Handling Constrained Multiobjective Optimization Problems with Constraints in Both the Decision and Objective Spaces

Zhi-Zhong Liu and Yong Wang, *Senior Member, IEEE*

***Abstract*—Constrained multiobjective optimization problems (CMOPs) are frequently encountered in real-world applications, which usually involve constraints in both the decision and objective spaces. However, current artificial CMOPs never consider constraints in the decision space (i.e., decision constraints) and constraints in the objective space (i.e., objective constraints) at the same time. As a result, they have a limited capability to simulate practical scenes. To remedy this issue, a set of CMOPs, named DOC, is constructed in this paper. It is the first attempt to consider both the *d*ecision and *o*bjective *c*onstraints simultaneously in the design of artificial CMOPs. Specifically, in DOC, various decision constraints (e.g., inequality constraints, equality constraints, linear constraints, and nonlinear constraints) are collected from real-world applications, thus making the feasible region in the decision space have different properties (e.g., nonlinear, extremely small, and multimodal). On the other hand, some simple and controllable objective constraints are devised to reduce the feasible region in the objective space and to make the Pareto front have diverse characteristics (e.g., continuous, discrete, mixed, and degenerate). As a whole, DOC poses a great challenge for a constrained multiobjective evolutionary algorithm (CMOEA) to obtain a set of well-distributed and well-converged feasible solutions. In order to enhance current CMOEAs' performance on DOC, a simple and efficient *two-p*hase framework, named ToP, is proposed in this paper. In ToP, the first phase is implemented to find the promising feasible area by transforming a CMOP into a constrained single-objective optimization problem. Then in the second phase, a specific CMOEA is executed to obtain the final solutions. ToP is applied to four state-of-the-art CMOEAs, and the experimental results suggest that it is quite effective.**

***Index Terms*—Constrained multiobjective optimization problems, decision space, objective space, constraint-handling technique, evolutionary algorithms**

I. Introduction

CONSTRAINED multiobjective optimization problems (CMOPs) refer to the multiobjective optimization problems including constraints. CMOPs are not far away from our daily life. For example, if someone wants to buy a car, he/she may consider two essential issues: the minimum cost and at the same time the maximum comfort. Besides, he/she may have some other requirements. For example, the number of seats should be more than five, the car should be a German car, and the cost of the car should be less than \$50K. This car selection problem can be regarded as a CMOP since it contains two conflicting objectives (i.e., the minimum cost and the maximum comfort) and three constraints (i.e., $Seats \geq 5$, $Country = Germany$, and $Cost \leq$ \$50K).

Without loss of generality, a CMOP can be expressed as:

$$\begin{aligned} \min \quad & \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))^T \in \mathbb{F} \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, l \\ & h_j(\mathbf{x}) = 0,\ j = l+1, \ldots, n \\ & \mathbf{x} = (x_1, x_2, \ldots, x_D)^T \in \mathbb{S} \end{aligned} \tag{1}$$

where $\mathbf{x}$ is a D-dimensional decision vector, $\mathbb{S}$ is the decision space, $\mathbf{F}(\mathbf{x})$ consists of m real-valued objective functions, $\mathbb{F}$ is the objective space, $f_i(\mathbf{x})$ is the ith objective function, $g_j(\mathbf{x})$ is the jth inequality constraint, $h_j(\mathbf{x})$ is the $(j-l)$th equality constraint, and l and $(n-l)$ are the number of inequality and equality constraints, respectively.

Indeed, CMOPs are frequently encountered in many science and engineering disciplines. Many real-world applications can be formulated as CMOPs, such as the web service location allocation [1], the risk-constrained energy and reserve procurement [2], the optimal scheduling in microgrids [3], the optimal demand response strategies to mitigate oligopolistic behavior [4], and the deployment optimization of near space communication [5]. These real-world CMOPs usually contain multiple objective functions[1] and diverse constraints.

In terms of constraints, they usually have various forms and features [6], [7], [8], [9]. According to our observation, constraints in real-world CMOPs can be roughly classified into two categories: the decision constraints and the objective constraints. In this paper, the decision constraints refer to the constraints which cannot be expressed by objective functions and thus cannot be described explicitly in the objective space. Most constraints in real-world applications belong to the decision constraints. For the objective constraints, they are formulated by objective functions due to the fact that objective functions may be mutually restricted. Overall, the decision and objective constraints are easier to be interpreted in the decision and objective spaces, respectively. To make a clear explanation of these two kinds of constraints, we take the

This work was supported in part by the Innovation-Driven Plan in Central South University under Grant 2018CX010, in part by the National Natural Science Foundation of China under Grant 61673397, in part by the Hunan Provincial Natural Science Fund for Distinguished Young Scholars (Grant No. 2016JJ1018), in part by the Beijing Advanced Innovation Center for Intelligent Robots and Systems under Grant 2018IRS06, and in part by the Graduate Innovation Fund of Hunan Province of China under Grant CX2017B062. (*Corresponding author: Yong Wang*).

Z.-Z. Liu and Y. Wang are with the School of Automation, Central South University, Changsha 410083, China. (Email: zhizhongliu@csu.edu.cn; ywang@csu.edu.cn)

[1]At least two of them are conflicting with each other.

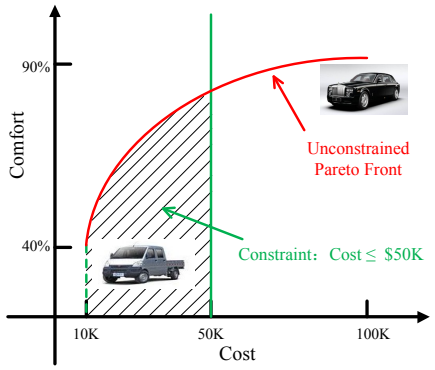


Fig. 1. Illustration of the objective constraint in the objective space. It is obvious that the constraint $Cost \leq$ \$50K is an objective constraint in the objective space. However, the constraints $Seats \geq 5$ and $Country = Gemany$ cannot be described in a straightforward way in the objective space.

above car selection problem as an example. As shown in Fig. 1, the constraint $Cost \leq$ \$50K can be directly described in the objective space while the constraints $Seats \geq 5$ and $Country = Gemany$ cannot. Therefore, $Cost \leq$ \$50K is an objective constraint and $Seats \geq 5$ and $Country = Gemany$ are two decision constraints. In general, combining these two kinds of constraints together causes great difficulties and challenges to the solution of CMOPs.

Although many researchers have a consensus that constrained multiobjective evolutionary algorithms (CMOEAs) are a promising way to deal with CMOPs [10], little effort has been devoted to developing CMOEAs in the community of evolutionary computation although both multiobjective EAs (MOEAs) [11] and constraint-handling techniques [12] have been extensively investigated individually. To boost the development of CMOEAs, constrained multiobjective optimization benchmark functions are always required. Note, however, that real-world CMOPs are usually not suitable to be benchmark functions, since the computational simulation of them may require special hardware or software [13]. Hence, it is a very important topic to devise some representative artificial CMOPs to simulate real-world scenes, which can definitely help researchers to analyze and understand the performance of CMOEAs and encourage users to select the desired ones.

Unfortunately, current artificial CMOPs do not consider the decision constraints and the objective constraints simultaneously. For some artificial CMOPs, such as CTP [8], C-DTLZ [14], NCTPs [15] and DAS-CMOPs [16], only the objective constraints are considered while the decision constraints are neglected in the design process. In addition, with respect to CFs [17], the decision and objective constraints are considered individually. Due to the fact that the decision and objective constraints widely coexist in real-world CMOPs, the capability of current artificial CMOPs to simulate practical scenes is limited. Moreover, according to the report in [13], for many current artificial CMOPs, even an MOEA without any constraint-handling technique can find well-approximated feasible solutions. The above phenomenon means that the effectiveness of these artificial CMOPs is questionable. Therefore, it is necessary to carry out an in-depth investigation on the construction of artificial CMOPs.

Based on the above considerations, we design a set of CMOPs in this paper, named DOC, which takes both the *d*ecision and *o*bjective *c*onstraints into account simultaneously. In DOC, a variety of decision constraints is collected from real-world applications (e.g., inequality constraints, equality constraints, linear constraints, and nonlinear constraints), which can help to construct the feasible region in the decision space and make the feasible region have many properties (e.g., nonlinear, extremely small, and multimodal). Meanwhile, some simple and controllable objective constraints are designed to restrict the feasible region in the objective space and to make the Pareto front associated with various characteristics, such as continuous, discrete, mixed, and degenerate.

Due to the above complicated properties of DOC, the performance of current CMOEAs is poor based on our experiments. When solving DOC, it is very challenging for current CMOEAs to approach the Pareto optimal set. To improve the performance of current CMOEAs, we propose a brand-new *two-p*hase framework, referred to as ToP. In the first phase, a CMOP in DOC is transformed into a constrained single-objective optimization problem. This transformation not only leads to faster convergence speed but also alleviates the premature convergence inside the feasible region. Moreover, the objective function information can be flexibly utilized because of the single objective function. Afterward, a specific CMOEA is implemented in the second phase to obtain the final solutions. Because of its simple structure, ToP can be integrated with many current CMOEAs. In this paper, we have successfully applied ToP to four state-of-the-art CMOEAs: NSGA-II-CDP [10], IDEA [18], CMOEAD [19], and MOEA/D-CDP [20].

The main contributions of this paper are summarized as follows:

- This paper constructs a set of novel artificial CMOPs, called DOC. To the best of our knowledge, it is the first attempt to consider both decision and objective constraints simultaneously in the design of artificial CMOPs. Moreover, both equality and inequality constraints are involved in DOC. Note that at present, few artificial CMOPs consider equality constraints [21].
- A novel two-phase framework named ToP is proposed in this paper. The unique feature of ToP is its first phase, in which we transform a CMOP into a constrained single-objective optimization problem by making use of the weighted sum approach. In addition, current CMOEAs can be directly applied in the second phase. These two phases aim at discovering the promising feasible area and achieving the Pareto optimal solutions, respectively. It should be noted that we do not design any new weighted sum approach, constraint-handling technique, search engine, and CMOEA. It is because we focus on the reason why the performance of current CMOEAs is poor on DOC and how to enhance the performance of current CMOEAs on DOC through a two-phase perspective. Moreover, we would like to keep our framework simple to understand and easy to implement.
- Systematic experiments have been conducted on DOC

to verify the effectiveness of ToP. The experimental results suggest that ToP can significantly improve the performance of four well-established CMOEAs.

The rest of this paper is organized as follows. Section II introduces the related work. The details of DOC are given in Section III. Subsequently, ToP is presented in Section IV. The experimental setup is introduced in Section V and the experiments and discussions are carried out in Section VI. Finally, Section VII concludes this paper.

II. Related Work

In this section, we will introduce some basic definitions in CMOPs and give a brief introduction to current artificial CMOPs and CMOEAs.

A. Basic Definitions in CMOPs

- *Pareto Dominance:* Considering the m objective functions of a CMOP in (1) and two decision vectors $\mathbf{x}_u$ and $\mathbf{x}_v$, if $\forall i \in \{1,2,\ldots,m\}$, $f_i(\mathbf{x}_u) \leq f_i(\mathbf{x}_v)$ and $\exists i \in \{1,2,\ldots,m\}$, $f_i(\mathbf{x}_u) < f_i(\mathbf{x}_v)$, then $\mathbf{x}_u$ is said to Pareto dominate $\mathbf{x}_v$, denoted as $\mathbf{x}_u \prec \mathbf{x}_v$.
- *Feasible Region:* The feasible region of a CMOP in (1) is defined as $\mathbb{O} = \{\mathbf{x} \in \mathbb{S} | CV(\mathbf{x}) = 0\}$, where

$$CV(\mathbf{x}) = \sum_{i=1}^{n} CV_i(\mathbf{x}) \tag{2}$$

is the degree of constraint violation on all the constraints and $CV_i(\mathbf{x})$ is the degree of constraint violation on the ith constraint:

$$CV_i(\mathbf{x}) = \begin{cases} max(0, g_i(\mathbf{x})), & \text{if } i \leq l \\ max(0, |h_i(\mathbf{x})| - \eta), & \text{otherwise} \end{cases}, \ i = 1, \ldots, n \tag{3}$$

In (3), η is a very small positive value (e.g., $\eta = 10^{-4}$).

- *Pareto Optimal Solution:* A solution $\mathbf{x}_u \in \mathbb{O}$ is called a Pareto optimal solution of a CMOP if and only if $\neg\exists \mathbf{x}_v \in \mathbb{O}$, $\mathbf{x}_v \prec \mathbf{x}_u$.
- *Pareto Optimal Set:* The Pareto optimal set of a CMOP is defined as $\mathcal{PS} = \{\mathbf{x}_u \in \mathbb{O} | \neg\exists \mathbf{x}_v \in \mathbb{O}, \mathbf{x}_v \prec \mathbf{x}_u\}$.
- *Pareto Front:* The Pareto front of a CMOP is defined as $\mathcal{PF} = \{\mathbf{F}(\mathbf{x}_u) | \mathbf{x}_u \in \mathcal{PS}\}$, which is the image of the Pareto optimal set in the objective space.

B. A Brief Introduction to Current Artificial CMOPs

To date there exist a few artificial CMOPs to test CMOEAs' performance. In this paper, we classify them into two categories, according to the way of constructing constraints.

The first category only constructs the objective constraints yet ignores the decision constraints. Most of current artificial CMOPs belong to this category, including CTPs [8], C-DTLZ [14], DAS-CMOPs [16], and NCTPs [15]. CTPs were proposed by Deb *et al.* in 2000 [8], which are the most commonly used artificial CMOPs. This test set contains seven CMOPs and all of them are two-objective optimization problems. Overall, CTPs provide two different types of difficulties to a CMOEA: 1) the difficulty in the vicinity of the Pareto front, and 2) the difficulty in the entire search space. In the first type, the constraints can make the unconstrained Pareto optimal solutions infeasible and divide the Pareto front into a number of discrete regions. In the second type, the constraints are designed to reduce the feasible region in the entire search space. Note that the difficulties in CTPs are tunable. In terms of the C-DTLZ problems [14], they are constructed by adding constraints to the DTLZ problems [22]. This test suite includes five CMOPs and each of them can be scalable to more than 15 objectives. The C-DTLZ problems can be divided into three types. Type I introduces difficulties in converging to the Pareto front, type II introduces infeasibility to a part of the Pareto front, and in type III, multiple constraints are involved and portions of the added constraint surfaces form the Pareto front. As for DAS-CMOPs, it was presented by Fan *et al.* [16] in 2016. This test suite contains nine CMOPs and all of them are constructed by using a novel toolkit. This toolkit considers three primary types of difficulties to characterize the constraint functions in CMOPs, including feasibility-hardness, convergence-hardness, and diversity-hardness. Afterward, this toolkit constructs three types of parameterized constraint functions according to the proposed three primary types of difficulties. By combining the constraint functions with different parameters, a variety of CMOPs, whose difficulty can be adjustable and scalable, is generated. Very recently, Li *et al.* proposed NCTPs [15], which are based on CTPs [8]. As pointed out in [15], CTPs have some weaknesses, such as the low dimension and large feasible region. To overcome these issues, 18 test instances are devised. Compared with the original CTPs, NCTPs exhibit more complex characteristics. To be specific, in NCTPs, different test instances have different shapes of the Pareto front, different dimensions of the search space, and different sizes of the feasible region.

In the second category, the decision and objective constraints are developed individually. CFs, proposed by Zhang *et al.* [17] in 2008, are a representative in this category. They consist of ten CMOPs, in which seven are two-objective optimization problems and three are three-objective optimization problems. The construction of problems $1-3$ and $8-10$ is inspired by the method introduced in [8], in which only the objective constraints are considered. While for the others, they are constructed by the authors themselves, in which only the decision constraints are involved.

C. A Brief Introduction to Current CMOEAs

In essence, a CMOEA involves two key components: a MOEA and a constraint-handling technique. In terms of the former, the current CMOEAs can be roughly grouped into three classes [23]: the dominance-based approaches, the decomposition-based approaches, and the indicator-based approaches. Since very few attempts are made toward the indicator-based CMOEAs, in this subsection, we only introduce the dominance-based and decomposition-based CMOEAs.

In the dominance-based CMOEAs, the dominance rule is employed to rank the individuals. The most famous one in this class is NSGA-II-CDP, which is an extension of NSGA-II [10] for solving CMOPs. In NSGA-II-CDP, the constraint-domination principle (CDP) is proposed to sort the individuals.

Specifically, any feasible individuals dominate any infeasible individuals, and for two infeasible individuals, the individual with a smaller constraint violation is preferred. IDEA [18] is another popular CMOEA which explicitly maintains a small percentage of infeasible solutions during the evolution. It is claimed that the presence of infeasible solutions is beneficial for IDEA to search the Pareto optimal solutions from both the feasible and infeasible regions. When solving CMOPs, the adaptive tradeoff model [15] divides the whole evolutionary process into three situations according to the feasibility proportion of the current population. These three situations are the infeasible situation, the semi-feasible situation, and the feasible situation. In different situations, different constraint-handling techniques are designed to cope with constraints. In [24], Woldesenbet and Yen proposed a CMOEA based on an adaptive penalty function and a distance measure, which can not only search for the Pareto optimal solutions in the feasible region, but also exploit the important information provided by the infeasible individuals with better objective function values and lower constraint violations. In [25], Young presented a CMOEA, which can cross the infeasible regions of the objective space and find the true constrained Pareto front. The main idea of this algorithm is to blend an individual's rank in the objective space with its rank in the constraint space. In [26], a CMOEA is proposed in which the Pareto dominance concept is combined with a constraint-handling technique and a diversity mechanism. In [27], a new constraint-handling technique based on Pareto-optimality and niching concept is presented for handling CMOPs. In [28] and [29], the simulated annealing and immune system model are used for solving CMOPs, respectively, while in [30] and [31], differential evolution (DE) is applied to cope with CMOPs. In [32], Jiao *et al.* introduced a modified objective function method to lead a dominance checking, and adopted a feasible-guiding strategy to repair the infeasible individuals. Recognizing the limitation of a single constraint-handling technique, a novel CMOEA is proposed in [33], which makes use of an ensemble of constraint-handling techniques to solve CMOPs.

In the decomposition-based CMOEAs, the original CMOP is decomposed into a set of constrained single-objective optimization problems, and then these problems are optimized in a collaborative way [34]. In [19], a decomposition-based CMOEA named CMOEAD is proposed, which can be regarded as an extension of MOEA/D [35] by adding a novel constraint-handing technique. In this constraint-handling technique, the violation threshold is adaptively adjusted based on the type of constraints, the size of the feasible space, and the search outcome. The aim of this constraint-handling technique is to add selection pressure, and to make the infeasible solutions with violations less than the identified threshold at par with the feasible solutions. In [20], the extended/modified versions of stochastic ranking and CDP are implemented under the MOEA/D framework. The experimental results suggest that CDP works better than stochastic ranking. In [36], an improved epsilon constraint-handling technique is applied to MOEA/D, in which the epsilon level is dynamically adapted according to the feasibility ratio in the current population. Recently, a push and pull search (PPS) framework is embedded into MOEA/D for solving CMOPs [37]. In PPS, the search process is divided into two different stages: the push stage and the pull stage. In the push stage, the constraints are ignored, aiming to get across the infeasible regions in front of the unconstrained Pareto front. Then in the pull stage, the constrains are considered, and an improved epsilon constraint-handling technique is employed to pull the solutions obtained in the push stage toward the feasible and nondominated area. Very recently, in [38], a constraint-handling technique named angle-based constrained dominance principle is incorporated into MOEA/D for solving CMOPs.

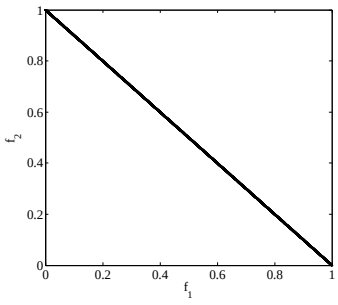


Fig. 2. Pareto front of the objective functions in (4).

III. Proposed CMOPs

From the above introduction, it is clear that current artificial CMOPs never include constraints in both the decision and objective spaces simultaneously. In this paper, we seek to remedy this issue by constructing a new set of CMOPs, called DOC. Next, we will introduce the principle, instances, and characteristics of DOC.

A. Principle of DOC

In general, a CMOP with constraints in both the decision and objective spaces should comprise of three aspects: objective functions, objective constraints, and decision constraints. In DOC, three steps are implemented to construct these three aspects, respectively.

It is well-known that in unconstrained multiobjective optimization, many excellent benchmark functions have been put forward during the past two decades [17], [22], [39], [40]. Thus, in the first step, it is not a difficult task to construct objective functions for CMOPs, since we can benefit from the benchmark functions of unconstrained multiobjective optimization. For example, based on the idea in [40], it is easy to construct the following two objective functions:

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})\left(1 - \frac{f_1}{g(\mathbf{x})}\right) \end{cases} \tag{4}$$

Suppose that $g(\mathbf{x}) = 1 + x_2 + x_3$, $0 \leq x_1 \leq 1$, and $0 \leq x_2, x_3 \leq 10$. Under this condition, $g(\mathbf{x})$ can reach the minimum value (i.e., 1), when $x_2 = 0$ and $x_3 = 0$. It is necessary to note that the Pareto front of (4) is dependent on the minimum value of $g(\mathbf{x})$ [40]. Based on the minimum value

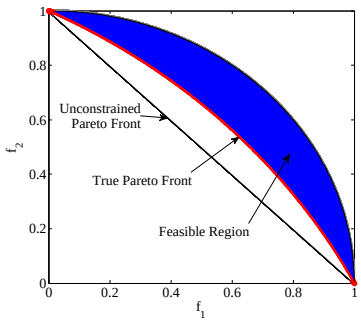


Fig. 3. Illustration of the effect of the objective constraints. The red arc is the Pareto front, and the blue region is the feasible region in the objective space. As can be seen, the objective constraints cut down the feasible region in the objective space and modify the Pareto front.

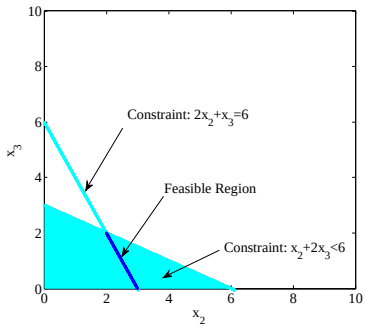


Fig. 4. Illustration of the effect of the decision constraints. The blue line segment is the feasible region. We can see that the decision constraints reduce the feasible region in the decision space significantly.

of $g(\mathbf{x})$, we can obtain the Pareto front, as shown in Fig. 2:

$$\begin{aligned} &f_2 = 1 - f_1 \\ &0 \le f_1 \le 1 \end{aligned} \tag{5}$$

The aim of the second step is to add objective constraints to the objective functions. In principle, the objective constraints can cut down the feasible region in the objective space and produce a constrained Pareto front. It is noteworthy that the constrained Pareto front may be totally different from the unconstrained Pareto front, in terms of the shape and location. Actually, the construction of objective constraints is straightforward. We can make use of some simple functions to describe the relationship among the objective functions. Moreover, we can modify such functions to control the size and shape of the feasible region in the objective space, due to the fact that it is easy to visualize the objective constraints if the number of objective functions is less than four. For example, we add the following two objective constraints to (4):

$$\begin{aligned} &g_1 = f_1^2 + f_2^2 \le 1; \\ &g_2 = (f_1+1)^2 + (f_2+1)^2 \ge 5. \end{aligned} \tag{6}$$

It can be observed from Fig. 3 that the objective constraints result in a small feasible region in the objective space, and the Pareto front is changed to an arc:

$$\begin{aligned} &f_2 = \sqrt{4 - 2f_1 - f_1^2} - 1 \\ &0 \le f_1 \le 1 \end{aligned} \tag{7}$$

Finally, the decision constraints are constructed. Overall, this kind of constraints can reduce the feasible region in the decision space and make the feasible region have a variety of complex properties. In order to explain the construction of decision constraints, we still take (4) as an example and add the following two decision constraints:

$$\begin{aligned} &g_3 = x_2 + 2x_3 \le 6; \\ &h_1 = 2x_2 + x_3 = 6. \end{aligned} \tag{8}$$

As depicted in Fig. 4, the feasible region in the $x_2 - x_3$ plane is extremely small, which is a line segment:

$$\begin{aligned} &x_3 = -2x_2 + 6 \\ &2 \le x_2 \le 3 \end{aligned} \tag{9}$$

If we just consider the decision constraints in (8), $g(\mathbf{x}) = 1 + x_2 + x_3$ will reach the minimum value (i.e., 4) and the maximum value (i.e., 5), when $x_2 = 3$ and $x_3 = 0$ and when $x_2 = 2$ and $x_3 = 2$, respectively. Thus, the value of $g(\mathbf{x})$ ranges from 4 to 5. Under this condition, according to (4), the value of f_2 will range from $(4-f_1)$ to $(5-f_1)$. Due to the fact that $0 \le f_1 = x_1 \le 1$, the value of f_2 will range from 3 to 5. It is interesting to see that the objective constraint g_1 in (6) cannot be satisfied any more. It is because the absolute value of f_2 in g_1 should be less than 1. This phenomenon suggests that the decision constraints may be conflicting with the objective constraints, which may make the constructed CMOP have no feasible solution.

To address this problem, a simple way is to modify $g(\mathbf{x})$ to guarantee that there exist feasible solutions in both the decision and objective spaces. For example, if $g(\mathbf{x})$ is modified by subtracting 3, then $g(\mathbf{x}) = (1+x_2+x_3)-3 = -2+x_2+x_3$. As a result, the finally constructed CMOP is:

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})\left(1 - \frac{f_1}{g(\mathbf{x})}\right) \end{cases} \tag{10}$$

s.t.

$$\begin{aligned} &g_1 = f_1^2 + f_2^2 \le 1; \\ &g_2 = (f_1+1)^2 + (f_2+1)^2 \ge 5; \\ &g_3 = x_2 + 2x_3 \le 6; \\ &h_1 = 2x_2 + x_3 = 6. \end{aligned} \tag{11}$$

where $g(\mathbf{x}) = -2 + x_2 + x_3$, $0 \le x_1 \le 1$, and $0 \le x_2, x_3 \le 10$.

Remark 1: Compared with the decision constraints, the objective constraints can control the Pareto front of a CMOP in the objective space flexibly. Thus, we can easily design various CMOPs with known Pareto fronts. Under this condition, the performance of difference CMOEAs can be evaluated and compared via the designed CMOPs. In principle, the objective constraints and the decision constraints can be dealt with by the same constraint-handling technique during the evolution since both of them should be satisfied.

B. Instances of DOC

Following the above three steps, we have constructed nine instances of DOC, which are presented in the supplementary

TABLE I
INFORMATION OF DOC.

Instance	m	D	NOC	NDC	NIC	NEC	Feasibility Ratio	Properties of the Pareto Front	Properties of the Feasible Region
DOC-1	2	6	1	6	7	0	26.97%	Concave, Continuous	Nonlinear
DOC-2	2	16	2	5	7	0	0.00%	Convex, Disconnected	Very small, Nonlinear
DOC-3	2	10	4	6	6	4	0.00%	Concave, Disconnected, Multimodal	Very small, Nonlinear, Multimodal
DOC-4	2	8	2	4	6	0	0.53%	Linear, Disconnected	Nonlinear, Small
DOC-5	2	8	3	6	4	5	0.00%	Disconnected, Multimodal	Very small, Nonlinear
DOC-6	2	11	2	8	10	0	0.00%	Mixed, Multimodal	Very small, Nonlinear
DOC-7	2	11	3	3	3	3	0.00%	Mixed, Multimodal	Very small, Multimodal
DOC-8	3	10	1	6	7	0	0.00%	Linear, Disconnected	Very small, Nonlinear
DOC-9	3	11	1	13	14	0	0.00%	Degenerate, Multimodal	Very small, Nonlinear, Multimodal

file. To be specific, there are seven CMOPs with two objective functions and two CMOPs with three objective functions.

Although researchers put more emphasis on the objective constraints in the present study, it is more common to face the decision constraints in real-world applications. Moreover, the construction of decision constraints is much harder than that of objective constraints. The reasons are twofold: 1) the number of decision variables is usually significantly larger than that of objective functions; and 2) it is difficult to control the size and shape of the feasible region in the decision space due to the higher dimension. Fortunately, Liang *et al*. [41] collected 24 practical constrained single-objective optimization problems at IEEE CEC2006. These 24 problems only contain the decision constraints. In DOC, we borrow the decision constraints from [41] based on the following two considerations: 1) the decision constraints in [41] have been well-studied during the past 12 years and we have already carried out a series of work to ascertain their properties [42], [43], [44]; and 2) the minimum value of $g(\mathbf{x})$ under the decision constraints can be obtained by employing a powerful constrained single-objective EA, such as the method proposed in [45]. Note that once we get the minimum value of $g(\mathbf{x})$ under the decision constraints of each instance, we can obtain the Pareto front of each instance. Therefore, the Pareto front of each instance in DOC can be known *a priori*.

C. Characteristics of DOC

The main information of the proposed DOC is summarized in Table I. In this table, m denotes the number of objective functions, D is the dimension, NOC and NDC are the number of objective constraints and decision constraints, respectively, and NIC and NEC denote the number of inequality and equality constraints, respectively. For the feasibility ratio, it was estimated by computing the percentage of feasible solutions among 10^5 uniformly and randomly generated solutions from the search space, following the suggestion in [24].

Next, we give the following comments on DOC:

- It considers the objective and decision constraints at the same time. As a result, DOC exhibits better potential to simulate actual scenes, compared with other artificial CMOPs. Hence, it is believed that DOC is a better test bed.
- It contains both inequality and equality constraints. It is common to face equality constraints in actual CMOPs, which can result in a very small feasible region. Unfortunately, few current artificial CMOPs involve equality constraints.
- The Pareto fronts of DOC have various properties, such as continuous, disconnected, convex, concave, linear, mixed, degenerate, and multimodal.
- The feasible regions in the decision space also show many properties, such as nonlinear, very small, and multimodal.
- Because of the above diverse characteristics, it is expected that DOC can attract much attention from the evolutionary computation research community, thus further promoting the development of evolutionary constrained multiobjective optimization.

IV. PROPOSED FRAMEWORK

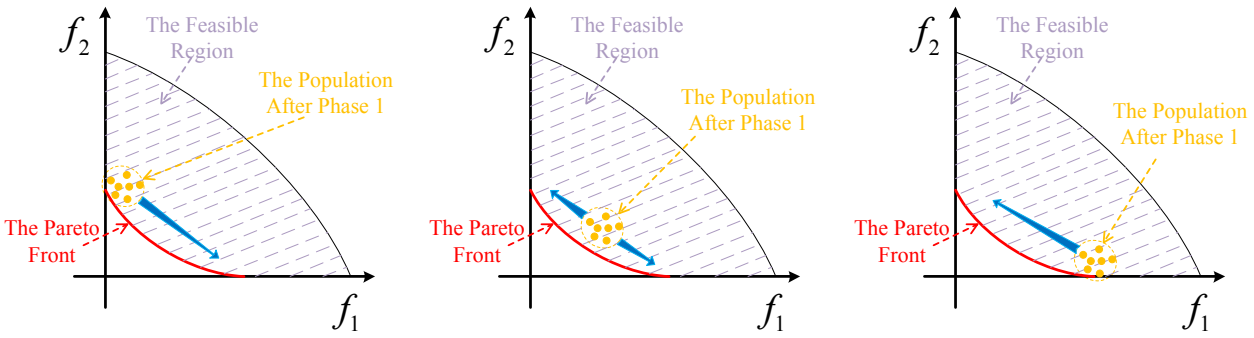
A. ToP

As introduced in Section III, DOC includes many complex properties in both the decision and objective spaces by constructing objective constraints, decision constraints, equality constraints, and inequality constraints, which cause grand difficulties for a CMOEA to firstly find the promising feasible area[2], and subsequently find the Pareto optimal solutions. The reasons why it is hard for a CMOEA to find the promising feasible area are explained as follows:

- Due to the fact that a CMOEA needs to balance all the objective functions in the feasible region, the convergence speed of the population is inevitably slow.
- It is expected that each individual in the population can become a Pareto optimal solution of a CMOP in the end. Suppose that a CMOP has a convex Pareto front. According to [46], we know that each Pareto optimal solution of (1) corresponds to the optimal solution of a constrained single-objective optimization problem with a weight vector $\mathbf{w} = \{w_1, w_2, \ldots, w_m\}^T$. By using the weighted sum approach, (1) can be transformed as:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} w_i * f_i(\mathbf{x}) \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, l \\ & h_j(\mathbf{x}) = 0,\ j = l+1, \ldots, n \\ & \mathbf{x} = (x_1, x_2, \ldots, x_D)^T \in \mathbb{S} \end{aligned} \tag{12}$$

[2]The promising feasible area denotes either the feasible area around the Pareto optimal solutions in the decision space or the feasible area around the Pareto front in the objective space.



(a) Top (b) Central (c) Bottom

Fig. 5. Extending the distribution of the population from different places. It is clear that the extension of the population from the central part of the Pareto front is easier than from both the top and bottom parts of the Pareto front.

where $w_i \geq 0$ for all $i = 1, 2, \ldots, m$, and $\sum_{i=1}^{m} w_i = 1$. Suppose that the population size is N. If there is a local optimal solution in the decision space, then all the N constrained single-objective optimization problems with N different weight vectors should jump out of this local optimal solution, with the purpose of finding the promising feasible area. Obviously, the above process will consume a considerable number of fitness evaluations. In principle, any CMOEA will face this dilemma. More importantly, if there are numerous local optimal solutions in the decision space, the efficiency and effectiveness of a CMOEA will drastically drop.

- In constrained optimization, the information provided by the objective function plays an important role in searching for the optimal solution, in particular, searching for the optimal solution located on the boundaries of the feasible region [45]. However, for a CMOP, it is not trivial to make use of the objective function information as it always includes several conflicting objective functions.

Recognizing the above three aspects, we propose a new framework called ToP, which divides the solution of a CMOP into two phases. In the first phase, a CMOP is transformed into a constrained single-objective optimization problem:

$$\begin{aligned} \min \quad & f'(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x}) \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0, \ j = 1, \ldots, l \\ & h_j(\mathbf{x}) = 0, \ j = l+1, \ldots, n \\ & \mathbf{x} = (x_1, x_2, \ldots, x_D)^T \in \mathbb{S} \end{aligned} \tag{13}$$

(13) is a special case of (12) with the weight vector $\mathbf{w} = \{w_1 = 1/m, w_2 = 1/m, \cdots, w_m = 1/m\}^T$. Compared with other weight vectors, the benefit of this weight vector is explained in the following. The image of the optimal solution of the constrained single-objective optimization problem with this weight vector may be approximately located in the center of the Pareto front of the original CMOP. As a result, after the first phase, the images of the high-quality candidate solutions may also scatter around the central part of the Pareto front of the original CMOP, which makes the extension of the distribution of the population easier (as explained in Fig. 5). Since we only modify the objective functions while keeping the constraints untouched, (1) and (13) share the same feasible region.

Algorithm 1 ToP

Input: a CMOP and the population size N
Output: $\mathcal{P}_{t+1}$

```
Initialization(P_0);
t ← 0
while the stopping criterion is not met do
    if TOP is in its first phase then
        P_{t+1} ← Constrained-Single-Objective-Optimization(P_t)
    else
        P_{t+1} ← Constrained-Multiobjective-Optimization(P_t)
    end if
    t ← t + 1;
end while
```

Compared with solving (1) directly, the above transformation provides the following technical advantages:

- In most cases, (13) only contains one optimal solution and under this condition an optimization method only needs to focus on this optimal solution [47]. As a consequence, the convergence speed of the population in the feasible region is faster.
- For a local optimal solution in the decision space, we only need to guide a constrained single-objective optimization problem, rather than N constrained single-objective optimization problems, to cross the local attraction basin.
- It is easy to obtain and utilize the objective function information. Moreover, such information can be integrated with a search engine to search for the optimal solution.

Because of the aforementioned advantages, the first phase has the good potential to probe the promising feasible area. Nevertheless, it is necessary to note that the ultimate goal is to find the Pareto optimal solutions of the original CMOP, the images of which are uniformly distributed over the Pareto front. To this end, it is necessary to make the images of the individuals produced in the first phase further approach the Pareto front and uniformly spread over the Pareto front simultaneously. Clearly, a CMOEA can be an effective tool to achieve this. Therefore, in the second phase, a specific

CMOEA is implemented.

The main framework of ToP is presented in **Algorithm 1**. Overall, the tasks of the first and second phases are solving the transformed constrained single-objective optimization problem and the original CMOP, respectively.

B. The First Phase—Constrained Single-Objective Optimization

When solving a constrained single-objective optimization problem, an optimization method should include two main components: a constraint-handling technique and a search engine. In addition, since the aim of the first phase is to provide high-quality candidate solutions for the second phase, it is necessary to design a stopping criterion for the first phase.

1) Constraint-handling Technique: The feasibility rule [48] serves as the constraint-handling technique in ToP to compare pairwise individuals. Specifically, given two individuals $\mathbf{x}_u$ and $\mathbf{x}_v$, $\mathbf{x}_u$ is said to be better than $\mathbf{x}_v$ if one of the following cases is satisfied:

- both $\mathbf{x}_u$ and $\mathbf{x}_v$ are the feasible solutions, and $f^{'}(\mathbf{x}_u) < f^{'}(\mathbf{x}_v)$;
- $\mathbf{x}_u$ is feasible yet $\mathbf{x}_v$ is infeasible;
- both $\mathbf{x}_u$ and $\mathbf{x}_v$ are the infeasible solutions, and $CV(\mathbf{x}_u) < CV(\mathbf{x}_v)$;

In general, the feasibility rule can motivate the population to approach or enter the feasible region promptly.

2) Search Engine: DE [49] is considered as the search engine in ToP. We employ two popular trial vector generation strategies of DE to generate offspring for each individual $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,D})^T (i \in \{1, \ldots, N\})$: DE/current-to-rand/1 and DE/rand-to-best/1/bin.

- DE/current-to-rand/1:

$$\mathbf{u}_i = \mathbf{x}_i + F * (\mathbf{x}_{r_1} - \mathbf{x}_i) + F * (\mathbf{x}_{r_2} - \mathbf{x}_{r_3}) \quad (14)$$

- DE/rand-to-best/1/bin:

$$\mathbf{v}_i = \mathbf{x}_{r_1} + F * (\mathbf{x}_{best} - \mathbf{x}_{r_1}) + F * (\mathbf{x}_{r_2} - \mathbf{x}_{r_3}) \quad (15)$$

$$u_{i,j} = \begin{cases} v_{i,j}, \text{if } rand_j < CR \text{ or } j = j_{rand} \\ x_{i,j}, \text{otherwise} \end{cases}, j = 1, \ldots, D. \quad (16)$$

where $\mathbf{v}_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,D})^T$ is the ith mutant vector, $\mathbf{u}_i = (u_{i,1}, u_{i,2}, \ldots, u_{i,D})^T$ is the ith trial vector, r_1, r_2 and r_3 are three mutually different integers randomly chosen from $[1, N]$ and also different from i, $\mathbf{x}_{best}$ denotes the individual with the smallest transformed objective function value in the current population, $rand_j$ is a uniformly distributed random number between 0 and 1 for each j, j_{rand} is a random integer in $[1, D]$, F is the scaling factor, and CR is the crossover control parameter. In DE/current-to-rand/1, the binomial crossover is not applied; thus, it is rotation-invariant.

In DE/current-to-rand/1, each individual learns the information from other randomly selected individuals. In contrast, in DE/rand-to-best/1/bin, the information of the best individual is exploited. Note that the best individual is determined based on the transformed objective function $f^{'}$. As analyzed in [45], before the population enters the feasible region, the individual with the smallest $f^{'}$ may change from generation to generation. Under this condition, the best individual is similar to a randomly selected individual. Therefore, both DE/current-to-rand/1 and DE/rand-to-best/1/bin are able to enhance the exploration ability of the population. After the population enters the feasible region, if the individual with the smallest $f^{'}$ is a feasible solution, then the population will be guided by this individual toward the optimal solution. However, if the individual with the smallest $f^{'}$ is an infeasible solution near the boundaries of the feasible region, it is very likely that the optimal solution is located on the boundaries of the feasible region. In this case, the information of the best individual can be used to search around the boundaries of the feasible region. Therefore, the information provided by the objective function is beneficial for exploration in the early stage of evolution and for exploitation in the later stage of evolution.

Algorithm 2 Constrained-Single-Objective-Optimization

Input: $\mathcal{P}_t = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$
Output: $\mathcal{P}_{t+1}$
1: $\mathcal{P}_{t+1} = \emptyset$
2: **for** $i = 1 : N$ **do**
3: **if** $rand < 0.5$ **then**
4: Generate the trial vector $\mathbf{u}_i$ according to (14);
5: **else**
6: Generate the trial vector $\mathbf{u}_i$ according to (15) and (16);
7: **end if**
8: Employ the feasibility rule to compare $\mathbf{u}_i$ and $\mathbf{x}_i$, and store the better one into $\mathcal{P}_{t+1}$;
9: **end for**

In this paper, DE/current-to-rand/1 and DE/rand-to-best/1/bin are applied to generate a trial vector $\mathbf{u}_i$ for each individual $\mathbf{x}_i$ with the same probability, i.e., 0.5. In addition, following the suggestion in [45], F and CR are randomly chosen from a scaling factor pool (i.e., $F_{pool} = [0.6, 0.8, 1.0]$) and a crossover control parameter pool (i.e., $CR_{pool} = [0.1, 0.2, 1.0]$), respectively. After the trial vector has been generated, the feasibility rule selects the better one between $\mathbf{x}_i$ and $\mathbf{u}_i$ for the next generation.

The framework of constrained single-objective optimization is presented in **Algorithm 2**.

3) Stopping Criterion: If we give plenty of fitness evaluations to the first phase, all the individuals in the population may converge to the optimal solution of the transformed single-objective optimization problem. Obviously, it is not desirable and the evolution should be halted before the population converge to a single point. In essence, when the first phase ends, we expect to obtain a number of high-quality feasible solutions, which are close to the Pareto optimal solutions of the original CMOP, but maintaining a good diversity. Under this circumstance, we consider that the population attains the promising feasible area. To obtain such high-quality feasible solutions, we design the following two conditions:

- *Condition 1*: The feasibility proportion (i.e., P_f) of the current population is larger than $1/3$.
- *Condition 2*: Suppose that $f_{max,j}$ and $f_{min,j}$ represent the maximum and minimum values of the jth objective function among all the discovered feasible solutions during the evolution, respectively. Afterward, the jth objec-

tive function of each feasible individual in the population (denoted as $\mathbf{y}_i$) is normalized as follows:

$$\overline{f}_j(\mathbf{y}_i) = \frac{f_j(\mathbf{y}_i) - f_{min,j}}{f_{max,j} - f_{min,j}} \tag{17}$$

Subsequently, we add up all the normalized objective function values and obtain $\overline{f}'(\mathbf{y}_i)$:

$$\overline{f}'(\mathbf{y}_i) = \sum_{j=1}^{m} \overline{f}_j(\mathbf{y}_i) \tag{18}$$

Finally, we sort the feasible solutions based on $\overline{f}'$, and calculate the biggest difference of $\overline{f}'$ among the first $1/3$ feasible solutions. If this difference (i.e., δ) is less than 0.2, the second condition is regarded to be satisfied.

The aim of condition 1 is to guarantee that a number of feasible individuals have been obtained. In addition, condition 2 denotes that some feasible solutions are of high quality and gradually converge to a small area. Therefore, the first phase should be terminated once both condition 1 and condition 2 are met, thus maintaining the quality of the feasible solutions and preventing the loss of the diversity.

C. The Second Phase—Constrained Multiobjective Optimization

Although the promising feasible area has been found, some individuals in the population may still be far away from the Pareto optimal solutions since we only utilize the best 1/3 feasible solutions in the population to test condition 2. Thus, the whole population should be further evolved toward the Pareto optimal solutions. On the other hand, due to the lack of explicit diversity preservation mechanism in constrained single-objective optimization, it is necessary to spread the image of the whole population throughout the Pareto front.

Fortunately, based on the high-quality candidate solutions produced in the first phase, it is not a difficult task for a CMOEA to achieve well-distributed and well-converged feasible solutions efficiently. The reason is simple: converging to the Pareto front and maintaining a well-distributed set of nondominated feasible solutions are two fundamental goals of a CMOEA. Thereby, in principle, any CMOEA is applicable in the second phase of ToP.

V. Experimental Setup

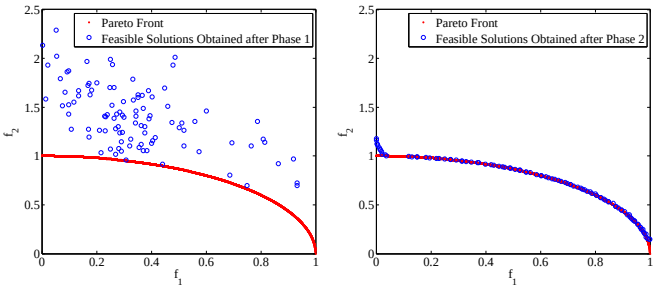
A. Test Instances and Performance Metrics

Our experiments were conducted on DOC which contains nine instances. These instances are denoted as DOC-1-DOC-9, which can be divided into two groups: 1) two-objective CMOPs: DOC-1-DOC-7; and 2) three-objective CMOPs: DOC-8 and DOC-9.

To compare the performance of different algorithms, three indicators were employed in our experiments.

- *Feasible Rate (FR):* Suppose that $FeasibleRuns$ denotes the number of runs where a CMOEA can find at least one feasible solution in the final population, and $TotalRuns$ denotes the number of total runs. Then, FR is defined as:

$$FR = \frac{FeasibleRuns}{TotalRuns} \tag{19}$$



(a) After the first phase (b) After the second phase

Fig. 6. Images of the feasible solutions obtained after the first and second phases, which are provided by ToP-NSGA-II-CDP on DOC-1 in a run.

The value of FR ranges from 0% to 100%, and the larger the value of FR, the higher the probability that a CMOEA enters the feasible region.

- *Inverted Generational Distance (IGD) [50]:* IGD has been widely used to evaluate a MOEA's performance. However, IGD may lose its effectiveness to evaluate a CMOEA owing to the existence of the infeasible solutions. Herein, we only keep the feasible solutions and compute the IGD value of them. Specifically, suppose that $\mathcal{P}$ is the set of images of the feasible solutions, and $\mathcal{P}^*$ is a set of nondominated points uniformly distributed on the Pareto front. Then, the IGD metric is calculated as:

$$IGD(\mathcal{P}) = \frac{1}{|\mathcal{P}^*|} \sum_{z^* \in \mathcal{P}^*} distance(z^*, \mathcal{P}) \tag{20}$$

where $distance(z^*, \mathcal{P})$ is the minimum Euclidean distance between z^* and all the feasible solutions in $\mathcal{P}$, and $|\mathcal{P}^*|$ is the cardinality of $\mathcal{P}^*$. The smaller the IGD value, the better the performance of a CMOEA.

- *Hypervolume (HV) [51]:* Similarly, the infeasible solutions should be deleted before the calculation of HV. Then, HV measures the volume enclosed by $\mathcal{P}$ and a specified reference point in the objective space [52]. HV has the capability to assess both convergence and diversity of $\mathcal{P}$. Usually, the larger the HV value, the better the performance of a CMOEA. In our experiments, the HV value is calculated by using the reference point which is set to 1.1 times of the upper bounds of the Pareto front.

B. Algorithms for Comparison and Parameter Settings

ToP was applied to improve the performance of both dominance-based and decomposition-based CMOEAs. In this paper, we chose two widely used dominance-based CMOPs: NSGA-II-CDP [10] and IDEA [18], and two state-of-the-art decomposition-based CMOPs: CMOEAD [19] and MOEA/D-CDP [20]. They have been introduced in Section II-C.

For the sake of convenience, if a specific CMOEA is under the framework of ToP, the name of this CMOEA will be modified by adding four letters "ToP-". For example, NSGA-II-CDP under our framework is named ToP-NSGA-II-CDP.

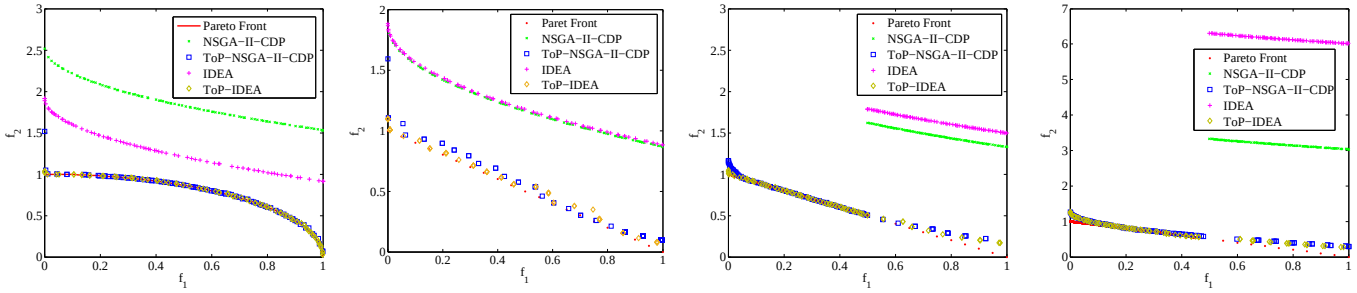
In our experiments, we adopted the following parameter settings:

TABLE II
EXPERIMENTAL RESULTS OF NSGA-II-CDP, TOP-NSGA-II-CDP, IDEA, AND TOP-IDEA OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, THE WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	NSGA-II-CDP		ToP-NSGA-II-CDP		IDEA		ToP-IDEA	
	FR	IGD	FR	IGD	FR	IGD	FR	IGD
DOC-1	100%	2.070e+0(1.37e+0)	100%	**6.925e-3(4.62e-4)**	100%	7.828e-1(1.03e+1)	100%	**7.831e-3(2.77e-4)**
DOC-2	0%	NA	**100%**	**1.671e-1(3.69e-3)**	5%	4.982e-1(0.00e+0)	**100%**	**1.800e-2(6.26e-3)**
DOC-3	0%	NA	**100%**	**1.270e-2(5.82e-2)**	15%	5.521e+2(2.21e+2)	**100%**	**1.157e-1(6.44e-2)**
DOC-4	100%	8.712e-1(6.76e-1)	100%	**4.820e-2(1.33e-2)**	100%	6.718e-1(5.38e-1)	100%	**3.841e-2(9.57e-3)**
DOC-5	0%	NA	**100%**	**1.294e-1(8.75e-2)**	100%	7.460e+1(3.90e+1)	100%	**7.421e-2(5.05e-2)**
DOC-6	100%	1.562e+0(1.54e+0)	100%	**4.654e-3(7.23e-4)**	100%	1.985e+0(2.22e+0)	100%	**4.990e-3(3.75e-4)**
DOC-7	100%	2.417e+0(8.70e-1)	100%	**1.732e-2(5.88e-3)**	75%	5.128e+0(1.55e+0)	**100%**	**1.428e-2(4.89e-3)**
DOC-8	35%	1.447e+1(1.50e+1)	**100%**	**2.827e-1(1.66e-2)**	100%	9.015e+1(5.44e+1)	100%	**2.421e-1(1.04e-1)**
DOC-9	100%	1.321e-1(6.20e-2)	100%	**3.773e-2(5.88e-3)**	100%	1.175e-1(1.05e-1)	100%	**4.431e-2(1.07e-3)**

TABLE III
EXPERIMENTAL RESULTS OF NSGA-II-CDP, TOP-NSGA-II-CDP, IDEA, AND TOP-IDEA OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND HV. FOR HV, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, THE WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	NSGA-II-CDP		ToP-NSGA-II-CDP		IDEA		ToP-IDEA	
	FR	HV	FR	HV	FR	HV	FR	HV
DOC-1	100%	1.776e-2(7.26e-2)	100%	**4.041e-1(6.55e-3)**	100%	1.549e-1(1.44e-1)	100%	**4.057e-1(3.88e-3)**
DOC-2	0%	NA	**100%**	**5.677e-1(4.76e-3)**	5%	1.916e-1(0.00e+0)	**100%**	**5.672e-1(7.46e-3)**
DOC-3	0%	NA	**100%**	**2.630e-1(4.95e-2)**	15%	0.000e+0(0.00e+0)	**100%**	**2.741e-1(5.92e-2)**
DOC-4	100%	8.989e-2(1.47e-1)	100%	**6.144e-1(1.96e-2)**	100%	1.704e-1(1.91e-1)	100%	**6.154e-1(6.32e-2)**
DOC-5	0%	NA	**100%**	**5.022e-1(5.65e-2)**	100%	2.394e-2(1.07e-1)	100%	**5.432e-1(3.41e-2)**
DOC-6	100%	7.732e-2(1.77e-1)	100%	**6.248e-1(2.60e-2)**	100%	3.554e-2(7.65e-2)	100%	**6.501e-1(1.10e-2)**
DOC-7	100%	0.000e+0(0.00e+0)	100%	**5.654e-1(3.24e-2)**	75%	0.000e+0(0.00e+0)	**100%**	**5.682e-1(3.33e-2)**
DOC-8	35%	0.000e+0(0.00e+0)	**100%**	**7.471e-1(1.53e-1)**	100%	0.000e+0(0.00e+0)	100%	**8.070e-1(1.42e-1)**
DOC-9	100%	2.897e-2(9.02e-3)	100%	**3.696e-2(1.14e-3)**	100%	2.484e-2(1.42e-2)	100%	**3.603e-2(2.43e-3)**



(a) DOC-1 (b) DOC-4 (c) DOC-6 (d) DOC-7

Fig. 7. Images of the feasible solutions provided by two dominance-based CMOEAs (NSGA-II-CDP and IDEA) and their augmented algorithms in a run on DOC-1, DOC-4, DOC-6, and DOC-7.

- *Population Size*: Following the suggestions in [35], for two-objective CMOPs, the population size in each algorithm was set to 100, while for three-objective CMOPs, it was set to 300.
- *Parameter Settings for Operators*: For four CMOEAs (i.e., NSGA-II-CDP, IDEA, CMOEAD, and MOEA/D-CDP), the simulated binary crossover (SBX) and polynomial mutation were used to produce offspring. The crossover probability and the mutation probability were set to 1.0 and $1/D$, respectively. The distribution indexes of both SBX and the polynomial mutation were set to 20 [23]. To make a fair comparison, when a CMOEA is under the framework of ToP, its offspring generation operators and parameter settings were kept untouched.
- *Number of Independent Runs and Termination Condition*: All algorithms were independently run 20 times on each instance, and terminated when a maximum of 200,000 and 400,000 fitness evaluations reached for two-objective CMOPs and three-objective CMOPs, respectively.
- *Parameter Settings for Algorithms*: To ensure the comparison fair, the other parameter settings of NSGA-II-CDP, IDEA, CMOEAD, and MOEA/D-CDP were identical with their original papers, and remained unchanged when they were under the framework of ToP.

VI. Results and Discussions

A. Analysis of Principle

First of all, we intended to ascertain whether ToP can find the promising feasible area in the first phase and obtain well-distributed and well-converged feasible solutions in the second phase. To answer this question, we took ToP-NSGA-II-CDP as an example, and tested it on DOC-1 whose Pareto front is non-convex. The images of the feasible solutions obtained after the first and second phases are presented in Fig. 6.

From Fig. 6(a), it can be seen that even for a non-convex CMOP, the images of the feasible solutions obtained after the first phase not only are close to the Pareto front, but also have a good diversity around the Pareto front. The above phenomenon suggests that the first phase of ToP succeeds in finding the promising feasible area and that it is also terminated at a proper stage. From Fig. 6(b), we can observe that the images of the feasible solutions obtained after the second phase distribute well along the Pareto front. Thus, ToP-NSGA-II-CDP is able to push the images of the high-quality candidate solutions obtained in the first phase toward the Pareto front from diverse directions. After these two phases, ToP-NSGA-II-CDP eventually produces a set of representative feasible solutions.

B. Applying ToP to Two Dominance-based CMOEAs

Subsequently, we applied ToP to two popular dominance-based CMOEAs, namely NSGA-II-CDP and IDEA. Tables II and III summarize the experimental results of NSGA-II-CDP, ToP-NSGA-II-CDP, IDEA, and ToP-IDEA over 20 independent runs on DOC in terms of FR, IGD, and HV. For IGD and HV, their average and standard deviation are recorded. In addition, The Wilcoxon's rank sum test at 0.05 significance level is performed between a CMOEA and its augmented version, and the better result is highlighted in boldface on each instance.

At our first glance from Tables II and III, ToP can significantly improve the performance of both NSGA-II-CDP and IDEA in terms of three indicators (i.e., FR, IGD, and HV). The detailed discussions are given below:

- NSGA-II-CDP and IDEA under the framework of ToP perform better than or similar to their original algorithms in terms of FR. For NSGA-II-CDP, its FR values on DOC-2, DOC-3, DOC-5, and DOC-8 are 0%, 0%, 0%, and 35%, respectively. With respect to IDEA, its FR values on DOC-2, DOC-3, and DOC-7 are 5%, 15%, and 75%, respectively. However, when NSGA-II-CDP and IDEA are under the framework of ToP, both of them can obtain 100% FR on all instances. Therefore, ToP is capable of helping them to find the feasible region.
- In terms of IGD, from Table II, ToP-NSGA-II-CDP and ToP-IDEA beat their original algorithms on all instances as ToP-NSGA-II-CDP and ToP-IDEA consistently obtain smaller IGD values.
- Regarding the HV indicator, from Table III, ToP-NSGA-II-CDP and ToP-IDEA provide higher HV values on each instance than their original algorithms. When NSGA-II-CDP is applied to solve DOC-7 and DOC-8, the HV values are zero. The reason is that the images of the obtained feasible solutions are far away from the Pareto front. In addition, the similar phenomenon occurs when IDEA is applied to solve DOC-3, DOC-7, and DOC-8.

The images of the feasible solutions provided by NSGA-II-CDP, ToP-NSGA-II-CDP, IDEA, and ToP-IDEA in the end of a run are plotted in Fig. 7 on four instances (i.e., DOC-1, DOC-4, DOC-6, and DOC-7).

C. Applying ToP to Two Decomposition-based CMOEAs

Thereafter, we investigated the effectiveness of ToP on two state-of-the-art decomposition-based CMOEAs, namely CMOEAD and MOEAD-CDP. The experimental results are presented in Tables IV and V.

As shown in Tables IV and V, both CMOEAD and MOEA/D-CDP fail to provide 100% FR on four instances, i.e., DOC-2, DOC-3, DOC-5, and DOC-7. However, under the framework of ToP, they can achieve 100% FR on all instances, which means that ToP is able to improve their capability in finding the feasible region. As far as IGD is concerned, according to Table IV, both ToP-CMOEAD and ToP-MOEA/D-CDP provide smaller values than their original algorithms on each instance. With respect to HV, as shown in Table V, the values derived from ToP-CMOEAD and ToP-MOEA/D-CDP are much larger than their original algorithms on all instances expect for DOC-9.

From the above comparison, one can conclude that CMOEAD and MOEA/D-CDP under the framework of ToP achieve better performance than the original algorithms in terms of three indicators (i.e., FR, IGD, and HV), which verifies the effectiveness of ToP on the decomposition-based CMOEAs. Fig. 8 plots the images of the feasible solutions resulting from CMOEAD, ToP-CMOEAD, MOEA/D-CDP, and ToP-MOEA/D-CDP when a run halts on DOC-8.

D. Benefit of the Two-Phase Optimization Mechanism

The aim of this subsection is to investigate the benefit of the two-phase optimization mechanism in ToP. To this end, we selected ToP-NSGA-II-CDP as the instance algorithm and considered its two variants, i.e., Former-NSGA-II-CDP and Latter-NSGA-II-CDP. In Former-NSGA-II-CDP, the constrained single-objective optimization in the first phase of ToP was implemented throughout the whole evolution. As for Latter-NSGA-II-CDP, NSGA-II-CDP was implemented during the whole evolution. It is obvious that Latter-NSGA-II-CDP is equivalent to the original NSGA-II-CDP. Herein, we tested ToP-NSGA-II-CDP and its two variants on DOC-1 and DOC-3, and the experimental results are presented in Figs. S-10 and S-11 of the supplementary file, respectively.

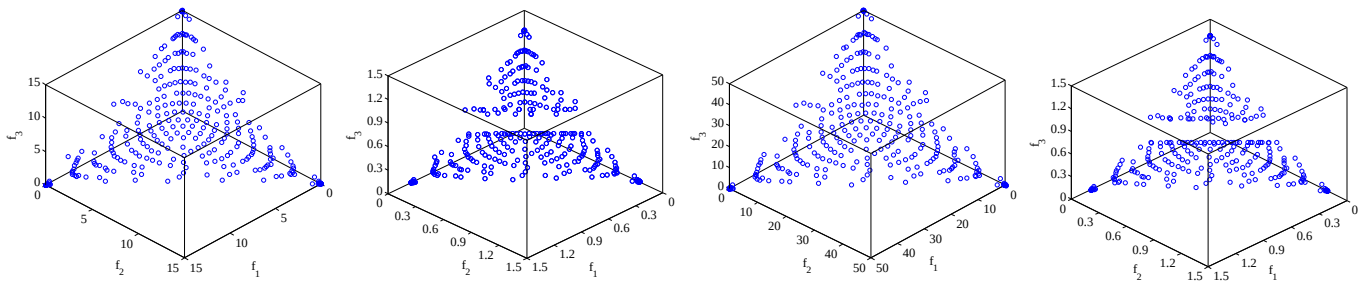
From Figs. S-10 and S-11, the images of the feasible solutions derived from Former-NSGA-II-CDP always cluster in a very small area of the Pareto front. It is not difficult to understand since the conflict among the objective functions of the original CMOP is ignored in Former-NSGA-II-CDP; thus, it converges toward the Pareto front from few directions. For Latter-NSGA-II-CDP, the images of the feasible solutions are distant from the Pareto front on DOC-1. It is probably

TABLE IV
EXPERIMENTAL RESULTS OF CMOEAD, TOP-CMOEAD, MOEA/D-CDP, AND TOP-MOEA/D-CDP OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, THE WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	CMOEAD		ToP-CMOEAD		MOEA/D-CDP		ToP-MOEA/D-CDP	
	FR	IGD	FR	IGD	FR	IGD	FR	IGD
DOC-1	100%	3.761e+2(2.15e+2)	100%	**2.600e-1(2.45e-1)**	100%	3.526e+2(2.02e+2)	100%	**2.010e-1(2.41e-1)**
DOC-2	0%	NA	**100%**	**3.871e-2(3.45e-2)**	0%	NA	**100%**	**4.794e-2(3.55e-2)**
DOC-3	25%	6.417e+2(1.60e+2)	**100%**	**1.739e-1(7.68e-2)**	25%	5.892e+2(2.91e+2)	**100%**	**2.627e-1(1.60e-1)**
DOC-4	100%	7.411e+0(7.62e+0)	100%	**5.789e-2(1.39e-2)**	100%	9.181e+1(1.07e+1)	100%	**5.936e-2(2.27e-2)**
DOC-5	0%	NA	**100%**	**1.971e-1(7.91e-2)**	0%	NA	**100%**	**2.258e-1(5.35e-2)**
DOC-6	100%	1.036e+2(2.71e+2)	100%	**8.929e-3(2.26e-2)**	100%	9.486e+1(2.82e+2)	100%	**4.825e-3(1.23e-3)**
DOC-7	75%	8.284e+0(2.87e+0)	**100%**	**3.106e-2(2.60e-2)**	65%	9.574e+0(2.03e+0)	**100%**	**3.055e-2(1.42e-2)**
DOC-8	100%	2.810e+1(3.26e+1)	100%	**1.590e-1(7.51e-2)**	100%	3.878e+1(3.93e+1)	100%	**1.752e-1(6.77e-2)**
DOC-9	100%	1.362e-1(9.59e-2)	100%	**5.811e-2(5.92e-3)**	100%	1.228e-1(7.76e-2)	100%	**6.002e-2(8.67e-3)**

TABLE V
EXPERIMENTAL RESULTS OF CMOEAD, TOP-CMOEAD, MOEA/D-CDP, AND TOP-MOEA/D-CDP OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND HV. FOR HV, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, THE WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	CMOEAD		ToP-CMOEAD		MOEA/D-CDP		ToP-MOEA/D-CDP	
	FR	HV	FR	HV	FR	HV	FR	HV
DOC-1	100%	0.000e+0(0.00e+0)	100%	**2.424e-1(1.37e-1)**	100%	0.000e+0(0.00e+0)	100%	**2.879e-1(1.37e-1)**
DOC-2	0%	NA	**100%**	**5.533e-1(1.80e-2)**	0%	NA	**100%**	**5.449e-1(1.659e-2)**
DOC-3	25%	0.000e+0(0.00e+0)	**100%**	**2.257e-1(5.83e-2)**	25%	0.000e+0(0.00e+0)	**100%**	**1.817e-1(8.12e-2)**
DOC-4	100%	0.000e+0(0.00e+0)	100%	**5.936e-1(3.19e-2)**	100%	0.000e+0(0.00e+0)	100%	**6.007e-1(2.98e-2)**
DOC-5	0%	NA	**100%**	**5.233e-1(1.67e-2)**	0%	NA	**100%**	**4.289e-1(2.78e-2)**
DOC-6	100%	1.224e-3(5.48e-3)	100%	**5.816e-1(5.41e-2)**	100%	6.040e-4(2.70e-3)	100%	**6.143e-1(3.54e-2)**
DOC-7	75%	0.000e+0(0.00e+0)	**100%**	**5.178e-1(4.35e-2)**	65%	0.000e+0(0.00e+0)	**100%**	**5.128e-1(6.25e-2)**
DOC-8	100%	0.000e+0(0.00e+0)	100%	**8.985e-1(1.12e-1)**	100%	0.000e+0(0.00e+0)	100%	**8.726e-1(9.92e-2)**
DOC-9	100%	2.117e-2(1.38e-2)	100%	**3.380e-2(1.67e-3)**	100%	2.308e-2(1.32e-2)	100%	**3.332e-2(2.69e-3)**



(a) CMOEAD (b) ToP-CMOEAD (c) MOEA/D-CDP (d) ToP-MOEA/D-CDP

Fig. 8. Images of the feasible solutions provided by two decomposition-based CMOEAs (CMOEAD and MOEA/D-CDP) and their augmented algorithms in the end of a run on DOC-8.

because Latter-NSGA-II-CDP runs the risk of getting stuck at a local optimal area in the feasible region. Moreover, Latter-NSGA-II-CDP cannot obtain any feasible solution on DOC-3. In contrast, the images of the feasible solutions provided by ToP-NSGA-II-CDP can scatter throughout the Pareto front well on DOC-1. When solving DOC-3, the images of its finally obtained feasible solutions can also approach the Pareto front with a good distribution.

The above comparison demonstrates the importance of the two-phase optimization mechanism in ToP, thus verifying the main motivation of this paper. Due to the lack of the first phase, a CMOEA may either not be able to enter the feasible region, or stall in a local attraction basin. On the other hand, without the second phase, a constrained single-objective EA is prone to converge to a small area of the Pareto optimal solutions.

E. Investigation to the Search Engine

As depicted in Tables II–V, the four CMOEAs (i.e., NSGA-II-CDP, IDEA, CMOEAD, and MOEA/D-CDP) fail to obtain promising results on most instances. Someone may attribute the poor performance to the search engine used in them. To investigate this guess, DE introduced in Section IV-B2 was utilized to produce offspring for these four CMOEAs. For

TABLE VI
EXPERIMENTAL RESULTS OF NSGA-II-CDP-DE, TOP-NSGA-II-CDP-DE, IDEA-DE, AND TOP-IDEA-DE OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, THE WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	NSGA-II-CDP-DE		ToP-NSGA-II-CDP-DE		IDEA-DE		ToP-IDEA-DE	
	FR	IGD	FR	IGD	FR	IGD	FR	IGD
DOC-1	100%	4.696e-1(1.99e-1)	100%	**6.437e-3(2.30e-4)**	100%	8.230e-3(3.53e-4)	100%	8.235e-3(4.53e-4)
DOC-2	100%	4.426e-1(1.47e-4)	100%	4.431e-1(1.67e-4)	0%	NA	**100%**	**1.013e-2(9.94e-4)**
DOC-3	100%	4.062e-1(2.78e-1)	100%	**8.466e-3(2.02e-3)**	0%	NA	**100%**	**1.129e-2(2.73e-3)**
DOC-4	100%	7.055e-1(2.32e-1)	100%	**9.392e-2(2.35e-1)**	100%	2.949e-2(3.71e-3)	100%	2.760e-2(3.39e-3)
DOC-5	100%	1.010e+2(4.79e+1)	100%	**2.968e-2(5.78e-3)**	90%	2.124e-1(2.79e-1)	**100%**	**4.718e-2(1.21e-2)**
DOC-6	100%	2.790e-3(1.21e-4)	100%	2.790e-3(9.93e-5)	100%	3.679e-3(1.90e-4)	100%	3.684e-3(1.48e-4)
DOC-7	100%	2.589e-3(7.35e-5)	100%	2.590e-3(8.99e-5)	0%	NA	**100%**	**3.398e-3(1.29e-4)**
DOC-8	100%	1.127e+0(6.86e-1)	100%	**1.188e-1(1.26e-2)**	20%	2.391e+1(2.05e+1)	**100%**	**1.659e-1(3.36e-2)**
DOC-9	100%	9.529e-2(1.02e-2)	100%	**8.599e-2(1.26e-2)**	100%	1.636e-1(1.53e-1)	100%	**9.681e-2(7.79e-3)**

TABLE VII
EXPERIMENTAL RESULTS OF CMOEAD-DE, TOP-CMOEAD-DE, MOEA/D-CDP-DE, AND TOP-MOEA/D-CDP-DE OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, THE WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	CMOEAD-DE		ToP-CMOEAD-DE		MOEA/D-CDP-DE		ToP-MOEA/D-CDP-DE	
	FR	IGD	FR	IGD	FR	IGD	FR	IGD
DOC-1	100%	1.009e-1(9.78e-2)	100%	**4.288e-3(1.04e-4)**	100%	1.437e-1(1.20e-1)	100%	**1.016e-1(2.01e-1)**
DOC-2	80%	5.258e-1(7.37e-2)	**100%**	**4.409e-1(3.35e-5)**	70%	5.080e-1(2.67e-2)	**100%**	**4.198e-1(9.46e-2)**
DOC-3	100%	2.912e+2(1.48e+2)	100%	**3.091e-2(5.28e-2)**	100%	3.662e+2(2.26e+2)	100%	**3.938e-2(5.43e-2)**
DOC-4	100%	3.670e-1(1.92e-1)	100%	**2.867e-2(8.43e-3)**	100%	4.381e-1(2.74e-1)	100%	**3.179e-2(9.13e-3)**
DOC-5	100%	8.516e+1(6.43e+1)	100%	**7.076e-2(6.46e-2)**	100%	6.842e+1(4.79e+1)	100%	**6.294e-2(3.08e-2)**
DOC-6	100%	3.165e-1(1.28e-1)	100%	**3.351e-3(1.66e-4)**	100%	3.328e-1(8.78e-2)	100%	**3.471e-3(2.23e-4)**
DOC-7	100%	2.082e+0(1.36e+0)	100%	**4.561e-3(1.39e-3)**	100%	2.385e+0(1.35e+0)	100%	**5.147e-3(1.27e-3)**
DOC-8	100%	**3.288e-2(6.19e-4)**	100%	3.638e-2(1.32e-3)	100%	1.094e-1(2.11e-2)	100%	**7.208e-2(1.73e-2)**
DOC-9	100%	5.468e-2(4.17e-3)	100%	**5.360e-2(4.30e-3)**	100%	5.437e-2(4.40e-3)	100%	5.431e-2(2.84e-3)

convenience, a CMOEA with DE is named CMOEA-DE, and a CMOEA-DE under the framework of ToP is denoted as ToP-CMOEA-DE. Note that in a CMOEA-DE, it is hard to define the best individual due to the multiple objective functions. Therefore, the best individual in (15) was replaced with a random individual in the current population. While for other parameters (i.e., F and CR), they were kept the same as in Section IV-B2. The experimental results are given in Tables VI and VII.

From Tables VI and VII, it is interesting to see that, overall, NSGA-II-CDP-DE, CMOEAD-DE, and MOEA/D-CDP-DE can obtain very good FR values on all instances, which indicates that DE is more powerful to find the feasible solutions for these three CMOEAs. For NSGA-II-CDP-DE, it prefers constraints to objective functions and a CMOP is regarded as an unconstrained single-objective optimization problem (the objective function is the degree of constraint violation) when the population is infeasible. In addition, for CMOEAD-DE and MOEA/D-CDP-DE, the original CMOP is decomposed into a set of constrained single-objective optimization problems. Considering that DE is a powerful search engine to solve single-objective optimization problems [45], [53], it is not a hard task for NSGA-II-CDP-DE, CMOEAD-DE, and MOEA/D-CDP-DE to find the feasible region. However, the performance of IDEA-DE is still unsatisfactory in terms of FR. The reason might be that IDEA-DE compares infeasible solutions based on $(m+1)$ objective functions, in which all constraints are considered as an additional objective function, in addition to the m original objective functions. As a result, IDEA-DE has a low selection pressure to steer the individuals from the infeasible region to the feasible region, especially for some instances in DOC which have very small feasible regions. The fact that ToP-IDEA-DE can achieve 100% FR suggests that the first phase of ToP, which transforms the original CMOP into a constrained single-objective optimization problem, is helpful for IDEA-DE.

In terms of IGD, from Tables VI and VII, ToP-NSGA-II-CDP-DE, ToP-IDEA-DE, ToP-CMOEAD-DE, and ToP-MOEA/D-CDP-DE perform better than their original algorithms on six, six, seven, and seven instances, respectively; while the original algorithms cannot beat their enhanced versions on more than one instance. Therefore, despite the powerful DE can help some CMOEAs (i.e., NSGA-II-CDP, CMOEAD and MOEA/D-CDP) to enter the feasible region, ToP can still further improve their performance inside the feasible region. The performance improvement should be attributed to the transformation from the original CMOP into a constrained single-objective optimization problem in ToP. This transformation can result in faster convergence speed in the feasible region due to concentrating on only one optimal solution and a stronger capability to cross the local optimal area in the feasible region as analysed in Section IV-A.

From the above discussions, the performance of CMOEAs can be improved through a powerful search engine. However, the transformation idea in ToP is also indispensable to the performance improvement of CMOEAs.

Remark 2: In the supplementary file, we also investigated the effect of the parameter settings in Section S-II-A, the generality of ToP on other CMOPs in Section S-II-B, the impact of the search engine in the first phase of ToP in Section S-II-C, the effect of the constraint-handling techniques in the second phase of ToP in Section S-II-D, the influence of the normalized process in Section S-II-E, the allocation of the number of fitness evaluations in the first and second phases in Section S-II-F, and the effectiveness of the best individual in ToP in Section S-II-G.

VII. Conclusions

In this paper, a set of artificial CMOPs, named DOC, was proposed. It contained seven two-objective CMOPs and two three-objective CMOPs. It was the first attempt to consider both decision and objective constraints simultaneously in the design of CMOPs. It was also one of the first artificial CMOP suites considering both inequality and equality constraints. In general, DOC posed a great challenge for a CMOEA to obtain a set of well-distributed and well-converged feasible solutions.

Subsequently, we proposed a two-phase framework called ToP to improve current CMOEAs' performance on DOC. The first phase was implemented to find the promising feasible area. To achieve this goal, the original CMOP was transformed into a constrained single-objective optimization problem by ignoring the conflict among the objective functions. We analyzed the advantages of the above transformation. After the promising feasible area has been discovered, a specific CMOEA was implemented in the second phase to approximate the Pareto front. ToP had a simple structure and could be applied to many current CMOEAs. In this paper, we applied ToP to four state-of-the-art CMOPs, and the experimental results suggested that ToP can improve their performance significantly.

In the future, we will design some other powerful constrained single-objective EAs and CMOEAs in the first and second phases of ToP, respectively. The new developments in constrained single-objective optimization and constrained multiobjective optimization can also be integrated into the framework of ToP. Moreover, we plan to apply the idea of ToP to deal with expensive constrained multiobjective optimization. For expensive constrained multiobjective optimization, the fast convergence is required since the computational resource is very limited. We expect that the ability of ToP rapidly locating the promising feasible area can benefit the solution of expensive constrained multiobjective optimization.

The Matlab source codes of DOC and ToP can be downloaded from Y. Wang's homepage: http://www.escience.cn/people/yongwang1/index.html

References

[1] B. Tan, H. Ma, Y. Mei, and M. Zhang, "Evolutionary multi-objective optimization for web service location allocation problem," *IEEE Transactions on Services Computing*, 2018, in press, DOI 10.1109/TSC.2018.2793266.

[2] N. G. Paterakis, M. Gibescu, A. G. Bakirtzis, and J. P. S. Catalão, "A multi-objective optimization approach to risk-constrained energy and reserve procurement using demand response," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3940–3954, 2018.

[3] H. Farzin, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, "A stochastic multi-objective framework for optimal scheduling of energy storage systems in microgrids," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 117–127, 2017.

[4] M. Shafie-Khah, P. Siano, and J. P. S. Catalao, "Optimal demand response strategies to mitigate oligopolistic behavior of generation companies using a multi objective decision analysis," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4264–4274, 2018.

[5] M. Gong, Z. Wang, Z. Zhu, and L. Jiao, "A similarity-based multi-objective evolutionary algorithm for deployment optimization of near space communication system," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 6, pp. 878–897, 2017.

[6] S. Zeng, R. Jiao, C. Li, X. Li, and J. S. Alkasassbeh, "A general framework of dynamic constrained multiobjective evolutionary algorithms for constrained optimization," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2678–2688, Sept 2017.

[7] H. Fukumoto and A. Oyama, "Benchmarking multiobjective evolutionary algorithms and constraint handling techniques on a real-world car structure design optimization benchmark problem," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, ser. GECCO '18, 2018, pp. 177–178.

[8] K. Deb, A. Pratap, and T. Meyarivan, "Constrained test problems for multi-objective evolutionary optimization," in *First International Conference on Evolutionary Multi-Criterion Optimization (EMO 2001)*, 2000, pp. 284–298.

[9] C. Li, T. T. Nguyen, S. Zeng, M. Yang, and M. Wu, "An open framework for constructing continuous optimization problems," *IEEE Transactions on Cybernetics*, 2018, in press. DOI: 10.1109/TCYB.2018.2825343.

[10] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[11] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang, "Multiobjective evolutionary algorithms: A survey of the state of the art," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 32–49, 2011.

[12] C. A. Coello Coello, "Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art," *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 11-12, pp. 1245–1287, 2002.

[13] R. Tanabe, A. Oyama, R. Tanabe, and A. Oyama, "A note on constrained multi-objective optimization benchmark problems," in *IEEE Congress on Evolutionary Computation*, 2017, pp. 1127–1134.

[14] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, Part II: Handling constraints and extending to an adaptive approach," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 602–622, 2014.

[15] J.-P. Li, Y. Wang, S. Yang, and Z. Cai, "A comparative study of constraint-handling techniques in evolutionary constrained multiobjective optimization," in *2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2016, pp. 4175–4182.

[16] Z. Fan, W. Li, X. Cai, H. Li, K. Hu, Q. Zhang, K. Deb, and E. D. Goodman, "Difficulty adjustable and scalable constrained multi-objective test problem toolkit," *arXiv preprint arXiv:1612.07603*, 2016.

[17] Q. Zhang, A. Zhou, S. Zhao, P. N. Suganthan, W. Liu, and S. Tiwari, "Multiobjective optimization test instances for the CEC 2009 special session and competition," *University of Essex, Colchester, UK and Nanyang technological University, Singapore, special session on performance assessment of multi-objective optimization algorithms, technical report*, vol. 264, 2008.

[18] T. Ray, H. K. Singh, A. Isaacs, and W. Smith, "Infeasibility driven evolutionary algorithm for constrained optimization," in *Constraint-Handling in Evolutionary Optimization*. Springer, 2009, pp. 145–165.

[19] M. Asafuddoula, T. Ray, R. Sarker, and K. Alam, "An adaptive constraint handling approach embedded MOEA/D," in *2012 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2012, pp. 1–8.

[20] M. A. Jan and R. A. Khanum, "A study of two penalty-parameterless constraint handling techniques in the framework of MOEA/D," *Applied Soft Computing*, vol. 13, no. 1, pp. 128–148, 2013.

[21] J. Fliege and I. A. Vaz, "A SQP type method for constrained multiobjective optimization," *Optimization Online*, pp. 1–35, 2015.

[22] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," *Evolutionary Multiobjec-*

tive Optimization. Theoretical Advances and Applications, pp. 105–145, 2005.

[23] Z. Fan, Y. Fang, W. Li, J. Lu, X. Cai, and C. Wei, "A comparative study of constrained multi-objective evolutionary algorithms on constrained multi-objective optimization problems," in *2017 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2017, pp. 209–216.

[24] Y. G. Woldesenbet, G. G. Yen, and B. G. Tessema, "Constraint handling in multiobjective evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 3, pp. 514–525, 2009.

[25] N. Young, "Blended ranking to cross infeasible regions in constrained-multiobjective problems," in *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on*, vol. 2. IEEE, 2005, pp. 191–196.

[26] F. Jimenez, A. F. Gómez-Skarmeta, G. Sánchez, and K. Deb, "An evolutionary algorithm for constrained multi-objective optimization," in *2002 Congress on Evolutionary Computation (CEC'02)*, vol. 2. IEEE, 2002, pp. 1133–1138.

[27] A. Oyama, K. Shimoyama, and K. Fujii, "New constraint-handling method for multi-objective and multi-constraint evolutionary optimization," *Transactions of the Japan Society for Aeronautical and Space Sciences*, vol. 50, no. 167, pp. 56–62, 2007.

[28] H. K. Singh, T. Ray, and W. Smith, "C-psa: Constrained pareto simulated annealing for constrained multi-objective optimization," *Information Sciences*, vol. 180, no. 13, pp. 2499–2513, 2010.

[29] S. Qian, Y. Ye, B. Jiang, and J. Wang, "Constrained multiobjective optimization algorithm based on immune system model," *IEEE Transactions on Cybernetics*, vol. 46, no. 9, pp. 2056–2069, 2016.

[30] X. Yu and Y. Lu, "A corner point-based algorithm to solve constrained multi-objective optimization problems," *Applied Intelligence*, vol. 48, no. 9, pp. 3019–3037, Sep 2018.

[31] X. Yu, X. Yu, Y. Lu, G. G. Yen, and M. Cai, "Differential evolution mutation operators for constrained multi-objective optimization," *Applied Soft Computing*, vol. 67, pp. 452–466, 2018.

[32] L. Jiao, J. Luo, R. Shang, and F. Liu, "A modified objective function method with feasible-guiding strategy to solve constrained multi-objective optimization problems," *Applied Soft Computing*, vol. 14, no. 1, pp. 363–380, 2014.

[33] B. Y. Qu and P. N. Suganthan, "Constrained multi-objective optimization algorithm with an ensemble of constraint handling methods," *Engineering Optimization*, vol. 43, no. 4, pp. 403–416, 2011.

[34] A. Trivedi, D. Srinivasan, K. Sanyal, and A. Ghosh, "A survey of multiobjective evolutionary algorithms based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 3, pp. 440–462, 2017.

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

[36] Z. Fan, H. Li, C. Wei, W. Li, H. Huang, X. Cai, and Z. Cai, "An improved epsilon constraint handling method embedded in MOEA/D for constrained multi-objective optimization problems," *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–8, Dec 2016.

[37] Z. Fan, W. Li, X. Cai, H. Li, C. Wei, Q. Zhang, K. Deb, and E. D. Goodman, "Push and pull search for solving constrained multi-objective optimization problems," *Swarm and Evolutionary Computation*, 2018, https://doi.org/10.1016/j.swevo.2018.08.017.

[38] Z. Fan, Y. Fang, W. Li, X. Cai, C. Wei, and E. Goodman, "MOEA/D with angle-based constrained dominance principle for constrained multi-objective optimization problems," *Applied Soft Computing, https://doi.org/10.1016/j.asoc.2018.10.027*, 2018.

[39] S. Huband, P. Hingston, L. Barone, and L. While, "A review of multiobjective test problems and a scalable test problem toolkit," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 5, pp. 477–506, 2006.

[40] E. Zitzler, K. Deb, and L. Thiele, "Comparison of multiobjective evolutionary algorithms: Empirical results," *Evolutionary computation*, vol. 8, no. 2, pp. 173–195, 2000.

[41] J. J. Liang, T. P. Runarsson, E. Mezura-Montes, M. Clerc, P. N. Suganthan, C. A. Coello Coello, and K. Deb, "Problem definitions and evaluation criteria for the CEC 2006 special session on constrained real-parameter optimization," *Journal of Applied Mechanics*, vol. 41, no. 8, pp. 8–31, 2006.

[42] Z. Cai and Y. Wang, "A multiobjective optimization-based evolutionary algorithm for constrained optimization," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 6, pp. 658–675, Dec 2006.

[43] Y. Wang, Z. Cai, Y. Zhou, and W. Zeng, "An adaptive tradeoff model for constrained evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp. 80–92, Feb 2008.

[44] Y. Wang, Z. Cai, G. Guo, and Y. Zhou, "Multiobjective optimization and hybrid evolutionary algorithm to solve constrained optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 3, pp. 560–575, June 2007.

[45] Y. Wang, B.-C. Wang, H.-X. Li, and G. G. Yen, "Incorporating objective function information into the feasibility rule for constrained evolutionary optimization," *IEEE Transactions on Cybernetics*, vol. 46, no. 12, pp. 2938–2952, 2016.

[46] K. Miettinen, *Nonlinear Multiobjective Optimization*. Springer Science & Business Media, 1999.

[47] M. Ehrgott, *Multicriteria Optimization*. Springer Science & Business Media, 2005.

[48] K. Deb, "An efficient constraint handling method for genetic algorithms," *Computer Methods in Applied Mechanics and Engineering*, vol. 186, no. 2, pp. 311–338, 2000.

[49] Z.-Z. Liu, Y. Wang, S. Yang, and Z. Cai, "Differential evolution with a two-stage optimization mechanism for numerical optimization," in *2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2016, pp. 3170–3177.

[50] C. A. Coello Coello, G. B. Lamont, D. A. Van Veldhuizen *et al.*, *Evolutionary Algorithms for Solving Multi-objective Problems*. Springer, 2007.

[51] E. Zitzler and L. Thiele, "Multiobjective optimization using evolutionary algorithms – A comparative case study," in *International Conference on Parallel Problem Solving from Nature*. Springer, 1998, pp. 292–301.

[52] M. Emmerich, N. Beume, and B. Naujoks, "An EMO algorithm using the hypervolume measure as selection criterion." in *EMO*, vol. 3410. Springer, 2005, pp. 62–76.

[53] B.-C. Wang, H.-X. Li, J.-P. Li, and Y. Wang, "Composite differential evolution for constrained evolutionary optimization," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018, in press, DOI: 10.1109/TSMC.2018.2807785.

Zhi-Zhong Liu received the B.S. degree in automation from Central South University, Changsha, China, in 2013, where he is currently pursuing the Ph.D. degree in control science and engineering. His current research interests include evolutionary computation, bioinformatics, swarm intelligence, constrained optimization, and multimodal optimization.

Yong Wang (M'08–SM'17) received the B.S. degree in automation from the Wuhan Institute of Technology, Wuhan, China, in 2003, and the M.S. degree in pattern recognition and intelligent systems and the Ph.D. degree in control science and engineering both from the Central South University, Changsha, China, in 2006 and 2011, respectively.

He is a Professor with the School of Automation, Central South University, Changsha, China. His current research interests include the theory, algorithm design, and interdisciplinary applications of computational intelligence.

Dr. Wang is an Associate Editor for the *Swarm and Evolutionary Computation*. He was a Web of Science highly cited researcher in Computer Science in 2017 and 2018.

Supplementary File for "Handling Constrained Multiobjective Optimization Problems with Constraints in Both the Decision and Objective Spaces"

S-I. Nine Instances of DOC

A. DOC-1

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})(1 - \sqrt{f_1}/g(\mathbf{x})) \end{cases}$$

  where $g(\mathbf{x}) = 5.3578547x_4^2 + 0.8356891x_2x_6 + 37.293239x_2 - 10125.6023282166$.
- *Objective constraint:*

$$g_1 = f_1^2 + f_2^2 - 1 \geq 0.$$

- *Decision constraints [1]:*

$$\begin{aligned}
g_2 &= +85.334407 + 0.0056858x_3x_6 + 0.0006262x_2x_5 - 0.0022053x_4x_6 \leq 92;\\
g_3 &= -85.334407 - 0.0056858x_3x_6 - 0.0006262x_2x_5 + 0.0022053x_4x_6 \leq 0;\\
g_4 &= +80.51249 + 0.0071317x_3x_6 + 0.0029955x_2x_3 + 0.0021813x_4^2 \leq 110;\\
g_5 &= -80.51249 - 0.0071317x_3x_6 - 0.0029955x_2x_3 - 0.0021813x_4^2 \leq -90;\\
g_6 &= +9.300961 + 0.0047026x_4x_6 + 0.0012547x_2x_4 + 0.0019085x_4x_5 \leq 25;\\
g_7 &= -9.300961 - 0.0047026x_4x_6 - 0.0012547x_2x_4 - 0.0019085x_4x_5 \leq -20.
\end{aligned}$$

- The search space is: $0 \leq x_1 \leq 1$, $78 \leq x_2 \leq 102$, $33 \leq x_3 \leq 45$, and $27 \leq x_4, x_5, x_6 \leq 45$.
- Its Pareto front is:

$$\begin{aligned}
&f_2 = \sqrt{1 - f_1^2};\\
&0 \leq f_1 \leq 1.
\end{aligned}$$

  which is illustrated in Fig. S-1.

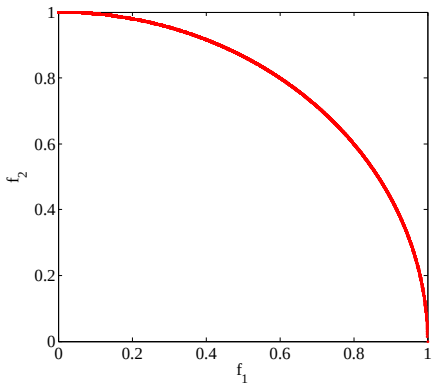


Fig. S-1. Pareto front of DOC-1.

TABLE S-I
PARAMETER SETTINGS FOR DOC-2

j	1	2	3	4	5
a_{1j}	-16	2	0	1	0
a_{2j}	0	-2	0	0.4	2
a_{3j}	-3.5	0	2	0	0
a_{4j}	0	-2	0	-4	-1
a_{5j}	0	-9	-2	1	-2.8
a_{6j}	2	0	-4	0	0
a_{7j}	-1	-1	-1	-1	-1
a_{8j}	-1	-2	-3	-2	-1
a_{9j}	1	2	3	4	5
a_{10j}	1	1	1	1	1
c_{1j}	30	-20	-10	32	-10
c_{2j}	-20	39	-6	-31	32
c_{3j}	-10	-6	10	-6	-10
c_{4j}	32	-31	-6	39	-20
c_{5j}	-10	32	-10	-20	30
d_j	4	8	10	6	2
e_j	-15	-27	-36	-18	-12

B. DOC-2

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})(1 - \sqrt[3]{f_1}/g(\mathbf{x})) \end{cases}$$

where $g(\mathbf{x}) = \sum_{j=1}^{5}\sum_{i=1}^{5} c_{ij}x_{11+i}x_{11+j} + 2\sum_{j=1}^{5} d_j x_{11+j}^3 - \sum_{i=1}^{10} b_i x_{i+1} - 31.6555929502$.

- *Objective constraints:*

$g_1 = \sqrt{f_1} + f_2 - 1 \geq 0$;

$g_2 = min\{(f_1 - \frac{1}{8})^2 + (f_2 - 1 + \frac{\sqrt{2}}{4})^2 - 0.0225, (f_1 - \frac{1}{2})^2 + (f_2 - 1 + \frac{\sqrt{2}}{2})^2 - 0.0225, (f_1 - \frac{7}{8})^2 + (f_2 - 1 + \frac{\sqrt{14}}{4})^2 - 0.0225\} \leq 0$.

- *Decision constraints [1]:*

$$g_{j+2} = -2\sum_{i=1}^{5} c_{ij}x_{11+i} - 3d_j x_{11+j}^2 - e_j + \sum_{i=1}^{10} a_{ij}x_{i+1} \leq 0, j = 1, 2, \ldots, 5.$$

where $\vec{b} = [-40, -2, -0.25, -4, -4, -1, -40, -60, 5, 1]$, and the remaining parameters are set in Table S-I.

- The search space is: $0 \leq x_1 \leq 1$, and $0 \leq x_i \leq 10, i = 2, 3, \ldots, 16$.
- Its Pareto front consists of three parts:

$$f_2 = 1 - \sqrt{f_1};$$
$$f_1 \in [0.050, 0.2202] \cup [0.3830, 0.6247] \cup [0.7440, 1].$$

which is illustrated in Fig. S-2.

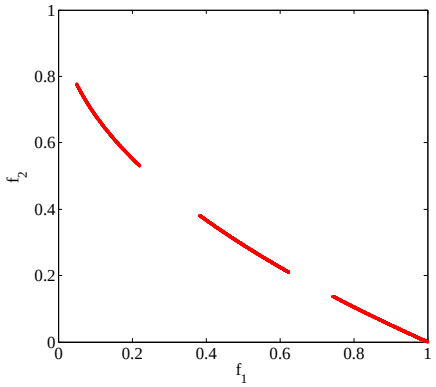


Fig. S-2. Pareto front of DOC-2.

C. DOC-3

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})(1 - f_1/g(\mathbf{x})) \end{cases}$$

  where $g(\mathbf{x}) = -9x_6 - 15x_9 + 6x_2 + 16x_3 + 10(x_7 + x_8) + 401.0551$.
- *Objective constraints:*

$$\begin{aligned} g_1 &= f_1^2 + f_2^2 - 1 \geq 0; \\ g_2 &= |f_1 - f_2 - 0.5| \geq 0.1; \\ g_3 &= |f_1 - f_2| \geq 0.1; \\ g_4 &= |f_1 - f_2 + 0.5| \geq 0.1. \end{aligned}$$

- *Decision constraints [2], [3]:*

$$\begin{aligned} g_5 &= x_{10}x_4 + 0.02x_7 - 0.025x_6 \leq 0; \\ g_6 &= x_{10}x_5 + 0.02x_8 - 0.015x_9 \leq 0; \\ h_1 &= x_2 + x_3 - x_4 - x_5 = 0; \\ h_2 &= 0.03x_2 + 0.01x_3 - x_{10}(x_4 + x_5) = 0; \\ h_3 &= x_4 + x_7 - x_6 = 0; \\ h_4 &= x_5 + x_8 - x_9 = 0. \end{aligned}$$

- The search space is: $0 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$, $0 \leq x_3 \leq 300$, $0 \leq x_4 \leq 100$, $0 \leq x_5 \leq 200$, $0 \leq x_6 \leq 100$, $0 \leq x_7 \leq 1$, $0 \leq x_8 \leq 100$, $0 \leq x_9 \leq 200$, and $0 \leq x_{10} \leq 0.03$.
- Its Pareto front consists of four parts:

$$\begin{aligned} &f_2 = \sqrt{1 - f_1^2}; \\ &f_1 \in [0, 0.3403] \cup [0.4782, 0.6553] \cup [0.7553, 0.8782] \cup [0.9403, 1]. \end{aligned}$$

  which is illustrated in Fig. S-3.

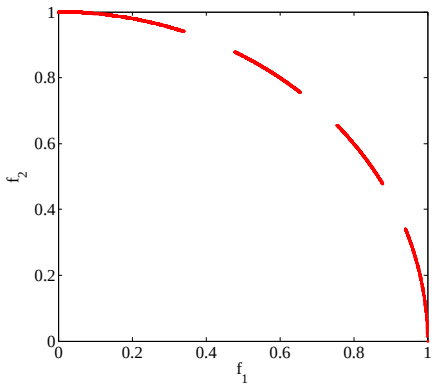


Fig. S-3. Pareto front of DOC-3.

D. DOC-4

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})(1 - \sqrt{f_1}/g(\mathbf{x})) \end{cases}$$

  where $g(\mathbf{x}) = (x_2 - 10)^2 + 5(x_3 - 12)^2 + x_4^4 + 3(x_5 - 11)^2 + 10x_6^6 + 7x_7^2 + x_8^4 - 4x_7x_8 - 10x_7 - 8x_8 - 679.6300573745$.
- *Objective constraints:*

$$\begin{aligned} g_1 &= f_1 + f_2 - 1 \geq 0; \\ g_2 &= f_1 + f_2 - 1 - |sin(10\pi(f_1 - f_2 + 1))| \geq 0. \end{aligned}$$

- *Decision constraints [4]:*

$$g_3 = -127 + 2x_2^2 + 3x_3^4 + x_4 + 4x_5^2 + 5x_6 \leq 0;$$
$$g_4 = -282 + 7x_2 + 3x_3 + 10x_4^2 + x_5 - x_6 \leq 0;$$
$$g_5 = -196 + 23x_2 + x_3^2 + 6x_7^2 - 8x_8 \leq 0;$$
$$g_6 = 4x_2^2 + x_3^2 - 3x_2x_3 + 2x_4^2 + 5x_7 - 11x_8 \leq 0.$$

- The search space is: $0 \leq x_1 \leq 1$, and $-10 \leq x_i \leq 10, i = 2, 3, \ldots, 8$.
- Its Pareto front consists of 21 points:

$$f_2 = 1 - f_1;$$
$$f_1 = i/20, i = 0, 1, \ldots, 20.$$

which is illustrated in Fig. S-4.

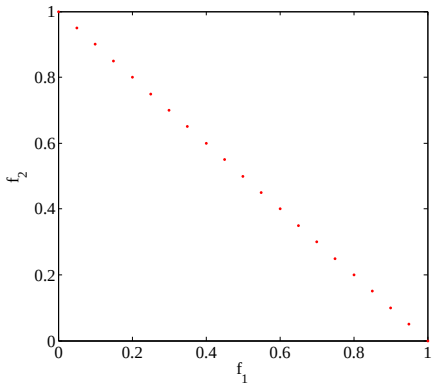


Fig. S-4. Pareto front of DOC-4.

E. DOC-5

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})(1 - \sqrt{f_1}/g(\mathbf{x})) \end{cases}$$

where $g(\mathbf{x}) = x_2 - 192.724510070035$.

- *Objective constraints:*

$$g_1 = f_1 + f_2 - 1 \geq 0;$$
$$g_2 = f_1 + f_2 - 1 - |sin(10\pi(f_1 - f_2 + 1))| \geq 0;$$
$$g_3 = (f_1 - 0.8)(f_2 - 0.6) \leq 0.$$

- *Decision constraints [2]:*

$$g_4 = -x_2 + 35x_3^{0.6} + 35x_4^{0.6} \leq 0;$$
$$h_1 = -300x_4 + 7500x_6 - 7500x_7 - 25x_5x_6 + 25x_5x_7 + x_4x_5 = 0;$$
$$h_2 = 100x_3 + 155.365x_5 + 2500x_8 - x_3x_5 - 25x_5x_8 - 15536.5 = 0;$$
$$h_3 = -x_6 + \log(-x_5 + 900) = 0;$$
$$h_4 = -x_7 + \log(x_5 + 300) = 0;$$
$$h_5 = -x_8 + \log(-2x_5 + 700) = 0.$$

- The search space is: $0 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1000$, $0 \leq x_3, x_4 \leq 40$, $100 \leq x_5 \leq 300$, $6.3 \leq x_6 \leq 6.7$, $5.9 \leq x_7 \leq 6.4$, and $4.5 \leq x_8 \leq 6.25$.
- Its Pareto front consists of 14 points:

$$f_2 = 1 - f_1;$$
$$f_1 = i/20, i = 0, 1, 2, 3, 4, 5, 6, 7, 8, 16, 17, 18, 19, 20.$$

which is illustrated in Fig. S-5.

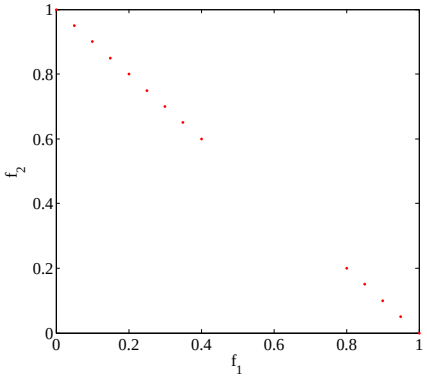


Fig. S-5. Pareto front of DOC-5.

F. DOC-6

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})(1 - \sqrt{f_1}/g(\mathbf{x})) \end{cases}$$

where $g(\mathbf{x}) = x_2^2 + x_3^2 + x_2x_3 - 14x_2 - 16x_3 + (x_4 - 10)^2 + 4(x_5 - 5)^2 + (x_6 - 3)^2 + 2(x_7 - 1)^2 + 5x_8^2 + 7(x_9 - 11)^2 + 2(x_{10} - 10)^2 + (x_{11} - 7)^2 + 21.693790931900001$.

- *Objective constraints:*

$$g_1 = f_1 + f_2 - 1 \geq 0;$$
$$g_2 = (f_1 - 0.5)(f_1 + f_2 - 1 - |sin(10\pi(f_1 - f_2 + 1))|) \geq 0.$$

- *Decision constraints [4]:*

$$g_3 = -105 + 4x_2 + 5x_3 - 3x_8 + 9x_9 \leq 0;$$
$$g_4 = 10x_2 - 8x_3 - 17x_8 + 2x_9 \leq 0;$$
$$g_5 = -8x_2 + 2x_3 + 5x_{10} - 2x_{11} - 12 \leq 0;$$
$$g_6 = 3(x_2 - 2)^2 + 4(x_3 - 3)^2 + 2x_4^2 - 7x_5 - 120 \leq 0;;$$
$$g_7 = 5x_2^2 + 8x_3 + (x_4 - 6)^2 - 2x_5 - 40 \leq 0;$$
$$g_8 = x_2^2 + 2(x_3 - 2)^2 - 2x_2x_3 + 14x_6 - 6x_7 \leq 0;$$
$$g_9 = 0.5(x_2 - 8)^2 + 2(x_3 - 4)^2 + 3x_6^2 - x_7 - 30 \leq 0;$$
$$g_{10} = -3x_2 + 6x_3 + 12(x_{10} - 8)^2 - 7x_{11} \leq 0.$$

- The search space is: $0 \leq x_1 \leq 1$, and $-10 \leq x_i \leq 10, i = 2, 3, \ldots, 11$.
- Its Pareto front is:

$$f_2 = 1 - f_1;$$
$$f_1 \in [0, 0.5] \cup f_1 = i/20, i = 11, \ldots, 20.$$

which is illustrated in Fig. S-6.

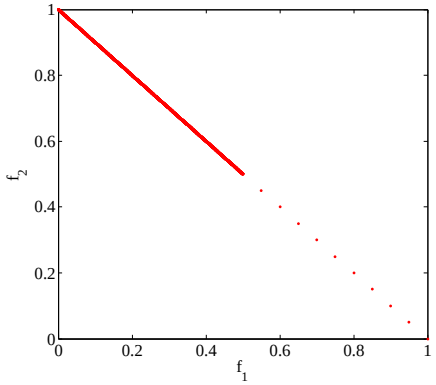


Fig. S-6. Pareto front of DOC-6.

G. DOC-7

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 \\ \min & f_2 = g(\mathbf{x})(1 - \sqrt{f_1}/g(\mathbf{x})) \end{cases}$$

where $g(\mathbf{x}) = \sum_{i=2}^{11} x_i(c_{i-1} + \ln \frac{x_i}{\sum_{j=2}^{11} x_j}) + 48.7648884$, $c_1 = -6.089$, $c_2 = -17.164$, $c_3 = -34.054$, $c_4 = -5.914$, $c_5 = -24.721$, $c_6 = -14.986$, $c_7 = -24.1$, $c_8 = -10.708$, $c_9 = -26.662$, and $c_{10} = -22.179$.

- *Objective constraints:*

$$\begin{aligned} g_1 &= f_1 + f_2 \geq 1; \\ g_2 &= (f_1 - 0.5)(f_1 + f_2 - |\sin(10\pi(f_1 - f_2 + 1))| + 1) \geq 0; \\ g_3 &= |f_1 - f_2| \geq 0.1. \end{aligned}$$

- *Decision constraints [5]:*

$$\begin{aligned} h_1 &= x_2 + 2x_3 + 2x_4 + x_7 + x_{11} = 2; \\ h_2 &= x_5 + 2x_6 + x_7 + x_8 = 1; \\ h_3 &= x_4 + x_8 + x_9 + 2x_{10} + x_{11} = 1. \end{aligned}$$

- The search space is: $0 \leq x_1 \leq 1$, and $0 \leq x_i \leq 10, i = 2, 3, \ldots, 11$.
- Its Pareto front is:

$$\begin{aligned} &f_2 = 1 - f_1; \\ &f_1 \in [0, 0.45] \cup f_1 = i/20, i = 11, 12, \ldots, 20. \end{aligned}$$

which is illustrated in Fig. S-7.

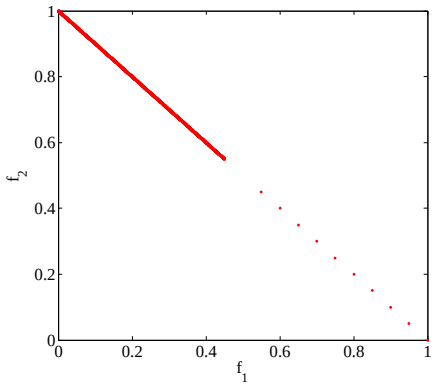


Fig. S-7. Pareto front of DOC-7.

H. DOC-8

- *Objective functions:*

$$\begin{cases} \min & f_1 = x_1 x_2 * g(\mathbf{x}) \\ \min & f_2 = x_1(1 - x_2) * g(\mathbf{x}) \\ \min & f_3 = (1 - x_1) * g(\mathbf{x}) \end{cases}$$

where $g(\mathbf{x}) = x_3 + x_4 + x_5 - 7048.2480205286$.

- *Objective constraint:*

$$g_1 = (f_3 - 0.4)(f_3 - 0.6) \geq 0.$$

- *Decision constraints [4]:*

$$\begin{aligned} g_2 &= -1 + 0.0025(x_6 + x_8) \leq 0; \\ g_3 &= -1 + 0.0025(x_7 + x_9 - x_6) \leq 0; \\ g_4 &= -1 + 0.01 * (x_{10} - x_7) \leq 0; \\ g_5 &= -x_3 x_8 + 833.33252 x_6 + 100 x_3 - 83333.333 \leq 0; \\ g_6 &= -x_4 x_9 + 1250 x_7 + x_4 x_6 - 1250 x_6 \leq 0; \\ g_7 &= -x_5 x_{10} + 1250000 + x_5 x_7 - 2500 x_7 \leq 0. \end{aligned}$$

- The search space is: $0 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$, $500 \leq x_3 \leq 1000$, $1000 \leq x_4 \leq 2000$, $5000 \leq x_5 \leq 6000$, and $100 \leq x_i \leq 500, i = 6, 7, \ldots, 10$.
- Its Pareto front is:

$$\begin{aligned} &f_1 = 1 - f_2 - f_3; \\ &0 \leq f_2 \leq 1; \\ &0 \leq f_3 \leq 0.4 \cup 0.6 \leq f_3 \leq 1. \end{aligned}$$

which is illustrated in Fig. S-8.

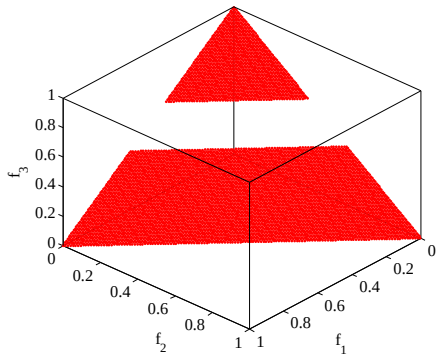


Fig. S-8. Pareto front of DOC-8.

I. DOC-9

- *Objective functions:*

$$\begin{cases} \min & f_1 = cos(0.5\pi x_1)cos(0.5\pi x_2) * g(\mathbf{x}) \\ \min & f_2 = cos(0.5\pi x_1)sin(0.5\pi x_2) * g(\mathbf{x}) \\ \min & f_3 = sin(0.5\pi x_2) * g(\mathbf{x}) \end{cases}$$

where $g(\mathbf{x}) = -0.5(x_3x_6 - x_4x_5 + x_5x_{11} - x_7x_{11} + x_7x_{10} - x_8x_9) + 1.8660254038$.
- *Objective constraint:*

$$g_1 = f_1^2 + f_2^2 \geq 1;$$

- *Decision constraints [1], [2]:*

$$\begin{aligned} g_2 &= x_5^2 + x_6^2 \leq 1; \\ g_3 &= x_{11}^2 \leq 1; \\ g_4 &= x_7^2 + x_8^2 \leq 1; \\ g_5 &= x_3^2 + (x_4 - x_{11})^2 \leq 1; \\ g_6 &= (x_3 - x_7)^2 + (x_4 - x_8)^2 \leq 1; \\ g_7 &= (x_3 - x_9)^2 + (x_4 - x_{10})^2 \leq 1; \\ g_8 &= (x_5 - x_7)^2 + (x_6 - x_8)^2 \leq 1; \\ g_9 &= (x_5 - x_9)^2 + (x_6 - x_{10})^2 \leq 1; \\ g_{10} &= x_9^2 + (x_{10} - x_{11})^2 \leq 1; \\ g_{11} &= x_4x_5 - x_3x_6 \leq 0; \\ g_{12} &= -x_5x_{11} \leq 0; \\ g_{13} &= x_7x_{11} \leq 0; \\ g_{14} &= x_8x_9 - x_7x_{10} \leq 0. \end{aligned}$$

- The search space is: $0 \leq x_1, x_2 \leq 1$, and $-1 \leq x_i \leq 10, i = 3, 4, \ldots, 11$.
- Its Pareto front is a line segment:

$$\begin{aligned} &f_2 = \sqrt{1 - f_1^2}; \\ &0 \leq f_1 \leq 1; \\ &f_3 = 0. \end{aligned}$$

which is illustrated in Fig. S-9.

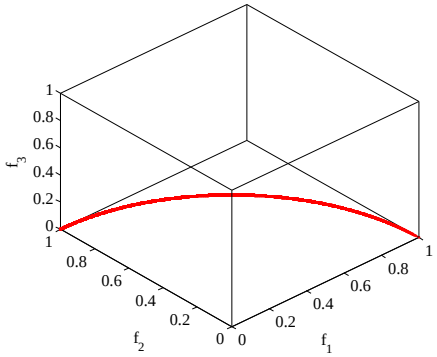
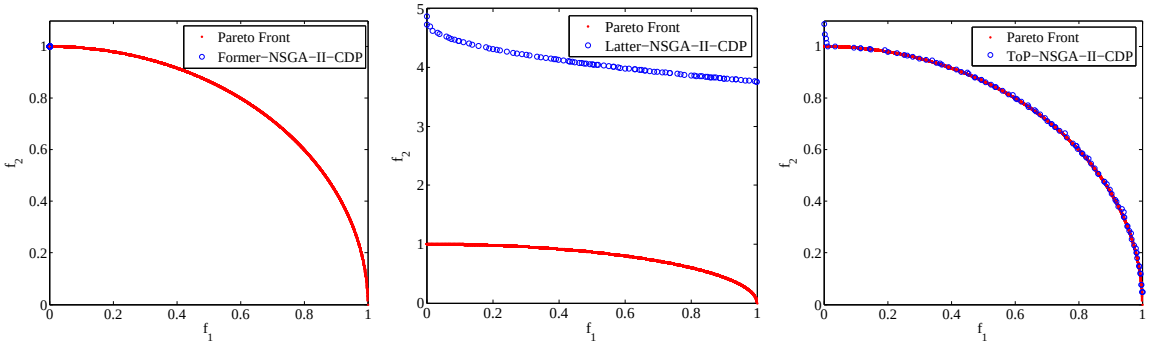
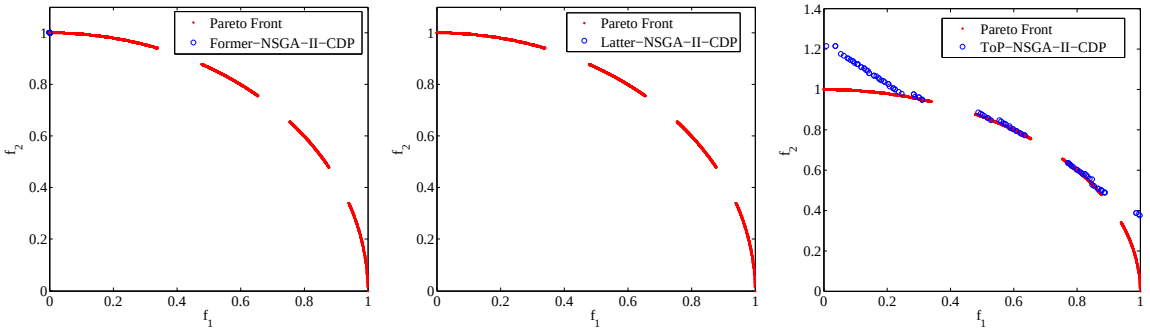


Fig. S-9. Pareto front of DOC-9.



(a) Former-NSGA-II-CDP (b) Latter-NSGA-II-CDP (c) ToP-NSGA-II-CDP

Fig. S-10. Images of the feasible solutions provided by ToP-NSGA-II-CDP and its two variants (Former-NSGA-II-CDP and Latter-NSGA-II-CDP) in the end of a run on DOC-1.



(a) Former-NSGA-II-CDP (b) Latter-NSGA-II-CDP (c) ToP-NSGA-II-CDP

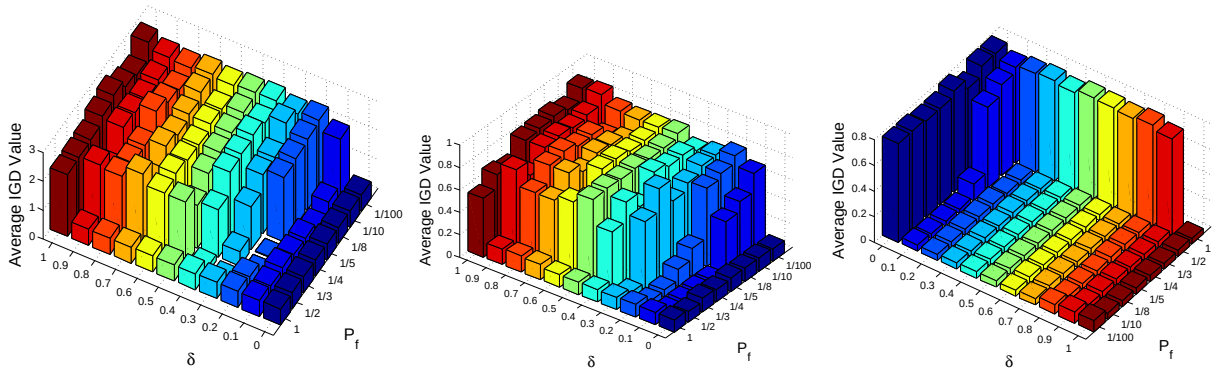
Fig. S-11. Images of the feasible solutions provided by ToP-NSGA-II-CDP and its two variants (Former-NSGA-II-CDP and Latter-NSGA-II-CDP) in the end of a run on DOC-3.

S-II. Additional Results and Discussions

A. *Effect of the Parameter Settings in ToP*

In this subsection, we conducted the sensitivity analysis of two parameters (i.e., P_f and δ) introduced in Section IV-B. These two parameters determined when to stop the first phase and enter the second phase. A large value of P_f and a small value of δ may result in the feasible solutions clustering in a very small area in the feasible region when the first phase ends. Whereas, with a small value of P_f and a large value of δ, the population may be distant from the Pareto front when the first phase terminates. Therefore, both P_f and δ should be set to a moderate value.

We selected ToP-NSGA-II-CDP as the instance algorithm and tested its performance on three DOC test instances: DOC-1, DOC-4, and DOC-9, with the aim of providing multi-facet insights. We chose eight different P_f values: 1/100, 1/10, 1/8, 1/5, 1/4, 1/3, 1/2, and 1, and 11 different δ values: 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. Fig. S-12 records



(a) DOC-1 (b) DOC-4 (c) DOC-9

Fig. S-12. Average IGD values provided by ToP-NSGA-II-CDP with 88 different combinations of P_f and δ on DOC-1, DOC-4, and DOC-9.

TABLE S-II
FEASIBILITY RATIOS OF CFS.

Instance	Feasibility Ratio
CF1	52.24%
CF2	99.42%
CF3	100.00%
CF4	50.00%
CF5	50.71%
CF6	30.73%
CF7	31.75%
CF8	0.44%
CF9	15.97%
CF10	0.00%

the average IGD values provided by ToP-NSGA-II-CDP with 88 different combinations of P_f and δ over 20 independent runs. From Fig. S-12, we can observe that, overall, ToP-NSGA-II-CDP exhibits better performance with $P_f \in [1/4, 1/3]$ and $\delta \in [0.2, 0.3]$. Therefore, in this paper, P_f was set to 1/3 and δ was set to 0.2.

B. Investigation to the Generality of ToP

The effectiveness of ToP has been verified on DOC whose feasibility ratios are very small. One might be interested in whether ToP can improve the performance of CMOEAs on other CMOPs whose feasibility ratios are large. To answer this question, we selected one dominance-based CMOEA (i.e., NSGA-II-CDP) and one decomposition-based CMOEA (i.e., MOEA/D-CDP) as two instance algorithms, and compared them with their variants under the framework of ToP (i.e., ToP-NSGA-II-CDP and ToP-MOEA/D-CDP) on the well-known CFs [6]. Note that CFs include ten instances for the Special Session & Competition on "Performance Assessment of Constrained/Bound Constrained Multi-Objective Optimization Algorithms" at IEEE CEC2009. The feasibility ratios of CFs are presented in Table S-II. It can be seen from Table S-II that all instances in CFs except for CF8 and CF10 have large feasible regions. For the parameter settings, they were set following the suggestions in Section V-B. The comparison results are presented in Table S-III and Table S-IV in terms of IGD and HV, respectively.

From Table S-III, ToP-NSGA-II-CDP and ToP-MOEA/D-CDP obtain better IGD values than NSGA-II-CDP and MOEA/D-CDP on nine and eight instances, respectively, while provide worse results on no more than one instance. Similarly, from Table S-IV, ToP-NSGA-II-CDP and ToP-MOEA/D-CDP outperform NSGA-II-CDP and MOEA/D-CDP on nine and eight instances, respectively, while lose on no more than one instance in terms of the HV metric. Thus, we can conclude that ToP has the capability to enhance the performance of NSGA-II-CDP and MOEA/D-CDP on CFs, which again validates the effectiveness of ToP.

C. Investigation to the Search Engine in the First Phase of ToP

In the first phase of ToP, both DE/current-to-rand/1 and DE/rand-to-best/1/bin are used to produce offspring. The aim of DE/current-to-rand/1 is to enhance the exploration ability of the population. As a classical DE version, DE/rand/1/bin also exhibits good exploration ability. Thus, a question which arises naturally is whether DE/current-to-rand/1 can be replaced with DE/rand/1/bin. To this end, we designed a variant of ToP-NSGA-II-CDP, named ToP-NSGA-II-CDP-1, in which DE/rand/1/bin is combined with DE/rand-to-best/1/bin in the first phase of ToP. The comparison results between ToP-NSGA-II-CDP-1 and ToP-NSGA-II-CDP on DOC are presented in Table S-V and Table S-VI in terms of IGD and HV, respectively.

TABLE S-III
EXPERIMENTAL RESULTS OF NSGA-II-CDP, TOP-NSGA-II-CDP, MOEA/D-CDP, AND TOP-MOEA/D-CDP OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD ON CFS. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	NSGA-II-CDP		ToP-NSGA-II-CDP		MOEA/D-CDP		ToP-MOEA/D-CDP	
	FR	IGD	FR	IGD	FR	IGD	FR	IGD
CF1	100%	4.952e-02(1.30e-02)	100%	4.948e-02(1.26e-02)	100%	8.230e-03(2.60e-03)	100%	**7.585e-03(2.51e-03)**
CF2	100%	1.702e-01(4.99e-02)	100%	**1.100e-01(2.15e-02)**	100%	1.350e-01(5.32e-02)	100%	**1.078e-01(3.43e-02)**
CF3	100%	4.849e-01(1.51e-01)	100%	**2.143e-01(6.23e-02)**	100%	4.398e-01(1.46e-01)	100%	**3.603e-01(1.76e-01)**
CF4	100%	1.590e-01(4.69e-02)	100%	**8.337e-02(3.53e-02)**	100%	2.029e-01(1.13e-01)	100%	**8.451e-02(2.85e-02)**
CF5	100%	3.761e-01(9.63e-02)	100%	**2.937e-01(1.48e-01)**	100%	3.945e-01(1.21e-01)	100%	**3.677e-01(1.34e-01)**
CF6	100%	1.254e-01(3.82e-02)	100%	**9.879e-02(2.77e-02)**	100%	1.546e-01(5.28e-02)	100%	**1.334e-01(4.57e-02)**
CF7	100%	3.763e-01(9.81e-02)	100%	**1.925e-01(9.03e-02)**	100%	3.803e-01(1.25e-01)	100%	**2.276e-01(1.09e-01)**
CF8	100%	3.453e-01(6.68e-02)	100%	**2.745e-01(4.19e-02)**	100%	**1.163e-01(3.17e-02)**	100%	1.560e-01(3.98e-02)
CF9	100%	1.966e-01(2.73e-02)	100%	**1.491e-01(2.41e-02)**	100%	1.058e-01(1.08e-02)	100%	1.060e-01(7.28e-03)
CF10	0%	NA	**100%**	**5.588e-01(9.19e-02)**	0%	NA	**100%**	**4.680e-01(1.09e-01)**

TABLE S-IV
EXPERIMENTAL RESULTS OF NSGA-II-CDP, TOP-NSGA-II-CDP, MOEA/D-CDP, AND TOP-MOEA/D-CDP OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND HV ON CFS. FOR HV, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN A CMOEA AND ITS AUGMENTED VERSION, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	NSGA-II-CDP		ToP-NSGA-II-CDP		MOEA/D-CDP		ToP-MOEA/D-CDP	
	FR	HV	FR	HV	FR	IGD	FR	HV
CF1	100%	6.133e-01(1.39e-02)	100%	6.131e-01(1.65e-02)	100%	6.721e-01(4.26e-03)	100%	**6.729e-01(4.05e-03)**
CF2	100%	6.917e-01(3.79e-02)	100%	**7.490e-01(2.74e-02)**	100%	7.448e-01(4.68e-02)	100%	**7.774e-01(1.94e-02)**
CF3	100%	1.871e-01(4.59e-02)	100%	**2.354e-01(6.44e-02)**	100%	1.582e-01(6.99e-02)	100%	**1.593e-01(7.00e-02)**
CF4	100%	4.249e-01(4.53e-02)	100%	**5.147e-01(5.62e-02)**	100%	4.179e-01(7.84e-02)	100%	**5.111e-01(4.14e-02)**
CF5	100%	2.746e-01(6.88e-02)	100%	**3.731e-01(1.10e-01)**	100%	2.905e-01(8.96e-02)	100%	**3.072e-01(9.18e-02)**
CF6	100%	7.196e-01(6.46e-02)	100%	**7.718e-01(1.60e-02)**	100%	7.383e-01(4.06e-02)	100%	**7.501e-01(3.18e-02)**
CF7	100%	4.310e-01(1.42e-01)	100%	**6.249e-01(8.07e-02)**	100%	4.690e-01(1.17e-01)	100%	**5.996e-01(8.90e-02)**
CF8	100%	2.698e-01(9.52e-02)	100%	**3.619e-01(6.18e-02)**	100%	**5.876e-01(3.29e-02)**	100%	5.495e-01(3.65e-02)
CF9	100%	4.460e-01(4.85e-02)	100%	**5.209e-01(4.59e-02)**	100%	6.465e-01(1.54e-02)	100%	6.464e-01(1.23e-02)
CF10	0%	NA	**100%**	**1.434e-01(5.06e-02)**	0%	NA	**100%**	**2.115e-01(8.98e-02)**

TABLE S-V
EXPERIMENTAL RESULTS OF TOP-NSGA-II-CDP-1 AND TOP-NSGA-II-CDP OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN TOP-NSGA-II-CDP-1 AND TOP-NSGA-II-CDP, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	ToP-NSGA-II-CDP-1		ToP-NSGA-II-CDP	
	FR	IGD	FR	IGD
DOC-1	100%	1.355e-01(2.30e-01)	100%	**6.859e-3(6.91e-4)**
DOC-2	100%	**2.444e-02(1.27e-02)**	100%	3.569e-2(1.47e-2)
DOC-3	100%	1.796e-01(9.16e-02)	100%	**9.264e-2(5.67e-2)**
DOC-4	100%	5.862e-02(2.18e-02)	100%	**4.477e-2(1.76e-2)**
DOC-5	100%	3.028e+01(5.02e+01)	100%	**1.634e-1(9.82e-2)**
DOC-6	100%	5.734e-03(2.38e-03)	100%	**4.550e-3(3.71e-4)**
DOC-7	100%	3.716e-02(5.55e-02)	100%	**1.716e-2(5.20e-3)**
DOC-8	100%	1.550e+00(2.09e+00)	100%	**1.045e-1(2.29e-2)**
DOC-9	100%	3.943e-02(7.05e-03)	100%	**3.162e-2(4.55e-3)**

From these two tables, it is evident that ToP-NSGA-II-CDP achieves superior performance on most of instances in terms of both IGD and HV, compared with ToP-NSGA-II-CDP-1. Therefore, we can conclude that DE/current-to-rand/1 is more effective than DE/rand/1/bin in the first phase of ToP. It may be because DE/current-to-rand/1 is a rotation-invariant DE version due to the fact that the binomial crossover is not applied, as pointed out in Section IV-B.

D. Effect of the Constraint-Handling Techniques in the Second Phase of ToP

In this subsection, we investigated the influence of different constraint-handling techniques in the second phase of ToP. Firstly, we incorporated CDP [7] and ATM [8] into NSGA-II and obtained NSGA-II-CDP and NSGA-II-ATM, respectively.

TABLE S-VI
EXPERIMENTAL RESULTS OF TOP-NSGA-II-CDP-1 AND TOP-NSGA-II-CDP OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND HV. FOR HV, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN TOP-NSGA-II-CDP-1 AND TOP-NSGA-II-CDP, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	ToP-NSGA-II-CDP-1		ToP-NSGA-II-CDP	
	FR	HV	FR	HV
DOC-1	100%	3.313E-01(1.30E-01)	100%	**4.041e-1(6.55e-3)**
DOC-2	100%	5.588E-01(1.46E-02)	100%	**5.677e-1(4.76e-3)**
DOC-3	100%	2.244E-01(7.14E-02)	100%	**2.630e-1(4.95e-2)**
DOC-4	100%	5.995E-01(2.99E-02)	100%	**6.144e-1(1.96e-2)**
DOC-5	100%	3.617E-01(2.47E-01)	100%	**5.022e-1(5.65e-2)**
DOC-6	100%	**6.312E-01(2.80E-02)**	100%	6.248e-1(2.60e-2)
DOC-7	100%	5.322E-01(7.67E-02)	100%	**5.654e-1(3.24e-2)**
DOC-8	100%	2.011E-01(2.44E-01)	100%	**7.471e-1(1.53e-2)**
DOC-9	100%	3.625E-02(2.03E-03)	100%	**3.696e-2(1.14e-3)**

TABLE S-VII
EXPERIMENTAL RESULTS OF TOP-NSGA-II-CDP AND TOP-NSGA-II-ATM OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN TOP-NSGA-II-CDP AND TOP-NSGA-II-ATM, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	ToP-NSGA-II-CDP		ToP-NSGA-II-ATM	
	FR	IGD	FR	IGD
DOC-1	100%	**6.859e-3(6.91e-4)**	100%	1.413e-01(2.39e-01)
DOC-2	100%	3.569e-2(1.47e-2)	100%	**1.889e-02(8.13e-03)**
DOC-3	100%	**9.264e-2(5.67e-2)**	100%	1.401e-01(1.01e-01)
DOC-4	100%	**4.477e-2(1.76e-2)**	100%	8.506e-02(1.21e-01)
DOC-5	100%	**1.634e-1(9.82e-2)**	100%	1.934e-1(1.82e-01)
DOC-6	100%	**4.550e-3(3.71e-4)**	100%	5.955e-03(4.52e-03)
DOC-7	100%	**1.716e-2(5.20e-3)**	100%	1.859e-02(8.06e-03)
DOC-8	100%	**1.045e-1(2.29e-2)**	100%	2.920e-01(1.50e-01)
DOC-9	100%	**3.162e-2(4.55e-3)**	100%	3.612e-02(4.66e-03)

TABLE S-VIII
EXPERIMENTAL RESULTS OF TOP-NSGA-II-CDP AND TOP-NSGA-II-ATM OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND HV. FOR HV, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN TOP-NSGA-II-CDP AND TOP-NSGA-II-ATM, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	ToP-NSGA-II-CDP		ToP-NSGA-II-ATM	
	FR	HV	FR	HV
DOC-1	100%	**4.041e-1(6.55e-3)**	100%	3.324e-01(1.32e-01)
DOC-2	100%	**5.677e-1(4.76e-3)**	100%	5.651e-01(9.71e-03)
DOC-3	100%	2.630e-1(4.95e-2)	100%	2.634e-01(5.95e-02)
DOC-4	100%	**6.144e-1(1.96e-2)**	100%	5.724e-01(1.22e-01)
DOC-5	100%	**5.022e-1(5.65e-2)**	100%	4.873e-01(1.86e-01)
DOC-6	100%	**6.248e-1(2.60e-2)**	100%	6.226e-01(2.85e-02)
DOC-7	100%	**5.654e-1(3.24e-2)**	100%	5.588e-01(2.88e-02)
DOC-8	100%	**7.471e-1(1.53e-2)**	100%	6.888e-01(2.28e-01)
DOC-9	100%	3.696e-2(1.14e-3)	100%	**3.717e-02(1.04e-03)**

Subsequently, we applied ToP to NSGA-II-CDP and NSGA-II-ATM and tested the performance of ToP-NSGA-II-CDP and ToP-NSGA-II-ATM on DOC. The comparison results are summarized in Table S-VII and Table S-VIII in terms of IGD and HV, respectively.

Regarding IGD, Table S-VII shows that ToP-NSGA-II-CDP obtains better performance than ToP-NSGA-II-ATM on eight instances yet worse performance on only one instance (i.e., DOC-2). With respect to HV, Table S-VIII suggests that ToP-NSGA-II-CDP beats ToP-NSGA-II-ATM on seven instances while loses on only one instance (i.e., DOC-9). It is thus concluded that ToP-NSGA-II-CDP is superior to ToP-NSGA-II-ATM on DOC. The reason is that in ToP-NSGA-II-ATM, the infeasible solutions with good objective function values may be better than the feasible solutions with poor objective function values, and a lot of infeasible solutions may be kept in the final population. However, such infeasible solutions are deleted before the calculation of IGD and HV as introduced in Section V-A. Due to the relatively less feasible solutions, ToP-NSGA-II-ATM provides worse IGD and HV values, compared with ToP-NSGA-II-CDP.

TABLE S-IX
EXPERIMENTAL RESULTS OF NORMALIZED-TOP-NSGA-II-CDP AND TOP-NSGA-II-CDP OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN NORMALIZED-TOP-NSGA-II-CDP-DE AND TOP-NSGA-II-CDP-DE, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	Normalized-ToP-NSGA-II-CDP		ToP-NSGA-II-CDP	
	FR	IGD	FR	IGD
DOC-1	100%	3.088e+00(2.95e+00)	100%	**6.859e-3(6.91e-4)**
DOC-2	100%	1.743e-01(1.67e-01)	100%	**3.569e-2(1.47e-2)**
DOC-3	100%	2.013e+02(7.13e+01)	100%	**9.264e-2(5.67e-2)**
DOC-4	100%	6.868e-01(3.93e-01)	100%	**4.477e-2(1.76e-2)**
DOC-5	100%	7.018e+01(6.09e+01)	100%	**1.634e-1(9.82e-2)**
DOC-6	100%	1.135e+00(6.34e-01)	100%	**4.550e-3(3.71e-4)**
DOC-7	100%	1.760e+00(5.92e-01)	100%	**1.716e-2(5.20e-3)**
DOC-8	100%	6.346e+01(3.83e+01)	100%	**1.045e-1(2.29e-2)**
DOC-9	100%	1.193e-01(4.39e-02)	100%	**3.162e-2(4.55e-3)**

TABLE S-X
EXPERIMENTAL RESULTS OF NORMALIZED-TOP-NSGA-II-CDP-DE AND TOP-NSGA-II-CDP-DE OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND HV. FOR HV, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN NORMALIZED-NSGA-II-CDP AND TOP-NSGA-II-CDP, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

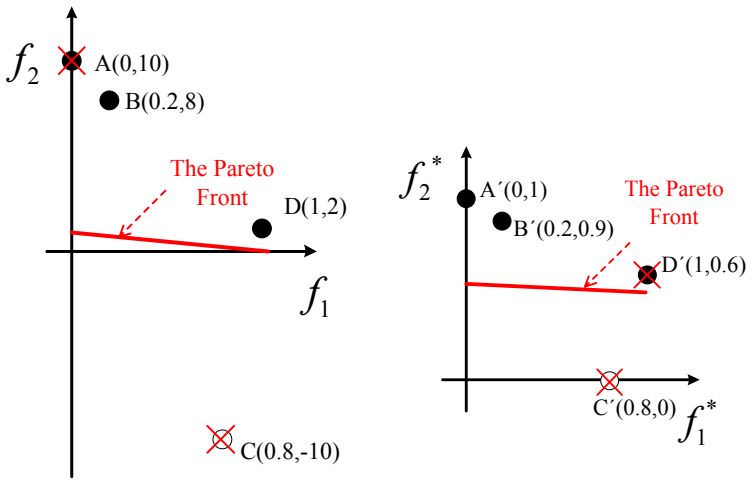
Instance	Normalized-ToP-NSGA-II-CDP		ToP-NSGA-II-CDP	
	FR	HV	FR	HV
DOC-1	100%	2.551e-02(8.79e-02)	100%	**4.041e-1(6.55e-3)**
DOC-2	100%	4.678e-01(9.25e-02)	100%	**5.677e-1(4.76e-3)**
DOC-3	100%	0.000e+00(0.00e+00)	100%	**2.630e-1(4.95e-2)**
DOC-4	100%	1.041e-01(1.22e-01)	100%	**6.144e-1(1.96e-2)**
DOC-5	100%	4.623e-02(1.19e-01)	100%	**5.022e-1(5.65e-2)**
DOC-6	100%	3.183e-02(8.68e-02)	100%	**6.248e-1(2.60e-2)**
DOC-7	100%	0.000e+00(0.00e+00)	100%	**5.654e-1(3.24e-2)**
DOC-8	100%	3.047e-04(1.36e-03)	100%	**7.471e-1(1.53e-2)**
DOC-9	100%	2.043e-02(7.48e-03)	100%	**3.696e-2(1.14e-3)**

E. Is the Normalized Process Necessary?

In the first phase of ToP, a CMOP is transformed into a constrained single-objective optimization problem. One may argue that each objective function should be normalized since the bias might occur among different objective functions. To investigate this issue, we selected ToP-NSGA-II-CDP as the instance algorithm and compared its performance with that of Normalized-ToP-NSGA-II-CDP on DOC. Normalized-ToP-NSGA-II-CDP is a variant of ToP-NSGA-II-CDP, whose objective functions are normalized. The comparison results are presented in Table S-IX and Table S-X in terms of IGD and HV, respectively.

As shown in Table S-IX and Table S-X, ToP-NSGA-II-CDP performs better than Normalized-ToP-NSGA-II-CDP on all instances in terms of both IGD and HV. The above comparison indicates that the normalized process is not necessary for ToP. It is probably because both the feasible and infeasible solutions are employed to normalize the objective functions. For some infeasible solutions, they might be far away from the feasible region but with small objective function values, then the usage of them in the normalized process may mislead the search of the population. Next, we take an example to explain it. Suppose that in a two-dimensional objective space, there are four individuals (i.e., **A**(0,10), **B**(0.2,8), **C**(0.8,-10), and **D**(1,2)) in the population and our task is to select two individuals into the next generation. For **A**, **B**, and **D**, they are feasible solutions, while for **C**, it is an infeasible solution which would be deleted first based on the feasibility rule in Section IV-B. If the normalized process is not implemented, for these three feasible solutions, **B** and **D** will be selected into the next generation as shown in Fig. 13(a), since $f^{'}(\mathbf{A}) = 0 + 10 = 10 > f^{'}(\mathbf{B}) = 8 + 0.2 = 8.2 > f^{'}(\mathbf{D}) = 1 + 2 = 3$. On the contrary, if the normalized process is conducted, then **A**, **B**, **C**, and **D** will be normalized as $\mathbf{A}^{'}$(0,1), $\mathbf{B}^{'}$(0.2,0.9), $\mathbf{C}^{'}$(0.8,0), and $\mathbf{D}^{'}$(1,0.6), respectively. Then **D** will be removed from the population since $f^{'}(\mathbf{A}^{'}) = 0+1 = 1 < f^{'}(\mathbf{B}^{'}) = 0.2+0.9 = 1.1 < f^{'}(\mathbf{D}^{'}) = 1+0.6 = 1.6$ as shown in Fig. 13(b). Note, however, that **D** is the closest individual to the Pareto front which should not be deleted. Thus we can conclude that the normalized process may mislead the selection in the first phase of ToP. Someone may argue why not only employ the feasible solutions to normalize the objective functions. Unfortunately, this way might be invalid, since all solutions in the population may be infeasible in the early stage of evolution.

Therefore, we did not normalize the objective functions in (13). The reason why ToP without the normalized process can obtain good performance may be because during the evolution, the bias among different objective functions will gradually decrease due to the fact that we minimize the sum of the objective functions in (13).



(a) Without the normalized process (b) With the normalized process

Fig. S-13. Illustration of the effects of the normalized process in the first phase of ToP. There are four individuals in the population, i.e., **A**, **B**, **C** and **D**, and the task is to select two promising individuals into the next generation. For **A**, **B**, and **D**, they are feasible solutions, while for **C**, it is an infeasible solution.

TABLE S-XI
EXPERIMENTAL RESULTS OF THE AVERAGE NUMBER OF FITNESS EVALUATIONS USED IN THE FIRST AND SECOND PHASES OF TOP-NSGA-II-CDP OVER 20 INDEPENDENT RUNS.

Instance	Phase 1		Phase 2	
	The average number of FEs	Percentage	The average number of FEs	Percentage
DOC-1	2.71e+04	13.53%	1.73e+05	86.47%
DOC-2	6.65e+04	33.23%	1.34e+05	66.78%
DOC-3	9.29e+04	46.45%	1.07e+05	53.55%
DOC-4	2.13e+04	10.67%	1.79e+05	89.34%
DOC-5	9.04e+04	45.18%	1.10e+05	54.82%
DOC-6	4.43e+04	22.17%	1.56e+05	77.83%
DOC-7	5.44e+04	27.19%	1.46e+05	72.81%
DOC-8	2.13e+05	53.26%	1.87e+05	46.74%
DOC-9	5.94e+04	14.84%	3.41e+05	85.16%

F. Allocation of the Number of Fitness Evaluations in the First and Second Phases

In this subsection, we are interested in investigating the allocation of the number of fitness evaluations (FEs) in the first and second phases of ToP for DOC. To this end, we selected ToP-NSGA-II-CDP as the instance algorithm and recorded the number of FEs allocated in the first and second phases of ToP-NSGA-II-CDP. The results are presented in Table S-XI.

From Table S-XI, it is obvious that for different instances, the numbers of FEs allocated in the first and second phases are different. It is the reason why we made use of two conditions, rather than a deterministic approach, to stop the first phase.

G. Effectiveness of the Best Individual in ToP

In Section IV-B, we mentioned that the objective function information is useful under the circumstance that the Pareto optimal solutions are exactly located on the boundaries of the feasible region. Moreover, the objective function information can also be used to accelerate the convergence in the feasible region. Indeed, the objective function information can be flexibly exploited in the first stage of ToP due to the single objective function, compared with the original multiple objective functions. In this paper, we made use of the objective function information via the best individual in (15). One may be interested in whether the objective function information is really helpful. To this end, we selected ToP-NSGA-II-CDP-DE as the instance algorithm, and tested its performance with or without the best individual in the first phase of ToP. The variant without the best individual is named Rand-ToP-NSGA-II-CDP-DE, in which the best individual is replaced with a randomly selected individual in the current population. The comparison results between Rand-ToP-NSGA-II-CDP-DE and ToP-NSGA-II-CDP-DE are presented in Table S-XII.

From Table S-XII, both ToP-NSGA-II-CDP and Rand-ToP-NSGA-II-CDP obtain satisfactory FR values (i.e, 100%) on all instances, which means that the difference between the utilization of the best individual and that of the random individual is not significant in finding the feasible region. It is because in the infeasible region, the best individual is similar to a random

TABLE S-XII
EXPERIMENTAL RESULTS OF RAND-NSGA-II-CDP-DE AND TOP-NSGA-II-CDP-DE OVER 20 INDEPENDENT RUNS IN TERMS OF FR AND IGD. FOR IGD, THE AVERAGE AND STANDARD DEVIATION ARE RECORDED. FOR EACH INSTANCE, WILCOXON'S RANK SUM TEST AT 0.05 SIGNIFICANCE LEVEL IS PERFORMED BETWEEN RAND-NSGA-II-CDP-DE AND TOP-NSGA-II-CDP-DE, AND THE BETTER RESULT IS HIGHLIGHTED IN BOLDFACE.

Instance	Rand-ToP-NSGA-II-CDP-DE		ToP-NSGA-II-CDP-DE	
	FR	IGD	FR	IGD
DOC-1	100%	8.804e-2(1.99e-1)	100%	**6.437e-3(2.30e-4)**
DOC-2	100%	4.441e-1(2.68e-4)	100%	4.431e-1(1.67e-4)
DOC-3	100%	9.911e-3(1.74e-3)	100%	**8.466e-3(2.02e-3)**
DOC-4	100%	9.388e-2(2.35e-1)	100%	9.392e-2(2.35e-1)
DOC-5	100%	2.094e+1(3.81e+1)	100%	**2.968e-2(5.78e-3)**
DOC-6	100%	2.817e-3(1.01e-4)	100%	2.790e-3(9.93e-05)
DOC-7	100%	2.637e-3(1.01e-4)	100%	**2.590e-3(8.99e-5)**
DOC-8	100%	1.669e-1(3.66e-2)	100%	**1.188e-1(1.26e-2)**
DOC-9	100%	8.882e-2(1.16e-2)	100%	**8.599e-2(1.26e-2)**

individual if the best individual is selected based on objective function, as analyzed in Section IV-B. In terms of IGD, ToP-NSGA-II-CDP-DE outperforms Rand-ToP-NSGA-II-CDP-DE on six instances (i.e., DOC-1, DOC-3, DOC-5, DOC-7, DOC-8, and DOC-9), and performs similar to Rand-ToP-NSGA-II-CDP-DE on the remaining three instances (i.e., DOC-2, DOC-4, and DOC-6). Thus, Rand-ToP-NSGA-II-CDP-DE cannot surpass ToP-NSGA-II-CDP-DE on any instance.

The above comparison demonstrates that it is beneficial to use the objective function information to guide the search in ToP.

REFERENCES

[1] D. M. Himmelblau, *Applied nonlinear programming*. McGraw-Hill Companies, 1972.

[2] J. J. Liang, T. P. Runarsson, E. Mezura-Montes, M. Clerc, P. N. Suganthan, C. A. Coello Coello, and K. Deb, "Problem definitions and evaluation criteria for the CEC 2006 special session on constrained real-parameter optimization," *International Journal of Computer Assisted Radiology and Surgery*, no. 2, 2005.

[3] Q. Xia, "Global optimization test problems," https://www.mat.univie.ac.at/~neum/glopt/xia.txt, 1996.

[4] W. Hock and K. Schittkowski, "Test examples for nonlinear programming codes," *Journal of Optimization Theory and Applications*, vol. 30, no. 1, pp. 127–129, 1980.

[5] C. A. Floudas, P. M. Pardalos, C. Adjiman, W. R. Esposito, Z. H. Gümüs, S. T. Harding, J. L. Klepeis, C. A. Meyer, and C. A. Schweiger, "Handbook of test problems in local and global optimization," *Nonconvex Optimization and Its Applications*, 1999.

[6] Q. Zhang, A. Zhou, S. Zhao, P. N. Suganthan, W. Liu, and S. Tiwari, "Multiobjective optimization test instances for the CEC 2009 special session and competition," *University of Essex, Colchester, UK and Nanyang technological University, Singapore, special session on performance assessment of multi-objective optimization algorithms, technical report*, vol. 264, 2008.

[7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[8] J.-P. Li, Y. Wang, S. Yang, and Z. Cai, "A comparative study of constraint-handling techniques in evolutionary constrained multiobjective optimization," in *2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2016, pp. 4175–4182.