort all $(\breve{x}_q^{(k)}; \breve{\sigma}_q^{(k)})$, so that $f(\breve{x}_a^{(k)}) \le f(\breve{x}_b^{(k)})$ for all $a, b = 1, \ldots, \mu + \lambda$.
6. Replace the individuals of the parent population

   $$(x_p^{(k+1)}; \sigma_p^{(k+1)}) = (\breve{x}_p^{(k)}; \breve{\sigma}_p^{(k)})$$

   for $p = 1, \ldots, \mu$.

   $k = k + 1$.
7. If stopping criterion is not true then return to step 3. else end.

As in traditional ESs, in PLES, at a given generation, there are μ parents, and λ offspring are generated by recombination and mutation. However, the recombination and mutation operators act jointly. Thus, all parents are recombined and mutated in order to generate new offspring. The $\mu+\lambda$ individuals are sorted according to their objective function values. Finally, the best μ of all the $\mu + \lambda$ members become the parents of the next generation. Thus, in PLES, the selection is similar to the $(\mu/\rho + \lambda)$-ES selection.

3.2 Recombination and Mutation

Each generation, λ offspring are generated by recombination and mutation of all pairs of two parents. Each combination of two parents generates an offspring. Thus, the offspring x_d, generated from parents x_a and x_b by recombination and mutation, is given by:

$$x_d = (x_{c,1}, \ldots, x_{c,n})$$

where $x_{c,i} = \frac{x_{a,i}+x_{b,i}}{2} + z_{d,i}$ with $z_{d,i} \sim N(0, \frac{\sqrt{\sigma_{a,i}^2+\sigma_{b,i}^2}}{2})$, for all $d = 1, \ldots, \lambda$ and $i = 1, \ldots, n$.

The new offspring inherits from parents a and b the following standard deviations:

$$\sigma_{d,i} = \frac{\sqrt{\sigma_{a,i}^2 + \sigma_{b,i}^2}}{2}$$

where $i = 1, \ldots, n$. The application of these operators in a single phase allows the implementation of a step size adaptation rule that takes into account the success of the parents in generating new better offspring.

3.3 Step size adaptation

During the search, the step sizes for mutation are adapted. The step sizes are adapted in two phases: firstly, the adaptation is based on the success of the parents to generate offspring; secondly, a non-isotropic adaptation rule is applied.

The success based adaption consists on, after the generation of an offspring from the recombination of two parents, to actualize the standard deviations of the parents. If the offspring is better than a parent (or both parents) then the standard deviations of the parent (or parents) is (are) actualized according to the distances between the parents and the offspring. Thus, for each decision variable i, the distances between each parent and the offspring are computed. This success based rule can be expressed by:

$$\begin{cases} \sigma_{a,i} = |D_{da,i}| & \text{if} \quad f(x_d) < f(x_a) \\ \sigma_{b,i} = |D_{db,i}| & \text{if} \quad f(x_d) < f(x_b) \end{cases}$$

where $D_{da,i} = x_{d,i} - x_{a,i}$ and $D_{db,i} = x_{d,i} - x_{b,i}$, for all $d = 1, \ldots, \lambda$ and $i = 1, \ldots, n$.

A non-isotropic adaptation scheme is also applied in order to actualize the standard deviations σ_i (for each decision variable) according to the equation:

$$\sigma_i^{(k+1)} = \sigma_i^{(k)} e^{z_i} e^{z} \qquad (6)$$

where $z_i \sim N(0, 1)$, $z \sim N(0, 1)$.

4 Results

In this section the results obtained by the PLES and a $(\mu/\rho + \lambda)$-ES are presented. The parameters of the algorithms, implemented in C language, were kept constant for all problems (no effort was made in finding the best parameter setting for each problem). The algorithms were executed in a PC with a Pentium(R) 4 (2.00 GHz) CPU and 256 MB of RAM (running the Windows ME operating system). The Table 1 presents the parameters considered for the $(\mu/\rho+\lambda)$-ES. The parent population size (μ) considered for PLES was 10 individuals.

Parameter	Value
Parents population size (μ)	10
Offspring population size (λ)	100
Number of recombinants (ρ)	10
Selection type	+
$\Delta\sigma$	$1/\sqrt(2n)$
$\Delta\sigma$‘	$1/\sqrt(2\sqrt(n))$

Table 1: The $(\mu/\rho + \lambda)$-ES parameters considered in experiments

4.1 Test Problems and Performance Evaluation Criteria

The 25 test problems considered are summarized in Table 3. These problems were collected by Suganthan and are available at the web site address: http://www.ntu.edu.sg/home/EPNSugan. This set of difficult test problems intends to constitute a standard test suite for global optimization, including problems with very distinct properties. In this paper only the results for the problems considering 10 decision variables are included (each problem was solved by each algorithm 25 times). Initial populations were uniformly generated at random within the search space (except for problems 7 and 25, for which specific initialization ranges are required). Each execution was terminated when the error ($|f(x) - f(x*)|$) becomes inferior to 10^{-8}. The maximum number of function evaluations (FES) allowed was 10000. In order to compare the performance of the algorithms during the search toward the optimum, the error was recorded for FES equal to 10^3 and 10^4.

Unimodal Functions:
F1: Shifted Sphere Function
F2: Shifted Schwefel's Problem 1.2
F3: Shifted Rotated High Conditioned Elliptic Function
F4: Shifted Schwefel's Problem 1.2 with Noise in Fitness
F5: Schwefel's Problem 2.6 with Global Optimum on Bounds
Multimodal Functions:
F6: Shifted Rosenbrock's Function
F7: Shifted Rotated Griewank's Function without Bounds
F8: Shifted Rotated Ackley's Function with Global Optimum on Bounds
F9: Shifted Rastrigin's Function
F10: Shifted Rotated Rastrigin's Function
F11: Shifted Rotated Weierstrass Function
F12: Schwefel's Problem 2.13
F13: Expanded Extended Griewank's plus Rosenbrock's Function (F8F2)
F14: Shifted Rotated Expanded Scaffer's F6
F15: Hybrid Composition Function
F16: Rotated Hybrid Composition Function
F17: Rotated Hybrid Composition Function with Noise in Fitness
F18: Rotated Hybrid Composition Function
F19: Rotated Hybrid Composition Function with a Narrow Basin for the Global Optimum
F20: Rotated Hybrid Composition Function with the Global Optimum on the Bounds
F21: Rotated Hybrid Composition Function
F22: Rotated Hybrid Composition Function with High Condition Number Matrix
F23: Non-Continuous Rotated Hybrid Composition Function
F24: Rotated Hybrid Composition Function
F25: Rotated Hybrid Composition Function without Bounds

Table 3: Test Problems

4.2 Discussion

Tables 2, 4 and 5 present the error values achieved when FES was set to 10^3 and 10^4. For each function, the best (1st), 7th, median (13th), 19th, worst (25th), mean and standard deviation of the error values achieved are presented. From these tables, taking into account the best error achieved when FES is 10^4, it can be ob-

FES		Alg.	1	2	3	4	5	6	7	8
10^3	1st	ES	1.7949E+03	4.5853E+03	1.0631E+07	9.3638E+03	7.4491E+03	1.5571E+07	1.0294E+02	2.0506E+01
		PLES	**9.1088E+00**	**5.7120E+02**	**9.1815E+05**	**4.4155E+02**	**7.1613E+02**	**7.6823E+03**	**4.0755E+00**	**2.0325E+01**
	7th	ES	8.2397E+03	1.2766E+04	8.6963E+07	1.5645E+04	1.2194E+04	5.0154E+08	2.5575E+02	2.0677E+01
		PLES	**3.6238E+01**	**1.7565E+03**	**5.0438E+06**	**3.4986E+03**	**1.2012E+03**	**7.4683E+04**	**3.1421E+01**	**2.0660E+01**
	13th	ES	1.1797E+04	1.8060E+04	1.5969E+08	2.2094E+04	1.5973E+04	1.0638E+09	4.1709E+02	**2.0744E+01**
		PLES	**9.0533E+01**	**2.6395E+03**	**7.7964E+06**	**5.9841E+03**	**1.8286E+03**	**4.9335E+05**	**6.8927E+01**	2.0786E+01
	19th	ES	1.5289E+04	2.4144E+04	2.7509E+08	2.5466E+04	1.7969E+04	3.1175E+09	6.2194E+02	**2.0814E+01**
		PLES	**3.5851E+02**	**3.8877E+03**	**1.6207E+07**	**8.6644E+03**	**2.5582E+03**	**1.3357E+06**	**9.0168E+01**	2.0906E+01
	25th	ES	3.5099E+04	5.0376E+04	4.2652E+08	5.2091E+04	2.4149E+04	2.2225E+10	9.4772E+02	**2.0914E+01**
		PLES	**9.4706E+03**	**1.1286E+04**	**2.8051E+07**	**1.7558E+04**	**6.5111E+03**	**1.3441E+08**	**3.1915E+02**	2.1077E+01
	mean	ES	1.2605E+04	1.9882E+04	1.7786E+08	2.2299E+04	1.5581E+04	2.5262E+09	4.4833E+02	**2.0738E+01**
		PLES	**6.1405E+02**	**3.2330E+03**	**1.0545E+07**	**6.6159E+03**	**2.2423E+03**	**8.2295E+06**	**7.5778E+01**	2.0762E+01
	std	ES	7.0079E+03	1.0686E+04	1.2536E+08	8.9026E+03	4.5958E+03	4.3882E+09	2.2861E+02	**9.8407E-02**
		PLES	**1.8803E+03**	**2.3321E+03**	**7.4163E+06**	**4.0028E+03**	**1.5144E+03**	**2.8481E+07**	**6.5894E+01**	1.8862E-01
10^4	1st	ES	1.7871E-06	2.5883E+01	1.6785E+05	**1.8235E+02**	3.2371E+02	7.4254E+00	3.3259E-01	2.0353E+01
		PLES	**4.0793E-09T**	**2.7034E-01**	**1.3234E+05**	2.7380E+02	**7.1163E-01**	**5.6220E-01**	**2.6615E-01**	**2.0052E+01**
	7th	ES	1.3839E-02	4.5342E+02	7.6943E+05	3.8105E+03	2.3440E+03	5.2294E+01	**7.9958E-01**	2.0452E+01
		PLES	**7.8396E-09T**	**3.1583E+00**	**4.5552E+05**	**3.3190E+03**	**6.7570E+01**	**7.7440E+00**	1.7361E+00	**2.0307E+01**
	13th	ES	3.8002E-01	1.0044E+03	3.2908E+06	7.2496E+03	4.3654E+03	4.0633E+02	**1.2580E+00**	2.0546E+01
		PLES	**8.5419E-09T**	**6.7311E+00**	**7.8938E+05**	**5.3960E+03**	**4.0588E+02**	**8.8934E+01**	2.3759E+00	**2.0497E+01**
	19th	ES	2.6012E+00	2.3194E+03	6.2816E+06	1.2039E+04	5.6344E+03	2.6639E+03	**2.2078E+00**	**2.0606E+01**
		PLES	**9.4716E-09T**	**2.3352E+01**	**1.7016E+06**	**8.5893E+03**	**1.6028E+03**	**3.3158E+02**	5.9660E+00	2.0801E+01
	25th	ES	2.5015E+03	1.3137E+04	6.9077E+07	2.1715E+04	1.3608E+04	1.4355E+05	**5.1871E+00**	**2.0648E+01**
		PLES	**9.9332E-09T**	**1.0480E+02**	**4.6729E+06**	**1.7484E+04**	**5.4705E+03**	**5.5121E+03**	1.7863E+01	2.1040E+01
	mean	ES	1.4741E+02	2.2683E+03	6.7116E+06	8.3023E+03	4.5769E+03	9.3256E+03	**1.7570E+00**	2.0528E+01
		PLES	**8.4020E-09T**	**2.0175E+01**	**1.1660E+06**	**6.0358E+03**	**9.3334E+02**	**8.8821E+02**	4.1339E+00	**2.0519E+01**
	std	ES	5.1400E+02	3.3490E+03	1.3654E+07	5.5439E+03	3.6313E+03	2.9937E+04	**1.3817E+00**	**8.9628E-02**
		PLES	**1.4061E-09T**	**2.7234E+01**	**1.1043E+06**	**3.8616E+03**	**1.2361E+03**	**1.7524E+03**	4.3759E+00	2.9356E-01

Table 2: Error values achieved for problems 1 to 8 ($n = 10$)

served that PLES performed better than $(\mu/\rho + \lambda)$-ES in almost all problems, except for problems 4, 13, 14, 22 and 25. It should be noted that, in some of these problems, PLES performed better than $(\mu/\rho + \lambda)$-ES initially (when FES is 10^3) but $(\mu/\rho + \lambda)$-ES outperformed PLES and achieved an inferior error value for greater values of FES. This fact can be explained by the premature loss of diversity that possibly occurs in population due to the step size adaptation rule implemented in PLES. On the other hand, it is clear the lower FES required be PLES to obtain the same accuracy levels of $(\mu/\rho + \lambda)$-ES.

5 Conclusion and Future Work

In this paper, a new Parameter-less Evolution Strategy for global optimization was presented. This approach incorporates the main features of traditional single objective Evolution Strategies, like real representation of the decision variables and self-adaptation of step sizes, but the initial values for parameters are set by the algorithm itself.

One of the main advantages of this algorithm is that relieves the user from having to set the initial values of the parameters. Usually, this setting requires some previous experimentation in order to choose the better values for the problem being solved. This experimentation is often time consuming and must take into account the interactions between parameters. Thus, since this task is difficult, it is important to reduce the number of parameters to set and simplify the usage of algorithms.

The proposed step size adaptation rule takes into account the success of parents to generate offspring. The distances between parents and offsprings are used to estimate suited step sizes. This rule combined with the one-phase recombination/mutation operator allows the achievement of performance of PLES superior to $(\mu/\rho + \lambda)$-ES on some of the test problems considered. However, more experimentation should be carried out in order to identify the strengths and drawbacks of the new approach. Generally, taking into account the results obtained, with the PLES the probability of to observe poor performance in a problem due to poor parameter settings is reduced.

Future work will concentrate on the study of the adaptation step size rule and the development of a recombination operator based on more than two parents. Moreover, since the non-isotropic step size rule may be extend to incorporate correlations between decision variables, some investigation will be carried in this subject.

Bibliography

[1] Goldberg, D.E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Massachusetts, 1989.

[2] Rechenberg, I., *Evolutionsstrategie ’94*, Frommann-Holzboog, Stuttgart, 1994.

[3] Schwefel, H.,-P, *Evolution and Optimum Seeking*, Wiley, New York, 1995.

[4] Costa, L. and Oliveira, P., Evolutionary algorithms approach to the solution of mixed integer non-linear programming problems, *Computers Chem. Engng.*, **25**, 2001, pp. 257–266.

[5] Bäck, T., *Evolutionary Algorithms in Theory and Practice*, Oxford University Press, Inc., NY, 1996.

FES		Alg.	9	10	11	12	13	14	15	16
	1st	ES	2.7057E+01	5.4136E+01	8.3729E+00	2.1072E+04	1.1915E+02	3.5590E+00	2.9790E+02	2.5069E+02
		PLES	**7.5067E+00**	**2.1358E+01**	**5.2205E+00**	**7.4932E+02**	**2.8158E+00**	**3.2407E+00**	**1.9347E+02**	**1.2240E+02**
	7th	ES	4.1113E+01	9.7312E+01	1.0585E+01	3.3877E+04	7.1901E+02	4.0491E+00	5.9457E+02	3.7744E+02
		PLES	**1.9733E+01**	**2.7741E+01**	**8.6920E+00**	**5.4597E+03**	**5.7107E+00**	**4.0123E+00**	**4.3607E+02**	**1.5615E+02**
	13th	ES	5.3037E+01	1.2001E+02	1.1564E+01	5.9259E+04	5.9269E+03	**4.2013E+00**	7.9756E+02	5.4425E+02
		PLES	**2.2108E+01**	**3.6522E+01**	**1.0305E+01**	**8.9372E+03**	**1.1359E+01**	4.2040E+00	**5.1719E+02**	**1.8247E+02**
10^3	19th	ES	6.5459E+01	1.3385E+02	1.2434E+01	9.0079E+04	2.2900E+04	**4.3105E+00**	9.2507E+02	6.1169E+02
		PLES	**2.8346E+01**	**4.4073E+01**	**1.1447E+01**	**1.8608E+04**	**1.8956E+01**	4.4791E+00	**5.6763E+02**	**2.1207E+02**
	25th	ES	1.0832E+02	1.9204E+02	**1.3654E+01**	2.5920E+05	9.6909E+04	**4.5187E+00**	1.2791E+03	8.9532E+02
		PLES	**6.6595E+01**	**5.4128E+01**	1.4147E+01	**6.0309E+04**	**2.7360E+02**	4.6975E+00	**6.1238E+02**	**2.3645E+02**
	mean	ES	5.6229E+01	1.1842E+02	1.1464E+01	7.3634E+04	1.7175E+04	**4.1490E+00**	7.4828E+02	5.3178E+02
		PLES	**2.6284E+01**	**3.6699E+01**	**9.9390E+00**	**1.3773E+04**	**3.0083E+01**	4.1802E+00	**4.7751E+02**	**1.8163E+02**
	std	ES	2.0323E+01	3.5049E+01	**1.4101E+00**	5.3413E+04	2.5349E+04	**2.4737E-01**	2.5661E+02	1.6404E+02
		PLES	**1.1958E+01**	**1.0306E+01**	2.1465E+00	**1.2994E+04**	**5.7279E+01**	3.3922E-01	**1.2424E+02**	**3.3074E+01**
	1st	ES	1.6914E+01	2.9849E+01	5.1131E+00	3.2862E+01	**6.9759E-01**	**2.6333E+00**	2.6209E+02	2.2160E+02
		PLES	**3.9798E+00**	**1.1940E+01**	**4.5567E+00**	**1.2109E+01**	8.6153E-01	3.2407E+00	**1.2129E+02**	**9.7223E+01**
	7th	ES	2.8854E+01	8.1724E+01	**7.9372E+00**	2.4425E+03	**2.4600E+00**	**3.5326E+00**	4.4027E+02	3.2551E+02
		PLES	**1.2259E+01**	**1.6914E+01**	8.1411E+00	**1.9924E+02**	2.4970E+00	3.9792E+00	**2.9810E+02**	**1.2610E+02**
	13th	ES	3.6813E+01	1.0248E+02	**9.1460E+00**	4.4323E+03	**3.7439E+00**	**3.8637E+00**	5.2672E+02	4.2561E+02
		PLES	**1.5919E+01**	**2.2884E+01**	9.4840E+00	**1.9340E+03**	4.2277E+00	4.0758E+00	**4.2669E+02**	**1.3641E+02**
10^4	19th	ES	5.2733E+01	1.2337E+02	**9.9777E+00**	1.1396E+04	**6.5318E+00**	**4.0127E+00**	7.2433E+02	5.0851E+02
		PLES	**1.9899E+01**	**3.2834E+01**	1.1447E+01	**4.3828E+03**	8.2262E+00	4.4688E+00	**4.5424E+02**	**1.6652E+02**
	25th	ES	1.0546E+02	1.8307E+02	1.1917E+01	4.6248E+04	1.0236E+02	**4.3405E+00**	9.6233E+02	7.4317E+02
		PLES	**4.2783E+01**	**4.8753E+01**	**1.4147E+01**	**1.6493E+04**	**7.8681E+01**	4.6975E+00	**5.2299E+02**	**2.0510E+02**
	mean	ES	4.4853E+01	1.0304E+02	**9.0910E+00**	8.7686E+03	1.0852E+01	**3.7017E+00**	5.8319E+02	4.3871E+02
		PLES	**1.6728E+01**	**2.5630E+01**	9.5264E+00	**4.1466E+03**	**9.0290E+00**	4.1392E+00	**3.7977E+02**	**1.4671E+02**
	std	ES	2.3113E+01	3.5427E+01	**1.7754E+00**	1.0182E+04	2.2368E+01	4.2005E-01	2.2561E+02	1.3786E+02
		PLES	**7.7682E+00**	**1.0305E+01**	2.3612E+00	**5.2894E+03**	**1.5951E+01**	**3.5645E-01**	**1.0553E+02**	**2.9549E+01**

Table 4: Error values achieved for problems 9 to 16 ($n = 10$)

FES		Alg.	17	18	19	20	21	22	23	24	25
	1st	ES	2.8553E+02	1.1067E+03	1.1065E+03	1.0789E+03	1.2987E+03	1.0406E+03	1.3113E+03	1.3257E+03	4.5353E+02
		PLES	**1.3854E+02**	**8.3929E+02**	**7.9570E+02**	**7.9570E+02**	**6.0682E+02**	**7.8899E+02**	**5.5947E+02**	**2.2092E+02**	**4.5296E+02**
	7th	ES	4.1902E+02	1.1938E+03	1.2279E+03	1.2348E+03	1.3970E+03	1.1422E+03	1.4071E+03	1.3791E+03	5.1444E+02
		PLES	**1.7866E+02**	**9.8381E+02**	**9.9097E+02**	**9.8344E+02**	**1.1433E+03**	**8.4256E+02**	**1.1974E+03**	**2.9861E+02**	**4.9868E+02**
	13th	ES	5.0622E+02	1.2963E+03	1.3001E+03	1.3001E+03	1.4450E+03	1.2477E+03	1.4396E+03	1.4216E+03	5.7657E+02
		PLES	**2.0330E+02**	**1.0446E+03**	**1.0455E+03**	**1.0350E+03**	**1.2352E+03**	**8.8961E+02**	**1.2651E+03**	**3.8930E+02**	**5.4300E+02**
10^3	19th	ES	7.6071E+02	1.3627E+03	1.3930E+03	1.3930E+03	1.4924E+03	1.6047E+03	1.4865E+03	1.4521E+03	**7.2705E+02**
		PLES	**2.5274E+02**	**1.0606E+03**	**1.0652E+03**	**1.0513E+03**	**1.2758E+03**	**9.4220E+02**	**1.2836E+03**	**5.9341E+02**	1.0567E+03
	25th	ES	1.1421E+03	1.5185E+03	1.6231E+03	1.6231E+03	1.5409E+03	1.9754E+03	1.5901E+03	1.5384E+03	**1.3341E+03**
		PLES	**3.3778E+02**	**1.1937E+03**	**1.1417E+03**	**1.1417E+03**	**1.2983E+03**	**9.9040E+02**	**1.3357E+03**	**1.3003E+03**	1.3803E+03
	mean	ES	5.8144E+02	1.2900E+03	1.3122E+03	1.3100E+03	1.4345E+03	1.3603E+03	1.4455E+03	1.4174E+03	**6.6478E+02**
		PLES	**2.1565E+02**	**1.0259E+03**	**1.0178E+03**	**1.0120E+03**	**1.1459E+03**	**8.9032E+02**	**1.1540E+03**	**4.6362E+02**	7.5600E+02
	std	ES	2.2707E+02	1.1529E+02	1.2561E+02	1.2987E+02	**7.0248E+01**	2.7670E+02	**6.9522E+01**	**5.6030E+01**	**2.2647E+02**
		PLES	**5.2522E+01**	**7.5220E+01**	**8.4824E+01**	**8.3478E+01**	2.0769E+02	**6.1677E+01**	2.3241E+02	2.5132E+02	3.4558E+02
	1st	ES	2.3407E+02	1.0304E+03	7.8285E+02	7.8285E+02	1.0754E+03	**3.0107E+02**	1.2411E+03	**2.0000E+02**	**4.1000E+02**
		PLES	**1.2007E+02**	**8.0001E+02**	**7.8089E+02**	**7.8089E+02**	**5.0000E+02**	7.8274E+02	**5.5947E+02**	**2.0000E+02**	4.3025E+02
	7th	ES	3.1683E+02	1.0751E+03	1.0764E+03	1.0755E+03	1.2677E+03	8.8553E+02	1.2819E+03	1.2963E+03	**4.1263E+02**
		PLES	**1.6083E+02**	**9.8164E+02**	**9.8336E+02**	**9.7745E+02**	**1.0287E+03**	**8.2676E+02**	**1.1404E+03**	**2.0000E+02**	4.4710E+02
	13th	ES	4.4399E+02	1.1166E+03	1.1303E+03	1.1303E+03	1.3255E+03	1.0522E+03	1.3546E+03	1.3354E+03	**4.1489E+02**
		PLES	**1.8788E+02**	**1.0333E+03**	**1.0286E+03**	**1.0279E+03**	**1.2020E+03**	**8.8961E+02**	**1.2650E+03**	**2.0000E+02**	4.6860E+02
10^4	19th	ES	5.5663E+02	1.1796E+03	1.1745E+03	1.1716E+03	1.3831E+03	1.1096E+03	1.4219E+03	1.3856E+03	**4.1997E+02**
		PLES	**2.1604E+02**	**1.0538E+03**	**1.0493E+03**	**1.0456E+03**	**1.2704E+03**	**9.4220E+02**	**1.2836E+03**	**2.0008E+02**	1.0547E+03
	25th	ES	9.3866E+02	1.3509E+03	1.3707E+03	1.3708E+03	1.4963E+03	1.9342E+03	1.5624E+03	1.5179E+03	**4.6031E+02**
		PLES	**3.2365E+02**	**1.1916E+03**	**1.1294E+03**	**1.1053E+03**	**1.2983E+03**	**9.9040E+02**	**1.3217E+03**	**1.2942E+03**	1.3714E+03
	mean	ES	4.6353E+02	1.1476E+03	1.1277E+03	1.1281E+03	1.3209E+03	1.0203E+03	1.3539E+03	1.2125E+03	**4.1841E+02**
		PLES	**1.9592E+02**	**1.0148E+03**	**1.0019E+03**	**9.9894E+02**	**1.0794E+03**	**8.8049E+02**	**1.1141E+03**	**2.8238E+02**	6.9232E+02
	std	ES	1.7297E+02	9.4503E+01	1.1569E+02	1.1491E+02	**1.0333E+02**	3.1895E+02	**8.2235E+01**	3.6895E+02	**1.0072E+01**
		PLES	**5.1283E+01**	**8.2949E+01**	**8.8570E+01**	**8.7189E+01**	2.8000E+02	**6.4666E+01**	2.7371E+02	**2.3622E+02**	3.6234E+02

Table 5: Error values achieved for problems 17 to 25 ($n = 10$)