Modifications for the Differential Evolution algorithm

Vasileios Charilogis, Ioannis G. Tsoulos, Alexandros Tzallas, Evangelos Karvounis

Department of Informatics and Telecommunications, University of Ioannina, Greece

Abstract

Differential Evolution (DE) is a method of optimization used in symmetrical optimization problems but also in problems that are not even continuous, are noisy and changed over time. DE optimizes a problem with a population of candidate solutions and creating new candidate solutions per generation in combination with existing rules according to discriminatory rules. The present work proposes two variations for this method. The first significantly improves the termination of the method by proposing an asymptotic termination rule, which is based on the differentiation of the average of the function values in the population of DE. The second modification proposes a new scheme for a critical parameter of the method, which scheme improves the method's ability to better explore the search space of the objective function. The proposed variations have been tested on a number of problems from the current literature and from the experimental results it appears that the proposed modifications render the method quite robust and faster even in large-scale problems.

1 Introduction

The location of the global minimum of a continuous and differentiable function $f : S \to R, S \subset R^n$ is formulated as

$$x^* = \arg\min_{x \in S} f(x) \tag{1}$$

where the set S is defined as

$$S = [a_1, b_1] \otimes [a_2, b_2] \otimes \ldots [a_n, b_n]$$

In the recent literature there are a plethora of real world problems that can formulated as global optimization problems such as problems from the physics area [1, 2, 3, 4], chemistry [5, 6, 7], economics [8, 9] etc. There are a variety of proposed methods to handle the global minimum problem such as Adaptive

Random Search [10], Competitive Evolution [11], Controlled Random Search [12], Simulated Annealing [13, 14, 16], Genetic Algorithms [17, 18], Bee optimization [19, 20], Ant Colony Optimization [21], Particle Swarm Optimization [23], Differential Evolution [24] etc. Recently many works have been appeared that take advantage of the GPU processing units to implement parallel global optimization methods [25, 26, 27]. This work introduces two major modifications for the Differential Evolution (DE) method, that aim to speed up the algorithm and to reduce the total number of function evaluations required by the method. The DE method initially creates a population of candidate solutions and through a series of iterations creates new solutions by combining the previous ones. The method does not require any prior knowledge of the derivative and is therefore quite fast and has low memory requirements.

After a literature review it was found that differential evolution is used in several areas and many modifications of the original algorithm have been introduced in the recent literature. More specifically, in the research of Zongjun et al [28], genetic and differential calculus algorithms were used to optimize the parameters of two models aimed at estimating evapotranspiration in three regions and it was found that the performance of evolution algorithms was better than the genetic algorithm. Another research focuses on a case study of a cellular neural network aimed at generating fractional classes of neurons. The best solutions provided by differential calculus and the use of accelerated particle swarm optimization (APSO) are presented concretely in the work of Tlelo-Cuautle et al [29]. Another article [30] proposes a regeneration framework based on space search adaptation (ARSA), which can be integrated into different variants of different evolution to address the problems of early convergence and population stability faced by different calculus. Also another interesting variation of the method is the Bernstain Search Differential Evolution Algorithm [31] for optimizing numerical functions.

The differential evolution method was applied also to energy science. Specifically, the article of Liang et al [32] evaluates the parameters of solar photovoltaic models through a self-adjusting differential evolution. Similarly, in the study of Peng et al [33], differential evolution is used for the prediction of electricity prices. Also, differential evolution was also incorporated in neural architecture search [34]. The DE method has been applied with success to neural network training [35, 36, 37], to the Traveling Salesman Problem [38, 39], training of RBF neural networks [40, 41, 42], optimization of the Lennard Jones potential [43, 44]. Also the DE method has been successfully combined with other techniques for machine learning applications such as classification [45, 46], feature selection [47, 48], deep learning [49, 50] etc.

The rest of this article is organized as follows: in section 2 the base DE algorithm as well as the proposed modifications are presented clearly, in section 3 the test functions used in the experiments are presented accompanied with the experimental results and finally in section 4 some conclusions are presented.

2 Modifications

This section starts with the detailed description of the DE method and continues with the modifications suggested in this article. The first modification is a new stopping rule which measures the difference of the mean of function values between the iterations of the algorithm and the second modification suggests a new scheme for a critical parameter of the DE algorithm called Differential Weight.

2.1 The base algorithm

The DE algorithm has been by various researchers in the recent literature such as the Compact Differential Evolution [51], a self adaptive DE [52] where the critical parameters of the method are adapted from previous generations, a fuzzy adaptive DE method [53] where fuzzy login is employed to adapt the parameters of the method, parallel Differential Evolution [54] with a self adaptation mechanism for the critical parameters of the DE method etc. A survey of the recent advances in differential evolution can be found in the work of Das et al [55]. The base DE algorithm has the steps described below

1. **Set** the population size NP ≥ 4, usually NP $= 10n$, where n is the dimension of the input problem.
2. **Set** the crossover probability CR $\in [0, 1]$. A typical value for this parameter is 0.9
3. **Set** the differential weight $F \in [0, 2]$. A typical value for this parameter is 0.8
4. **Initialize** all members of the population in the search space. The members of the population are called agents.
5. **Until** some stopping criterion is met repeat
   (a) **For** $i = 1 \ldots \mathrm{NP}$ **do**
      i. **Set** x as the agent i.
      ii. **Pick** randomly three agents a, b, c
      iii. **Pick** a random index $R \in \{1, \ldots, n\}$
      iv. **Compute** the trial vector $y = [y_1, y_2, \ldots, y_n]$ as follows
      v. **For** $j = 1, \ldots, n$ **do**
         A. **Set** $r_i \in [0, 1]$ a random number.
         B. **If** $r_j <$ CR **or** $j = R$ **then** $y_j = a_j + F \times (b_j - c_j)$ **else** $y_j = x_j$.
      vi. If $f(y) \leq f(x)$ then $x = y$.
      vii. **EndFor**
   (b) **EndFor**
6. **Return** the agent x_{best} in the population with the lower function value $f(x_{\mathrm{best}})$.

2.2 The new termination rule

Typically the DE method is terminated when a predefined number of iterations is reached. This can be extremely inefficient in some problems and in others it can lead to premature termination, ie termination before the total minimum is found. In the work of Ali et al [56] a different termination rule is proposed ie. terminate when

$$f_{\max} - f_{\min} \leq \epsilon \tag{2}$$

where $f_{\max}$ is the function value of the worst agent in the population, $f_{\min}$ is the function value of the best agent and $\epsilon$is a small positive number.

In the proposed termination rule, the average function value of the population is calculated in each iteration. If this value does not change significantly for a repetitive number of iterations, then it is very likely that the method may not discover a new global minimum and should therefore be terminated. Hence in every generation t we measure the quantity:

$$\delta^{(t)} = \left| \sum_{i=1}^{\mathrm{NP}} \left| f_i^{(t)} \right| - \sum_{i=1}^{\mathrm{NP}} \left| f_i^{(t-1)} \right| \right| \tag{3}$$

and the termination rule is defined as: **terminate** if $\delta^{(t)} \leq \epsilon$ for a predefined number of M generations.

2.3 The new differential weight

The differential weight initially proposed in the DE algorithm was a static value, which means that some tuning is required in order to discover the global minimum in every optimization function. Ali et al [56] proposed an adaptation mechanism for this parameter in order that the algorithm should search in larger spaces at the the first generations and to become more focused at latter generations. The mechanism proposed is expressed as:

$$F = \begin{cases} \max\left(l_{\min}, 1 - \left|\frac{f_{\max}}{f_{\min}}\right|\right) & , \quad \text{if} \quad \left|\frac{f_{\max}}{f_{\min}}\right| \leq 1 \\ \max\left(l_{\min}, 1 - \left|\frac{f_{\min}}{f_{\max}}\right|\right) & , \quad \text{otherwise} \end{cases} \tag{4}$$

The current work proposes a stochastic mechanism similar to the crossover operation of the Genetic algorithms. The proposed scheme is expressed as:

$$F = -\frac{1}{2} + 2 \times R \tag{5}$$

where $R \in [0, 1]$ is a random number. The proposed scheme, as with the randomness introduced by the method DE will be able to better explore the search space of the objective function and find with greater accuracy and speed the global minimum. In addition, this scheme has been used successfully in Genetic Algorithms.

3 Experiments

In order to determine the effectiveness of the proposed modifications, a series of experiments were performed on known functions from the relevant literature [57, 58]. The choice of these functions was made as they are widely used in the literature by many researchers [59, 60, 61, 62], they have quite a complex structure and in many cases have a large number of dimensions that make them ideal for study and testing.

The experiments were divided into two major categories. In the first category all the schemes for the Differential Weight were tested using the termination rule of Equation 2 and in the second category the same schemes were tested using the proposed termination criterion. Also, after every successful termination the local optimization method BFGS [63] was applied in order to get even closer to the global minimum.

3.1 Test functions

The description of the test functions used in the experiments has as follows:

- **Bf1** (Bohachevsky 1) function defined as:

$$f(x) = x_1^2 + 2x_2^2 - \frac{3}{10}\cos(3\pi x_1) - \frac{4}{10}\cos(4\pi x_2) + \frac{7}{10}$$

with $x \in [-100, 100]^2$. The value of global minimum is 0.0.

- **Bf2** (Bohachevsky 2) function defined as:

$$f(x) = x_1^2 + 2x_2^2 - \frac{3}{10}\cos(3\pi x_1)\cos(4\pi x_2) + \frac{3}{10}$$

with $x \in [-50, 50]^2$. The value of the global minimum is 0.0.

- **Branin** function. The function is defined by $f(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10$ with $-5 \leq x_1 \leq 10,\ 0 \leq x_2 \leq 15$. The value of global minimum is 0.397887.with $x \in [-10, 10]^2$. The value of global minimum is -0.352386.

- **CM** function. The Cosine Mixture function is given by the equation

$$f(x) = \sum_{i=1}^{n} x_i^2 - \frac{1}{10}\sum_{i=1}^{n}\cos(5\pi x_i)$$

where $x \in [-1, 1]^n$. For our experiments we used $n = 4$.

- **Camel** function. The function is given by

$$f(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4, \quad x \in [-5, 5]^2$$

- **Easom** function. The function is given by the equation

$$f(x) = -\cos(x_1)\cos(x_2)\exp\left((x_2-\pi)^2-(x_1-\pi)^2\right)$$

with $x \in [-100,100]^2$.

- **Exponential** function, defined as:

$$f(x) = -\exp\left(-0.5\sum_{i=1}^{n} x_i^2\right), \quad -1 \leq x_i \leq 1$$

The global minimum is located at $x^* = (0,0,...,0)$ with value -1. In our experiments we used this function with $n = 2,4,8,16,32$.

- **Goldstein and Price function**
The function is given by the equation

$$\begin{aligned} f(x) \quad = \quad & \Big[1+(x_1+x_2+1)^2 \\ & \left(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2\right)] \times \\ & [30+(2x_1-3x_2)^2 \\ & \left(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2\right)] \end{aligned}$$

With $x \in [-2,2]^2$. The global minimum is located at $x^* = (0,-1)$ with value 3.0

- **Griewank2** function. The function is given by

$$f(x) = 1 + \frac{1}{200}\sum_{i=1}^{2} x_i^2 - \prod_{i=1}^{2}\frac{\cos(x_i)}{\sqrt{(i)}}, \quad x \in [-100,100]^2$$

The global minimum is located at the $x^* = (0,0,...,0)$ with value 0.

- **Gkls** function. $f(x) = \text{Gkls}(x,n,w)$, is a function with w local minima, described in [64] with $x \in [-1,1]^n$ and n a positive integer between 2 and 100. The value of the global minimum is -1 and in our experiments we have used $n = 2,3$ and $w = 50,\ 100$.

- **Hansen** function. $f(x) = \sum_{i=1}^{5} i\cos[(i-1)x_1+i]\sum_{j=1}^{5} j\cos[(j+1)x_2+j]$, $x \in [-10,10]^2$.

- **Hartman 3** function. The function is given by

$$f(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{3} a_{ij}\left(x_j - p_{ij}\right)^2\right)$$

with $x \in [0,1]^3$ and $a = \begin{pmatrix} 3 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3 & 10 & 30 \\ 0.1 & 10 & 35 \end{pmatrix}$, $c = \begin{pmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{pmatrix}$ and

$$p = \begin{pmatrix} 0.3689 & 0.117 & 0.2673 \\ 0.4699 & 0.4387 & 0.747 \\ 0.1091 & 0.8732 & 0.5547 \\ 0.03815 & 0.5743 & 0.8828 \end{pmatrix}$$

- **Hartman 6** function.

$$f(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{6} a_{ij}\left(x_j - p_{ij}\right)^2\right)$$

with $x \in [0,1]^6$ and $a = \begin{pmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{pmatrix}$, $c = \begin{pmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{pmatrix}$

and

$$p = \begin{pmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4135 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6650 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{pmatrix}$$

- **Potential** function. The molecular conformation corresponding to the global minimum of the energy of N atoms interacting via the Lennard-Jones potential[65] is used as a test case here. The function to be minimized is given by:

$$V_{LJ}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{6}$$

In the current experiments three different cases were studied: $N = 3,\ 4,\ 5$

- **Rastrigin** function. The function is given by

$$f(x) = x_1^2 + x_2^2 - \cos(18x_1) - \cos(18x_2), \quad x \in [-1,1]^2$$

The global minimum is located at $x^* = (0,0)$ with value -2.0.

- ***Rosenbrock*** *function.*
This function is given by

$$f(x) = \sum_{i=1}^{n-1}\left(100\left(x_{i+1} - x_i^2\right)^2 + \left(x_i - 1\right)^2\right), \quad -30 \leq x_i \leq 30.$$

The global minimum is located at the $x^* = (0,0,...,0)$ with $f(x^*) = 0$. In our experiments we used this function with $n = 4,\ 8,\ 16$.

- **Shekel 7** function.

$$f(x) = -\sum_{i=1}^{7} \frac{1}{(x-a_i)(x-a_i)^T + c_i}$$

with $x \in [0,10]^4$ and $a = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \\ 2 & 9 & 2 & 9 \\ 5 & 3 & 5 & 3 \end{pmatrix}$, $c = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.4 \\ 0.4 \\ 0.6 \\ 0.3 \end{pmatrix}$. The value of global minimum is -10.342378.

- **Shekel 5** function.

$$f(x) = -\sum_{i=1}^{5} \frac{1}{(x-a_i)(x-a_i)^T + c_i}$$

with $x \in [0,10]^4$ and $a = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \end{pmatrix}$, $c = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix}$. The value of global minimum is -10.107749.

- **Shekel 10** function.

$$f(x) = -\sum_{i=1}^{10} \frac{1}{(x-a_i)(x-a_i)^T + c_i}$$

with $x \in [0,10]^4$ and $a = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \\ 2 & 9 & 2 & 9 \\ 5 & 5 & 3 & 3 \\ 8 & 1 & 8 & 1 \\ 6 & 2 & 6 & 2 \\ 7 & 3.6 & 7 & 3.6 \end{pmatrix}$, $c = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.4 \\ 0.4 \\ 0.6 \\ 0.3 \\ 0.7 \\ 0.5 \\ 0.6 \end{pmatrix}$. The value of global minimum is -10.536410.

- **Sinusoidal** function. The function is given by

$$f(x) = -\left(2.5 \prod_{i=1}^{n} \sin(x_i - z) + \prod_{i=1}^{n} \sin(5(x_i - z))\right), \quad 0 \le x_i \le \pi.$$

The global minimum is located at $x^* = (2.09435, 2.09435, ..., 2.09435)$ with $f(x^*) = -3.5$. In our experiments we used $n = 4, 8, 16, 32$ and $z = \frac{\pi}{6}$ and the corresponding functions are denoted by the labels SINU4, SINU8, SINU16 and SINU32 respectively.

- **Test2N** function. This function is given by the equation

$$f(x) = \frac{1}{2}\sum_{i=1}^{n} x_i^4 - 16x_i^2 + 5x_i, \quad x_i \in [-5, 5].$$

The function has 2^n in the specified range and in our experiments we used $n = 4, 5, 6, 7$. The corresponding values of global minimum is -156.664663 for $n = 4$, -195.830829 for $n = 5$, -234.996994 for $n = 6$ and -274.163160 for $n = 7$.

- **Test30N** function. This function is given by

$$f(x) = \frac{1}{10}\sin^2(3\pi x_1)\sum_{i=2}^{n-1}\Big((x_i-1)^2\left(1+\sin^2(3\pi x_{i+1})\right)\Big)+(x_n-1)^2\left(1+\sin^2(2\pi x_n)\right)$$

with $x \in [-10, 10]$, with 30^n local minima in the searc space. For our experiments we used $n = 3, 4$.

3.2 Experimental results

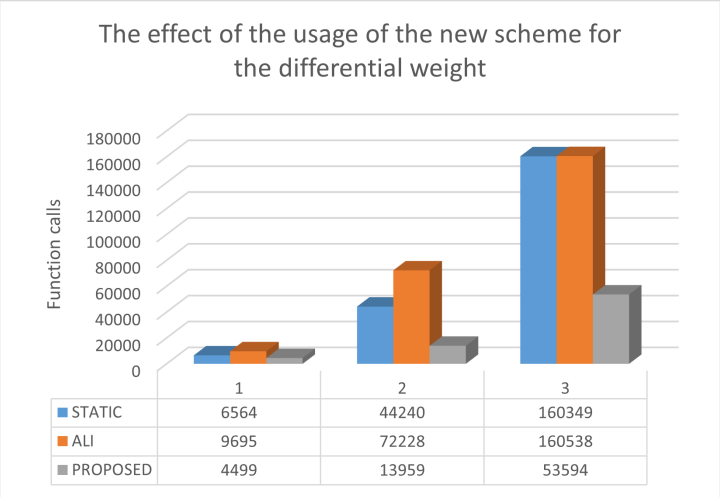
The experiments have been performed 30 times with different seed for the random generator each time for every test function and the avarage function calls is measured and reported. The code has been implemented in ANSI C++ and the random generator used was the function drand48() of the C programming languages. The execution environment was an Intel Xeon E5-2630 multi core machine. The parameters used in the experiments are shown in Table 1. The experiments where the stopping rule of Equation 2 was used are outlined in Table 2 and the experiments with the proposed stopping rule are listed in Table 3. The numbers in cells represent average function calls. The fraction in parentheses stands for the fraction of runs where the global optimum was found. If this number is missing then the global minimum was discovered in every independent run (100% success). The column STATIC represents the static value for the differential weight ($F = 0.8$), the column ALI stands for the mechanism given in the Equation 4 and lastly the column PROPOSED stands for the proposed scheme given in the Equation 5.

From the experiments performed we observe that the two proposed variations drastically reduce the required number of function calls. Also, the proposed changes do not seem to affect the average performance of the method, as it remains high in all cases. The effect of the proposed scheme for the differential weight is presented graphically in Figure 1, where we plot the average function calls for the functions ROSENBROCK4, ROSENBROCK8 and ROSENBROCK16 using the three schemes of the differential weight. Also, in the plot

Table 1: Experimental parameters.

PARAMETER	VALUE
NP	$10n$
F	0.8
CR	0.9
M	20
ϵ	10^{-4}

Figure 1: The effect of the usage of the new scheme for the differential weight.



of Figure 2, the average calls for the same functions are plotted using both the proposed scheme for the differential weight and the proposed termination rule. It is evident the the combination of both modifications reduced even more the average number of function calls required to locate the global minimum of the test functions. To show the effectiveness of the modifications in Figure 3 are presented the total times for 30 executions on an I7 computer with LINUX DEBIAN and with 16GB of memory. The comparison was made between the initial method with the ALI termination criterion, the proposed termination criterion (first modification) and the proposed termination criterion together with the proposed weight scheme. Obviously, the proposed modifications significantly reduce not only the number of calls but also the required execution time. To compare the proposed scheme for the differential weight with the other two methods, the Wilcoxon signed-rank test was used. The results obtained with this statistical test are shown in Figure 4

Table 2: Experiments with the termination rule of Ali.

FUNCTION	**STATIC**	**ALI**	**PROPOSED**
BF1	1142	1431	847
BF2	1164	1379	896
BRANIN	984	816	707
CM4	3590	7572	2079
CAMEL	1094	18849	685
EASOM	1707	2014	1327
EXP2	532	323	449
EXP4	2421	1019	1494
EXP8	15750	3670	5632
EXP16	160031	15150	21416
EXP32	320039	152548	77936
GKLS250	784	944	614
GKLS2100	772	1531	599(0.97)
GKLS350	1906(0.93)	3263	1275(0.93)
GKLS3100	1883	3539	1373
GOLDSTEIN	988	818	769
GRIEWANK2	1299(0.97)	1403	883(0.93)
HANSEN	2398	2968	1400
HARTMAN3	1448	836	1050
HARTMAN6	9489(0.97)	4015(0.97)	4667(0.80)
POTENTIAL3	90027	89776	21824
POTENTIAL4	120387(0.97)	120405(0.33)	45705(0.97)
POTENTIAL5	150073	150104	83342
RASTRIGIN	1246	1098(0.93)	871
ROSENBROCK4	6564	9695	4499
ROSENBROCK8	44240	72228	13959
ROSENBCROK16	160349(0.90)	160538(0.60)	53594
SHEKEL5	5524	3810	3057(0.83)
SHEKEL7	5266	3558	2992(0.87)
SHEKEL10	5319	3379	3076
TEST2N4	4200	1980	2592
TEST2N5	7357	2957	4055
TEST2N6	12074	4159	5836
TEST2N7	18872	5490	7904
SINU4	3270	1855	2216
SINU8	23108	6995	8135
SINU16	160092	36044	30943
SINU32	213757(0.70)	160536(0.53)	83369(0.80)
TEST30N3	1452	1732	959
TEST30N4	1917	2287	1378
TOTAL	1564515(0.97)	1062714(0.96)	506404(0.98)

Table 3: Experiments with the proposed termination rule.

FUNCTION	**STATIC**	**ALI**	**PROPOSED**
BF1	996	1124	889
BF2	926	1026	816
BRANIN	878	900	730
CM4	1148(0.70)	1991	1103
CAMEL	1049	904(0.93)	846
EASOM	447	448	446
EXP2	470	461	467
EXP4	915	903	892
EXP8	1797	3558	1796
EXP16	3578	7082	3521
EXP32	7082	14125	7022
GKLS250	498	576	493
GKLS2100	533	884(0.97)	515
GKLS350	823	1130(0.93)	814(0.97)
GKLS3100	858	1495(0.97)	829(0.93)
GOLDSTEIN	945	993	915
GRIEWANK2	947	921	826
HANSEN	2104	1949	1479
HARTMAN3	1017	1005	952
HARTMAN6	4679(0.90)	3744(0.97)	3128(0.87)
POTENTIAL3	21473	2284	8197
POTENTIAL4	44191(0.43)	3098(0.33)	24659(0.97)
POTENTIAL5	75910	3443	52664
RASTRIGIN	841	994	777
ROSENBROCK4	4934	7192	3300
ROSENBROCK8	29583	49696	10907
ROSENBCROK16	160349	160538(0.60)	38315
SHEKEL5	4389(0.97)	4266	2839(0.83)
SHEKEL7	3905	3685	2668
SHEKEL10	4049	3548	2629
TEST2N4	2785	2275	2221
TEST2N5	4481	3170	3122
TEST2N6	6852	4286	4296
TEST2N7	11971	5701	6267
SINU4	2322	1987	1755
SINU8	9990	6156	5113
SINU16	6892	3628(0.97)	16905
SINU32	7235(0.80)	7438(0.83)	7218
TEST30N3	1033	1098	951
TEST30N4	1355	1444	1285
TOTAL	432610(0.98)	321166(0.96)	224567(0.99)

Figure 2: Plot of function calls using the two modifications.

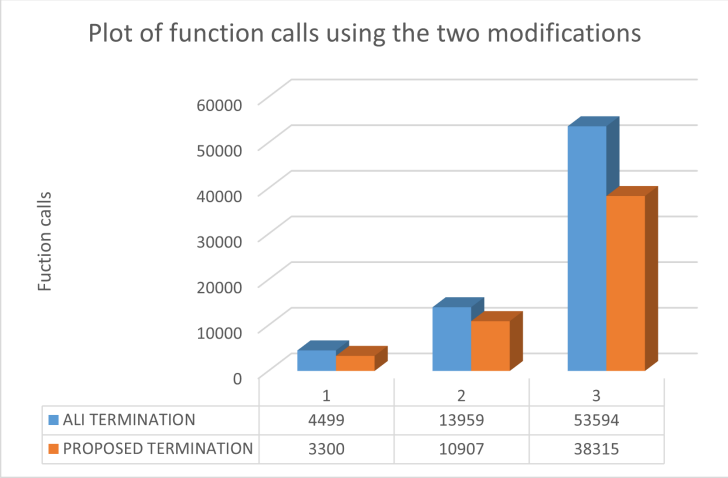


Figure 3: Time comparisons for a variety of test functions.

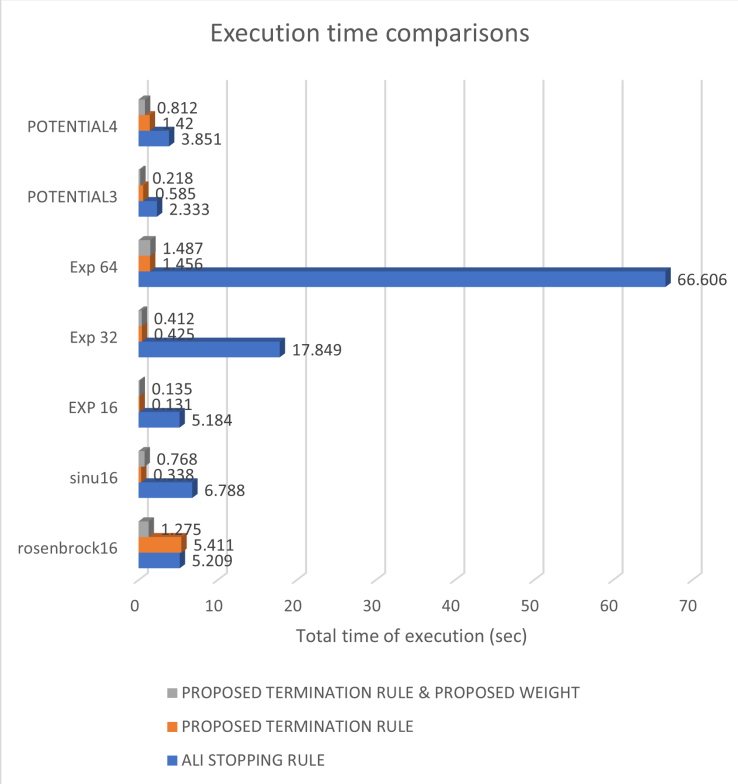
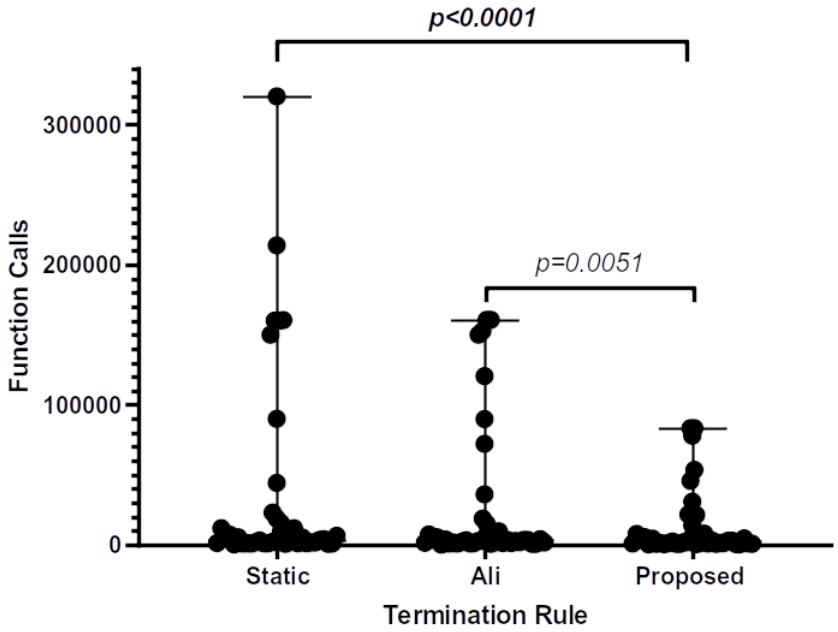


Figure 4: Box plot representation and Wilcoxon rank-sum test results of the comparison among the schemes for the differential weight. The stopping rule used is the proposed by Ali in equation 2. A p-value of less than 0.05 (2-tailed) was used for statistical significance and has been marked with bold.



4 Conclusions

In this text, two main additions to the DE method are presented. In the first an asymptotic termination rule was introduced and used successfully even in multidimensional problems. This rule is based on the observation that from one point onwards the average of the functional values of the agents does not change. This means that either the algorithm has already found the global minimum or its further continuation will have no meaning.

In the second case, a stochastic scheme was used to produce the differential weight. This scheme helped the algorithm to better explore the search space of the objective function without the need for more calls to the objective function.

The proposed modifications significantly speed up the original method in terms of function calls in most cases. Each of the proposed modifications can be used separately or together in the DE method. If used together there is a large reduction in the number of required function calls that reaches up to 80% without problems in the reliability of the method and its ability to successfully find the total minimum.

The DE method can also be used in cases of optimization problems with constraints as long as there is a change in the original problem so that the constraints are included in the function with the usage for example Langrage multipliers.

Compliance with Ethical Standards

All authors declare that they have no has no conflict of interest.

Acknowledgments

The experiments of this research work was performed at the high performance computing system established at Knowledge and Intelligent Computing Laboratory, Dept of Informatics and Telecommunications, University of Ioannina, acquired with the project "Educational Laboratory equipment of TEI of Epirus" with MIS 5007094 funded by the Operational Programme "Epirus" 2014-2020, by ERDF and national finds.

References

[1] Z.A. Kudyshev, A.V. Kildishev, V.M. Shalaev and A. Boltasseva, Machine learning–assisted global optimization of photonic devices, Nanophotonics **10**, pp. 371-383, 2021.

[2] X.L. Ding, Z.Y. Li, J.H. Meng, Y.X. Zhao and G.H. Sheng, Density-functional global optimization of $(LA_2O_3)_n$ clusters , J. Chem. Phys. **137**, 2012.

[3] S. Morita and N. Naoki, Global optimization of tensor renormalization group using the corner transfer matrix, Phys. Rev. B **103**, 2021.

[4] S. Heiles and R.L. Johnston, Global optimization of clusters using electronic structure methods, Int. J. Quantum Chem. **113**, pp. 2091-2109, 2013.

[5] Y. Yang, T. Pan and J. Zhang, Global Optimization of Norris Derivative Filtering with Application for Near-Infrared Analysis of Serum Urea Nitrogen, American Journal of Analytical Chemistry **10**, pp. 143-152, 2019.

[6] C. Grebner, J. Becker, D. Weber and B. Engels, Tabu search based global optimization algorithms for problems in computational chemistry, J Cheminform **4**, pp. 10, 2012.

[7] M. Dittner, J. Müller, H.M. Aktulga and B.J. Hartke, Efficient global optimization of reactive force-field parameters, Comput. Chem. **36**, pp. 1550–1561, 2015.

[8] W. Zhao, L. Wang and Z. Zhang, Supply-Demand-Based Optimization: A Novel Economics-Inspired Algorithm for Global Optimization, IEEE Access **7**, pp. 73182-73206, 2019.

[9] S.K. Mishra, Global Optimization of Some Difficult Benchmark Functions by Host-Parasite Co-Evolutionary Algorithm, Economics Bulletin **33**, pp. 1-18, 2013.

[10] H.A. Bremermann, A method for unconstrained global optimization, Mathematical Biosciences **9**, pp. 1-15, 1970.

[11] R.A. Jarvis, Adaptive global search by the process of competitive evolution, IEEE Trans. on Syst., Man and Cybergenetics **75**, pp. 297-311, 1975.

[12] W.L. Price, Global Optimization by Controlled Random Search, Computer Journal **20**, pp. 367-370, 1977.

[13] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science **220**, pp. 671-680, 1983.

[14] P. J. M. van Laarhoven, E. H. L. Aarts, Simulated Annealing: Theory and Applications, 1987, D. Riedel, Boston.

[15] A. Corana, M. Marchesi, C. Martini, S. Ridella, Minimizing Multimodal Functions of Continuous Variables with the Simulated Annealing Algorithm, ACM Transactions on Mathematical Software **13**, pp. 262-280, 1987.

[16] W.L. Goffe, G.D. Ferrier, J. Rogers, Global Optimization of Statistical Functions with Simulated Annealing, J. Econometrics **60**, pp. 65-100, 1994.

[17] D. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Publishing Company, Reading, Massachussets, 1989.

[18] Z. Michaelewicz, Genetic Algorithms + Data Structures = Evolution Programs. Springer - Verlag, Berlin, 1996.

[19] B. Akay and D. Karaboga, A modified Artificial Bee Colony algorithm for real-parameter optimization, Information Sciences **192**, pp. 120-142, 2012.

[20] G. Zhu and S. Kwong, Gbest-guided artificial bee colony algorithm for numerical function optimization, Applied Mathematics and Computation **217**, pp. 3166-3173, 2010.

[21] M. Dorigo, M. Birattari and T. Stutzle, Ant colony optimization, IEEE Computational Intelligence Magazine **1**, pp. 28-39, 2006.

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

[23] J. Kennedy, R.C. Everhart, Particle Swarm Optimization. In Proceedings of the 1995 IEEE International Conference on Neural Networks, Vol. 4, pp. 1942-1948, IEEE Press, 1995.

[24] R. Storn, On the usage of differential evolution for function optimization, In: Proceedings of North American Fuzzy Information Processing, pp. 519-523, 1996.

[25] Y. Zhou and Y. Tan, GPU-based parallel particle swarm optimization, In: 2009 IEEE Congress on Evolutionary Computation, pp. 1493-1500, 2009.

[26] L. Dawson and I. Stewart, Improving Ant Colony Optimization performance on the GPU using CUDA, In: 2013 IEEE Congress on Evolutionary Computation, pp. 1901-1908, 2013.

[27] Barkalov, K., Gergel, V. Parallel global optimization on GPU, J Glob Optim **66**, pp. 3–20, 2016.

[28] Z. Wu, N. Cui, L. Zhao, L. Han, X. Hu, H. Cai, D. Gong, L. Xing, X. Chen, B. Zhu, M. lv, S. Zhu, Q. Liu, Estimation of maize evapotranspiration in semi-humid regions of Northern China Using Penman-Monteith model and segmentally optimized Jarvis model, Journal of Hydrology **22**, pp 127483, 2022.

[29] E. Tlelo-Cuautle, A.M. González-Zapata, J.D. Díaz-Muñoz, L.G. de la Fraga and I. Cruz-Vega, Optimization of fractional-order chaotic cellular neural networks by metaheuristics, The European Physical Journal Special Topics, 2022.

[30] G. Sun, C. Li and L. Deng, An adaptive regeneration framework based on search space adjustment for differential evolution, Neural Computing and Applications **33**, pp. 9503–9519, 2021

[31] P. Civiciogluan and E. Besdok, Bernstain-search differential evolution algorithm for numerical function optimization, Expert Systems with Applications **138**, 2019.

[32] J. Liang, K. Qiao, K. Yu, S.Ge ,B. Qu, R. Xu and K. Li, Parameters estimation of solar photovoltaic models via a self-adaptive ensemble-based differential evolution, Solar Energy **207**, pp. 336-346, 2020.

[33] L. Peng, S. Liu, R. Liu and Lin Wang, Effective long short-term memory with differential evolution algorithm for electricity price prediction, Energy **162**, pp. 1301-1314, 2018.

[34] N. Awad and N. Mallik and F. Hutter, Differential Evolution for Neural Architecture Search, In Proceedings of the 1st workshop on neural architecture search, 2020.

[35] J. Ilonen, J.K. Kamarainen and J. Lampinen, Differential Evolution Training Algorithm for Feed-Forward Neural Networks, Neural Processing Letters **17**, pp. 93–105, 2003.

[36] A. Slowik, Application of an Adaptive Differential Evolution Algorithm With Multiple Trial Vectors to Artificial Neural Network Training, IEEE Transactions on Industrial Electronics **58**, pp. 3160-3167, 2011.

[37] L. Wang, Y. Zeng and T. Chen, Back propagation neural network with adaptive differential evolution algorithm for time series forecasting, Expert Systems with Applications **42**, pp. 855-863, 2015.

[38] X. Wang, G. Xu, Hybrid Differential Evolution Algorithm for Traveling Salesman Problem, Procedia Engineering **15**, pp. 2716-2720, 2011.

[39] I.M.Ali , D. Essam and K. Kasmarik, A novel design of differential evolution for solving discrete traveling salesman problems, Swarm and Evolutionary Computation **52**, 2020.

[40] J. Liu and J. Lampinen, A differential evolution based incremental training method for RBF networks, In Proceedings of the 7th annual conference on Genetic and evolutionary computation (GECCO '05), pp. 881–888, 2005.

[41] B. O'Hora, J. Perera and A. Brabazon, Designing Radial Basis Function Networks for Classification Using Differential Evolution, In: The 2006 IEEE International Joint Conference on Neural Network Proceedings, pp. 2932-2937, 2006.

[42] N. Naveen , V. Ravi , C. Raghavendra Rao and N. Chauhan, Differential evolution trained radial basis function network: application to bankruptcy prediction in banks, International Journal of Bio-Inspired Computation **2**, pp. 222-232, 2010.

[43] Z. Chen, X. Jiang, J. Li, S. Li and L. Wang, PDECO: Parallel differential evolution for clusters optimization, J. Comput. Chem. **34**, pp. 1046-1059. 2013.

[44] A. Ghosh, R. Mallipeddi, S. Das and A. K. Das, A Switched Parameter Differential Evolution with Multi-donor Mutation and Annealing Based Local Search for Optimization of Lennard-Jones Atomic Clusters, In: 2018 IEEE Congress on Evolutionary Computation (CEC), pp. 1-8, 2018.

[45] Y. Zhang, H. Zhang, J. Cai, B. Yang, A Weighted Voting Classifier Based on Differential Evolution, Abstract and Applied Analysis, vol. 2014, Article ID 376950, 6 pages, 2014.

[46] U. Maulik and I. Saha, Automatic Fuzzy Clustering Using Modified Differential Evolution for Image Classification, IEEE Transactions on Geoscience and Remote Sensing **48**, pp. 3503-3510, 2010.

[47] E. Hancer, Differential evolution for feature selection: a fuzzy wrapper–filter approach, Soft Comput **23**, pp. 5233–5248, 2019.

[48] T. Vivekanandan, N. Ch Sriman Narayana Iyengar, Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease, Computers in Biology and Medicine 90, pp. 125-136, 2017.

[49] W. Deng, H. Liu, J. Xu, H. Zhao and Y. Song, An Improved Quantum-Inspired Differential Evolution Algorithm for Deep Belief Network, IEEE Transactions on Instrumentation and Measurement **69**, pp. 7319-7327, 2020.

[50] T. Wu, X. Li, D. Zhou, N. Li, J. Shi, Differential Evolution Based Layer-Wise Weight Pruning for Compressing Deep Neural Networks, Sensors **21**, 2021.

[51] E. Mininno, F. Neri, F. Cupertino and D. Naso, Compact Differential Evolution, IEEE Transactions on Evolutionary Computation **15**, pp. 32-54, 2011.

[52] A. K. Qin, V. L. Huang and P. N. Suganthan, Differential Evolution Algorithm With Strategy Adaptation for Global Numerical Optimization, IEEE Transactions on Evolutionary Computation **13**, pp. 398-417, 2009.

[53] J. Liu, J. Lampinen, A Fuzzy Adaptive Differential Evolution Algorithm, Soft Comput **9**, pp. 448–462, 2005.

[54] H. Wang, S. Rahnamayan and Z. Wu, Parallel differential evolution with self-adapting control parameters and generalized opposition-based learning for solving high-dimensional optimization problems, Journal of Parallel and Distributed Computing **73**, pp. 62-73, 2013.

[55] S. Das, S.S. Mullick and P.N. Suganthan, Recent advances in differential evolution – An updated survey, Swarm and Evolutionary Computation **27**, pp. 1-30, 2016.

[56] M.M. Ali and A. Törn, Population set-based global optimization algorithms: some modifications and numerical studies, Computers & Operations Research **31**, pp. 1703-1725, 2004.

[57] M. Montaz Ali, Charoenchai Khompatraporn, Zelda B. Zabinsky, A Numerical Evaluation of Several Stochastic Algorithms on Selected Continuous Global Optimization Test Problems, Journal of Global Optimization **31**, pp 635-672, 2005.

[58] C.A. Floudas, P.M. Pardalos, C. Adjiman, W. Esposoto, Z. Gümüs, S. Harding, J. Klepeis, C. Meyer, C. Schweiger, Handbook of Test Problems in Local and Global Optimization, Kluwer Academic Publishers, Dordrecht, 1999.

[59] M.M. Ali and P. Kaelo, Improved particle swarm algorithms for global optimization, Applied Mathematics and Computation **196**, pp. 578-593, 2008.

[60] H. Koyuncu, R. Ceylan, A PSO based approach: Scout particle swarm algorithm for continuous global optimization problems, Journal of Computational Design and Engineering **6**, pp. 129–142, 2019.

[61] Patrick Siarry, Gérard Berthiau, François Durdin, Jacques Haussy, ACM Transactions on Mathematical Software **23**, pp 209–228, 1997.

[62] I.G. Tsoulos, I.E. Lagaris, GenMin: An enhanced genetic algorithm for global optimization, Computer Physics Communications **178,** pp. 843-851, 2008.

[63] M.J.D Powell, A Tolerant Algorithm for Linearly Constrained Optimization Calculations, Mathematical Programming **45**, pp. 547-566, 1989.

[64] M. Gaviano, D.E. Ksasov, D. Lera, Y.D. Sergeyev, Software for generation of classes of test functions with known local and global minima for global optimization, ACM Trans. Math. Softw. **29**, pp. 469-480, 2003.

[65] J.E. Lennard-Jones, On the Determination of Molecular Fields, Proc. R. Soc. Lond. A **106**, pp. 463–477, 1924.