A Scalability Test for Accelerated DE Using Generalized Opposition-Based Learning

Hui Wang, Zhijian Wu, Shahryar Rahnamayan, and Lishan Kang

Abstract—In this paper a scalability test over eleven scalable benchmark functions, provided by the current workshop (Evolutionary Algorithms and other Metaheuristics for Continuous Optimization Problems - A Scalability Test), are conducted for accelerated DE using generalized opposition-based learning (GODE). The average error of the best individual in the population has been reported for dimensions 50, 100, 200, and 500 in order to compare with the results of other algorithms which are participating in this workshop. Current work is based on opposition-based differential evolution (ODE) and our previous work, accelerated PSO by generalized OBL.

I. INTRODUCTION

Differential Evolution (DE) is an effective robust optimization algorithm which was proposed by Price and Storn in 1997 [1]. In this paper, an enhanced DE, based on generalized OBL (GOBL) is proposed to accelerate the convergence rate of classical DE. The GOBL was introduced in our previous work [10] which presented a general model for opposition-based learning (OBL). Accelerated DE by GOBL has been called GODE in this paper. In order to verify the performance of GODE, current work provides a scalability test over 11 benchmark functions, provided by the current workshop, for dimensions 50, 100, 200, and 500.

The rest of the paper is organized as follows. In Section II, the classical DE algorithm is briefly reviewed. The GOBL technique and its analysis are presented in Section III. Section IV gives an implementation of the proposed algorithm, GODE. In Section V, the test functions, parameters besides a comprehensive set of scalability experiments are provided. Finally, the work is summarized in Section VI.

II. A BRIEF REVIEW OF DIFFERENTIAL EVOLUTION

DE is a population-based stochastic search algorithm, and has been successfully applied to solve complex problems including linear and nonlinear, unimodal and multimodal functions. It has been shown that DE is faster and more robust on these functions than many other evolutionary algorithms [2].

There are several variants of DE [1], where the most popular variant is shown by “DE/rand/1/bin” which is called classical version. The proposed algorithm is also based on this DE scheme. Let us assume that $X_i(t)(i = 1, 2, \ldots, N_p)$ is the $i$th individual in population $P(t)$, where $ps$ is the population size, $t$ is the generation index, and $P(t)$ is the population in the $t$th generation. The main idea of DE is to generate trial vectors. Mutation and crossover are used to produce new trial vectors, and selection determines which of the vectors will be successfully selected into the next generation.

Mutation—For each vector $X_i(t)$ in Generation $t$, a mutant vector $V$ is generated by

$$V_i(t) = X_i(t) + F (X_j(t) - X_k(t)),$$

where $i \neq j \neq k \neq i$, $i = 1, 2, \ldots, N_p$ and $i_1, i_2, i_3$ are mutually different random integer indices within $[1, N_p]$. The population size $N_p$ should be satisfied $N_p \geq 4$ because $i, i_1, i_2, i_3$ are different. $F \in [0, 2]$ is a real number that controls the amplification of the difference vector ($X_{i_2}(t) - X_{i_3}(t)$).

Crossover—Like genetic algorithms, DE also employs a crossover operator to build trial vectors by recombining two different vectors. The trial vector is defined as follows:

$$U_i(t) = (U_{i,1}(t), U_{i,2}(t), \ldots, U_{i,n}(t)),\quad (2)$$

where $j = 1, 2, \ldots, n$ and

$$U_{i,j}(t) = \begin{cases} V_{i,j}(t), & \text{if } \text{rand}_j(0, 1) \leq CR \land j = l \\ X_{i,j}(t), & \text{otherwise} \end{cases}\quad (3)$$

$CR \in (0, 1)$ is the predefined crossover probability, and $\text{rand}_j(0, 1)$ is a random number within $[0, 1)$ for the $j$th dimension, and $l \in \{1, 2, \ldots, n\}$ is a random parameter index.
Selection—A greedy selection mechanism is used as follows:

\[ X_i(t) = \begin{cases} 
U_i(t), & \text{if } f(U_i(t) < f(X_i(t))) \\
X_i(t), & \text{otherwise} 
\end{cases} \quad (4) \]

Without loss of generality, this paper only considers minimization problem. If, and only if, the trial vector \( U_i(t) \) is better than \( X_i(t) \), then \( X_i(t) \) is set to \( U_i(t) \); otherwise, the \( X_i(t) \) is unchanged.

III. GENERALIZED OPPOSITION-BASED LEARNING (GOBL)

A. Opposition-Based Learning

Opposition-based Learning (OBL) [18], [19] is a new concept in computational intelligence, and has been proven to be an effective method to EAs in many optimization problems [8], [20]-[24]. When evaluating a solution \( x \) to a given problem, simultaneously computing its opposite solution will provide another chance for finding a solution more closer to the global optimum [9].

Opposite Number [20]–Let \( x \in [a, b] \) be a real number. The opposite number of \( x^* \) is defined by:

\[ x^* = a + b - x. \quad (5) \]

Similarly, the definition is generalized to higher dimensions as follows.

Opposite Point [20]–Let \( X = (x_1, x_2, \ldots, x_D) \) be a point in a \( D \)-dimensional space, where \( x_1, x_2, \ldots, x_D \in R \) and \( x_j \in [a_j, b_j], j \in 1, 2, \ldots, D \). The opposite point \( X^* = (x_1^*, x_2^*, \ldots, x_D^*) \) is defined by:

\[ x_j^* = a_j + b_j - x_j. \quad (6) \]

By applying the definition of opposite point, the opposition-based optimization can be defined as follows.

Opposition-based Optimization [20]–Let \( X = (x_1, x_2, \ldots, x_D) \) be a point in a \( D \)-dimensional space (i.e., a candidate solution). Assume \( f(X) \) is a fitness function which is used to evaluate the candidate’s fitness. According to the definition of the opposite point, \( X^* = (x_1^*, x_2^*, \ldots, x_D^*) \) is the opposite of \( X = (x_1, x_2, \ldots, x_D) \). If \( f(X^*) \) is better than \( f(X) \), then update \( X \) with \( X^* \); otherwise keep the current point \( X \). Hence, the current point and its opposite point are evaluated simultaneously in order to continue with the fitter one.

B. Generalization of OBL Model

Based on the concept of OBL, we propose a generalized OBL as follows [10]. Let \( x \) be a solution in current search space \( S \), \( x \in [a, b] \). The new solution \( x^* \) in the opposite space \( S^* \) is defined by [10]:

\[ x^* = \Delta - x, \quad (7) \]

where \( \Delta \) is a computable value and \( x^* \in [\Delta - b, \Delta - a] \). It is obvious that the differences between current search space \( S \) and opposite search space \( S^* \) are the center positions of search space. Because the size of search range (indicates the size of interval boundaries) of \( S \) and \( S^* \) are \( b - a \), and the center of current search space moves from \( a + b \) to \( 2\Delta - a - b \) after using GOBL.

Similarly, the definition of GOBL is generalized to a \( D \)-dimensional search space as follows:

\[ x_j^* = \Delta - x_j, \quad (8) \]

where \( j = 1, 2, \ldots, D \).

By applying the GOBL, we not only evaluate the current candidate \( x \), but also calculate its opposite candidate \( x^* \). This will provide more chance of finding solutions closer to the global optimum. So it is important to investigate when the GOBL is beneficial.

Assume that the current candidate is \( x \), and the corresponding candidate in the opposite space is \( x^* \). The opposite candidate \( x^* \) is closer to the global optimum than the current candidate \( x \), if and only if

\[ |x^* - x^0| < |x - x^0|. \quad (9) \]

Hence

\[ (x^* - x^0)^2 - (x - x^0)^2 < 0 \]
\[ \Rightarrow (x^* + x - 2x^0)(x^* - x) < 0 \]
\[ \Rightarrow (\Delta - 2x^0)(\Delta - 2x) < 0. \]

It is obvious that

\[ x < \frac{\Delta}{2}, \quad \text{if } x^0 > \frac{\Delta}{2}, \quad (11) \]
\[ x > \frac{\Delta}{2}, \quad \text{if } x^0 < \frac{\Delta}{2}. \quad (12) \]

That is, the opposite candidate \( x^* \) is closer to the global optimum \( x^0 \) than the current candidate \( x \), when \( x^0 \) and \( x \) are located at the different sides of \( \frac{\Delta}{2} \).

However, according to No-Free-Lunch theorem [17], the GOBL could not be suitable for all kinds of optimization problems. For instance, the opposite candidate may be far away from the global optimum when solving multimodal problems. To avoid this case, a new elite selection mechanism based on population is used after the opposition. The specific use method of GOBL are described in Section V.
C. Four Different Schemes of Generalized OBL

Let $\Delta = k(a + b)$, where $k$ is a real number. The generalized OBL model is defined by:

$$x^* = k(a + b) - x.$$  \hfill (13)

So the opposite candidate $x^*$ is closer to the global optimum $x^g$ than the current candidate $x$, when $x^g$ and $x$ are located at the different sides of $k(a+b)$. Let us consider four typical GOBL schemes with different values of $k$ as follows.

1) $k = 0$ (Symmetrical Solutions in GOBL, GOBL-SS) The GOBL-SS model is defined by

$$x^* = -x,$$  \hfill (14)

where $x \in [a, b]$ and $x^* \in [-b, -a]$. The current solution $x$ and opposite solution $x^*$ are on the symmetry of origin.

2) $k = \frac{1}{2}$ (Symmetrical Interval in GOBL, GOBL-SI) The GOBL-SI model is defined by

$$x^* = \frac{a + b}{2} - x,$$  \hfill (15)

where $x \in [a, b]$ and $x^* \in [-\frac{b-a}{2}, \frac{b-a}{2}]$. The interval of the opposite space is on the symmetry of origin.

3) $k = 1$ (Opposition-based learning, OBL) When $k = 1$, the GOBL model is identical to Eq.5, where $x \in [a, b]$ and $x^* \in [a, b]$.

4) $k = \text{rand}(0, 1)$ (Random GOBL, GOBL-R) The GOBL-SI model is defined by

$$x^* = k(a + b) - x,$$  \hfill (16)

where $k$ is a random number within $[0, 1]$, $x \in [a, b]$ and $x^* \in [k(a+b) - b, k(a+b) - a]$. The center of the opposite space is at a random position between $\frac{a+b}{2}$ and $\frac{a+b}{2}$.

For a given problem, it is possible that the opposite candidate may jump out of the definition domain $[X_{min}, X_{max}]$. When this happens, the GOBL will be invalid, because the opposite candidate is infeasible. To avoid this case, the opposite candidate is assigned to a random value as follows.

$$x^* = \text{rand}(a, b), \text{ if } x^* < X_{min} \text{ or } x^* > X_{max},$$  \hfill (17)

where $\text{rand}(a, b)$ is a random number within $[a, b]$, and $[a, b]$ is the interval boundaries of current search space.

IV. GOBL-BASED OPTIMIZATION

If the interval of the current search space is symmetric with respect to the origin ($a = -b$), then $\Delta = a + b = 0$. According to Eq.10, the GOBL is beneficial when $x^0 \cdot x < 0$ is satisfied. If the global optimum $x^g = 0$, then the GOBL is invalid because there does not exist a candidate $x$ which satisfies $0 \cdot x < 0$. The global optimum $x^g$ is located on the origin in many optimization problems [3], [7], [8], [11]. To have an asymmetric opposition, the interval boundaries $[a_j(t), b_j(t)]$ is dynamically updated according to the size of current search space. The new dynamic GOBL model is defined by [8]

$$X_{i,j}^* = k[a_j(t) + b_j(t)] - X_{i,j}$$  \hfill (18)

$$a_j(t) = \min(X_{i,j}(t)), \quad b_j(t) = \max(X_{i,j}(t))$$  \hfill (19)

$$X_{i,j}^* = \text{rand}(a_j(t), b_j(t)), \text{ if } X_{i,j}^* < X_{min} \text{ or } X_{i,j}^* > X_{max}$$  \hfill (20)

where $X_{i,j}$ is the $j$th vector of the $i$th candidate in the population, $X_{i,j}^*$ is the opposite candidate of $X_{i,j}$, $a_j(t)$ and $b_j(t)$ are the minimum and maximum values of the $j$th dimension in current search space respectively, $\text{rand}(a_j(t), b_j(t))$ is a random number within $[a_j(t), b_j(t)]$, $[X_{min}, X_{max}]$ is the definition domain, $N_p$ is the population size, $\text{rand}(0, 1)$ is a random number within $[0, 1]$, and $t = 1, 2, \ldots$ indicates the generations.

V. GENERALIZED OPPOSITION-BASED DIFFERENTIAL EVOLUTION (GODE)

In our previous work [10], GOBL was applied to PSO and the experimental results showed that the GOBL model with random $k$ works better than the other three models in many benchmark function problems. So, the proposed approach GODE is also based on the random GOBL model in this paper.

Like ODE, the GODE uses the GOBL method to initialize population and produce new candidates in evolutionary generations. The original DE is chosen as a parent algorithm and the proposed GOBL model is embedded in DE to improve its performance. However, the embedded strategy of GODE is different from ODE. In the ODE, the opposition occurs with a probability, and the classical DE executes every generation. But in the GODE, if $\text{rand}(0, 1) \leq p_o$, then execute the GOBL; otherwise execute the classical DE.

The pseudo-code of GODE is shown in Algorithm 1, where $P$ is the current population, $GOP$ is the transformed population after using GOBL, $P_i$ is the $i$th individual in $P$, $GOP_i$ is the $i$th individual in $GOP$, $k$ is a random number within $[0, 1]$, $p_o$ is the probability of GOBL, $N_p$ is the population size, $n$ is the dimension size, $a_j(t), b_j(t)$ is the interval boundaries.
Algorithm 1: Accelerated Differential Evolution

1. Randomly initialize each individual in population \( P \);
2. double \( k = \text{rand}(0, 1) \);
3. for \( i = 1 \) to \( N_p \) do
   4. for \( j = 1 \) to \( n \) do
      5. \( GOP_{i,j} = k(a_{ij} + b_{ij}) - P_{i,j} \);
      6. if \( GOP_{i,j} \) is out of the definition domain then
         7. \( GOP_{i,j} = \text{rand}(a_{ij}, b_{ij}) \);
      end
   end
   8. Calculate the fitness value of \( GOP_i \);
   9. NE++;
10. Select \( N_p \) fittest individuals from \( \{ P, GOP \} \) as an initial population \( P \);
11. while \( NE \leq \text{MAXNE} \) and BFV > VTR do
   12.   if \( \text{rand}(0, 1) \leq p_s \) then
      13.      Update the dynamic interval boundaries \( [a_{ij}(t), b_{ij}(t)] \) in \( P \)
      14.      according to Eq.19;
   15.   end
   16.   for \( i = 1 \) to \( N_p \) do
      17.      for \( j = 1 \) to \( n \) do
         18.         \( GOP_{i,j} = k(a_{ij} + b_{ij}) - P_{i,j} \);
         19.         if \( GOP_{i,j} \) is out of the definition domain then
            20.            \( GOP_{i,j} = \text{rand}(a_{ij}, b_{ij}) \);
         end
      end
      21.      Calculate the fitness value of \( GOP_i \);
      22.      NE++;
   23. end
   24. Select \( N_p \) fittest individuals from \( \{ P, GOP \} \) as current population \( P \);
   25. end
   26. else
      27.     for \( i = 1 \) to \( N_p \) do
         28.      Randomly select 3 parents \( P_{12} \), \( P_{13} \) and \( P_{23} \) from \( P \),
         29.      where \( i \neq i1 \neq i2 \neq i3 \);
         30.      for \( j = 1 \) to \( n \) do
            31.            if \( \text{rand}(0, 1) < CR \) then
               32.              \( V_{i,j} = P_{1,j} + F(P_{2,j} - P_{3,j}) \);
            else
               33.              \( U_{i,j} = P_{1,j} \);
            end
         end
         34.      end
         35.      Calculate the fitness value of \( U_i \);
         36.      NE++;
         37.      if \( f(U_i) < f(P_i) \) then
            38.            \( P_i = U_i \);
         else
            39.            \( P_i = P_i \);
         end
      end
   end
end

of current population, \( \text{rand}(a_{ij}(t), b_{ij}(t)) \) is a random number within \( [a_{ij}(t), b_{ij}(t)] \). BFV is the best fitness value, VTR is the value-to-reach [8]. NE is the number of evaluations, and MAXNE is the maximum number of evaluations.

VI. Conducted Scalability Tests

A. Experimental Framework

For the experiments, the following eleven scalable benchmark problems have been considered:

1. F1-F6 of the CEC’2008 Special Session and Competition on Large Scale Global Optimization test suite [25].
2. Schwefel’s Problem 2.22 (F7), Schwefel’s Problem 1.2 (F8), Extended f10 (F9), Bohachevsky (F10), and Schaffer (F11), see [26] for their descriptions.

The requirements on the simulation procedure are the followings [26]:

1. Each algorithm is run 25 times for each test function, and the average error of the best individual of the population is computed. For a solution \( x \), this measure is defined as: \( f(x) - f(op) \), where \( op \) is the optimum of the function.
2. The study has been made with dimensions 50, 100, 200, and 500. The maximum number of fitness evaluations is \( 5000 \times D \). Each run stops when the maximal number of evaluations is achieved.

B. Setting Control Parameters

Parameter settings for all conducted experiments are as follows (the same setting has been used in literature cited after of each parameter):
- Population size, \( N_p = 100 \) [8]
- Differential amplification factor, \( F = 0.5 \) [8]
- Crossover probability constant, \( C_r = 0.9 \) [8]
- Probability of GOBL, \( p_o = 0.4 \) (selected based on our previous experiments)
- Maximum number of function calls, \( \text{MAXNFC} = 5000 \times D \) (proposed by the current workshop) [26]
- Mutation Strategy: DE/rand/1/bin (classical DE) [8]

C. Numerical Results

Results for dimensions 50, 100, 200, and 500 are summarized in Tables I, II, III, and IV, respectively. In addition, in order to conduct some statistic tests during the workshop, the author will provide to the organizers an Excel file with all these results.

D. Results Analysis

As seen, for functions F1-F6 (shifted problems), when the dimension of the problems increases the average error \( f(x) - f(op) \) increases as well, which looks logical. But for functions F7-F11 (unshifted problems), GODE performs in a reverse manner. The main reason is that we use the same population size \( (N_p) \) for problems with different dimensions. This is unfair to lower dimensional problems, because higher dimensional problems have more maximum generation \( (\text{MAXNE}/N_p) \) than lower dimensional ones. Another reason is that GODE is not sensitive to F7-F11 with different dimensions. If given
GODE results for $D=50$. Each algorithm is run 25 times for each test function, and the average of error ($f(x) - f(op)$, $op$: the optimum the function) for the best individual of the population is computed.

<table>
<thead>
<tr>
<th>Function</th>
<th>$f(x) - f(op)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>1.6872e - 014</td>
</tr>
<tr>
<td>F2</td>
<td>20.1736</td>
</tr>
<tr>
<td>F3</td>
<td>64.9945</td>
</tr>
<tr>
<td>F4</td>
<td>332.358</td>
</tr>
<tr>
<td>F5</td>
<td>6.90133e - 004</td>
</tr>
<tr>
<td>F6</td>
<td>1.65117e - 008</td>
</tr>
<tr>
<td>F7</td>
<td>6.76794e - 067</td>
</tr>
<tr>
<td>F8</td>
<td>6.54375e - 080</td>
</tr>
<tr>
<td>F9</td>
<td>5.25139e - 032</td>
</tr>
<tr>
<td>F10</td>
<td>6.04275e - 136</td>
</tr>
<tr>
<td>F11</td>
<td>5.85362e - 032</td>
</tr>
</tbody>
</table>

GODE results for $D=100$. Each algorithm is run 25 times for each test function, and the average of error ($f(x) - f(op)$, $op$: the optimum the function) for the best individual of the population is computed.

<table>
<thead>
<tr>
<th>Function</th>
<th>$f(x) - f(op)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>3.12245e - 014</td>
</tr>
<tr>
<td>F2</td>
<td>47.134</td>
</tr>
<tr>
<td>F3</td>
<td>252.605</td>
</tr>
<tr>
<td>F4</td>
<td>466.972</td>
</tr>
<tr>
<td>F5</td>
<td>0.00614746</td>
</tr>
<tr>
<td>F6</td>
<td>0.734186</td>
</tr>
<tr>
<td>F7</td>
<td>3.22251e - 140</td>
</tr>
<tr>
<td>F8</td>
<td>1.9967e - 149</td>
</tr>
<tr>
<td>F9</td>
<td>5.55917e - 066</td>
</tr>
<tr>
<td>F10</td>
<td>5.64935e - 268</td>
</tr>
<tr>
<td>F11</td>
<td>5.83833e - 067</td>
</tr>
</tbody>
</table>

GODE results for $D=200$. Each algorithm is run 25 times for each test function, and the average of error ($f(x) - f(op)$, $op$: the optimum the function) for the best individual of the population is computed.

<table>
<thead>
<tr>
<th>Function</th>
<th>$f(x) - f(op)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>2.37789e - 010</td>
</tr>
<tr>
<td>F2</td>
<td>71.0062</td>
</tr>
<tr>
<td>F3</td>
<td>748.518</td>
</tr>
<tr>
<td>F4</td>
<td>411.156</td>
</tr>
<tr>
<td>F5</td>
<td>0.0338863</td>
</tr>
<tr>
<td>F6</td>
<td>2.92399</td>
</tr>
<tr>
<td>F7</td>
<td>6.1449e - 277</td>
</tr>
<tr>
<td>F8</td>
<td>1.19597e - 293</td>
</tr>
<tr>
<td>F9</td>
<td>0</td>
</tr>
<tr>
<td>F10</td>
<td>0</td>
</tr>
<tr>
<td>F11</td>
<td>0</td>
</tr>
</tbody>
</table>

As seen, the average error increases by dimension and these results confirm our reasoning at least for F7. Population size can play a crucial role to obtain more accurate solutions. The overall performance of GODE will be more clear during comparison with other algorithms participating in the workshop.

VII. Summary

In this paper, Differential Evolution Based on generalized opposition-based learning (GODE) is proposed. The GOBL is an enhanced opposition-based learning, which transforms candidates in current search space to a new search space. By simultaneously evaluating solutions in current search space and transformed space, we can provide more chance of finding better solutions. A scalability test over 11 scalable benchmark functions, provided for the current workshop, are conducted. Results for the dimensions 50, 100, 200 and 500 are reported.

References
