Opposition-based Ensemble Micro-Differential Evolution

Hojjat Salehinejad*, Shahryar Rahnamayan†, and Hamid R. Tizhoosh‡
*Department of Electrical & Computer Engineering, University of Toronto, Toronto, Canada
†Department of Electrical & Computer Engineering, University of Ontario Institute of Technology, Oshawa, Canada
‡KIMIA Lab, University of Waterloo, Waterloo, Canada
hojjat.salehinejad@mail.utoronto.ca, shahryar.rahnamayan@uoit.ca, tizhoosh@uwaterloo.ca

Abstract—Differential evolution (DE) algorithm with a small population size is called Micro-DE (MDE). A small population size decreases the computational complexity but also reduces the exploration ability of DE by limiting the population diversity. In this paper, we propose the idea of combining ensemble mutation scheme selection and opposition-based learning concepts to enhance the diversity of population in MDE at mutation and selection stages. The proposed algorithm enhances the diversity of population by generating a random mutation scale factor per individual and per dimension, randomly assigning a mutation scheme to each individual in each generation, and diversifying individuals selection using opposition-based learning. This approach is easy to implement and does not require the setting of mutation scheme selection and mutation scale factor. Experimental results are conducted for a variety of objective functions with low and high dimensionality on the CEC Black-Box Optimization Benchmarking 2015 (CEC-BBOB 2015). The results show superior performance of the proposed algorithm compared to the other micro-DE algorithms.

I. INTRODUCTION

Population size plays an important role in performance of the population-based algorithms such as ant colony optimization [1] and differential evolution (DE) [2]. DE with a small population size is called micro-DE (MDE) [3]. A small population size decreases the exploration ability of DE. The degrading performance is mainly due to the generation of fewer candidate individuals by the mutation scheme. This is while a large population size provides higher exploration ability, at the cost of more computational time.

The lack of diversity in the population of MDE to explore the problem landscape can cause stagnation or premature convergence problems. The stagnation problem refers to a diverse and scattered population that cannot converge to a local optimum [3]. From the other side, the lack of diversity in generating new individuals causes trapping of individuals in local optimaums, generally referred to as premature convergence. A trade-off between the population size and evolution method of population improves the performance of algorithm [4].

The standard version of DE uses a fixed value for mutation scale factor and a fixed mutation scheme. The MDE algorithm mostly suffers from lack of diversity in the population to generate new individuals [2]. By vectorized randomizing the mutation scale factor (called MDEVM), which is a random mutation scale factor for each dimension of each individual, the micro-DE performs much better than utilizing a static mutation scale factor [3]. Methods such as opposition-based DE (ODE) [5] and centroid DE [6–8] have solely improved the performance of DE algorithm by proper diversifying the population in DE algorithm. These methods can also increase the performance of MDE.

In this paper, we enhance the MDEVM algorithm [2], [3], [9], [10], by utilizing the idea of ensemble mutation scheme and opposition-based learning (OBL) [11] to diversify the population. OBL generates the opposite of the current individuals (candidate solutions) and selects the best individuals with respect to objective function values [12]. At the mutation stage, each individual randomly selects a mutation scheme at each generation, called ensemble mutation scheme [4]. Combination of vectorized randomized mutation scale factor, ensemble mutation scheme, and OBL, has highly enhanced the performance of MDE. We call this method opposition-based ensemble micro-DE (OEMDE).

II. RELATED WORKS

The main control parameters of DE algorithm are population size, mutation scale factor, crossover rate, and mutation scheme. Proper selection of these parameters helps to improve the performance of MDE.

A. Mutation Scale Factor

DE is about the difference of two or more vectors. Amplification of the difference vector is adjustable by the mutation scale factor $F$ which controls the evolving rate of the population. A small mutation scale factor decreases the exploration and may result in premature convergence. This is while a large mutation scale factor increases the exploration which results in a longer convergence time. The optimal value of $F$ is usually set based on the nature of the problem and experimental observations [13]. MDEVM algorithm [3] relaxes static selection of $F$ as a hyper-parameter by generating a random $F \in (0.1, 1.5)$ for each dimension of an individual in the population [3]. This randomness adds diversity to the population by amplifying the differences vectors in various scales. The experimental results in [2] have demonstrated outstanding performance of this technique for diversity enhancement.

B. Mutation Scheme and Crossover Rate

The recommended range for the crossover rate is [0.3,0.9] in the literature [14]. However, the optimal value is problem-dependent.
dependent. Adaptive parameters can change as the evolution proceeds, such as self-adaptive DE (SDE) \cite{15} and memetic algorithm.

It is possible to use a pool of mutation strategies and a pool of values for each control parameter of the algorithm \cite{16}. Ensemble algorithms use a pool of mutation schemes with two members, DE/rand/1 and DE/best/1. The strategy is to combine random parameters setting and mutation schemes to enhance convergence rate while maintaining the high diversity of population \cite{17}. The proposed method in \cite{18} switches the scale factor and crossover rate in a uniform-random strategy for each individual of the population. The switching is performed within a particular range of the variables. A pool of mutation schemes is also utilized and the individuals use either the DE/rand/1 or DE/best/1 scheme. In order to implement the DE on systems with limited resources for real-time applications, such as embedded systems, a version of DE algorithm, called compact DE (cDE), is proposed \cite{19,20}. This algorithm uses probabilistic representation of the population, instead of the whole population for optimization. The logic behind cDE is producing candidates by sampling an evolving distribution in every generation. However, cDE algorithm suffers from premature convergence problem, particularly for high dimensional problems. This refers to the shrinkage of the probability function which models the population \cite{19}. An ensemble version of cDE is proposed in \cite{19}.

The memetic algorithms are hybrid evolutionary algorithms. These algorithms can solve optimization problems by utilizing deterministic local search within the evolutionary process. This technique along with the cDE algorithm is used to develop optimization algorithms on control cards \cite{21}. The objective is to design an optimization algorithm that can function in absence of full power computational systems.

A novel type of ensemble DE, called MPEDE, is proposed in \cite{22}. It is a multi-population based approach which deploys a dynamic ensemble of multiple mutation strategies. It controls parameters of the algorithms such as mutation scale factor and crossover rate \cite{23}. This method uses the “current-to-topbest/1”, “current-to-rand/1”, and “rand/1” mutation schemes \cite{22}. Another method for unconstrained continuous optimization problems is μJADE \cite{24}. It is an adaptive DE with a small population size approach. The μJADE uses a new mutation operator, called current-by-rand-to-pbest \cite{24}.

Random perturbation method utilizes the mutation idea from the genetic algorithm (GA) to randomly change the population vector parameters with a fixed probability \cite{25}. A modified DE with smaller population size can add disturbances to the mutation operation \cite{26}. An adaptive system controls the intensity of disturbance based on the performance improvement during generations. A combination of modified Breeder GA mutation scheme and a random mutation scheme can help DE to avoid stagnation and/or premature convergence \cite{27}.

C. Selection Scheme and Population Size Adaptation

Large population size in DE algorithm causes high computational cost and adds more exploration ability to the population. The small population size in MDE algorithm reduces the number of function evaluations (per generation) with the cost of less exploration capability and a higher risk of premature convergence \cite{3}.

A population size adaptation method is proposed in \cite{28} which measures the Euclidean distance between individuals and finds out if the population diversity is poor and/or it moves toward stagnation. Depending on the search situation, it generates new individuals. A population size reduction method for cDE is proposed in \cite{29}, where the population size gradually reduces during the evolution. Cumu-DE is an adaptive DE and the effective population size is adapted automatically using mechanism based on a probabilistic model \cite{30}. The term effective refers to the effective part of the population where its size shrinks as the algorithm has more successful trials. Gradual reduction of the population size is another approach that has demonstrated higher robustness and efficiency compared to generic DE \cite{31}.

OBL has Type-I and Type-II schemes \cite{11,12}. The Type-I scheme has enhanced the performance of micro-ODE for image thresholding \cite{32}. This approach showed a better performance compared to the MDE algorithm.

Small-size cooperative sub-populations are capable of finding sub-components of the original problem concurrently \cite{33}. Combination of sub-components through their cooperation constructs a complete solution for the problem \cite{34}. A MDE version of this method is proposed to evolve an indirect representation of the bin packing problem \cite{34}. The idea of self-adaptive population size is carried out to test absolute and relative encoding methods for DE \cite{35}. The reported simulation results on 20 benchmark problems denote the self-adaptive population size with relative encoding outperforms the absolute encoding method and the DE algorithm \cite{35}.

III. DIFFERENTIAL EVOLUTION

In general, during a black-box optimization process, an optimizer does not have a priori knowledge about the problem landscape. Therefore, optimizers normally start with randomly generated candidate solutions and then they try to get intelligence about the problem landscape during the optimization procedure. The DE algorithm, similar to other population-based algorithms, uses random initial vectors with uniform distribution at the beginning, called initial population. Through a number of generations, it tries to improve the initial candidate individuals toward an optimal solution. The population \( P = \{X_1, ..., X_{N_P}\} \) has \( N_P \) vectors in generation \( g \), where \( X_i \) is a \( D \)-dimensional vector defined as \( X_i = (x_{i,1}, ..., x_{i,D}) \). The classic version of DE algorithm is consisted of the following three major evolutionary operations: mutation, crossover, and selection.

**Mutation:** This step selects three vectors randomly from the population such that \( i_1 \neq i_2 \neq i_3 \neq i \), where \( i \in \{1, ..., N_P\} \) s.t. \( N_P \geq 4 \) for each vector \( X_i \). The mutant vector is calculated as

\[
V_i = X_{i_1} + F(X_{i_2} - X_{i_3}),
\]

(1)
where the mutation factor $F \in (0.1, 1.5)$ is a real constant number that controls the amplification of the added differential variation of $(\mathbf{X}_{\text{i}_2} - \mathbf{X}_{\text{i}_1})$. The exploration of DE increases by selecting higher values for $F$.

**Crossover:** The crossover operation increases diversity of the population by shuffling the mutant and parent vectors as follows

$$U_{i,d} = \begin{cases} V_{i,d}, & \text{rand}_{d}(0, 1) \leq Cr \text{ or } d_{\text{rand}} = d, \\ x_{i,d}, & \text{otherwise} \end{cases}$$

where $d = [1, ..., D]$, $Cr \in [0, 1]$ is the crossover rate control parameter, and rand$(a, b)$ generates a random number in the interval $[a, b]$ with a uniform distribution. Therefore, the trial vector $U_i \forall i \in \{1, ..., N_P\}$ can be generated as

$$U_i = (U_{i,1}, ..., U_{i,D}).$$

**Selection:** The $U_i$ and $X_i$ vectors are evaluated and compared with respect to their fitness values; The one with better fitness is selected for the next generations.

**IV. OPPOSITION-BASED LEARNING**

The typical idea in population-based algorithms is to generate some random initial candidates and then try to guide them toward an optimal solution [12]. By considering candidates and corresponding opposite candidates simultaneously, the search algorithm can have a more exploration coverage that increases the chance of finding fitter candidates. This idea can be supported by the probability theory, where for a specific problem, the opposite of the current candidate solution has a higher chance to be closer to the solution than a second random number [36], [37]. This phenomenon can accelerate convergence rate of the optimizer. The concept of OBL can be categorized into Type-I and Type-II schemes [11], [12].

The Type-I OBL deals with the relationship among candidate solutions, based on their attributes, without considering the actual objective landscape. Generally speaking, the Type-I OBL computation is easy due to its linear definition in the variable space. Type-I OBL can be considered as an approximation of the genuine intelluction of opposite computing, which is the Type-II OBL. Type-II opposition scheme requires a profound understanding of the objective space. This concept is typically difficult to be utilized in real-world applications, since for the Type-I approach the variable space is already known, while in the Type-II scenario a priori knowledge about the landscape is required, which may not be available for black-box problems.

Figure 1 shows the Type-I and Type-II OBL schemes on a one dimensional problem landscape. OBL was initially defined with the Type-I trend with a min-max oppositional computing scheme. This approach has received more attention due to availability of variables’ boundaries. The other method is centroid opposition computing which works based on the centroid point of population and does not require variables’ boundaries. The min-max Type-I OBL is defined for a $D$-dimensional point $x_d \in [X_{\text{min},d}, X_{\text{max},d}]$ as

$$\bar{x}_d = X_{\text{min},d} + X_{\text{max},d} - x_d$$

where $x_d$ is the opposite of $x_d$ for $d = [1, ..., D]$. By considering $X = (x_1, ..., x_D)$ as a point in the $D$-dimensional space and $\bar{X} = (\bar{x}_1, ..., \bar{x}_D)$ as the opposite point of $X$, if $f(X) \leq f(\bar{X})$, where $f(\cdot)$ is the fitness value, $X$ is considered as the better solution for a maximization problem. In order to have a general definition, if $\psi \in \Psi$ in a $D$-dimensional space and $\Phi : \Psi \rightarrow \Psi$ is a one-to-one mapping function, which defines an oppositional relationship between two unique elements $\psi_1$ and $\psi_2$ of $\Psi$, the opposition concept for OBL Type-I is $\bar{\psi} = \Phi(\psi)$. In addition, if $\Phi(\psi_1) = \psi_2$ and $\Phi(\psi_2) = \psi_1$, the oppositional relationship is symmetric in this case. These approaches are discussed in [4] and [12].

For the point $X$ in Type-II OBL we have $f(X) \in [Y_{\text{min}}, Y_{\text{max}}]$ [38], [5]. Therefore, the min-max-based opposition for Type-II OBL is defined as

$$\bar{f}(X) = Y_{\text{min}} + Y_{\text{max}} - f(X).$$

In practice, since the landscape boundaries are unknown, the boundaries $Y_{\text{min}}$ and $Y_{\text{max}}$ can be estimated using sampling as $y_{\text{min}}$ and $y_{\text{max}}$, respectively. The other approach can be using the centroid of known limits. However, it is proved in [12] that the centroid-based method can demonstrate better estimation compared to min-max-based opposition.

**V. PROPOSED ALGORITHM**

In the classical DE, the mutation scale factor is set to a constant value and the mutation scheme is identical for all individuals during generations. In the proposed OEMDE algorithm, we enhance the exploration ability of MDE by proposing a combination of three strategies: 1) vectorized randomization of mutation scale factor; 2) utilization of ensemble mutation scheme; 3) population initialization and generation jumping based on the opposition-based learning.

The OEMDE algorithm is presented in Algorithm 1, where $BVF$ is the best value so far, $VTR$ is the value to reach, $EVTR$
Algorithm 1: Opposition-based Ensemble Micro-Differential Evolution (OEMDE)

1: Procedure OEMDE
2: \( g = 0 \) // generation counter
3: //Initial Population Generation
4: for \( i = 1 \rightarrow N_P \) do
5: \( X_{i,d} = x_{d}^{\min} + \text{rand}(0, 1) \times (x_{d}^{\max} - x_{d}^{\min}) \)
6: \( \text{end for} \)
7: //Compute opposite of generated population
8: for \( i = 1 \rightarrow N_P \) do
9: \( V_{i,d} = x_{d}^{\min} + \text{rand}(0, 1) \times (x_{d}^{\max} - x_{d}^{\min}) - X_{i,d} \)
10: \( \text{end for} \)
11: //End of compute opposite of generated population
12: //Initial Population Generation
13: \( S = \{ \min \{ f(X) \cup f(X) \} : |S| = N_P \} \)
14: \( P^g = \{ X_i | f(X_i) \in S \} // P \) is the population
15: //End of Initial Population Generation
16: // Mutation
17: Select \( J \) (number of individuals in the chosen mutation scheme) individuals from \( P^g \) where \( (i_1 \neq i_2 \neq i_3 \neq i) \)
18: for \( d = 1 \rightarrow D \) do
19: \( F_{i,d} = \text{rand}(0, 1, 1, 5) \)
20: \( V_{i,d} = f_i(F_{i,d}, X_{i_1,d}, \ldots, X_{i_J,d}) \)
21: \( \text{end for} \)
22: //End of Mutation
23: // Crossover
24: \( U_{i,d} = V_{i,d} \)
25: else
26: \( U_{i,d} = x_{i,d} \)
27: \( \text{end if} \)
28: //End of Crossover
29: // Selection
30: if \( f(U_i) \leq f(X_i) \) then
31: \( X_i' = U_i \)
32: else
33: \( X_i' = X_i \)
34: \( \text{end if} \)
35: //End of Selection
36: end for
37: //Compute opposite of population
38: for \( i = 1 \rightarrow N_P \) do
39: \( \hat{X}_{i,d} = x_{d}^{\max} + x_{d}^{\min} - X_{i,d} \)
40: \( \text{end for} \)
41: //End of compute opposite of generated population
42: \( S = \{ \min \{ f(X) \cup f(X) \} : |S| = N_P \} \)
43: \( g = g + 1 \)
44: \( P^g = \{ X_i | f(X_i) \in S \} // P \) is the population
45: end while

A. Vectorized Random Mutation Factor

The mutation scale factor proposed in \cite{3} is generated in a vectorized random scheme such that it is a different value for each dimension of an individual. Therefore, the mutation factor can be defined for each individual \( i \) as

\[
F_i = (F_{i,1}, \ldots, F_{i,D}), \forall i \in 1, \ldots, N_P,
\]

where \( F_{i,d} = \text{rand}(0,1,1,5) \) \cite{3}.

B. Ensemble Mutation Scheme

The OEMDE algorithm randomly assigns a mutation scheme to each individual \( i \) of the population from a pool of mutation schemes, Table 1. The mutation scheme is a function of mutation scale factor and \( J \) selected parents for each individual such as

\[
V_{i,d} = f_i(F_{i,d}, X_{i_1,d}, \ldots, X_{i_J,d})
\]

This function adds diversity into the population by using vectorized random mutation scale factor as well as using a random mutation scheme for each individual in each generation. Random selection of mutation scheme from a uniform distribution results in enhancing the diversity in the population and therefore, more exploration capability of the DE algorithm. Diversity in mutation scheme selection is particularly important for MDE algorithm, where reducing the population size to a micro-level results in a loss of diversity in the population. This is one of the main causes of premature convergence and stagnation.

C. Opposition-based Learning

The OBL not only considers the current population, but also gives a chance to the opposite individuals of the population to cooperate in the evolution. We use OBL in population initialization as well as for generation jumping similar to ODE algorithm \cite{5}. As Algorithm 1 shows, the opposite of individual \( i \) at dimension \( d \) is computed as

\[
\hat{X}_{i,d} = x_{d}^{\max} + x_{d}^{\min} - X_{i,d}
\]

TABLE I: Pool of mutation schemes.

| Mutation Scheme | Equation |
|-----------------|---------|
| DE/rand/1       | \( V_i = X_i + F(X_i - X_p) \) |
| DE/rand/2       | \( V_i = X_i + F(X_i - X_p) \) |
| DE/target-to-best/1 | \( V_i = X_i + F(X_{best} - X_i) + F(X_i - X_p) \) |
| DE/rand/2       | \( V_i = X_i + F(X_i - X_p) \) |
| DE/target-to-best/2 | \( V_i = X_i + F(X_{best} - X_i) + F(X_i - X_p) \) |

These values of the opposite population give a chance to the opposite individuals of the population from a pool with opposition in initialization (OIEMDE).
The population size is set to \( N \). The mutation scheme pool is given as in Table I. Since one of the algorithms compared with the OEMDE are set based on the best setting reported in their corresponding papers. Settings of OIEMDE is similar to OEMDE.

C. Performance Analysis

The performance results of the OEMDE versus EMDE, OIEMDE, PAjDE, \( \mu \)JADE, and ODE for \( N_P = 6 \) for \( D \in \{30, 50, 100, 200, 500, 1000\} \) are presented in Table II. The Wilcoxon statistical test is performed to find significant results. Due to space limitations, the error and standard deviation values per objective function for all dimensions except 30 are presented in Tables IV-VII.

Results in Table II show that as the problem dimensionality increases OEMDE achieves more success rates compared to EMDE, OIEMDE, and PAjDE. We can observe similar success versus \( \mu \)JADE and ODE, but with more neutral rates, particularly for \( D = 1000 \). For this dimension, OEMDE and ODE have similar performance, with 7 neutral significance and 7 positive significance for ODE and OEMDE. This is while OEMDE has much better performance comparing to PAjDE, with 22 positive significance. Error value of algorithms in Table II is plotted versus number of function calls in Figures 2 and 3 for \( D \in \{30, 1000\} \) and some objective functions \( f_1, f_6, f_{15}, f_{21} \) (due to lack of space). The plots clearly show that the OEMDE converges to a better candidate solution much faster than the other methods. For example, for \( D = 30 \) and \( f_1 \) the \( \mu \)JADE and PAjDE algorithms are trying to decrease the error value while the ODE and OEMDE already have arrived to a much lower error value. As problem dimensionality increases, similar convergence pattern is observable. For example, Figure 3 for \( f_6, f_{15}, \) and \( f_{21} \).
TABLE III: Error and standard deviation of the proposed OEMDE algorithm versus state of the art. The \( W \) represents the Wilcoxon statistical test results compared to OEMDE. The signs ‘+’ and ‘−’ show significance of the OEMDE over the competitive algorithm and vice-versa, respectively; ‘=’ shows neither is significant. \( D=50 \) and \( N_F=6 \).

| P   | OEMDE  | OEMDE | OEMDE | PAWE | pPAWE | ODE  | Error | W | Error | W | Error | W | Error | W | Error | W |
|-----|--------|--------|--------|-------|--------|------|-------|---|-------|---|-------|---|-------|---|-------|---|
| 1   | 4.98e+08 ± 2.81e+05 | 6.11e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 4.98e+08 ± 2.81e+05 | 6.11e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 2   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 3   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 4   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |

TABLE IV: Error and standard deviation of the proposed OEMDE algorithm versus state of the art. The \( W \) represents the Wilcoxon statistical test results compared to OEMDE. The signs ‘+’ and ‘−’ show significance of the OEMDE over the competitive algorithm and vice-versa, respectively; ‘=’ shows neither is significant. \( D=100 \) and \( N_F=6 \).

| P   | OEMDE  | OEMDE | OEMDE | PAWE | pPAWE | ODE  | Error | W | Error | W | Error | W | Error | W | Error | W |
|-----|--------|--------|--------|-------|--------|------|-------|---|-------|---|-------|---|-------|---|-------|---|
| 1   | 4.98e+08 ± 2.81e+05 | 6.11e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 4.98e+08 ± 2.81e+05 | 6.11e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 2   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 3   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 4   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |

TABLE V: Error and standard deviation of the proposed OEMDE algorithm versus state of the art. The \( W \) represents the Wilcoxon statistical test results compared to OEMDE. The signs ‘+’ and ‘−’ show significance of the OEMDE over the competitive algorithm and vice-versa, respectively; ‘=’ shows neither is significant. \( D=200 \) and \( N_F=6 \).

| P   | OEMDE  | OEMDE | OEMDE | PAWE | pPAWE | ODE  | Error | W | Error | W | Error | W | Error | W | Error | W |
|-----|--------|--------|--------|-------|--------|------|-------|---|-------|---|-------|---|-------|---|-------|---|
| 1   | 4.98e+08 ± 2.81e+05 | 6.11e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 4.98e+08 ± 2.81e+05 | 6.11e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 2   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
| 3   | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 9.95e+08 ± 2.81e+05 | 6.41e+08 ± 3.06e+05 | 3.80e+08 ± 1.76e+05 | 5.29e+08 ± 3.08e+05 | + | 4.14e+08 ± 2.13e+05 | + | 5.09e+08 ± 2.85e+05 | + | 5.29e+08 ± 3.08e+05 | + |
shows gradual exploration and exploitation of PAJDE and μJADE algorithms. This is while the increased diversity in OEMDE helps a faster convergence.

### VII. Conclusion

In this paper, an opposition-based ensemble version of micro-differential evolution (OEMDE) algorithm using a pool of mutation schemes is proposed. The OEMDE uses randomized mutation scale factor, such that a different mutation scale factor is assigned to each dimension of each individual. The OEMDE utilizes different mutation schemes for each individual of population, rather than a fixed mutation scheme. It enhances the diversity of population by giving a chance to opposite vectors of the population and involving them during the evolution. The objective is to increase the exploration ability of micro-DE algorithm and to decrease the chance of pre-mature convergence and stagnation by providing more diversity in the generated mutant vectors. The OEMDE is simple to implement and its performance is superior to the classical micro-DE algorithm. Its performance is on a par with ODE. However, OEMDE uses a smaller population size and does not require the setting of mutation scale factor and mutation scheme, which requires less trial and error for parameter setting.

Diversity enhancement is an important factor in micro-DE. Other possibilities to increase diversity needs to be studied at different levels of algorithm for different population sizes and problem dimensionalities. The proposed work can also be extended to multi-objective and many-objective problems.

### References

[1] H. Salehinejad and S. Talebi, “Dynamic fuzzy logic-ant colony system-based route selection system,” Applied Computational Intelligence and Soft Computing, 2010.

[2] H. Salehinejad, S. Rahnamayan, and H. R. Tizhoosh, “Micro-differential evolution: Diversity enhancement and a comparative study,” Applied Soft Computing, vol. 52, pp. 812–833, 2017.

[3] H. Salehinejad, S. Rahnamayan, H. R. Tizhoosh, and S. Y. Chen, “Micro-differential evolution with vectorized random mutation factor,” in Evolutionary Computation (CEC), 2014 IEEE Congress on, pp. 2055–2062. IEEE, 2014.
[4] H. Salehinejad, S. Rahnamayan, and H. R. Tizhoosh, “Exploration enhancement in ensemble micro-differential evolution,” in *Evolutionary Computation (CEC)*, 2016 IEEE Congress on, pp. 63–70, IEEE, 2016.

[5] S. Rahnamayan, H. R. Tizhoosh, and M. M. Salama, “Opposition versus randomness in soft computing techniques,” *Applied Soft Computing*, vol. 8, no. 2, pp. 906–918, 2008.

[6] S. Rahnamayan, J. Jesuthasan, F. Bourennani, H. Salehinejad, and G. F. Naterer, “Computing opposition by involving entire population,” in *Evolutionary Computation (CEC)*, 2014 IEEE Congress on, pp. 1800–1807, IEEE, 2014.

[7] S. Rahnamayan, J. Jesuthasan, F. Bourennani, G. F. Naterer, and H. Salehinejad, “Centroid opposition-based differential evolution,” *International Journal of Applied Metaheuristic Computing (IJAMC)*, vol. 5, no. 4, pp. 1–25, 2014.

[8] H. Salehinejad and S. Rahnamayan, “Effects of centralized population initialization in differential evolution,” in *Computational Intelligence (SSCI)*, 2016 IEEE Symposium Series on, pp. 1–8, IEEE, 2016.

[9] H. Salehinejad, R. Zadeh, R. Liscano, and S. Rahnamayan, “3d localization in large-scale wireless sensor networks: A micro-differential evolution approach,” in *Personal, Indoor, and Mobile Radio Communication (PIMRC)*, 2014 IEEE 25th Annual International Symposium on, pp. 1824–1828, IEEE, 2014.

[10] H. Salehinejad, *Micro-differential evolution: diversity enhancement and comparative study*. PhD thesis, 2014.

[11] H. R. Tizhoosh, “Opposition-based learning: a new scheme for machine intelligence,” in *CIMCA*, vol. 1, pp. 695–701, IEEE, 2005.

[12] H. Salehinejad, S. Rahnamayan, and H. R. Tizhoosh, “Type-II opposition-based differential evolution,” in *Evolutionary Computation (CEC)*, 2014 IEEE Congress on, pp. 1768–1775, IEEE, 2014.

[13] V. Feoktistov, *Differential evolution*. Springer, 2006.

[14] R. Gämperle, S. D. Müller, and P. Koumoutsakos, “A parameter study for differential evolution,” *Advances in intelligent systems, fuzzy systems, evolutionary computation*, vol. 10, pp. 293–298, 2002.

[15] A. Salman, A. P. Engelbrecht, and M. G. Omran, “Empirical analysis of self-adaptive differential evolution,” *European Journal of operational research*, vol. 183, no. 2, pp. 785–804, 2007.

[16] R. Mallipeddi, P. N. Suganthan, Q.-K. Pan, and M. F. Tasgetiren, “Differential evolution algorithm with ensemble of parameters and mutation strategies,” *Applied Soft Computing*, vol. 11, no. 2, pp. 1679–1696, 2011.

[17] X. Yu, X. Wang, J. Cao, and M. Cai, “An ensemble differential evolution for numerical optimization,” *International Journal of Information Technology & Decision Making*, vol. 14, no. 04, pp. 915–942, 2015.

[18] S. Das, A. Ghosh, and S. S. Mulkic, “A switched parameter differential evolution for large scale global optimization—simpler may be better,” in *Mendel 2015*, pp. 103–129, Springer, 2015.

[19] R. Mallipeddi, G. Iacca, P. N. Suganthan, F. Neri, and E. Mininno, “Ensemble strategies in compact differential evolution,” in *Evolutionary Computation (CEC)*, 2011 IEEE Congress on, pp. 1972–1977, IEEE, 2011.

[20] E. Mininno, F. Neri, F. Cupertino, and D. Naso, “Compact differential evolution,” *Evolutionary Computation, IEEE Transactions on*, vol. 15, no. 1, pp. 32–54, 2011.

[21] F. Neri and E. Mininno, “Memetic compact differential evolution for cartesian robot control,” *Computational Intelligence Magazine, IEEE*, vol. 5, no. 2, pp. 54–65, 2010.

[22] G. Wu, R. Mallipeddi, P. Suganthan, R. Wang, and H. Chen, “Differential evolution with multi-population based ensemble of mutation strategies,” *Information Sciences*, vol. 329, pp. 329–345, 2016.

[23] J. Zhang and A. C. Sanderson, “Jade: adaptive differential evolution with optional external archive,” *Evolutionary Computation, IEEE Transactions on*, vol. 13, no. 5, pp. 945–958, 2009.

[24] C. Brown, Y. Jin, M. Leach, and M. Hodgson, “‘mu jade: adaptive differential evolution with a small population,” *Soft Computing*, pp. 1–10, 2015.

[25] I. Fajfar, T. Tuma, J. Puhar, J. Olenšek, and A. Bürmen, “Towards smaller populations in differential evolution,” *J Microelectron Electron Compon Mater*, vol. 42, pp. 152–163, 2012.

[26] X. Ren, Z.-z. Chen, and Z. Ma, “Differential evolution using smaller population,” in *2010 Second International Conference on Machine Learning and Computing*, pp. 76–80, IEEE, 2010.

[27] A. W. Mohamed, H. Z. Sabry, and M. Khorsheid, “An alternative differential evolution algorithm for global optimization,” *Journal of advanced research*, vol. 3, no. 2, pp. 149–165, 2012.

[28] M. Yang, Z. Cai, C. Li, and J. Guan, “An improved adaptive differential evolution algorithm with population adaptation,” in *Proceedings of the 15th annual conference on Genetic and evolutionary computation*, pp. 145–152, ACM, 2013.

[29] G. Iacca, R. Mallipeddi, E. Mininno, F. Neri, and P. N. Suganthan, “Super-fit and population size reduction in compact differential evolution,” in *Memetic Computing (MC)*, 2011 IEEE Workshop on, pp. 1–8, IEEE, 2011.

[30] J. Aalto and J. Lampinen, “A population adaptation mechanism for differential evolution algorithm,” in *Computational Intelligence, 2015 IEEE Symposium Series on*, pp. 1514–1521, IEEE, 2015.

[31] J. Brest and M. S. Maučec, “Population size reduction for the differential evolution algorithm,” *Applied Intelligence*, vol. 29, no. 3, pp. 228–247, 2008.

[32] S. Rahnamayan and H. R. Tizhoosh, “Image thresholding using micro opposition-based differential evolution (micro-ode),” in *Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence)*, IEEE Congress on, pp. 1409–1416, IEEE, 2008.

[33] K. E. Parasopoulos, “Cooperative micro-differential evolution for high-dimensional problems,” in *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, pp. 531–538, ACM, 2009.

[34] M. A. Sotelo-Figueroa, H. J. P. Soberanes, J. M. Carpio, H. J. F. Huacuja, L. C. Reyes, and J. A. S. Alcaraz, “Evolving bin packing heuristic using micro-differential evolution with indirect representation,” in *Recent Advances on Hybrid Intelligent Systems*, pp. 349–359, Springer, 2013.

[35] N. S. Teng, J. Teo, and M. H. A. Hijazi, “Self-adaptive population sizing for a tune-free differential evolution,” *Soft Computing*, vol. 13, no. 7, pp. 709–724, 2009.

[36] W. Wang, H. Wang, H. Sun, and S. Rahnamayan, “Using opposition-based learning to enhance differential evolution: A comparative study,” in *Evolutionary Computation (CEC)*, 2016 IEEE Congress on, pp. 71–77, IEEE, 2016.

[37] H. Liu, Z. Wu, H. Wang, S. Rahnamayan, and C. Deng, “Improved differential evolution with adaptive opposition strategy,” in *Evolutionary Computation (CEC)*, 2014 IEEE Congress on, pp. 1776–1783, IEEE, 2014.

[38] S. Rahnamayan, G. G. Wang, and M. Ventresca, “An intuitive distance-based explanation of opposition-based sampling,” *Applied Soft Computing*, vol. 12, no. 9, pp. 2828–2839, 2012.

[39] “http://coco.gforge.inria.fr/doku.php?id=cec-bb-ob-2015,”