Comparison of Adaptive Differential Evolution Algorithms on the MOEA/D-DE Framework

Kei Nishihara
Graduate School of Engineering Science
Yokohama National University
Yokohama, Japan
nishihara-kei-jv@ynu.jp

Masaya Nakata
Graduate School of Engineering Science
Yokohama National University
Yokohama, Japan
nakata-masaya-tb@ynu.ac.jp

***Abstract*—Existing works have reported that adaptive differential evolution algorithms, i.e., adaptive DEs, improve the MOEA/D-DE algorithm, but this result is limited to small-scale multi-objective optimization problems. This paper compares four popular adaptive DEs on the MOEA/D-DE framework to evaluate their scalability to the number of decision variables and objectives. Specifically, we employ jDE, JADE, EPSDE, and SaDE in this paper. Our experimental results provide the following novel observations. MOEA/D-DE with JADE derives the best average rank on small-scale problems. However, the performances of MOEA/D-DE with JADE, EPSDE, and SaDE gradually degrade with the increase of the problem scale. In contrast, jDE stably improves the performance of MOEA/D-DE on large-scale problems employed in this paper (i.e., 11 objectives and 100 decision variables). Thus, we find a critical tradeoff among adaptive DEs in terms of the scalability of the MOEA/D-DE framework; a statistical adaption like JADE is suitable for small-scale problems, but a randomization adaptation like jDE is effective with the increase of the problem scale. Our results also suggest that parameter-only adaptation can be suitable for MOEA/D-DE regardless of the problem scale.**

***Index Terms*—algorithmic configuration adaptation, MOEA/D-DE, many-objective optimization**

I. Introduction

Although a large number of multi-objective evolutionary algorithms (MOEAs) have been proposed in the literature [1], [2], we often suffer to decide their algorithmic configurations suitable for solving a given problem [3]. The performances of evolutionary algorithms, including MOEAs, highly depend on their algorithmic configurations, e.g., hyper-parameter settings and genetic operators [4], [5]. Consequently, adaptation methods of algorithmic configurations can be a proper approach to improve the performances [6], [7]. The adaptation methods have been mainly studied in single-objective evolutionary algorithms, especially in differential evolution algorithms (DEs) [8], e.g., LSHADE-Epsin [9] and jSO [10].

In MOEAs, MOEA/D-DE [11] has been frequently employed as a basis of the adaptation methods. Specifically, existing methods, i.e., adaptive MOEA/D-DEs, adapt the DE parameters and/or the DE genetic operators both implemented in MOEA/D-DE. For instance, in [12], the scaling factor F and the crossover rate CR are adapted based on two popular adaptive DEs: jDE [13] and SHADE [14]. In [15], the parameter-adaptation method of F is customized for MOEA/D-DE based on jDE, JADE [16], and MDE_pBX [17]. The mutation strategy is also adapted in MOEA/D-FRRMAB [18] and ADEMO/D [5]. MOEA/D-CDE [19] adapts both the mutation strategy and F. Those three specific algorithms employ the credit assignment strategy to adapt the configurations.

TABLE I
Classification of four adaptive DEs employed in this paper

	F	CR	mutation strategy
jDE [13]	randomization	randomization	–
JADE [16]	statistical	statistical	–
EPSDE [20]	randomization	randomization	randomization
SaDE [21]	statistical	statistical	statistical

However, there is a lack of fundamental insights to develop adaptation methods for multi-objective optimization problems (MOPs) compared to single-objective optimization problems (SOPs). Specifically, the following issues remain unclear:

- The scalability of adaptive MOEA/D-DEs to the number of decision variables, objectives, and configurations to be adapted is not investigated. The most of existing adaptive MOEA/D-DEs have been tested on small-scale problems with the number of decision variables $D \leq 30$ and the number of objectives $M \leq 3$. Besides, existing works have not intensively compared adaptive MOEA/D-DEs in terms of the number of algorithmic configurations to be adapted.
- Adaptation strategies employed in adaptive DEs can be roughly classified into randomization adaptation (e.g., jDE and EPSDE [20]) and statistical adaptation (e.g., JADE and SaDE [21]) [15]. Many literature have reported that the statistical adaptation derives superior performance in SOPs [14], [16]. However, such a tendency has not yet been observed in MOPs due to the lack of comparative studies. In both [15] and [12], adaptation strategies were compared, but no clear superiority was observed.

Note that the randomization adaptation indicates that algorithmic configurations are randomly generated regardless of adaptation results during the run, while the statistical adaptation utilizes the adaptation results to sample new configurations

978-1-7281-8393-0/21/$31.00 ©2021 IEEE

using statistical distributions.

Thus, this paper conducts an intensive comparison of the following four popular adaptive DEs on the MOEA/D-DE framework; jDE, JADE, EPSDE, and SaDE. These adaptive DEs are basic algorithms often used in comparative studies for SOPs [22], [23]. TABLE I summarizes a rough classification of the four adaptive DEs employed in this paper. Note that SaDE is conceptually designed to adapt both the mutation and crossover strategies, but it virtually uses the binomial crossover as implemented in this paper. Note also that some adaptive DEs, e.g., CoDE [24] and DE-CAT [25], adapt both the mutation and crossover strategies in SOPs, but only mutation strategy is typically controlled in existing adaptive MOEA/D-DEs. We also set different scale problems with $D = \{20, 50, 100\}$ and $M = \{3, 7, 11\}$.

This paper is organized as follows. Section II briefly introduces adaptive MOEA/D-DEs with jDE, JADE, EPSDE, and SaDE, denoted as MOEA/D-jDE, MOEA/D-JADE, MOEA/D-EPSDE, and MOEA/D-SaDE, respectively. In Section III, we test those four adaptive MOEA/D-DEs on DTLZ [26] and WFG [27] benchmark problems with different problem settings. We also demonstrate adaptation results to analyze our experimental results. Finally, Section IV gives a summary of this paper with future work.

II. MOEA/D-DE AND ITS ADAPTATION

This section introduces a generalized framework of the adaptive MOEA/D-DE. Then, the detail frameworks of the four adaptive MOEA/D-DEs are described.

As a basic introduction, MOEA/D-DE is a population-based optimization algorithm, where a population $\mathcal{P}$ consists of N individuals $\boldsymbol{x}_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,D}]$, $i \in \{1, 2, \ldots, N\}$; $x^u_{i,j}$ and $x^l_{i,j}$ are the upper and lower values of $x_{i,j}$, respectively. An MOP with M objective functions $f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_M(\boldsymbol{x})$ is decomposed into N single-objective sub-problems; each i-th sub-problem is paired with a scalarization function $g(\boldsymbol{x}|\boldsymbol{\lambda}_i)$. Specifically, a weight vector is defined as $\boldsymbol{\lambda}_i = [\lambda_{i,1}, \ldots, \lambda_{i,M}]^\mathsf{T}$ with the condition $\sum_{j=1}^{M} \lambda_{i,j} = 1$, and we use the Tchebycheff function as the scalarization function to be minimized [11], [28], given by;

$$g(\boldsymbol{x}_i|\boldsymbol{\lambda}_i, \boldsymbol{z}^*) = \max_{j \in \{1,2,\ldots,M\}} \{\lambda_j |f_j(\boldsymbol{x}_i) - z^*_j|\}, \qquad (1)$$

where $\boldsymbol{z}^*$ denotes a set of ideal reference points used to determine the search direction; and $\boldsymbol{x}_i$ is the individual assigned to i-th sub-problem. In this paper, we use a provisional minimum value as the reference point $\boldsymbol{z}$; i.e., $z_j = \min_{\boldsymbol{x} \in \mathcal{P}} \{f_j(\boldsymbol{x})\}$, $j \in \{1, 2, \ldots, M\}$.

A. Generalized framework of adaptive MOEA/D-DE

Algorithm 1 shows a generalized framework of the adaptive MOEA/D-DE employed in this paper. Note that jDE, JADE, EPSDE, and SaDE can be classified to an individual-based adaptation method, that is, each individual $\boldsymbol{x}_i$ is paired with its own algorithmic configuration. This paper describes an algorithmic configuration for $\boldsymbol{x}_i$ as $\boldsymbol{\theta}_i = [\theta_{v,i}, \theta_{F,i}, \theta_{CR,i}]$;

Algorithm 1 Generalized framework of adaptive MOEA/D-DE

1: $t = 0$
2: Initialize $\mathcal{P} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$
3: Initialize $\Theta = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_N\}$ ($\boldsymbol{\theta}_i = [\theta_{v,i}, \theta_{F,i}, \theta_{CR,i}]$)
4: **for** $i = 1$ **to** N **do**
5: Get neighborhood indices list $\mathcal{B}_i = \{b_1, b_2, \ldots, b_T\}$
6: **for** $j = 1$ **to** M **do**
7: Calculate $z_j = \min_{\boldsymbol{x} \in \mathcal{P}} \{f_j(\boldsymbol{x})\}$
8: **while** *termination criteria* are not met **do**
9: $t = t + 1$
10: **for** $i = 1$ **to** N **do**
11: Get indices list $\mathcal{L}_i = \begin{cases} \mathcal{B}_i & \textbf{if } rand[0,1] \leq \delta \\ \{1, 2, \ldots, N\} & \textbf{otherwise} \end{cases}$
12: $\boldsymbol{\theta}^{t-1}_i = \boldsymbol{\theta}_i$
13: $\boldsymbol{\theta}_i \leftarrow$ *Sample*$(\boldsymbol{\theta}^{t-1}_i)$
14: $\boldsymbol{v}_i \leftarrow$ *DE-Mutation*$(\mathcal{P}, \theta_{v,i}, \theta_{F,i})$
15: $\boldsymbol{u}_i \leftarrow$ *DE-Crossover*$(\boldsymbol{x}_i, \boldsymbol{v}_i, \theta_{CR,i})$
16: $\boldsymbol{u}_i \leftarrow$ *Polynomial-Mutation*$(\boldsymbol{u}_i, p_m, \eta)$
17: **for** $j = 1$ **to** M **do**
18: Update $z_j = \min\{z_j, f_j(\boldsymbol{u}_i)\}$
19: $ct = 0$
20: **while** $ct < n_r$ **and** $\mathcal{L}_i \neq \emptyset$ **do**
21: $j \leftarrow$ Select an index randomly from $\mathcal{L}_i$, $\mathcal{L}_i = \mathcal{L}_i \backslash \{j\}$
22: **if** $g(\boldsymbol{u}_i|\boldsymbol{\lambda}_j, \boldsymbol{z}) \leq g(\boldsymbol{x}_j|\boldsymbol{\lambda}_j, \boldsymbol{z})$ **then**
23: $\boldsymbol{x}_j = \boldsymbol{u}_i, ct = ct + 1$
24: $\boldsymbol{\theta}_i \leftarrow$ *Update*$(\boldsymbol{\theta}_i, \boldsymbol{\theta}^{t-1}_i)$

$\theta_{v,i}$ is a categorical variable, which indicates an index of predefined mutation strategies; $\theta_{F,i}$ and $\theta_{CR,i}$ are real-value variables, which indicate settings of F and CR, respectively. All algorithmic configurations are contained in a configuration set Θ. Note that $\theta_{v,i}$ is always set to 1 as the default value if we do not adapt the mutation strategy. For instance, we can set and fix $\boldsymbol{\theta}_i$ to $[1, 0.5, 1.0], \forall i \in \{1, 2, \ldots, N\}$ for (non-adaptive) MOEA/D-DE [11], where $\theta_{v,i} = 1$ indicates the *current/1* mutation strategy; $\boldsymbol{v}_i = \boldsymbol{x}_i + \theta_{F,i}(\boldsymbol{x}_{r_1} - \boldsymbol{x}_{r_2})$, where $\boldsymbol{x}_{r_1}$ and $\boldsymbol{x}_{r_2}$ are mutually exclusive individuals randomly selected from the current population $\mathcal{P}$ and $\boldsymbol{v}_i$ is a mutant individual. Thus, adaptive MOEA/D-DEs control $\boldsymbol{\theta}_i$ during the run.

As the initialization process, the adaptive MOEA/D-DE defines N single-objective sub-problems with Eq. (1), and it generates N initial individuals $\boldsymbol{x}_i \in \mathcal{P}$. In addition, all N algorithmic configurations $\boldsymbol{\theta}_i \in \Theta$ are also initialized according to the default settings of an employed adaptive DE. As in the MOEA/D framework, a neighborhood indices list $\mathcal{B}_i = \{b_1, b_2, \ldots, b_T\}$ is prepared, where each b denotes the index of a sub-problem and T is a hyper-parameter.

Next, for each sub-problem, the adaptive MOEA/D-DE chooses an indices list $\mathcal{L}_i$ to be used in the updating process (lines 20-24, Algorithm 1). Specifically, $\mathcal{L}_i$ is set to the neighborhood $\mathcal{B}_i$ of i-th sub-problem with the probability δ; otherwise $\mathcal{L}_i$ is set to a set of all possible indices, i.e., $\{1, 2, \ldots, N\}$. Next, the adaptive MOEA/D-DE stores the current configuration $\boldsymbol{\theta}_i$ as $\boldsymbol{\theta}^{t-1}_i$, and then it newly samples $\boldsymbol{\theta}_i$ with the defined adaptation method, denoted by a pseudo function *Sample*$(\boldsymbol{\theta}^{t-1}_i)$. Then, it generates a new individual $\boldsymbol{u}_i$ based on the DE procedures with $\boldsymbol{\theta}_i$; a mutant individual $\boldsymbol{v}_i$ is generated by the $\theta_{v,i}$-th mutation strategy with a scaling factor $\theta_{F,i}$ and then a trial individual $\boldsymbol{u}_i$ is generated by applying a DE crossover strategy to $\boldsymbol{u}_i$ and $\boldsymbol{x}_i$ with the crossover rate $\theta_{CR,i}$, and subsequently the polynomial mutation is further applied to $\boldsymbol{u}_i$ as;

$$u_{i,j} = \begin{cases} u_{i,j} + \sigma_k(x^u_{i,j} - x^l_{i,j}) & \text{if } rand[0,1] \leq p_m, \\ u_{i,j} & \text{otherwise,} \end{cases} \tag{2}$$

$$\sigma_k = \begin{cases} (2rand[0,1])^{\frac{1}{\eta+1}} - 1 & \text{if } rand[0,1] \leq 0.5, \\ 1 - (2 - 2rand[0,1])^{\frac{1}{\eta+1}} & \text{otherwise,} \end{cases} \tag{3}$$

where p_m and η are the mutation probability and the distribution index, respectively. Note that we use the *binomial* crossover frequently used in MOEA/D-DE variants [5], [12], [15], [19], [29], given by;

$$u_{i,j} = \begin{cases} v_{i,j} & \text{if } rand[0,1] \leq \theta_{CR,i} \text{ or } j = j_{rand}, \\ x_{i,j} & \text{otherwise,} \end{cases} \tag{4}$$

where $j_{rand} \in [1, D]$ is an integer randomly sampled from uniform distribution.

After $\boldsymbol{u}_i$ is generated, the adaptive MOEA/D-DE calculates the fitness values of $\boldsymbol{u}_i$ and then it updates the reference point $\boldsymbol{z}$. At the end of generation process (lines 20-24), the adaptive MOEA/D-DE replaces $\boldsymbol{x}_j \in \mathcal{P}$ with $\boldsymbol{u}_i$ using an index j randomly selected from $\mathcal{L}_i$, if $\boldsymbol{u}_i$ improves j-th sub-problem $g(|\boldsymbol{\lambda}_j, \boldsymbol{z})$. Note that $\boldsymbol{\theta}_i$ may be further updated with a pseudo function *Update*($\boldsymbol{\theta}_i, \boldsymbol{\theta}_i^{t-1}$) dependent on the detail framework of the adaptive DEs. For instance, $\boldsymbol{\theta}_i$ is replaced with $\boldsymbol{\theta}_i^{t-1}$ if $\boldsymbol{u}_i$ does not improve any individual $\boldsymbol{x}$ on jDE.

In short, our generalized framework of the adaptive MOEA/D-DE adds two new pseudo functions *Sample*($\boldsymbol{\theta}_i^{t-1}$) and *Update*($\boldsymbol{\theta}_i, \boldsymbol{\theta}_i^{t-1}$) to the MOEA/D-DE framework.

B. Detail framework of the four adaptive MOEA/D-DEs

TABLE II summarizes the initialization of $\boldsymbol{\theta}_i$ and the two pseudo functions *Sample*($\boldsymbol{\theta}_i^{t-1}$) and *Update*($\boldsymbol{\theta}_i, \boldsymbol{\theta}_i^{t-1}$) employed in MOEA/D-jDE, MOEA/D-JADE, MOEA/D-EPSDE, and MOEA/D-SaDE, respectively. Note that we do not use the best individual-based mutation strategies (e.g., *best/1* and *current-to-pbest/1*) since these strategies cannot be plausibly defined in MOPs. Thus, MOEA/D-JADE employs the *current/1* mutation strategy instead of *current-to-pbest/1*. In addition, MOEA/D-jDE also employs *current/1* (instead of *rand/1*) to investigate a pure impact of the jDE's adaptation algorithm compared to JADE. In the similar manner, different from the original frameworks of EPSDE and SaDE, our MOEA/D-EPSDE and MOEA/D-SaDE use *current/1* and *rand/1* as candidates of the mutation strategy.

1) MOEA/D-jDE: jDE adapts only hyper-parameters F and CR by the randomization adaptation. This paper adopts *current/1* mutation strategy. Specifically, $\theta_{F,i}$ and $\theta_{CR,i}$ for an individual $\boldsymbol{x}_i$ are randomly re-sampled with the probabilities τ_F and τ_{CR}, respectively;

$$\theta_{F,i} = \begin{cases} rand[0.1, 1] & \text{if } rand[0,1] \leq \tau_F, \\ \theta_{F,i}^{t-1} & \text{otherwise,} \end{cases} \tag{5}$$

$$\theta_{CR,i} = \begin{cases} rand[0, 1] & \text{if } rand[0,1] \leq \tau_{CR}, \\ \theta_{CR,i}^{t-1} & \text{otherwise.} \end{cases} \tag{6}$$

Note that for SOPs, jDE resets $\boldsymbol{\theta}_i$ to $\boldsymbol{\theta}_i^{t-1}$ if $f(\boldsymbol{u}_i) > f(\boldsymbol{x}_i)$. In the adaptive MOEA/D-DE framework, we reset $\boldsymbol{\theta}_i$ to $\boldsymbol{\theta}_i^{t-1}$ if no individual is updated with $\boldsymbol{u}_i$ (in lines 20-24, Algorithm 1). Technically, *Update*($\boldsymbol{\theta}_i, \boldsymbol{\theta}_i^{t-1}$) is designed to replace $\boldsymbol{\theta}_i$ with $\boldsymbol{\theta}_i^{t-1}$ if $ct = 0$.

2) MOEA/D-JADE: JADE adapts only hyper-parameters F and CR by the statistical adaptation. Again, the original JADE framework (for SOPs) uses the *current-to-pbest/1* mutation strategy; however, this paper uses the *current/1* mutation strategy. In the JADE framework, two meta-parameters μ_F and μ_{CR} are defined to specify the Cauchy distribution $\mathcal{C}(\mu_F, 0.1)$ and the normal distribution $\mathcal{N}(\mu_{CR}, 0.1)$ for sampling $\theta_{F,i}$ and $\theta_{CR,i}$, respectively. Those two meta-parameters are updated as follows;

$$\mu_F = (1 - c)\mu_F + c \cdot mean_L(S_F), \tag{7}$$

$$\mu_{CR} = (1 - c)\mu_{CR} + c \cdot mean(S_{CR}), \tag{8}$$

where $c \in [0, 1]$ is the learning rate; S_F and S_{CR} denote sets of $\theta_{F,i}$ and $\theta_{CR,i}$ which have succeeded in updating the individual in a generation, respectively. Here, $mean(S_{CR})$ calculates the mean of all values stored in S_{CR}; and $mean_L(S_F)$ returns the second-order Lehmer mean of all values stored S_F, i.e., $\left(\sum_{\theta_{F,i} \in S_F} \theta_{F,i}^2 / \sum_{\theta_{F,i} \in S_F} \theta_{F,i}\right)$.

3) MOEA/D-EPSDE: EPSDE adapts the mutation strategy in addition to F and CR by the randomization adaptation. EPSDE defines candidates of configurations as; $P_F = \{0.4, 0.5, \cdots, 0.9\}$, $P_{CR} = \{0.1, 0.2, \cdots, 0.9\}$, $P_v = \{$*current/1*, *rand/1*$\}$. The *rand/1* mutation strategy can be denoted as $\boldsymbol{v}_i = \boldsymbol{x}_{r_1} + \theta_{F,i}(\boldsymbol{x}_{r_2} - \boldsymbol{x}_{r_3})$, where $\boldsymbol{x}_{r_3}$ is also a mutually exclusive and randomly sampled individual. Note that the original EPSDE prepares *best/2* as a candidate of the mutation strategy. At the initializing process, $\theta_{F,i}$, $\theta_{CR,i}$, and $\theta_{v,i}$ for i-th individual are randomly selected from the predefined candidates. Then, EPSDE updates $\theta_{F,i}$, $\theta_{CR,i}$, and $\theta_{v,i}$ by randomly selecting the candidates if $f(\boldsymbol{u}_i) > f(\boldsymbol{x}_i)$.

4) MOEA/D-SaDE: SaDE adapts the mutation strategy, F, and CR by the statistical adaptation. In SaDE, $\theta_{F,i}$ is always sampled from $\mathcal{N}(0.5, 0.3)$; $\theta_{CR,i}$ and the mutation strategy $\theta_{v,i}$ are updated based on a success history information of the individual updates in past LP generations. Note that the original SaDE framework adapts both the mutation and crossover strategies, but this paper adapts only the mutation strategy to investigate a pure impact of the SaDE's statistical adaptation compared to the EPSDE's randomization adaptation. In particular, our MOEA/D-SaDE defines *current/1*

TABLE II

SUMMARY OF INITIALIZATION, *Sample*($\boldsymbol{\theta}_i^{t-1}$), AND *Update*($\boldsymbol{\theta}_i, \boldsymbol{\theta}_i^{t-1}$) EMPLOYED IN MOEA/D-JDE, MOEA/D-JADE, MOEA/D-EPSDE, AND MOEA/D-SADE

Method	Initialization of $\boldsymbol{\theta}_i = [\theta_{v,i}, \theta_{F,i}, \theta_{CR,i}] \in \Theta$ (line 3 in Algorithm 1)	*Sample*($\boldsymbol{\theta}_i^{t-1}$) (line 13 in Algorithm 1)	*Update*($\boldsymbol{\theta}_i, \boldsymbol{\theta}_i^{t-1}$) (line 24 in Algorithm 1)
jDE	$\theta_{v,i} = 1$ (*current/1*) $\theta_{F,i} = 0.5$ $\theta_{CR,i} = 1.0$	$\theta_{v,i} = 1$ (*current/1*) $\theta_{F,i} =$ Eq. (5) $\theta_{CR,i} =$ Eq. (6)	$\theta_{v,i} = 1$ (*current/1*) $\theta_{F,i} = \theta_{F,i}^{t-1}$ (if failed in individual update) $\theta_{CR,i} = \theta_{CR,i}^{t-1}$ (if failed in individual update)
JADE	$\theta_{v,i} = 1$ (*current/1*) $\mu_F = 0.5$ $\mu_{CR} = 0.5$	$\theta_{v,i} = 1$ (*current/1*) $\theta_{F,i} = \mathcal{C}(\mu_F, 0.1)$ $\theta_{CR,i} = \mathcal{N}(\mu_{CR}, 0.1)$	–
EPSDE	$\theta_{v,i}$: randomly select from $P_v = \{$*current/1*, *rand/1*$\}$ $\theta_{F,i}$: randomly select from P_F of original EPSDE $\theta_{CR,i}$: randomly select from P_{CR} of original EPSDE	–	$\theta_{v,i}$: re-sample (if failed in individual update) $\theta_{F,i}$: re-sample (if failed in individual update) $\theta_{CR,i}$: re-sample (if failed in individual update)
SaDE	$\theta_{v,i}$: randomly select from $\{$*current/1*, *rand/1*$\}$ (thus $k \in \{1, 2\}, K = 2$) $CRm_k = 0.5$	$\theta_{v,i}$: select with probability p_k in Eq. (9) $\theta_{F,i} = \mathcal{N}(0.5, 0.3)$ $\theta_{CR,i} = \mathcal{N}(CRm_k, 0.1)$	–

and *rand/1* as candidates of the mutation strategy. Then, the selection probability for k-th candidate is calculated as;

$$p_{k,t} = \frac{S_{k,t}}{\sum_{k=1}^{K} S_{k,t}}, \tag{9}$$

$$S_{k,t} = \frac{\sum_{g=t-LP}^{t-1} ns_{k,g}}{\sum_{g=t-LP}^{t-1} ns_{k,g} + \sum_{g=t-LP}^{t-1} nf_{k,g}} + \epsilon, \tag{10}$$

where $ns_{k,t}$ and $nf_{k,t}$ are the number of success and failure updates at generation t, respectively; ϵ is set to a constant value to avoid zero division. For $\theta_{CR,i}$, the values of successful individual updates in the past LP generations are stored in CR_{Memory_k} for each strategy, and new $\theta_{CR,i}$ is sampled from the normal distribution $\mathcal{N}(CRm_k, 0.1)$ with standard deviation of 0.1 and mean CRm_k, the median value of CR_{Memory_k}.

III. COMPARISON

This section evaluates the scalability of the following five algorithms to the number of decision variables D and objectives M; MOEA/D-DE as a baseline, MOEA/D-jDE, MOEA/D-JADE, MOEA/D-EPSDE, and MOEA/D-SaDE.

A. Experimental design

We use DTLZ 1-7 [26] and WFG 1-9 [27] benchmark problems with $D = \{20, 50, 100\}$ and $M = \{3, 7, 11\}$. Thus, 144 experimental cases are conducted. For the WFG problems, the number of position variables k is set to $k = M - 1$; and the number of distance variables l is set to $l = D - k$ [27].

Specific parameter settings of the four adaptive MOEA/D-DEs are summarized in TABLE II except for the following settings; $\tau_F = \tau_{CR} = 0.1$ for MOEA/D-jDE [13], $c = 0.1$ for MOEA/D-JADE [16], and $\epsilon = 0.01$ for MOEA/D-SaDE [21]. As common parameter settings, we use $p_m = 1/D$, $\eta = 20$, $n_r = 2$, and $\delta = 0.9$ [11]. In addition, we set $N = \{91, 91, 77\}$ and $T = \{10, 10, 8\}$ both for $M = \{3, 7, 11\}$, where $T = ceil(N/10)$. Note that N is determined with the two-layered approach [30], i.e., $(H_1, H_2) = \{(12, 0), (3, 1), (2, 1)\}$ for $M = \{3, 7, 11\}$ in this paper. The maximum number of fitness evaluations is strictly set to $100,000$, i.e., we forcedly termimate each run when the number of fitness evaluations reaches $100,000$.

We use the inverted generational distance (IGD) [31] with $10,000$ reference points to evaluate the performances of the algorithms. IGD values at the maximum number of fitness evaluations are reported as average values of 30 independent runs.In this paper, we do not use hypervolume (HV) [32] because the computation time to calculate HV increases exponentially with M [33]. All experiments are conducted on the PlatEMO software [34].

B. Result

TABLE III shows the IGD values of the five algorithms with different problem dimensions for $M = \{3, 7, 11\}$, respectively. Note that the best and worst values are highlighted with green (and bold) and pink, respectively. Figs. 1 and 2 summarize the average ranks of the five algorithms for $D = \{20, 50, 100\}$ and $M = \{3, 7, 11\}$, respectively.

As shown in the figures, for $M = 3$, the four adaptive MOEA/D-DEs outperform MOEA/D-DE on small-scale problems with $D = 20, M = 3$. This tendency is consistent with the existing works [5], [12], [15], [18], [19]; and all the adaptive MOEA/D-DEs outperform MOEA/D-DE even when D is further increasing to 50 and 100. However, as a general trend, the difference of average ranks between MOEA/D-DE and each adaptive MOEA/D-DE except for MOEA/D-jDE gradually becomes smaller with the increase of D and M. In terms of the differences among the adaptive MOEA/D-DEs, MOEA/D-JADE derives the best average ranks for $D \leq 50$ and $M \leq 7$; however, its average ranks clearly degrade when D and M are further increasing to 100 and 11, respectively. In contrast, the average rank of MOEA/D-jDE gradually improves with the increase of the D and M. MOEA/D-EPSDE and MOEA/D-SaDE never derive the best average rank.

Thus, our experiments show that the parameter-only adaptation like jDE and JADE contributes to boost the performance of MOEA/D-DE; the statistical adaptation like JADE effectively outperforms the randomization adaptation like jDE

TABLE III

THE IGD VALUES OF THE FIVE ALGORITHMS WITH $D = \{20, 50, 100\}$ FOR $M = 3$ (TOP), 7 (CENTER), 11 (BOTTOM); MOEA/D-DE, MOEA/D-JDE, MOEA/D-JADE, MOEA/D-EPSDE, AND MOEA/D-SADE ARE BRIEFLY DENOTED AS -DE, -JDE, -JADE, -EPSDE, AND -SADE, RESPECTIVELY. THE BEST AND WORST VALUES ARE HIGHLIGHTED WITH GREEN (AND BOLD) AND PINK, RESPECTIVELY.

a) $M = 3$

	$D = 20$					$D = 50$					$D = 100$				
	-DE	-jDE	-JADE	-EPSDE	-SaDE	-DE	-jDE	-JADE	-EPSDE	-SaDE	-DE	-jDE	-JADE	-EPSDE	-SaDE
DTLZ1	9.73E+00	5.71E-01	**5.09E-02**	8.45E+00	2.57E+00	4.06E+01	4.64E+01	**2.16E+01**	1.36E+02	1.07E+02	**1.54E+02**	1.57E+02	2.19E+02	3.63E+02	3.74E+02
DTLZ2	7.53E-02	7.54E-02	**7.52E-02**	7.55E-02	7.53E-02	7.73E-02	7.54E-02	**7.46E-02**	7.52E-02	7.53E-02	9.95E-02	7.54E-02	**7.53E-02**	7.54E-02	7.55E-02
DTLZ3	9.90E+00	3.27E-01	**1.11E-01**	7.86E+00	3.41E+00	1.67E+02	9.44E+01	**4.26E+01**	3.98E+02	2.60E+02	**4.33E+02**	4.53E+02	5.23E+02	1.10E+03	9.98E+02
DTLZ4	1.49E-01	7.57E-02	**7.47E-02**	7.57E-02	7.55E-02	1.75E-01	7.60E-02	**7.59E-02**	**7.59E-02**	**7.59E-02**	1.97E-01	**7.71E-02**	7.91E-02	7.74E-02	7.72E-02
DTLZ5	**1.45E-02**	1.46E-02	1.46E-02	1.46E-02	1.46E-02	1.69E-02	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	3.66E-02	**1.47E-02**	1.50E-02	1.49E-02	1.48E-02
DTLZ6	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**	**1.46E-02**
DTLZ7	2.02E-01	2.27E-01	**1.99E-01**	2.03E-01	2.27E-01	2.15E-01	2.23E-01	**2.05E-01**	2.10E-01	2.11E-01	3.89E-01	2.11E-01	2.06E-01	**2.05E-01**	**2.05E-01**
WFG1	1.42E+00	5.53E-01	5.16E-01	6.53E-01	**5.04E-01**	1.53E+00	1.03E+00	9.08E-01	1.17E+00	9.36E-01	1.54E+00	1.32E+00	**1.25E+00**	1.39E+00	1.26E+00
WFG2	3.22E-01	3.02E-01	3.04E-01	**2.95E-01**	3.02E-01	3.62E-01	3.04E-01	3.14E-01	**3.00E-01**	3.11E-01	4.03E-01	3.12E-01	3.28E-01	**3.09E-01**	3.22E-01
WFG3	1.90E-01	**1.09E-01**	1.13E-01	**1.09E-01**	1.11E-01	3.69E-01	1.53E-01	1.65E-01	**1.47E-01**	1.52E-01	4.41E-01	**2.23E-01**	2.48E-01	2.40E-01	2.31E-01
WFG4	3.86E-01	3.56E-01	**3.52E-01**	3.59E-01	3.54E-01	4.06E-01	3.65E-01	**3.58E-01**	3.71E-01	3.62E-01	4.22E-01	3.75E-01	**3.64E-01**	3.80E-01	3.68E-01
WFG5	3.53E-01	3.50E-01	**3.44E-01**	3.51E-01	3.47E-01	3.50E-01	3.48E-01	**3.41E-01**	3.50E-01	3.45E-01	3.53E-01	3.48E-01	**3.42E-01**	3.48E-01	3.43E-01
WFG6	3.88E-01	3.91E-01	3.90E-01	3.93E-01	**3.84E-01**	3.61E-01	3.61E-01	**3.60E-01**	3.62E-01	3.61E-01	3.56E-01	3.57E-01	3.55E-01	3.57E-01	**3.54E-01**
WFG7	3.61E-01	3.57E-01	3.57E-01	**3.56E-01**	3.59E-01	3.77E-01	**3.58E-01**	3.59E-01	**3.58E-01**	3.59E-01	4.29E-01	**3.59E-01**	3.60E-01	**3.59E-01**	3.61E-01
WFG8	4.00E-01	**3.71E-01**	3.72E-01	3.73E-01	3.73E-01	4.50E-01	3.79E-01	**3.78E-01**	3.79E-01	**3.78E-01**	4.95E-01	3.92E-01	**3.86E-01**	3.94E-01	3.89E-01
WFG9	**3.56E-01**	3.64E-01	3.63E-01	3.72E-01	3.67E-01	**3.45E-01**	3.55E-01	3.50E-01	3.48E-01	3.54E-01	**3.41E-01**	3.51E-01	3.45E-01	3.43E-01	3.52E-01

b) $M = 7$

	$D = 20$					$D = 50$					$D = 100$				
	-DE	-jDE	-JADE	-EPSDE	-SaDE	-DE	-jDE	-JADE	-EPSDE	-SaDE	-DE	-jDE	-JADE	-EPSDE	-SaDE
DTLZ1	5.05E+00	1.11E+00	**1.76E-01**	7.34E+00	2.39E+00	**5.05E+01**	6.41E+01	5.72E+01	1.31E+02	1.01E+02	**1.54E+02**	2.51E+02	2.79E+02	3.88E+02	3.20E+02
DTLZ2	5.84E-01	5.21E-01	5.21E-01	**5.19E-01**	5.22E-01	7.25E-01	5.36E-01	**5.23E-01**	5.31E-01	5.24E-01	8.71E-01	5.40E-01	**5.36E-01**	5.53E-01	5.43E-01
DTLZ3	1.23E+01	7.74E-01	**5.19E-01**	1.53E+01	1.84E+00	1.62E+02	1.58E+02	**1.50E+02**	4.01E+02	3.24E+02	**4.54E+02**	7.84E+02	1.04E+03	1.21E+03	1.14E+03
DTLZ4	6.14E-01	5.24E-01	**5.18E-01**	5.22E-01	5.23E-01	7.14E-01	5.87E-01	**5.35E-01**	5.82E-01	5.51E-01	8.06E-01	6.82E-01	**5.76E-01**	6.29E-01	6.38E-01
DTLZ5	5.27E-02	5.25E-02	**5.24E-02**	5.28E-02	5.30E-02	6.45E-02	5.25E-02	5.23E-02	**5.20E-02**	5.25E-02	8.63E-02	5.26E-02	5.45E-02	**5.25E-02**	5.33E-02
DTLZ6	5.23E-02	**5.22E-02**	**5.22E-02**	**5.22E-02**	**5.22E-02**	5.25E-02	**5.23E-02**	**5.23E-02**	5.25E-02	5.24E-02	5.28E-02	**5.24E-02**	5.26E-02	**5.24E-02**	**5.24E-02**
DTLZ7	**1.02E+00**	1.35E+00	1.42E+00	1.28E+00	1.31E+00	1.24E+00	1.24E+00	1.41E+00	**1.05E+00**	1.17E+00	3.61E+00	9.37E-01	1.31E+00	**8.84E-01**	9.51E-01
WFG1	1.82E+00	**1.40E+00**	**1.40E+00**	1.51E+00	1.44E+00	2.30E+00	**1.99E+00**	2.00E+00	2.00E+00	**1.99E+00**	2.58E+00	2.13E+00	2.21E+00	**2.09E+00**	2.21E+00
WFG2	1.61E+00	1.42E+00	**1.39E+00**	1.44E+00	1.43E+00	**1.59E+00**	1.65E+00	1.69E+00	1.70E+00	1.68E+00	1.68E+00	**1.62E+00**	1.70E+00	1.68E+00	1.67E+00
WFG3	2.29E+00	**2.22E+00**	2.23E+00	2.25E+00	2.23E+00	**2.16E+00**	2.29E+00	2.28E+00	2.30E+00	2.34E+00	**2.23E+00**	**2.23E+00**	2.30E+00	2.25E+00	2.32E+00
WFG4	4.66E+00	4.13E+00	4.38E+00	**4.00E+00**	4.34E+00	4.82E+00	4.60E+00	4.49E+00	**4.46E+00**	4.49E+00	4.90E+00	4.77E+00	4.83E+00	4.86E+00	**4.72E+00**
WFG5	4.20E+00	4.17E+00	**4.02E+00**	4.04E+00	4.17E+00	4.29E+00	**4.16E+00**	4.17E+00	4.18E+00	4.36E+00	4.36E+00	**4.15E+00**	4.22E+00	4.23E+00	4.41E+00
WFG6	**4.71E+00**	4.93E+00	4.95E+00	4.94E+00	4.87E+00	**4.77E+00**	4.89E+00	4.95E+00	4.93E+00	4.87E+00	**4.73E+00**	4.88E+00	4.96E+00	4.90E+00	4.78E+00
WFG7	**4.48E+00**	4.85E+00	4.79E+00	4.86E+00	4.82E+00	4.86E+00	4.52E+00	4.51E+00	4.42E+00	**4.39E+00**	4.76E+00	4.68E+00	4.82E+00	4.67E+00	**4.64E+00**
WFG8	4.89E+00	4.83E+00	**4.73E+00**	4.81E+00	4.77E+00	**4.78E+00**	4.84E+00	4.81E+00	4.85E+00	4.83E+00	**4.82E+00**	4.83E+00	4.84E+00	**4.86E+00**	4.84E+00
WFG9	4.43E+00	4.51E+00	**4.35E+00**	4.46E+00	4.52E+00	**4.38E+00**	4.44E+00	**4.38E+00**	4.44E+00	4.43E+00	**4.37E+00**	4.40E+00	**4.37E+00**	4.43E+00	4.49E+00

c) $M = 11$

	$D = 20$					$D = 50$					$D = 100$				
	-DE	-jDE	-JADE	-EPSDE	-SaDE	-DE	-jDE	-JADE	-EPSDE	-SaDE	-DE	-jDE	-JADE	-EPSDE	-SaDE
DTLZ1	4.19E+00	2.00E-01	**1.88E-01**	1.30E+00	2.91E-01	4.82E+01	4.75E+01	**3.38E+01**	9.35E+01	6.66E+01	**1.86E+02**	2.16E+02	1.93E+02	3.04E+02	2.56E+02
DTLZ2	8.20E-01	**7.86E-01**	7.97E-01	7.94E-01	7.93E-01	9.66E-01	7.90E-01	7.88E-01	7.90E-01	**7.80E-01**	1.08E+00	8.14E-01	8.03E-01	8.19E-01	**7.80E-01**
DTLZ3	4.26E+00	**7.61E-01**	**7.61E-01**	4.04E+00	7.64E-01	1.48E+02	1.09E+02	**8.16E+01**	3.76E+02	2.51E+02	**4.55E+02**	7.01E+02	8.74E+02	1.33E+03	1.18E+03
DTLZ4	8.30E-01	7.95E-01	**7.76E-01**	7.77E-01	7.79E-01	8.94E-01	8.52E-01	**8.18E-01**	8.36E-01	8.32E-01	9.73E-01	9.28E-01	**8.72E-01**	8.97E-01	9.10E-01
DTLZ5	**7.72E-02**	9.69E-02	9.64E-02	1.04E-01	9.55E-02	8.37E-02	**8.01E-02**	8.69E-02	8.48E-02	8.56E-02	1.25E-01	**7.47E-02**	8.95E-02	8.00E-02	8.12E-02
DTLZ6	**6.78E-02**	6.79E-02	6.89E-02	6.80E-02	6.95E-02	6.82E-02	**6.79E-02**	6.84E-02	6.94E-02	6.85E-02	7.14E-02	**6.87E-02**	7.16E-02	6.88E-02	7.12E-02
DTLZ7	**1.57E+00**	5.79E+00	6.62E+00	6.36E+00	5.68E+00	**1.68E+00**	5.73E+00	6.47E+00	5.96E+00	5.50E+00	**3.56E+00**	5.17E+00	6.20E+00	4.85E+00	4.65E+00
WFG1	2.36E+00	**2.04E+00**	**2.04E+00**	2.15E+00	2.07E+00	2.97E+00	2.64E+00	2.69E+00	**2.58E+00**	**2.58E+00**	3.36E+00	2.87E+00	2.99E+00	**2.75E+00**	2.82E+00
WFG2	2.24E+00	2.20E+00	2.20E+00	2.44E+00	**2.15E+00**	2.61E+00	2.37E+00	2.44E+00	2.44E+00	**2.36E+00**	2.65E+00	**2.33E+00**	2.39E+00	2.38E+00	2.37E+00
WFG3	**3.25E+00**	3.54E+00	3.51E+00	3.65E+00	3.59E+00	**3.28E+00**	3.39E+00	3.59E+00	3.46E+00	3.67E+00	**3.28E+00**	3.30E+00	3.43E+00	3.39E+00	3.49E+00
WFG4	**8.95E+00**	**8.95E+00**	9.39E+00	9.03E+00	9.19E+00	8.94E+00	8.90E+00	**8.13E+00**	8.45E+00	8.50E+00	**8.94E+00**	9.08E+00	9.09E+00	9.28E+00	9.15E+00
WFG5	8.63E+00	8.45E+00	8.40E+00	**8.26E+00**	8.45E+00	8.85E+00	**8.33E+00**	8.36E+00	8.34E+00	8.58E+00	9.10E+00	**8.25E+00**	8.48E+00	8.36E+00	8.59E+00
WFG6	**9.42E+00**	1.01E+01	1.01E+01	1.02E+01	9.86E+00	**9.62E+00**	1.00E+01	1.02E+01	1.02E+01	1.01E+01	**9.66E+00**	1.00E+01	1.02E+01	1.01E+01	9.91E+00
WFG7	**9.15E+00**	9.70E+00	9.58E+00	1.01E+01	9.70E+00	9.96E+00	9.01E+00	**8.88E+00**	8.89E+00	8.92E+00	9.79E+00	9.17E+00	9.21E+00	9.20E+00	**9.10E+00**
WFG8	1.06E+01	1.03E+01	1.03E+01	**1.01E+01**	1.04E+01	1.05E+01	1.00E+01	**9.69E+00**	9.77E+00	9.96E+00	1.02E+01	9.66E+00	**9.41E+00**	9.44E+00	9.86E+00
WFG9	**8.78E+00**	9.03E+00	9.55E+00	8.86E+00	9.01E+00	9.20E+00	**9.03E+00**	9.89E+00	9.16E+00	9.49E+00	**9.00E+00**	9.52E+00	1.01E+01	9.66E+00	9.92E+00

on small-scale problems; however, the randomization adaptation improves the scalability to the number of decision variables and objectives. This is an interesting trend compared to existing results for SOPs. Specifically, many works have empirically revealed that the statistical adaptation (e.g., JADE and SHADE) outperforms the randomization adaptation (e.g., jDE) on SOPs [14], [16]; this trend can be also observed for $D = 100$ [9]. In fact, the randomization adaptation realizes a well-scalable adaptive MOEA/D-DE to D and M on the MOPs. Thus, we find a critical tradeoff among adaptive DEs in terms of the scalability of the MOEA/D-DE framework.

We further analyze the tradeoff between MOEA/D-jDE and MOEA/D-JADE in terms of the diversity of configurations and the final individuals. Figs. 3, 4, and 5 show the adaptation results of $\theta_{F,i}$ and $\theta_{CR,i}$ obtained by MOEA/D-jDE and MOEA/D-JADE on WFG6 with $D = 100, M = \{3, 7, 11\}$, respectively. Figs. 6 and 7 show the distributions of position variables of the final individuals obtained by MOEA/D-jDE and MOEA/D-JADE on WFG6 for all $M = \{3, 7, 11\}$ with $D = \{50, 100\}$, respectively. Note that k position variables defined in WFG problems have an impact to decide the diversity of Pareto individuals. Note also that the variables are compressed into two dimensions by t-SNE. From Fig. 3, 4, and 5, MOEA/D-jDE tends to derive relatively well-distributed values of $\theta_{F,i}$ and $\theta_{CR,i}$, while MOEA/D-JADE has a biased distribution for all $M = \{3, 7, 11\}$. These tendencies cause the difference of the distribution of the final individuals. Specifically, in a) and b) in Fig. 6 and a) and b) in Fig. 7,

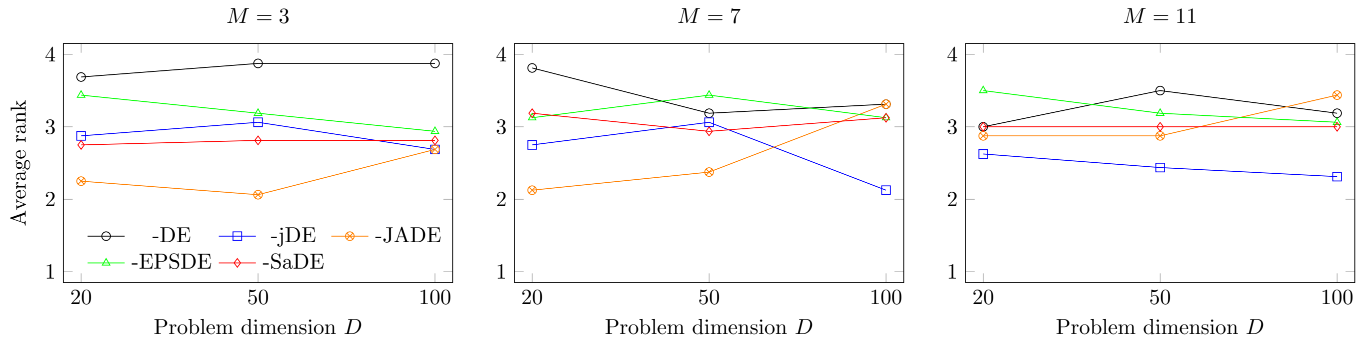


Fig. 1. The average ranks of the five algorithms for $D = \{20, 50, 100\}$.

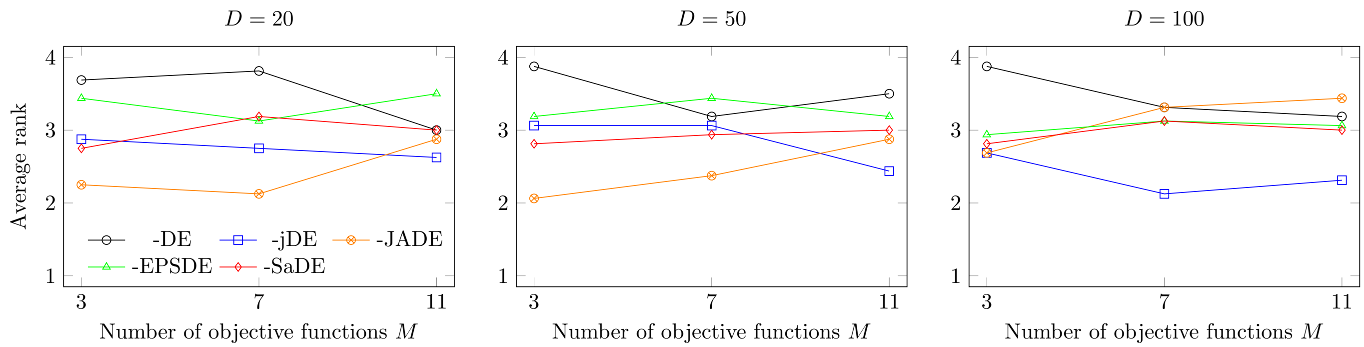


Fig. 2. The average ranks of the five algorithms for $M = \{3, 7, 11\}$.

both MOEA/D-jDE and MOEA/D-JADE generate position variables that have approximately the same values. However, in c) in Fig. 6 and c) in Fig. 7, position variables derived by MOEA/D-jDE cover almost the same region as in the position variables derived by MOEA/D-JADE. In addition, MOEA/D-jDE also has position variables existed in other regions. In other words, as D and M increase to 100 and 11, it becomes more pronounced that MOEA/D-jDE has more diverse hyper-parameters than MOEA/D-JADE, which makes the position variables of MOEA/D-jDE more diverse than those of MOEA/D-JADE. Thus, although MOEA/D-JADE derives better IGD in small-scale WFG6, MOEA/D-jDE outperforms MOEA/D-JADE on WFG6 with increased D and M by more diverse Pareto individuals of MOEA/D-jDE.

IV. Conclusion

This paper compared jDE, JADE, EPSDE, and SaDE on the MOEA/D-DE framework to evaluate their scalability to the number of decision variables D and objectives M. Our experimental results provided the following observations. Firstly, the adaptation of F and CR effectively improves the performance of MOEA/D-DE compared to the adaptation of the mutation strategy. However, the differences of the type of algorithmic configurations to be adapted have less impact on the scalability of the adaptive MOEA/D-DEs. Secondly, the difference of the adaptation strategies (i.e., the randomization adaptation and the statistical adaptation) can be a critical factor to affect the scalability to the number of decision variables and objectives. The JADE-like statistical adaptation outperforms the jDE-like randomization adaptation on DTLZ and WFG problems with $D \leq 50$ and $M \leq 7$. However, the jDE-like randomization adaptation can contribute to realize a well-scalable adaptive MOEA/D-DE. From those observations, we suggest that a hybridization of the statistical and randomization adaptations may be important to further improve adaptive MOEA/D-DEs. The investigation of this suggestion can be our future work.

References

[1] A. Trivedi, D. Srinivasan, K. Sanyal, and A. Ghosh, "A survey of multiobjective evolutionary algorithms based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 440–462, 2016.

[2] J. Wang, Y. Su, Q. Lin, L. Ma, D. Gong, J. Li, and Z. Ming, "A survey of decomposition approaches in multiobjective evolutionary algorithms," *Neurocomputing*, vol. 408, pp. 308–330, 2020.

[3] N. M. Hamza, D. L. Essam, and R. A. Sarker, "Constraint consensus mutation-based differential evolution for constrained optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 447–459, 2015.

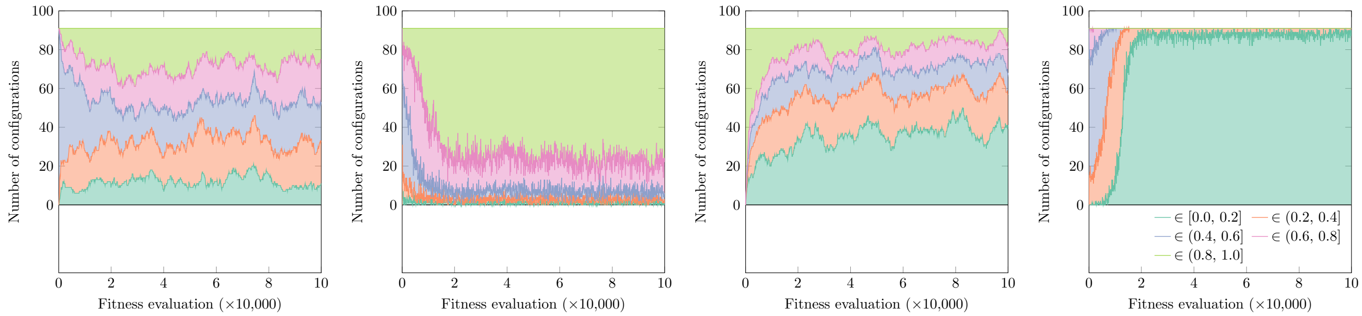
[4] Á. E. Eiben, R. Hinterding, and Z. Michalewicz, "Parameter control in evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 3, no. 2, pp. 124–141, 1999.

[5] S. M. Venske, R. A. Gonçalves, and M. R. Delgado, "ADEMO/D: Multiobjective optimization by an adaptive differential evolution algorithm," *Neurocomputing*, vol. 127, pp. 65–77, 2014.

[6] A. E. Eiben, Z. Michalewicz, M. Schoenauer, and J. E. Smith, "Parameter control in evolutionary algorithms," in *Parameter setting in evolutionary algorithms*. Springer, 2007, pp. 19–46.

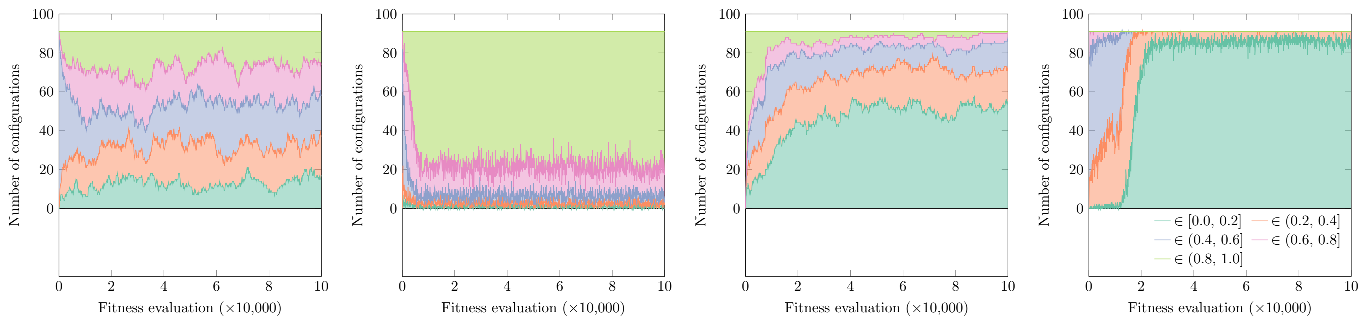
[7] G. Karafotias, M. Hoogendoorn, and Á. E. Eiben, "Parameter control in evolutionary algorithms: Trends and challenges," *IEEE Trans. Evol. Comput.*, vol. 19, no. 2, pp. 167–187, 2014.

[8] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.



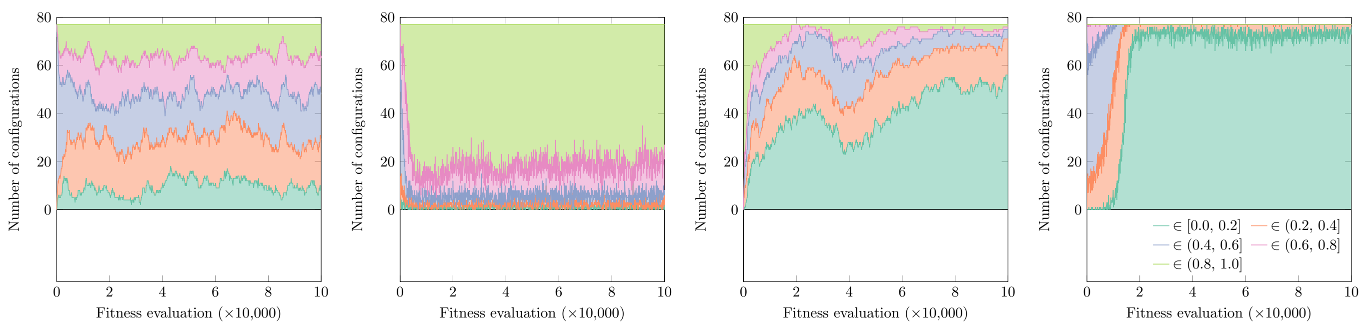
a) MOEA/D-jDE $\theta_{F,i}$ b) MOEA/D-JADE $\theta_{F,i}$ c) MOEA/D-jDE $\theta_{CR,i}$ d) MOEA/D-JADE $\theta_{CR,i}$

Fig. 3. Examples of the hyper-parameter $\theta_{F,i}$ and $\theta_{CR,i}$ obtained by MOEA/D-jDE and MOEA/D-JADE on WFG6 with $D = 100, M = 3$. Note that figures are reported with stacked curves; each curve indicates the number of configurations which employ values in the corresponding range.



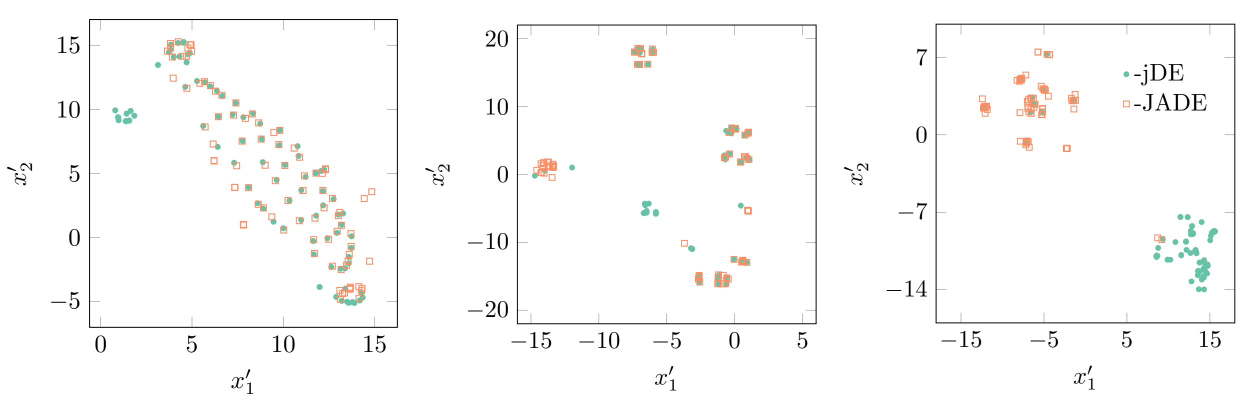
a) MOEA/D-jDE $\theta_{F,i}$ b) MOEA/D-JADE $\theta_{F,i}$ c) MOEA/D-jDE $\theta_{CR,i}$ d) MOEA/D-JADE $\theta_{CR,i}$

Fig. 4. Examples of the hyper-parameter $\theta_{F,i}$ and $\theta_{CR,i}$ obtained by MOEA/D-jDE and MOEA/D-JADE on WFG6 with $D = 100, M = 7$.



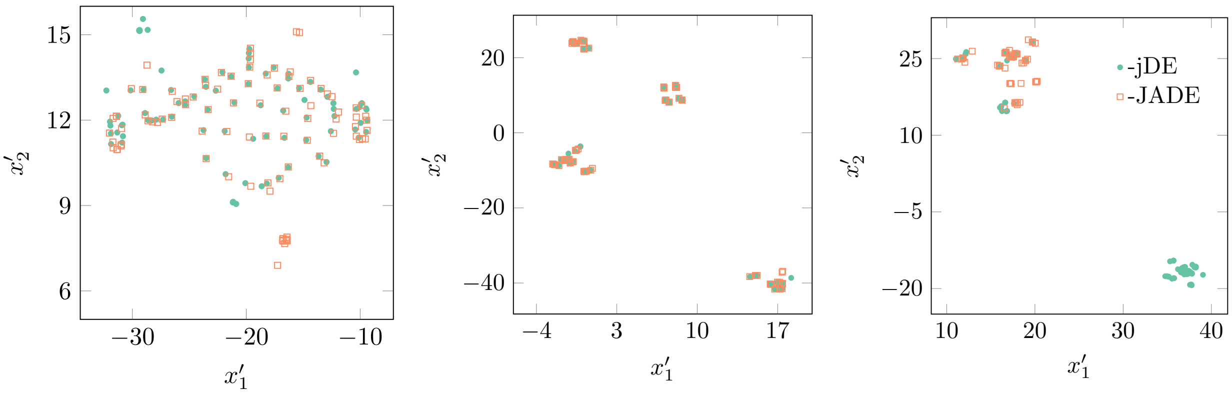
a) MOEA/D-jDE $\theta_{F,i}$ b) MOEA/D-JADE $\theta_{F,i}$ c) MOEA/D-jDE $\theta_{CR,i}$ d) MOEA/D-JADE $\theta_{CR,i}$

Fig. 5. Examples of the hyper-parameter $\theta_{F,i}$ and $\theta_{CR,i}$ obtained by MOEA/D-jDE and MOEA/D-JADE on WFG6 with $D = 100, M = 11$.



a) $M = 3$ b) $M = 7$ c) $M = 11$

Fig. 6. Distributions of position variables obtained by MOEA/D-jDE and MOEA/D-JADE on WFG6 with $D = 50, M = \{3, 7, 11\}$, at the maximum fitness evaluation. MOEA/D-jDE and MOEA/D-JADE are briefly denoted as -jDE and -JADE, respectively.



a) $M = 3$ b) $M = 7$ c) $M = 11$

Fig. 7. Distributions of position variables obtained by MOEA/D-jDE and MOEA/D-JADE on WFG6 with $D = 100, M = \{3, 7, 11\}$, at the maximum fitness evaluation.

[9] N. H. Awad, M. Z. Ali, P. N. Suganthan, and R. G. Reynolds, "An ensemble sinusoidal parameter adaptation incorporated with L-SHADE for solving CEC2014 benchmark problems," in *Proc. IEEE Congr. Evol. Comput.* IEEE, 2016, pp. 2958–2965.

[10] J. Brest, M. S. Maučec, and B. Bošković, "Single objective real-parameter optimization: algorithm jSO," in *Proc. IEEE Congr. Evol. Comput.* IEEE, 2017, pp. 1311–1318.

[11] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 284–302, 2008.

[12] S. Akhmedova and V. Stanovov, "Success-History Based Parameter Adaptation in MOEA/D Algorithm," in *Int. Conf. on Swarm Intell.* Springer, 2020, pp. 455–462.

[13] J. Brest, S. Greiner, B. Boskovic, M. Mernik, and V. Zumer, "Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems," *IEEE Trans. Evol. Comput.*, vol. 10, no. 6, pp. 646–657, 2006.

[14] R. Tanabe and A. Fukunaga, "Success-history based parameter adaptation for differential evolution," in *Proc. IEEE Congr. Evol. Comput.* IEEE, 2013, pp. 71–78.

[15] X. Qiu, W. Xu, J.-X. Xu, and K. C. Tan, "A new framework for self-adapting control parameters in multi-objective optimization," in *Proc. Annu. Conf. Genet. Evol. Comput.*, 2015, pp. 743–750.

[16] J. Zhang and A. C. Sanderson, "JADE: adaptive differential evolution with optional external archive," *IEEE Trans. Evol. Comput.*, vol. 13, no. 5, pp. 945–958, 2009.

[17] S. M. Islam, S. Das, S. Ghosh, S. Roy, and P. N. Suganthan, "An adaptive differential evolution algorithm with novel mutation and crossover strategies for global numerical optimization," *IEEE Trans. Syst. Man Cybern.*, vol. 42, no. 2, pp. 482–500, 2011.

[18] K. Li, A. Fialho, S. Kwong, and Q. Zhang, "Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 18, no. 1, pp. 114–130, 2013.

[19] Q. Lin, Z. Liu, Q. Yan, Z. Du, C. A. C. Coello, Z. Liang, W. Wang, and J. Chen, "Adaptive composite operator selection and parameter control for multiobjective evolutionary algorithm," *Inform. Sci.*, vol. 339, pp. 332–352, 2016.

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

[21] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 398–417, 2008.

[22] S. Das, S. S. Mullick, and P. N. Suganthan, "Recent advances in differential evolution–an updated survey," *Swarm and Evol. Comput.*, vol. 27, pp. 1–30, 2016.

[23] R. Tanabe and A. Fukunaga, "Reviewing and benchmarking parameter control methods in differential evolution," *IEEE Trans. Cybern.*, vol. 50, no. 3, pp. 1170–1184, 2019.

[24] Y. Wang, Z. Cai, and Q. Zhang, "Differential evolution with composite trial vector generation strategies and control parameters," *IEEE Trans. Evol. Comput.*, vol. 15, no. 1, pp. 55–66, 2011.

[25] K. Nishihara and M. Nakata, "Competitive-Adaptive Algorithm-Tuning of Metaheuristics inspired by the Equilibrium Theory: A Case Study," in *Proc. IEEE Congr. Evol. Comput.* IEEE, 2020, pp. E–24 156, (8).

[26] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization*. Springer, 2005, pp. 105–145.

[27] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 477–506, 2006.

[28] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, 2007.

[29] W.-x. Wang, K.-s. Li, X.-z. Tao, and F.-h. Gu, "An improved MOEA/D algorithm with an adaptive evolutionary strategy," *Inform. Sci.*, vol. 539, pp. 1–15, 2020.

[30] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, 2013.

[31] C. A. C. Coello and N. C. Cortés, "Solving multiobjective optimization problems using an artificial immune system," *Genetic Programming and Evolvable Machines*, vol. 6, no. 2, pp. 163–190, 2005.

[32] E. Zitzler and L. Thiele, "Multiobjective optimization using evolutionary algorithms—a comparative case study," in *Int. Conf. on Parallel Problem Solving from Nature*. Springer, 1998, pp. 292–301.

[33] R. Tanabe and H. Ishibuchi, "An analysis of quality indicators using approximated optimal distributions in a 3-D objective space," *IEEE Trans. Evol. Comput.*, vol. 24, no. 5, pp. 853–867, 2020.

[34] Y. Tian, R. Cheng, X. Zhang, and Y. Jin, "PlatEMO: A MATLAB platform for evolutionary multi-objective optimization [educational forum]," *IEEE Comput. Intell. Magazine*, vol. 12, no. 4, pp. 73–87, 2017.