METHODOLOGIES AND APPLICATION

Cellular direction information based differential evolution for numerical optimization: an empirical study

Jingliang Liao1 · Yiqiao Cai¹ · Tian Wang1 · Hui Tian1 · Yonghong Chen1

© Springer-Verlag Berlin Heidelberg 2015

Abstract Differential evolution (DE) is a well-known evolutionary algorithm which has been successfully applied in many scientific and engineering fields. In most DE algorithms, the base and difference vectors for mutation are randomly selected from the current population. That is, the useful population information cannot be fully exploited to guide the search of DE through mutation. Furthermore, the selection of parents in mutation has been verified to be critical for the DE performance. Therefore, to alleviate this drawback and improve the performance of DE, a novel DE algorithm with a directional mutation based on cellular topology is proposed in this study. This proposed algorithm is named as Cellular Direction Information based DE (DE-CDI). In DE-CDI, the cellular topology is employed first to define a neighborhood for each individual of population and then the direction information based on the neighborhood is incorporated into the mutation operator. In this way, DE-CDI not only utilizes the neighborhood information to exploit the regions of better individuals and accelerate convergence, but also introduces the direction information to guide the search to the promising area. To evaluate the performance of the proposed method, DE-CDI is applied to the original DE algorithms, as well as several advanced DE variants. Experimental results demonstrate the high performance of DE-CDI for most DE algorithms studied.

Communicated by V. Loia.

B Yiqiao Cai yiqiao00@163.com; caiyq@hqu.edu.cn

¹ College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China

Keywords Differential evolution · Cellular topology · Neighborhood information · Direction information · Mutation operator · Numerical optimization

1 Introduction

Differential evolution (DE), proposed by Storn and Price [\(Storn and Price 1997](#page-26-0); [Storn et al. 2005](#page-26-1)), is a simple yet efficient evolutionary algorithm for global numerical optimization. It has many attractive characteristics, such as ease to use, compact structure, robustness and speediness. Recently, DE has been extended to handle large-scale, multiobjective, constrained, dynamic and uncertain optimization problems [\(Das and Suganthan 2011](#page-25-0)). Furthermore, DE has been successfully applied to diverse domains of science and engineering, such as pattern recognition [\(Das et al.](#page-25-1) [2008](#page-25-1); [Omran et al. 2005](#page-26-2)[\),](#page-26-3) [chemical](#page-26-3) [engineering](#page-26-3) [\(](#page-26-3)Wang and Jang [2000](#page-26-3); [Lampinen 1999](#page-26-4)[\),](#page-26-5) [engineering](#page-26-5) [design](#page-26-5) [\(](#page-26-5)Joshi and Sanderson [1999](#page-26-5); [Rogalsky et al. 1999](#page-26-6)), signal processing [\(Das and Konar 2006](#page-25-2)[\),](#page-26-7) [satellite](#page-26-7) [communications](#page-26-7) [\(](#page-26-7)Wang and Cai [2015](#page-26-7)) and so on.

In DE, two main factors significantly influence the performance of DE: the control parameters (i.e., population size *NP*, scaling factor *F* and crossover rate *Cr*) and the evolutionary operators (i.e., mutation, crossover and selection). During the past decade, there are many enhanced DE variants proposed in the literature. According to [Neri and Tirronen](#page-26-8) [\(2010](#page-26-8)), these advanced DE variants can be divided into two categories: algorithms with additional components and algorithms with a modified DE structure. In these DE variants, modifications mostly focus on devising the new mutation oper[ators](#page-26-10) [\(Zhang and Sanderson 2009](#page-26-9)[;](#page-26-10) [Das et al. 2009;](#page-25-3) Wang et al. [2014\)](#page-26-10), employing the self-adaptive strategies for control parameters [\(Qin et al. 2009](#page-26-11); [Brest et al. 2006;](#page-25-4) [Yang et al.](#page-26-12)

[2015\)](#page-26-12), proposing the ensemble strategies [\(Qin et al. 2009](#page-26-11); [Wang et al. 2011](#page-26-13); [Tang et al. 2015\)](#page-26-14), developing the hybrid DE with other optimization methods [\(Sun et al. 2005](#page-26-15)), etc.

In DE, the mutant vector, generated by the mutation operator, can be treated as a lead individual to explore the search space and is constructed by adding a difference vector to a base vector. However, these two vectors (i.e., the base and difference vectors) in most DE variants are always randomly selected from the current population. That is, the useful population information cannot be fully utilized to guide the search through mutation.

In order to alleviate this drawback and enhance the performance of DE, a new DE algorithm with a directional mutation based on cellular topology is proposed in this study, which is named as Cellular Direction Information based DE (DE-CDI). In DE-CDI, a cellular topology is employed first to define a neighborhood for each individual of population. Then, the neighbors of each individual are divided into the superior and inferior groups according to their fitness. Finally, the direction information is incorporated into mutation by selecting two vectors from the superior and inferior groups, respectively. Here, all the parents for mutation are selected from the neighbors of the current vector. In this way, DE-CDI not only utilizes the information of neighboring individuals to exploit the regions of minima and accelerate convergence, but also incorporates the direction information of population to guide the search to the promising area. Therefore, the population information composed by neighborhood and direction information can be simultaneously and fully utilized in DE-CDI to guide the search.

To evaluate the effectiveness of the proposed method, DE-CDI is applied to six original DE algorithms, as well as several advanced DE variants. With the analysis of the extensive experiments on a set of benchmark functions, we can clearly find that DE-CDI is able to improve the performance of most DE algorithms studied.

The main contributions of this study include the following:

- Both neighborhood and direction information are fully and simultaneously utilized in selecting parents for mutation to guide the search of DE.
- DE-CDI provides a simple yet powerful approach for improving the search ability of DE. In addition, DE-CDI can be easily applied to other DE variants.
- The extensive experiments are carried out to show the effectiveness of DE-CDI. The results demonstrate that CDI is able to enhance the performance of most DE algorithms studied.

The rest of this paper is organized as follows: In Sect. [2,](#page-1-0) the original DE is introduced. Section [3](#page-2-0) briefly reviews some related work. The proposed DE-CDI is presented in detail in Sect. [4.](#page-3-0) In Sect. [5,](#page-5-0) experimental results are reported. Finally, the conclusions are drawn in Sect. [6.](#page-24-0)

2 DE

In this study, DE is for solving the numerical optimization problem [\(Storn and Price 1997\)](#page-26-0). Without loss of generality, the optimization problem which is considered to be minimized is $f(X)$, $X \in R^D$, and *D* is the dimension of the decision variables. DE evolves a population of vectors representing the candidate solutions. Every individual is denoted as $X_{i,G} = (x_{i,G}^1, x_{i,G}^2, \dots, x_{i,G}^D)$, where $i = 1, 2, \dots, NP$, *NP* is the population size and *G* is the current generation.

The DE algorithmic schemes can be classified using the notation "DE/x/y/z", where DE means differential evolution algorithm, *x* means the method of selecting the parent that constitutes the base vector, *y* means the number of difference vectors that are used to perturb *x* and *z* stands for the crossover type employed.

2.1 Initialization

In DE, the initial population should cover the entire search space as much as possible by uniformly randomizing individuals within the search space. Here, the *j*th parameter in the *i*th individual is initialized as follows:

$$
x_{i,G}^j = L_j + \text{rand}(0, 1) \times (U_j - L_j), \tag{1}
$$

where $\text{rand}(0, 1)$ represents a uniformly distributed random number within the range [0, 1], and L_i and U_j represent the lower and upper bounds of the *j*th variable, respectively.

2.2 Mutation

DE employs the mutation strategy to generate a mutant vector $V_{i,G}$ with respect to each individual $X_{i,G}$. Here, several widely used mutation strategies are listed as follows:

• DE/rand/1

$$
V_{i,G} = X_{r1,G} + F \times (X_{r2,G} - X_{r3,G})
$$
 (2)

• DE/rand/2

$$
V_{i,G} = X_{r1,G} + F \times (X_{r2,G} - X_{r3,G}) + F \times (X_{r4,G} - X_{r5,G})
$$
(3)

• DE/best/1

$$
V_{i,G} = X_{\text{best},G} + F \times (X_{r1,G} - X_{r2,G})
$$
 (4)

• DE/best/2

$$
V_{i,G} = X_{\text{best},G} + F \times (X_{r1,G} - X_{r2,G}) + F \times (X_{r3,G} - X_{r4,G})
$$
(5)

• DE/current-to-best/1

$$
V_{i,G} = X_{i,G} + F \times (X_{\text{best},G} - X_{i,G}) + F \times (X_{r1,G} - X_{r2,G})
$$
(6)

• DE/rand-to-best/1

$$
V_{i,G} = X_{r1,G} + F \times (X_{\text{best},G} - X_{r1,G}) + F \times (X_{r2,G} - X_{r3,G}),
$$
 (7)

where F is called the mutation scaling factor, and the indices *r*1, *r*2, *r*3, *r*4 and *r*5 are mutually exclusive integers randomly generated within the range [1, *NP*], which are also different from the index *i*. $X_{\text{best},G}$ is the best individual vector at generation *G*. More details of them can be found in [Storn and Price](#page-26-0) [\(1997](#page-26-0)), [Storn et al.](#page-26-1) [\(2005](#page-26-1)[\)](#page-25-0) [and](#page-25-0) Das and Suganthan [\(2011](#page-25-0)).

2.3 Crossover

The crossover operator is applied to each pair of *Xi*,*^G* and $V_{i,G}$ to generate a trial vector $U_{i,G}$. There are two kinds of crossover scheme: binomial and exponential [\(Storn and Price](#page-26-0) [1997;](#page-26-0) [Storn et al. 2005](#page-26-1)). The binomial crossover, as it is widely used, is outlined as follows:

$$
u_{i,G}^j = \begin{cases} v_{i,G}^j & \text{if } \text{rand}(0,1) \le Cr \text{ or } j = j_{\text{rand}} \\ x_{i,G}^j & \text{otherwise,} \end{cases}
$$
 (8)

where $Cr \in [0, 1]$ is called the crossover rate. *j*_{rand} is a randomly chosen integer in the range [1, *D*]. If $u_{i,G}^j$ is out of the boundary, it will be reinitialized within the range $[L_i, U_j]$.

2.4 Selection

The selection operator selects the better one from each pair of $X_{i,G}$ and $U_{i,G}$ according to their fitness values for the next generation. The selection operator is given as follows:

$$
X_{i,G} = \begin{cases} U_{i,G} & \text{if } f(U_{i,G}) \le f(X_{i,G}) \\ X_{i,G} & \text{otherwise} \end{cases}
$$
(9)

3 Related work

This section briefly reviews some recently improved DE variants that utilize the population information, especially neighborhood and direction information, in the mutation operators.

The neighborhood concepts are usually used to enhance the performance of DE by the following ways: offspring generation operator, structured population and local search [\(Epitropakis et al. 2011\)](#page-26-16). There are two main types of neighborhood information: index-based and distance-based. For the index-based neighborhood, the neighbors of each individual do not necessary lie in the vicinity of its topological region in the search space. Many DE variants utilize the index-based neighborhood information with the structured population. In these DE variants, the individuals for mutation are selected according to a neighbor list constructed from the population topologies. Two main canonical kinds of structured population in DE is employed in literature, i.e., [cellular](#page-26-18) [DE](#page-26-18) [\(cDE\)](#page-26-18) [\(Noman and Iba 2011](#page-26-17)[;](#page-26-18) Noroozi et al. [2011](#page-26-18); [Dorronsoro and Bouvry 2010](#page-26-19)) and distributed DE (dDE) [\(Weber et al. 2011,](#page-26-20) [2010](#page-26-21); [Neri et al. 2011](#page-26-22)). In [De Falco et al.](#page-25-5) [\(2014](#page-25-5)), the impact of several network topologies on the performance of distributed differential evolution was investigated. By using the ring topology, a neig[hborhood-based](#page-25-3) [mutation](#page-25-3) [operator](#page-25-3) [was](#page-25-3) [proposed](#page-25-3) [in](#page-25-3) Das et al. [\(2009\)](#page-25-3). In [Omran et al.](#page-26-23) [\(2006\)](#page-26-23), the self-adaptive DE was enhanced using a ring neighborhood topology. In [Omran et al.](#page-26-24) [\(2009](#page-26-24)), by employing the concept of index neighborhoods, the bare bones DE algorithm was proposed. To balance the exploration and exploitation, an adaptive memetic DE was proposed [\(Piotrowski 2013](#page-26-25)), where the ring topology was used to construct the neighborhood of individuals for local model. In [Hu et al.](#page-26-26) [\(2014\)](#page-26-26), the ring neighborhood topology was employed to improve a memetic DE algorithm. Recently, several population topologies, e.g. distributed, cellular, ring, small-world, have been introduced in DE to improve its performance [\(Dorronsoro and Bouvry 2011](#page-26-27)). For the distance-based neighborhood, the individuals are regarded as the neighbors of an individual when they locate in the vicinity of its topological region in the search space. In [Epitropakis et al.](#page-26-16) [\(2011\)](#page-26-16), a proximity-based DE framework (ProDE) was proposed using an affinity matrix based on the Euclidean distance to select the individuals for mutation. For improving the performance of DE, the learning-enhanced DE (LeDE) was proposed in [Cai et al.](#page-25-6) [\(2012\)](#page-25-6). In LeDE, the neighbors of each individual were defined based on the identified clusters. In [Liang et al.](#page-26-28) [\(2014](#page-26-28)), a novel DE variant that employed a modified Fitness Euclidean-distance Ratio technique was proposed. In [Sarkar et al.](#page-26-29) [\(2015\)](#page-26-29), an effective new grouping strategy namely adaptive clustering and reclustering was proposed based on Fuzzy *c*-means clustering technique and was integrated with a hybrid of crowding niching technique and DE. To enhance the performance on multi-modal problems, an ensemble crowding DE with neighborhood mutation was proposed [\(Hui and Suganthan](#page-26-30) [2013](#page-26-30)).

Due to that difference vectors in mutation of DE are always constructed in a random manner, the direction concept is incorporated into many DE variants to enhance the exploration ability. In [Wang and Xiang](#page-26-31) [\(2008\)](#page-26-31), a new mutation strategy, named as DE/rand/±mean, was proposed. In this strategy, the population is partitioned into two groups and the difference vector is constructed by randomly selecting the vecto[rs](#page-26-32) [from](#page-26-32) [the](#page-26-32) [two](#page-26-32) [groups](#page-26-32) [respectively.](#page-26-32) [In](#page-26-32) Fan and Lampinen [\(2003](#page-26-32)), a trigonometric mutation DE (TDE) was proposed with a probabilistic triangle mutation strategy that incorporates the direction information into DE. Recently, a novel DE framework, DE with neighborhood and direction information (NDi-DE), was proposed by designing three types of direction information for mutation to guide the search [\(Cai and Wang 2013\)](#page-25-7). After that, adaptive operator selection mechanism was introduced into NDi-DE for different mutation strategies to dynamically balance the expl[oration](#page-25-9) [and](#page-25-9) [exploitation](#page-25-9) [\(Cai et al. 2015](#page-25-8)[\).](#page-25-9) [In](#page-25-9) Bi and Xiao [\(2011\)](#page-25-9), using the direction information with the current best solution and the best previous solution of each individual, a classification-based self-adaptive DE was proposed. In [Iorio and Li](#page-26-33) [\(2006](#page-26-33)) and [Liu et al.](#page-26-34) [\(2009\)](#page-26-34), the direction information was also employed to enhance the performance of D[E](#page-26-35) [for](#page-26-35) [the](#page-26-35) [multi-objective](#page-26-35) [optimization.](#page-26-35) [In](#page-26-35) Zhang and Yuen [\(2015](#page-26-35)), with a difference vector pool containing good direction information, a directional mutation DE (DMDE) was proposed to speed up the convergence of DE. Based a distributed topology, a directional strategy was employed to enhance convergence speed by using the direction information in [Gou et al.](#page-26-36) [\(2015](#page-26-36)).

4 DE-CDI

In this section, the proposed algorithm, i.e., DE-CDI, is described in detail. First, the motivations of this study are given. Second, two main components of CDI based mutation, i.e., cellular topology-based neighborhood and neighborhood-based directional mutation, are presented. Third, the complete proposed framework of DE-CDI is shown. Finally, the complexity of DE-CDI is discussed.

4.1 Motivations

As mentioned in Sect. [3,](#page-2-0) many attempts with neighborhood or/and direction information of population are effective to guide the search of DE. Furthermore, the selection of parents in mutation has been verified to be critical for the DE performance [\(Noman and Iba 2011](#page-26-17); [Noroozi et al. 2011](#page-26-18)[;](#page-26-19) Dor-ronsoro and Bouvry [2010](#page-26-19); [Weber et al. 2011,](#page-26-20) [2010](#page-26-21); [Neri et al.](#page-26-22) [2011;](#page-26-22) [Das et al. 2009;](#page-25-3) [Omran et al. 2006](#page-26-23), [2009;](#page-26-24) [Cai et al.](#page-25-6) [2012;](#page-25-6) [Wang and Xiang 2008](#page-26-31); [Fan and Lampinen 2003](#page-26-32)[;](#page-25-7) Cai and Wang [2013](#page-25-7); [Bi and Xiao 2011\)](#page-25-9). However, in most DE algorithms, on the one hand, the base and difference vectors are always selected randomly or locally. That it, both neighborhood and direction information are not simultaneously and fully utilized during the evolutionary process of DE. On the other hand, the individuals are often guided only with the best individual of population, even if the best individual is far away from the global optimum. By this way, the individuals may be trapped in local optimum more frequently, especially for the multimodal problems.

Based on these considerations, a directional mutation based on cellular topology is introduced into the mutation operator of DE to exploit the neighborhood and direction information of population simultaneously. Thus, a simple and effective framework, DE-CDI, is proposed. The primary idea of DE-CDI is to incorporate cellular topology and direction information into mutation to enhance the performance of DE. In CDI-based mutation, there are two main components: cellular topology-based neighborhood and neighborhood-based directional mutation.

4.2 Cellular topology-based neighborhood

In order to define a neighborhood for each individual, the cellular topology is introduced into the population. In the neighborhood-based on cellular topology, the population is spatially structured in a two-dimensional toroidal grid. Each grid-point contains exactly one individual, and the neighborhood of an individual is defined by the surrounding grid-points [\(Noman and Iba 2011](#page-26-17)). Here, the cellular topology-based neighborhood is called Cn, where *n* indicates the number of total individuals in the neighborhood (including the current individual itself and its closest neighbors). Figure [1](#page-3-1) illustrates the concept with C9.

When employing Cn in DE-CDI, the population is structured in a regular grid of two dimensions and each individual can only interact with its surrounding neighboring solutions during the evolutionary process. As shown in Fig. [1,](#page-3-1) the individual at the center position in the grid $(i.e., X_i)$ can generate an offspring only with its eight closest neighbors (i.e., the solid circles). That is, the interactions among individuals are

restricted to the closest ones, which will lead to slow diffusion of information throughout the population.

In this way, the neighborhood information based on cellular topology will be exploited fully to guide the search of DE-CDI through selecting parents from the neighbors. Although various neighborhood topologies, e.g., distributed, ring or small-worl[d](#page-26-27) [topology,](#page-26-27) [are](#page-26-27) [proposed](#page-26-27) [for](#page-26-27) [DE](#page-26-27) [\(](#page-26-27)Dorronsoro and Bouvry [2011\)](#page-26-27), the cellular topology can obtain the more attractive and robust results in our preliminary tests. In the future work, the detailed comparisons of the DE-CDI variants with other neighborhood topologies will be made.

4.3 Neighborhood based directional mutation

Based on the defined neighborhood with cellular topology, direction information is introduced into mutation of DE by selecting neighbors to construct difference vectors. The new proposed mutation operator is named as neighborhood-based directional mutation (NDM). In NDM, all the parents in mutation are selected from the neighbors and the difference vectors are constructed by directing at a better neighbor from a worse one. Concretely, according to the fitness of the base vector, all of the neighbors are partitioned into the superior and inferior groups, and the difference vectors are constructed by randomly selecting two neighbors from the superior and inferior groups respectively.

In order to illustrate how NDM works, a simple example for DE/rand/1 with direction information based on C9 is shown in Fig. [2.](#page-4-0) As illustrated in Fig. [2,](#page-4-0) the base vector, $X_{r1,G}$, is randomly selected from the neighbors of $X_{i,G}$ first. Then, based on the fitness of $X_{r1,G}$, all the eight neighbors of *Xi*,*^G* are partitioned into the superior and inferior groups. Finally, the terminal point of the difference vector, $X_{r2,G}$, is randomly selected from the superior group, and the start point, $X_{r3,G}$, is randomly selected from the inferior group. In this way, the direction information is incorporated into mutation for guiding the search of DE to the promising area.

4.4 The framework of DE-CDI

The framework of DE-CDI is illustrated in Fig. [3.](#page-4-1) The pseudo-code of DE-CDI with DE/rand/1 (named as DE-

Fig. 2 DE/rand/1 with direction information based on C9

Fig. 3 Flowchart of DE-CDI

CDI/rand/1) is also shown in Algorithm 1 where the differences with respect to DE/rand/1 are highlighted with "∗". It is clear that DE-CDI only affects the mutation operator and it can be directly and easily applied to most DE algorithms.

When applying DE-CDI to the mutation operator that employs the best individual (e.g., DE/best/1, DE/ currentto-best/1 and DE/rand-to-best/1), the best neighbor of the current individual will be served as the best individual for mutation. That is, the best individual is refined in these mutation operators. As for constructing the difference vector, when the base vector is the best or worst vector in the neighborhood, two vectors will be randomly selected from the neighborhood of the base vector and the difference vector will be constructed by directing at the better vector from the worse one.

4.5 The complexity of DE-CDI

Compared with the original DE algorithm, the additional computation of DE-CDI depends on the selection of parents for mutation, i.e., steps 5–7 in Algorithm 1. During one generation, selecting individuals to construct difference vector based on Cn will take *n* − 1 times of comparison with the base vector. Therefore, the additional computation of DE-CDI is $O(n)$. Since the complexity of the original DE algorithm is $O(G_{\text{max}} \times NP \times D)$ where G_{max} is the maximal number of generation, the total complexity of DE-CDI is $O(G_{\text{max}} \times NP \times \text{max}(D, n)).$

Algorithm 1 DE-CDI/rand/1

- 1: Generate the initial population P^G and set $G = 1$;
- 2: Evaluate the fitness for each individual in P^G ;
- 3: **While** the terminated condition is not satisfied **do**
- 4: **For** each individual $X_{i,t}$ **do**
5: * Randomly select the base
- 5: ∗ Randomly select the base vector *Xr*¹ from the neighborhood of $X_{i,G}$;
- 6: ∗ Partition all the neighbors of *Xi*,*^G* into the superior and inferior groups by comparing their fitness with *Xi*,*G*;
- 7: ∗ Randomly select *Xr*² and *Xr*³ from the superior and inferior groups respectively;
- 8: Use Eq. [\(2\)](#page-1-1) to generate a mutant vector;
- 9: Use Eq. [\(8\)](#page-2-1) to generate a trial vector;
- 10: Use Eq. [\(9\)](#page-2-2) to determine the survived vector:
- 11: **End For**
- 12: Set $G = G + 1$;
- 13: **End while**

5 Experimental result and analysis

In order to evaluate the performance of DE-CDI, 25 benchmark functions from the CEC2005 special session on realparameter optimization [\(Suganthan et al. 2005](#page-26-37)) and three real-w[orld](#page-25-10) [problems](#page-25-10) [\(Eshelman et al. 1997](#page-26-38)[;](#page-25-10) Das and Suganthan [2010](#page-25-10)) are used. In this section, the benchmark functions are presented first. Then, the experimental setup is given. Finally, the simulation results are analyzed and discussed.

5.1 Benchmark functions

25 benchmark functions are used, denoted as *F*1–*F*25, which are from the special session on real-parameter optimization of the 2005 IEEE Congress on Evolutionary Computation (CEC 2005) [\(Suganthan et al. 2005](#page-26-37)). According to [Suganthan et al.](#page-26-37) [\(2005](#page-26-37)), these functions can be categorized into four groups: unimodal functions (*F*1–*F*5), basic multimodal functions (*F*6–*F*12), expanded multimodal functions (*F*13–*F*14) and hybrid composition functions (*F*15–*F*25). More details of them can be found in [Rogalsky et al.](#page-26-6) [\(1999\)](#page-26-6).

5.2 Parameter settings

For a fair comparison, the same random initial population is used to evaluate the performance of different algorithms. The parameters are set as follows unless a change is mentioned:

- Dimension of each function (*D*): 30 and 50.
- Population size (*NP*): 100.
- Mutation factor (F) : 0.5.
- Crossover factor (*Cr*): 0.9.
- Number of neighbors (*n*): 13 (i.e., C13).
- Maximum number of function evaluations (MNFEs): $10^4 \times D$.
- Number of runs (NumR): 25.

In the experimental study, comparisons between six original DE algorithms (i.e., DE/rand/1, DE/rand/2, DE/best/1, DE/best/2, DE/current-to-best/1 and DE/ rand-to-best/1) and their corresponding DE-CDI algorithms are made first. Then, we compare the performance of DE-CDI with advanced DE variants, i.e., CoDE [\(Wang et al. 2011\)](#page-26-13), jDE [\(Brest et al.](#page-25-4) [2006](#page-25-4)), ODE [\(Rahnamayan et al. 2008](#page-26-39)) and SaDE [\(Qin et al.](#page-26-11) [2009](#page-26-11)). All the parameters of these DE variants are set as their original papers except *NP* in CoDE and SaDE. For comparing the algorithms with the same number of neighbors, e.g., $n = 25$ (C25) or 49 (C49), *NP* in CoDE and SaDE is set to 100. Simulations are carried out on an Intel Core i3 duo PC with 3.30-GHz CPU and 4 GB RAM.

To show the significant differences among the competitors, several nonparametric statistical tests [\(García et al.](#page-26-40) [2009](#page-26-40); [Derrac et al. 2011](#page-26-41)) are also carried out by the KEEL software [\(Alcalá-Fdez et al. 2015](#page-25-11)). The results of singleprob[lem](#page-26-40) [Wilcoxon](#page-26-40) [signed-rank](#page-26-40) [test](#page-26-40) [at](#page-26-40) $\alpha = 0.05$ (García et al. [2009;](#page-26-40) [Derrac et al. 2011\)](#page-26-41) are shown in the tables as "+/ = /-", which means that DE-CDI wins, ties and loses on the specific function when compared with its competitor.

5.3 Comparison with original DE algorithms

Six original DE mutation operators (see Eqs. [2–](#page-1-1)[7\)](#page-2-3) are used here. The results for all the functions at 30*D* and 50*D* are shown in Tables [1](#page-6-0) and [2,](#page-8-0) respectively. The better values in terms of mean solution error and standard deviation compared between DE and its corresponding DE-CDI variants are highlighted in boldface.

For the functions at 30*D*, the results of Table [1](#page-6-0) show that DE-CDI can provide significantly better results than the corresponding original DE algorithms on most functions. Specifically, for DE/best/1, DE-CDI exhibits substantial performance improvements on 24 out of 25 functions. For DE/current-to-best/1 and DE/rand-to-best/1, DE-CDI is significantly better on 21 and 21 functions, respectively. For DE/rand/2, DE-CDI can effectively enhance its exploitative ability and yield significant improvements on 22 functions. For DE/rand/1 and DE/best/2, DE-CDI is significantly better on 11 and 9 functions, respectively.

For the functions at 50*D*, the results of Table [2](#page-8-0) also show that DE-CDI is consistently superior to most of the corresponding DE algorithms. DE-CDI/best/1 is significantly better than DE/best/1 on 22 functions, while DE-CDI/currentto-best/1 and DE-CDI/rand-to-best/1 are significantly better than their corresponding original DE algorithms on 21 and 21 functions, respectively. For DE/rand/1 and DE/rand/2, DE-CDI significantly outperforms them on 10 and 22 functions, respectively. For DE/best/2, DE-CDI is also significantly better on 12 functions.

Table 1 continued **Table 1** continued

 $\underline{\textcircled{\tiny 2}}$ Springer

Table 2 continued

Fig. 4 Convergence graphs of the original DE algorithm and the corresponding DE-CDI for the selected functions at 30*D*

Figures [4](#page-10-0) and [5](#page-10-1) also show that the convergence speed of DE-CDI is better than the corresponding original DE algorithms for most selected functions at 30*D* and 50*D*.

Furthermore, to show the significant differences between DE-CDI and its corresponding DE algorithm, the multiproblem Wilcoxon signed-rank test [\(García et al. 2009](#page-26-40); [Derrac et al. 2011](#page-26-41)) is also carried out on all the functions at 30*D* and 50*D*. The results are shown in Tables [3](#page-10-2) and [4,](#page-10-3) respectively. For the functions at 30*D*, it is clear that DE-CDI can obtain the higher *R*+ values than *R*− values in all the cases. In addition, the *p* values in most cases are less

Fig. 5 Convergence graphs of the original DE algorithm and the corresponding DE-CDI for the selected functions at 50*D*

than 0.05, which means that DE-CDI is significantly better than most original DE algorithms overall. For the functions at 50*D*, the similar results can also be obtained from Table [4.](#page-10-3)

5.4 Comparison with advanced DE algorithms

In order to evaluate the effectiveness of DE-CDI for the advanced DE variants, four DE variants are used, i.e., CoDE [\(Wang et al. 2011](#page-26-13)[\),](#page-26-39) [jDE](#page-26-39) [\(Brest et al. 2006](#page-25-4)[\),](#page-26-39) [ODE](#page-26-39) [\(](#page-26-39)Rahnamayan et al. [2008](#page-26-39)) and SaDE [\(Qin et al. 2009](#page-26-11)). The results

Algorithm	$+/-$	$R+$	$R-$	<i>p</i> value	$\alpha = 0.05$	$\alpha = 0.1$
DE-CDI/rand/1 vs DE/rand/1	11/10/4	228	97	7.58E-02	$=$	
DE-CDI/rand/2 vs DE/rand/2	22/3/0	315	10	$3.60E - 0.5$	$^+$	$^+$
DE-CDI/best/1 vs DE/best/1	22/2/1	318		$2.70E - 0.5$	$^+$	
DE-CDI/best/2 vs DE/best/2	9/11/5	200.5	124.5	$3.00E - 01$	$=$	
DE-CDI/current-to-best/1 vs DE/current-to-best/1	21/2/2	319	6	$2.40E - 0.5$	$^+$	
DE-CDI/rand-to-best/1 vs DE/rand-to-best/1	21/3/1	295	5	$3.20E - 05$	$^+$	

Table 3 Results of the multi-problem Wilcoxon's test for DE-CDI versus the original DE algorithm for all the functions at 30*D*

Table 4 Results of the multi-problem Wilcoxon's test for DE-CDI versus the original DE algorithm for all the functions at 50*D*

Algorithm	$+/-$	$R+$	$R-$	<i>p</i> value	$\alpha = 0.05$	$\alpha = 0.1$
DE-CDI/rand/1 vs DE/rand/1	10/5/10	161.5	138.5	$7.32E - 01$	$=$	
DE-CDI/rand/2 vs DE/rand/2	22/3/0	323.5	1.5	$1.30E - 0.5$	$^{+}$	
DE-CDI/best/1 vs DE/best/1	23/1/1	298	2	$2.10E - 0.5$	$^{+}$	
DE-CDI/best/2 vs DE/best/2	12/9/4	221	79	$4.11E - 02$	$^+$	
DE-CDI/current-to-best/1 vs DE/current-to-best/1	22/1/2	296	4	$2.80E - 0.5$	$^+$	
DE-CDI/rand-to-best/1 vs DE/rand-to-best/1	22/2/1	296	4	$2.80E - 0.5$	$^+$	

Table 5 Mean and standard deviation of the best error values obtained by DE-CDI and the advanced DE algorithm on all the functions at 30 D

Table 6 Mean and standard deviation of the best error values obtained by DE-CDI and the advanced DE algorithm on all the functions at 50 D

Fig. 6 Convergence graphs of the advanced DE variants and the corresponding DE-CDI for the selected functions at 30*D*

In Table [5,](#page-11-0) it is obvious that DE-CDI exhibits significant improvements for most advanced DE variants. Specifically, for the DE variants with ensemble strategies, CoDE-CDI is significantly better than CoDE on 22 functions, while SaDE-CDI significantly outperforms SaDE on five functions. For jDE, DE-CDI significantly outperforms it on eight functions and is outperformed by it on two functions. For ODE, DE-CDI is significantly better on nine functions and is worse on two functions.

For the functions at 50*D* in Table [6,](#page-12-0) DE-CDI also significantly improves the performance of most DE variants. For CoDE and SaDE, DE-CDI is significantly better on 22 and 7 functions, respectively. For jDE, DE-CDI is significantly

Fig. 7 Convergence graphs of the advanced DE variants and the corresponding DE-CDI for the selected functions at 50*D*

better on five functions. For ODE, DE-CDI can obtain the significant better results on six functions.

From Figs. [6](#page-13-0) and [7,](#page-13-1) it is clear that DE-CDI is better than the advanced DE variants in terms of the convergence speed for most selected functions at 30*D* and 50*D*.

Furthermore, the multi-problem Wilcoxon signed rank tests at $\alpha = 0.05$ and $\alpha = 0.1$ are employed to show the significant differences between DE-CDI and its corresponding DE variant, and the results are shown in Tables [7](#page-13-2) and [8.](#page-13-3) It is obvious that DE-CDI can obtain the higher R + values than *R*− values in most cases. These results indicate that DE-CDI is able to improve the performance of most advanced DE variants overall.

Summarily, the results of Tables [5,](#page-11-0) [6,](#page-12-0) [7](#page-13-2) and [8](#page-13-3) indicate that DE-CDI can also bring the beneficial to most advanced DE variants studied.

Table 7 Results of the multi-problem Wilcoxon's test for DE-CDI versus the advanced DE algorithm for all the functions at 30*D*

Algorithm	$+/-$	$R+$	$R-$	<i>p</i> value	$\alpha = 0.05$	$\alpha = 0.1$
CoDE-CDI vs CoDE	22/3/0	320		$1.60E - 05$	$\hspace{0.1mm} +\hspace{0.1mm}$	
jDE-CDI vs jDE	8/15/2	197	128	$3.30E - 01$	$=$	
ODE-CDI vs ODE	9/14/2	269.5	55.5	$3.71E - 03$		
SaDE-CDI vs SaDE	5/15/5	166.5	133.5	$6.27E - 01$	$=$	

Table 10 Mean and standard deviation of the best error values obtained by DE-CDI and NDi-DE on all the functions at 50D **Table 10** Mean and standard deviation of the best error values obtained by DE-CDI and NDi-DE on all the functions at 50*D*

Table 11 Results of the multi-problem Wilcoxon's test	Algorithm	$+/-$	$R+$	$R-$	p value	$\alpha = 0.05$	$\alpha = 0.1$
for DE-CDI versus NDi-DE for	DE-CDI/rand/1 vs NDi-DE/rand/1	3/13/9	133	192	$1.00e + 00$	$=$	
all the functions at $30D$	DE-CDI/best/1 vs NDi-DE/best/1	13/11/1	182	118	$3.53E - 01$	$=$	$=$
	CoDE-CDI vs NDi-CoDE	21/3/1	285.5	14.5	$6.70E - 05$	$+$	$+$
	ODE-CDI vs NDi-ODE	1/16/8	101.5	198.5	$1.00e + 00$	$=$	
Table 12 Results of the							
	Algorithm	$+/-$	$R+$	$R-$	p value	$\alpha = 0.05$	$\alpha = 0.1$
multi-problem Wilcoxon's test for DE-CDI versus NDi-DE for	DE-CDI/rand/1 vs NDi-DE/rand/1	0/14/11	63.5	261.5	$1.00e + 00$	$\overline{}$	
all the functions at $50D$	DE-CDI/best/1 vs NDi-DE/best/1	21/3/1	298	2	$1.70E - 0.5$	$+$	$^{+}$
	CoDE-CDI vs NDi-CoDE	20/2/3	271.5	53.5	$3.11E - 03$	$+$	$+$

Table 13 Average computational time (in seconds) used by all the DE algorithms and their corresponding NDi-DE and DE-CDI variants

Ratio1 means the value that the cost of DE-CDI is divided by that of the original DE algorithm. Ratio2 means defined as the value that the cost of NDi-DE is divided by that of the DE-CDI

Table 14 Mean and standard deviation of the best error values obtained by DMDE, cDMDE and DE-CDI on all the functions at 30D **Table 14** Mean and standard deviation of the best error values obtained by DMDE, cDMDE and DE-CDI on all the functions at 30*D*

Table 15 Results of the multi-problem Wilcoxon's test between DMDE, cDMDE and DE-CDI for all the functions at 30*D*

Upper diagonal of level significance $\alpha = 0.9$, lower diagonal level of significance $\alpha = 0.95$ a The method in the row improves the method of the column

^b The method in the column improves the method of the row

5.5 Comparison with NDi-DE

For investigating the effectiveness of DE-CDI with both neighborhood and direction information, a DE framework, i.e., NDi-DE [\(Cai and Wang 2013](#page-25-7)), is considered. Two original DE algorithms, i.e., DE/rand/1 and DE/best/1, and two advanced DE variants, i.e., CoDE and ODE, are used here for comparison. The results for all the functions at 30*D* and 50*D* are shown in Tables [9](#page-14-0) and [10,](#page-15-0) respectively.

In Table [9,](#page-14-0) we can find that DE-CDI can obtain the comparable results to NDi-DE for the functions at 30*D*. For DE/rand/1 and DE/best/1, DE-CDI is significantly better than NDi-DE on 3 and 13 functions, respectively. For CoDE and ODE, DE-CDI can obtain the significant better results on 21 and 1 functions, respectively.

For the functions at 50*D*, the results of Table [10](#page-15-0) show that DE-CDI consistently perform as well as NDi-DE. Specifically, for DE/rand/1 and DE/best/1, DE-CDI is significantly better than NDi-DE on 0 and 21 functions, respectively. For CoDE and ODE, DE-CDI significantly outperforms NDi-DE on 20 and 6 functions, respectively.

Furthermore, the multi-problem Wilcoxon signed rank tests at $\alpha = 0.05$ and $\alpha = 0.1$ are carried out between DE-CDI and NDi-DE. Tables [11](#page-16-0) and [12](#page-16-1) present the results for functions at 30*D* and 50*D*, respectively. For the functions at 30*D*, it is clear that DE-CDI can obtain the higher *R*+ values than *R*− values in the cases of DE/best/1 and CoDE, while the lower *R*+ values than *R*− values in DE/rand/1 and ODE. In addition, the *p* values in the case of CoDE are less than 0.05, which means that DE-CDI is significantly better than NDi-DE overall. For the functions at 50*D*, the results of the multi-problem Wilcoxon signed rank tests show that DE-CDI is significantly better than NDi-DE in the cases of DE/best/1 and CoDE and is outperformed by it in the case of DE/rand/1.

As discussed in Sect. [4.5,](#page-4-2) the complexity of DE-CDI is $O(G_{\text{max}} \times NP \times \text{max}(D, n))$. In [Cai and Wang](#page-25-7) [\(2013\)](#page-25-7), the complexity of NDi-DE is $O(G_{\text{max}} \times NP \times NP \times D)$. To show the efficiency of DE-CDI, DE-CDI is compared with NDi-DE in terms of the average runtime on all the functions at 30*D* and 50*D*. The results are shown in Table [13.](#page-16-2) In Table [13,](#page-16-2) the runtime value for each function means the average overhead of all the DE algorithms considered in this study on the corresponding function. In Table [13,](#page-16-2) the value of Ratio1 means the cost of DE-CDI is divided by that of the original DE algorithm, and the value of Ratio2 means the cost of NDi-DE is divided by that of DE-CDI. The last column "Avg." means the average value for all the functions.

From Table [13,](#page-16-2) it is clear that the Ratio1 values are close to 1 for all the functions at 30*D* and 50*D*. It indicates that the additional computational cost of DE-CDI is trivial in all the cases when compared with the original DE algorithms. When compared with NDi-DE, the Ratio2 values are higher than 2 for most of *F*1–*F*14, which means that DE-CDI is more efficient than NDi-DE for these functions. For the functions *F*15–*F*25, most Ratio2 values are lower than 1.1. Different from *F*1–*F*14, the evaluations of *F*15–*F*25 are costly. Thus, the overhead of additional computational cost of NDi-DE becomes trivial, which has been verified in [Cai and Wang](#page-25-7) [\(2013](#page-25-7)).

Overall, compared with NDi-DE, on the one hand, DE-CDI does not need to select the type of direction information and the scaling factor for different mutation operators. On the other hand, the results in Tables [9,](#page-14-0) [10,](#page-15-0) [11,](#page-16-0) [12](#page-16-1) and [13](#page-16-2) clearly show that DE-CDI can obtain the better or comparable results to NDi-DE with low complexity.

5.6 Comparison with DMDE

In order to further show the high performance of DE-CDI, a recently proposed DE variant with a directional mutation operator (DMDE) [\(Zhang and Yuen 2015](#page-26-35)), is considered for comparison. In DMDE, a pool of difference vectors between the child and the parent individuals is calculated when the fitness is improved at a generation, and the difference vectors are randomly selected and incorporated into the mutation operator [\(Zhang and Yuen 2015](#page-26-35)). In addition, to deeply study the effectiveness of DE-CDI with different implementations

Table 16 continued

 $\underline{\textcircled{\tiny 2}}$ Springer

The average ranking values (ARV) by Friedman

Algorithm	(1)	(2)	(3)	(4)	Algorithm	(1)	(2)	(3)	(4)
DE/best/1(1)		193.0	$295.0^{\rm a}$	7.0 ^b	$DE/rand-to-best/1$ (1)	-	$313.5^{\rm a}$	1.0 ^b	5.0 ^b
$DE-CELL/best/1(2)$	107.0	-	318.0 ^a	4.0 ^b	DE-CELL/rand-to-best/1 (2)	11.5^{b}	$\overline{}$	4.0 ^b	3.0 ^b
$DE-DIR/best/1$ (3)	5.0 ^b	7.0 ^b	-	5.0 ^b	$DE-DIR/rand-to-best/1(3)$	299.0^a	$296.0^{\rm a}$	$\overline{}$	7.0 ^b
DE -CDI/best/1 (4)	318.0 ^a	296.0 ^a	320.0 ^a	$\overline{}$	DE -CDI/rand-to-best/1 (4)	$295.0^{\rm a}$	297.0 ^a	$293.0^{\rm a}$	
CoDE(1)		39.5^{b}	18.5^{b}	5.0 ^b	ODE(1)	-	117.0	117.0	55.5^{b}
$CoDE-CELL(2)$	$285.5^{\rm a}$	$\overline{}$	50.5^{b}	1.5^{b}	ODE-CELL (2)	183.0	$\overline{}$	170.0	115.5
$CoDE-DIR(3)$	$306.5^{\rm a}$	$249.5^{\rm a}$	$\qquad \qquad$	13.0 ^b	ODE-DIR (3)	208.0	155.0	$\overline{}$	140.5
$CoDE-CDI(4)$	320.0 ^a	$298.5^{\rm a}$	312.0 ^a	-	ODE-CDI (4)	$269.5^{\rm a}$	184.5	159.5	

Table 17 Results of the multi-problem Wilcoxon's test between original DE, DE-CELL, DE-DIR and DE-CDI for all the functions at 30*D*

Upper diagonal of level significance $\alpha = 0.9$, lower diagonal level of significance $\alpha = 0.95$ ^a The method in the row improves the method of the column

^bThe method in the column improves the method of the row

OOOOOOO	. a a a a a a	0000000	0.0000000
. OO OO OO O	○ ○ ○ ● ○ ○ ○	$\circ\bullet\bullet\bullet\bullet\circ$	
$000 \bullet 000$	0000000	$\circ\bullet\bullet\bullet\bullet\circ$	
$\circ\circ\bullet\bullet\bullet\circ\circ$	$\circ\bullet\bullet\bullet\bullet\circ$	$\circ\bullet\bullet\bullet\bullet\circ$	$\bullet\bullet\bullet\blacksquare\bullet\bullet\bullet$
$000 \bullet 000$	ററ●●●ററ	0 	
. 000000	റററ●റററ	$\circ\bullet\bullet\bullet\bullet\circ$	
0000000	0000000	0000000	
$C5(n=5)$	$Cl3(n=13)$	$C25(n=25)$	$C49(n=49)$

Fig. 8 Cellular topology-based neighborhood with different number of neighbors

of directional mutation, an improved DMDE with cellular topology, named cDMDE, is also presented here. In cDMDE, the cellular topology is incorporated into the population of DMDE. In this section, four DE algorithms, i.e., DE/rand/2 and DE/best/1, CoDE and ODE, are used. The results for all the functions at 30*D* are shown in Table [14.](#page-17-0) Furthermore, the results of the multi-problem Wilcoxon signed rank tests are also presented in Table [15.](#page-19-0)

In Table [14,](#page-17-0) it is clear that DE-CDI can obtain the better results than DMDE in all the cases. For DE/rand/2, DE-CDI is significantly better than DMDE on 22 functions. For DE/best/1, DE-CDI is significantly better than DMDE on 22 functions, while is outperformed by it on one function. For CoDE and ODE, DE-CDI can obtain the significant better results than DMDE on 18 and 8 functions, respectively. From the results of Friedman test, DE-CDI obtains the better average ranking values than DMDE in all the cases. The results of the multi-problem Wilcoxon signed rank tests in Table [15](#page-19-0) also show that DE-CDI obtains the higher R + values than *R*− values in all the cases, when compared with DMDE. These results clearly indicate that DE-CDI is more effective than DMDE to enhance the performance of DE.

When compared with cDMDE, on the one hand, DE-CDI is consistently superior to it in most cases, except for DE/best/1. On the other hand, it is interesting to find that cDMDE can obtain the better results than DMDE in all the cases. Therefore, it is clear that the neighborhood information benefits DMDE. Furthermore, the effectiveness of DE-CDI is also demonstrated when it is implemented with different directional mutation operators.

5.7 Benefit of CDI components

In DE-CDI, the CDI-based mutation operator is composed by cellular topology-based neighborhood and neighborhoodbased directional mutation. The experimental results presented above have shown that DE-CDI is effective for improving the performance of DE. In this section, to identify the benefit of these two components, two DE-CDI variants are considered: DE-CELL that only incorporates the cellular topology-based neighborhood into DE and DE-DIR that only introduces the neighborhood-based directional mutation into DE. In DE-CELL, all the individuals for mutation are selected from the neighborhood of the current individual. In DE-DIR, based on the randomly selected base vector, the whole population is partitioned into the superior and inferior groups and the difference vector is constructed as that in DE-CDI. Four DE algorithms, DE/best/1, DE/rand-to-best/1, CoDE and ODE, are used for comparison. The results for the functions at 30*D* are shown in Table [16.](#page-20-0)

From Table [16,](#page-20-0) both DE-CELL and DE-DIR perform worse than the corresponding original DE algorithm in two cases, while DE-CDI can obtain significantly better results than DE in all the cases. Specifically, for DE/best/1, DE-CELL and DE-DIR are significantly outperformed by the original DE algorithm on 2 and 20 functions, respectively, while DE-CDI is significantly better than the original DE algorithm on 22 functions. For DE/rand-to-best/1, DE-CELL and DE-DIR are significantly better on zero and 17 functions, respectively, while DE-CDI is significantly better on 21 functions. For CoDE and ODE, all the DE-CDI variants are significantly better on most functions. From the results of Friedman test [\(García et al. 2009](#page-26-40); [Derrac et al. 2011\)](#page-26-41), DE-CDI obtains the best average ranking values in all the cases. In addition, both DE-CELL and DE-DIR perform better than the advanced DE algorithms overall.

To further show the advantage of combining the neighborhood and direction information in DE-CDI, the multiTable 18 Mean and standard deviation of the best error values obtained by DE/rand/1 and DE-CDI/rand/1 with different type of cellular topology on all the functions at 30*D*
Expression

problem Wilcoxon signed rank tests are also carried out and the results are presented in Table [17.](#page-22-0) It can be found that DE-CDI obtains the higher *R*+ values than *R*− values in all the cases. According to the multi-problem Wilcoxon signed rank test at $\alpha = 0.05$, DE-CDI is significantly better than DE-CELL and DE-DIR overall in most cases. These results clearly show that DE-CDI is better than the corresponding DE-CELL and DE-DIR overall.

From the results in Tables [16](#page-20-0) and [17,](#page-22-0) we can obtain some interesting observations. First, both the neighborhood information and direction information are beneficial to enhancing the performance of most DE algorithms studied. Second, compared with the DE algorithm only with single information (i.e., neighborhood or direction information), DE-CDI can utilize the population information more effectively to guide the search.

5.8 Parameter study

In DE-CDI, there is a control parameter (i.e., *n* in C*n*) that decides the neighborhood size of each individual. To test the influence of the neighborhood size on the performance of DE-CDI, the experiment studies with different *n* values are carried out. DE/rand/1 is employed for comparison here. Different number of neighbors, i.e., $n = 5, 9, 13, 25$ and 49, are considered. Figure [8](#page-22-1) illustrates the cellular topology based neighborhood with different number of neighbors. The results are shown in Tables [18](#page-23-0) and [19.](#page-24-1)

From Table [18,](#page-23-0) it is clear that DE-CDI with different *n* values is better than DE/rand/1 in all the cases except $n = 5$. In the cases of $n = 9$, 13, 25 and 49, DE-CDI is significantly better than DE/rand/1 on 13, 15, 13 and 11 functions, respectively, and is worse on 2, 2, 3 and zero functions, respectively. In addition, DE-CDI/rand/1 with different *n* values can obtain better average ranking value than DE/rand/1. According to the results of Table [19,](#page-24-1) in the cases of $n = 9, 13, 25$ and 49, DE-CDI obtains the higher *R*+ values than *R*− values. It indicates that DE-CDI with different *n* values significantly outperforms DE/rand/1 overall. Furthermore, DE-CDI with $n = 13$ and $n = 9$ obtains the best and second best results. In the case of $n = 5$, DE-CDI obtains the worst results. The reason may lie in that the neighborhood size (e.g., $n = 5$) for each individual is too small to be partitioned into the superior and inferior groups, and the direction information cannot be utilized effectively to construct the difference vector.

According to the results in Tables [18](#page-23-0) and [19,](#page-24-1) DE-CDI with $n = 13$ is a good choice for the benchmark functions in this study. In order to choose the appropriate neighborhood size for DE-CDI, adaptive or self-adaptive parameter control techniques (e.g., [Zhang and Sanderson 2009;](#page-26-9) [Qin et al. 2009](#page-26-11); [Islam et al. 2012\)](#page-26-42) will be studied in the future work.

5.9 Application to real-world problems

In order to test the effectiveness of DE-CDI on real-world problems, three problems are selected from [Eshelman et al.](#page-26-38) [\(1997](#page-26-38)) and [Das and Suganthan](#page-25-10) [\(2010\)](#page-25-10). The first two problems are from the CEC 2011 competition on testing EA on realworld numerical optimization problems [\(Das and Suganthan](#page-25-10) [2010](#page-25-10)). One is parameter estimation for frequency modulated sound waves (FMP) and the other is spread spectrum radar poly phase code design (SRP). FMP is a highly complex multi-modal problem with strong epistasis and SRP is with numerous local optima and has proven to be an*NP*-hard problem [\(Das and Suganthan 2010\)](#page-25-10). The last problem is systems of linear equations problem (LEP) which has proven to be quite difficult for the optimizers [\(Suganthan et al. 2005\)](#page-26-37). For each problem, MNFEs is set to 150,000 [\(Das and Suganthan](#page-25-10) [2010](#page-25-10)). Table [20](#page-25-12) presents the results.

As Table [20](#page-25-12) shows, it is clear that DE-CDI can obtain the better results than the corresponding DE algorithm in most cases. Specifically, for LEP and SRP, DE-CDI is better than the corresponding DE algorithms in 10 out of 12 cases. For FMP, DE-CDI is better than the corresponding DE algorithms in seven cases.

In sum, the results of Table [20](#page-25-12) indicate that DE-CDI is able to enhance the performance of DE on the real-world problems considered.

6 Conclusions

In this study, a simple and effective framework, DE with cellular direction information (DE-CDI), has been presented for numerical optimization. In DE-CDI, the cellular topology is employed to define a neighborhood for each individual

and the direction information is introduced into the mutation operator by constructing difference vector with the neighbors. In this way, the neighborhood and direction information can be simultaneously and effectively utilized to guide the search of DE. According to the extensive experimental study, it is obvious that DE-CDI is able to enhance the performance of most DE algorithms considered.

In the future, adaptive or self-adaptive methods for choosing number of neighbors in the cellular topology-based neighborhood will be investigated first. Then, the application of DE-CDI to other real-world problems will be studied. Finally, DE-CDI will also be extended to the multi-objective and large-scale optimization problems.

Acknowledgments This work was supported in part by the National Natural Science Foundation of China (61305085, 61202468), the Natural Science Foundation of Fujian Province of China (2014J05074, 2014J01240), the Support Program for Innovative Team and Leading Talents of Huaqiao University (2014KJTD13) and the Fundamental Research Funds for the Central Universities (12BS216).

References

- Alcalá-Fdez J, Sánchez L, García S (2015) KEEL: a software tool to assess evolutionary algorithms for data mining problems (online). <http://www.keel.es/>
- Bi XJ, Xiao J (2011) Classification-based self-adaptive differential evolution with fast and reliable convergence performance. Soft Comput 15(8):1581–1599
- Brest J, Greiner S, Boskovic B, Mernik M, Zumer V (2006) Selfadapting control parameters in differential evolution: a comparative study on numerical benchmark problems. IEEE Trans Evol Comput 10(6):646–657
- Cai Y, Wang J (2013) Differential evolution with neighborhood and direction information for numerical optimization. IEEE Trans Cybern 43(6):2202–2215
- Cai Y, Wang J, Yin J (2012) Learning-enhanced differential evolution for numerical optimization. Soft Comput 16(2):303–330
- Cai Y, Wang J, Chen Y et al (2015) Adaptive direction information in differential evolution for numerical optimization. Soft Comput (in press)
- Das S, Konar A (2006) Design of two dimensional IIR filters with modern search heuristics: a comparative study. Int J Comput Intell Appl 6(3):329–355
- Das S, Suganthan PN (2010) Problem definitions and evaluation criteria for CEC 2011 competition on testing evolutionary algorithms on real world optimization problems. In: Technical report, Jadavpur University, West Bengal. Nanyang Technological University, Singapore
- Das S, Suganthan PN (2011) Differential evolution: a survey of the state-of-the-art. IEEE Trans Evol Comput 15(1):4–31
- Das S, Abraham A, Konar A (2008) Adaptive clustering using improved differential evolution algorithm. IEEE Trans Syst Man Cybern A 38(1):218–237
- Das S, Abraham A, Chakraborty UK, Konar A (2009) Differential evolution using a neighborhood-based mutation operator. IEEE Trans Evolut Comput 13(3):526–553
- De Falco I, Della Cioppa A, Maisto D, Scafuri U, Taranino E (2014) Impact of the topology on the performance of distributed differential evolution. Appl Evol Comput 8602:75–85
- Derrac J, García S, Molina D, Herrera F (2011) A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. Swarm Evol Comput 1(1):3–18
- Dorronsoro B, Bouvry P (2010) Differential evolution algorithms with cellular populations. In: Proceedings of the 11th PPSN, pp 320– 330
- Dorronsoro B, Bouvry P (2011) Improving classical and decentralized differential evolution with new mutation operator and population topologies. IEEE Trans Evol Comput 15(1):67–98
- Epitropakis MG, Tasoulis DK, Pavlidis NG, Plagianakos VP, Vrahatis MN (2011) Enhancing differential evolution utilizing proximity based mutation operators. IEEE Trans Evol Comput 15(1):99–119
- Eshelman LJ, Mathias KE, Schaffer JD (1997) Convergence controlled variation. In: Belew R, Vose M (eds) Foundations of genetic algorithms 4. Morgan Kaufmann, SanMateo, pp 203–224
- Fan H, Lampinen J (2003) A trigonometric mutation operation to differential evolution. J Global Optim 27(1):105–129
- García S, Fernandez A, Luengo J, Herrera F (2009) A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability. Soft Comput 13(10):959– 977
- Gou J, Guo W, Hou F, Wang C, Cai Y (2015) Adaptive differential evolution with directional strategy and cloud model. Appl Intell 42(2):369–388
- Hu Z, Cai X, Fan Z (2014) An improved memetic algorithm using ring neighborhood topology for constrained optimization. Soft Comput 18:2023–2041
- Hui S, Suganthan PN (2013) Ensemble crowding differential evolution with neighborhood mutation for multimodal optimization. In: Proceedings of the IEEE symposium on differential evolution (SDE), pp 135–142. IEEE, New York
- Iorio A, Li X (2006) Incorporating directional information within a differential evolution algorithm for multi-objective optimization. In: Proceedings of the 8th annual conference on genetic evolutionary computational, pp 691–698
- Islam SM, Das S, Ghosh S, Roy S, Suganthan PN (2012) An adaptive differential evolution algorithm with novel mutation and crossover strategies for global numerical optimization. IEEE Trans Syst Man Cybern B Cybern 42(2):482–500
- Joshi R, Sanderson AC (1999) Minimal representation multi-sensor fusion using differential evolution. IEEE Trans Syst Man Cybern Part A 29(1):63–76
- Lampinen J (1999) A bibliography of differential evolution algorithm. In: Technical report, Laboratory of Information Processing, Department of Information Technology, Lappeenranta University of Technology. <http://www.lut.fi/jlampine/debiblio.htm> (online)
- Liang JJ, Qu B-Y, Mao X, Niu B, Wang D (2014) Differential evolution based on fitness Euclidean-distance ratio for multimodal optimization. Neurocomputing 137(5):252–260
- Liu J, Fan Z, Goodman E (2009) SRDE: an improved differential evolution based on stochastic ranking. In: Proceedings of the 1st ACM/SIGEVO, pp 345–352
- Neri F, Tirronen V (2010) Recent advances in differential evolution: a survey and experimental analysis. Artif Intell Rev 33(1/2):61–106
- Neri F, Iacca G, Mininno E (2011) Disturbed exploitation compact differential evolution for limited memory optimization problems. Inf Sci 181(12):2469–2487
- Noman N, Iba H (2011) Cellular differential evolution algorithm. In: Proceedings of the AII advanced artificial intelligence, pp 293–302
- Noroozi V, Hashemi A, Meybodi M (2011) CellularDE: a cellular based differential evolution for dynamic optimization problems. In: Proceedings of the adapting natural computational algorithms, pp 340–349
- Omran M, Engelbrecht AP, Salman A (2005) Differential evolution methods for unsupervised image classification. In: Proceedings of

the 7th congress on evolutionary computation (CEC-2005), vol 2, pp 966–973. IEEE Press, Piscataway

- Omran M, Engelbrecht A, Salman A (2006) Using the ring neighborhood topology with self-adaptive differential evolution. In: Jiao L, Wang L, Gao X-B, Liu J, Wu F (eds) Advances in natural computation. Springer, Berlin, pp 976–979
- Omran M, Engelbrecht A, Salman A (2009) Bare bones differential evolution. Eur J Oper Res 196(1):128–139
- Piotrowski AP (2013) Adaptive memetic differential evolution with global and local neighborhood-based mutation operators. Inf Sci 241(20):164–194
- Qin A, Huang V, Suganthan PN (2009) Differential evolution algorithm with strategy adaptation for global numerical optimization. IEEE Trans Evol Comput 13(2):398–417
- Rahnamayan S, Tizhoosh HR, Salama MMA (2008) Opposition based differential evolution. IEEE Trans Evol Comput 12(1):64–79
- Rogalsky T, Derksen RW, Kocabiyik S (1999) Differential evolution in aerodynamic optimization. In: Proceedings of the 46th annual conference on Canadian Aeronautics and Space Institute, pp 29–36
- Sarkar S, Mukherjee R, Biswas S, Kundu R, Das S (2015) An adaptive clustering and re-clustering based crowding differential evolution for continuous multi-modal optimization. In: Proceedings of the 18th Asia Pacific symposium on intelligent and evolutionary systems, vol 1, pp 373–388. Springer, New York
- Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11(4):341–359
- Storn R, Price KV, Lampinen J (2005) Differential evolution—a practical approach to global optimization. Springer, Berlin
- Suganthan PN, Hansen N, Liang JJ, Deb K, Chen Y-P, Auger A, Tiwari S (2005) Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization. In: KanGAL Report No. 2005005, Nanyang Technological University, Singapore. IIT Kanpur, India
- Sun J, Zhang Q, Tsang EPK (2005) DE/EDA: a new evolutionary algorithm for global optimization. Inf Sci 169(3):249–262
- Tang L, Dong Y, Liu J (2015) Differential evolution with an individualdependent mechanism. IEEE Trans Evol Comput (in press)
- Wang F-S, Jang H-J (2000) Parameter estimation of a bio-reaction model by hybrid differential evolution. In: Proceedings of the IEEE congress on evolutionary computation, vol 1, pp 410–417. IEEE Press, Piscataway
- Wang YX, Xiang QL (2008) Exploring new learning strategies in differential evolution algorithm. In: Proceedings of the IEEE congress on evolutionary computational, pp 204–209
- Wang J, Cai Y (2015) Multiobjective evolutionary algorithm for frequency assignment problem in satellite communications. Soft Comput 19(5):1229–1253
- Wang Y, Cai Z, Zhang Q (2011) Differential evolution with composite trial vector generation strategies and control parameters. IEEE Trans Evol Comput 15(1):55–66
- Wang J, Liao J, Zhou Y, Cai Y (2014) Differential evolution enhanced with multiobjective sorting based mutation operators. IEEE Trans Cybern 46(12):2792–2805
- Weber M, Tirronen V, Neri F (2010) Scale factor inheritance mechanism in distributed differential evolution. Soft Comput 14(11):1187– 1207
- Weber M, Neri F, Tirronen V (2011) A study on scale factor in distributed differential evolution. Inf Sci 181(12):2488–2511
- Yang M, Li C, Cai Z, Guan J (2015) Differential evolution with autoenhanced population diversity. IEEE Trans Cybern (in press)
- Zhang J, Sanderson A (2009) JADE: adaptive differential evolution with optional external archive. IEEE Trans Evol Comput 13(5):945–958
- Zhang X, Yuen SY (2015) A directional mutation operator for differential evolution algorithms. Appl Soft Comput 30:529–548