Global Pareto Optimality of Cone Decomposition of Bi-objective Optimization

Weiqin Ying*†, Bin Wu*, Yuehong Xie*, Weipeng He*
*School of Software Engineering
South China University of Technology
Guangzhou 510006, China
Email: yingweiqin@scut.edu.cn
†State Key Laboratory for Novel Software Technology
Nanjing University
Nanjing 210023, China

Yu Wu
School of Computer Science and Educational Software
Guangzhou University
Guangzhou 510006, China

***Abstract*—The idea of decomposition is becoming increasingly successful and popular in evolutionary multi-objective optimization. An efficient cone decomposition approach was further developed to divide the bi-objective space and associate each subproblem with an exclusive conical subregion in the conical area evolutionary algorithm (CAEA). This approach improves the runtime efficiency and population diversity of decomposition-based algorithms effectively for bi-objective optimization in practice. In this paper, it is proved that the optimum of any conical subproblem in cone decomposition possesses the global Pareto optimality, apart from previously known the local Pareto optimality, in the presence of continuous frontier segment within the associated subregion. Experiments are conducted on 5 bi-objective benchmark problems among CAEA and the other four state-of-the-art algorithms. And the experimental results indicate that the solutions of conical subproblems obtained by CAEA not only acquire obviously better Pareto optimality in terms of C-metric , but also constitute higher qualities of frontiers in terms of IGD$^+$ metric.**

***Keywords*-Evolutionary algorithm, bi-objective optimization, Pareto optimality, Pareto frontier, decomposition.**

I. Introduction

Many real-life applications in business, management, and engineering can be naturally modeled as multi-objective optimization problems (MOPs) which need to optimize multiple and often conflicting objectives simultaneously [1]. For example, product designing may involve several objectives such as minimizing cost, maximizing performance, maximizing reliability and so on. A MOP can be generally formulated as the following form:

$$\begin{array}{ll} minimize & \mathbf{y}=\mathbf{f}(\mathbf{x})=(f_1(\mathbf{x}),f_2(\mathbf{x}),\dots f_m(\mathbf{x}))^T \\ subject\ \ to & \mathbf{x}=(x_1,x_2,\cdots,x_n)\in\Omega. \end{array} \quad (1)$$

where $\mathbf{x}\in\Omega$ is a decision vector and $\Omega\subseteq\mathbb{R}^n$ is the decision space, while $\mathbf{y}\in\mathbb{R}^m$ is the objective vector and $\mathbb{R}^m$ is the objective space. When the number of objectives m is equals to 2, a MOP becomes a *bi-objective optimization problem* (BOP) which is the most elementary case of MOP.

Since the natural confliction existing among objectives, the improvement of one among them often results in the deterioration of others. Consequently, there generally exists no single perfect solution which simultaneously optimizes all these objective functions. Instead, there is a set of best trade-off solutions, known as non-dominated solutions for MOPs. Let $\mathbf{u}$,$\mathbf{v}\in\mathbb{R}^m$, $\mathbf{u}$ is said to *dominate* $\mathbf{v}$, written as $\mathbf{u}\prec\mathbf{v}$, if and only if $\forall\, i\in[1..m]: u_i\leq v_i \wedge \exists j\in[1..m]: u_j<v_j$. A solution $\mathbf{x}^*\in\Omega$ is Pareto optimal or non-dominated, if there is no solution $\mathbf{x}\in\Omega$ such that $\mathbf{f}(\mathbf{x})$ dominates $\mathbf{f}(\mathbf{x}^*)$. The set S^* of all Pareto optimality solutions in the decision space, also called as *Pareto Set* (PS), can be mapped to a *Pareto Frontier* (PF) $\mathcal{F}=\mathbf{f}(S^*)=\{\mathbf{f}(\mathbf{x}^*)|\mathbf{x}^*\in S^*\}$ in the objective space [2].

As population-based stochastic search approaches, evolutionary algorithms (EAs) are particularly suitable to solve MOPs for the reason that their inherent parallelism makes them able to find a set of Pareto non-dominated solutions in a single run[3], instead of a series separate runs using traditional techniques. The majority of existing multi-objective optimization EAs (MOEAs), such as the popular *non-dominated sorting genetic algorithm II* (NSGA-II)[4], are based on Pareto dominance. NSGA-II sorts individuals in a population according to their dominance depths to approach the PF and prefers the individuals with larger crowding distances to preserve the diversity of distribution along the PF. Thereafter, a MOEA using reference-point based non-dominated sorting (NSGA-III) [2] was proposed, which employs a predefined set of reference points to keep the diversity of population instead of the assistant strategy based on crowding distance in NSGA-II. Nevertheless, there has been still no effective algorithm for non-dominance checking in a population so far. As a result, the runtime efficiencies of dominance-based MOEAs decrease seriously as the population size N increases. In recent years, several more efficient decomposition-based MOEAs, such as *multi-objective evolutionary algorithm based on decomposion* (MOEA/D) [5], [6], are built by borrowing the idea of decomposition widely used in traditional mathematical programming methods. MOEA/D avoids the expensive non-domination sorting operation in domination-based MOEAs by converting a MOP into a series of scalar optimization subproblems. Thereafter, apart from introducing a differential evolution (DE) operator to generate offsprings, MOEA/D-DE [7] was proposed to apply n_r to limit the maximal number of solutions replaced by offspring instead of using the neighborhood size T. Moreover, MOEA/D-AGR [8] adopts a global replacement scheme for selection in MOEA/D to find the most appropriate subproblem in view of objective functions for updating.

To effectively enhance the runtime efficiency and the population diversity of decomposition-based algorithms for bi-objective optimization, a *conical area evolutionary algorithm* (CAEA) [9] has been further developed, which employs an efficient cone decomposition approach. In this approach, not only a BOP is decomposed into N scalar conical subproblems, but also the objective space is also divided into N conical subregions. In addition, each conical sub-problem utilizes a specially designed conical area indicator as the scalar objective to obtain a local optimal solution in its own conical sub-region. Importantly, it has been proved that the cone decomposition approach has a very good property of local Pareto optimality. However, whether the cone decomposition approach possesses the better property of global Pareto optimality, that is to say, whether the optimal solution of each conical subproblem is Pareto non-dominated in the entire bi-objective space, is still an open question till now. This paper theoretically proves and empirically validates that the cone decomposition approach has an excellent property of global Pareto optimality under certain continuity assumptions which are easy to fulfill.

The remainder of this paper is organized as follows. Section II briefly introduces the cone decomposition approach as well as its local Pareto optimality. In Section III, we give in detail a strict proof of the global Pareto optimality of cone decomposition under certain continuity assumptions. Section IV provides the experimental results and analysis on 5 benchmark bi-objective problems. Finally, Section V concludes this paper.

II. Preliminaries

A. Cone Decomposition Approach

A *utopia point* and a *nadir point* are acquired to divide the objective space. Given a decision space Ω for a BOP, the utopia point and the nadir point over Ω are determined as $\mathbf{f}^{\triangle}(\Omega) = (f_1^{\triangle}(\Omega), f_2^{\triangle}(\Omega))$ and $\mathbf{f}^{\bigtriangledown}(\Omega) = (f_1^{\bigtriangledown}(\Omega), f_2^{\bigtriangledown}(\Omega))$ respectively, where $\mathbf{f}_i^{\triangle}(\Omega) = \min_{\mathbf{x}\in\Omega} f_i(\mathbf{x})$ and $f_i^{\bigtriangledown}(\Omega) = \max_{\mathbf{x}\in\Omega} f_i(\mathbf{x}), i = 1, 2$. For the sake of brevity, the utopian point in objective space can be easily transformed as the origin of a new coordinate system for any objective vector by following $\mathbf{y}' = \mathbf{y} - \mathbf{f}^{\triangle}(\Omega)$. And then all objective vectors mentioned later in this paper are transformed objective vectors without loss of generality. Thereafter, two endpoints of the Pareto frontier, called *anchor points*, fall on two objective axes of the new coordinate system in the objective space, as shown in Fig.1(a).

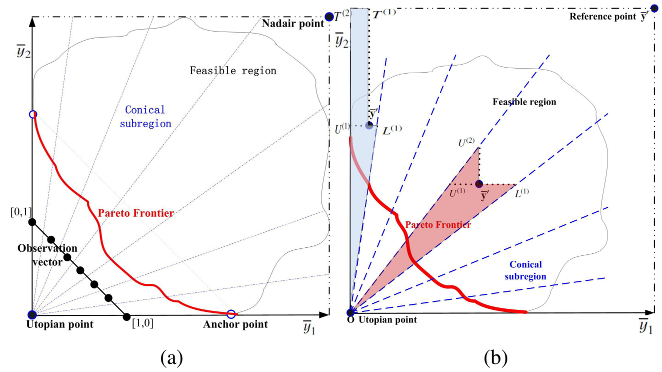

Figure 1. Conical partition (a) and conical area (b).

In addition, the *observation vector* $\mathbf{V}(\mathbf{y})$ of any transformed objective vector $\mathbf{y} = (y_1, y_2)$ is also required for division of the bi-objective space, which is defined as: $\mathbf{V}(\mathbf{y}) = (v_1(\mathbf{y}), v_2(\mathbf{y}))$, where $v_i(\mathbf{y}) = \frac{y_i}{y_1+y_2}$, $i = 1, 2$. Any observation vector defined above has the following characteristics: $v_i(\mathbf{y}) \geq 0$ and $v_1(\mathbf{y}) + v_2(\mathbf{y}) = 1$. As shown in Fig. 1(a), all observation vectors fall on a particular line segment between $(0, 1)$ and $(1, 0)$. Given a division number N, a series of uniformly distributed central observation vectors $\mathbf{V}^{(k)}$ can be defined as $\mathbf{V}^{(k)} = (\frac{k}{N-1}, 1 - \frac{k}{N-1}), k = 0, ..., N-1$.

Furthermore, the region $C = \{\mathbf{y} = (y_1, y_2)|y_1 \geq 0 \wedge y_2 \geq 0\}$ can be divided into N conical sub-regions $C^{(k)} = \{\mathbf{y} \in C|\frac{k-0.5}{N-1} \leq v_1(\mathbf{y}) < \frac{k+0.5}{N-1}\}$, $k = 0, ..., N-1$, where k is the index of sub-region $C^{(k)}$. Correspondingly, the k-th decision sub-set of the decision space Ω is defined as $\Omega^{(k)} = \{\mathbf{x} \in \Omega|\mathbf{f}(\mathbf{x}) - \mathbf{f}^{\triangle}(\Omega) \in C^{(k)}\}$. The upper boundary $L^{k,\triangleright}$ and the lower boundary $L^{k,\triangleleft}$ of the k-th conical sub-region $C^{(k)}$ can be represented as $L^{k,\triangleright} = \{\mathbf{y} \in C|v_1(\mathbf{y}) = max(0, \frac{k-0.5}{N-1})\}$ and $L^{k,\triangleleft} = \{\mathbf{y} \in C|v_1(\mathbf{y}) = min(\frac{k+0.5}{N-1}, 1)\}$ respectively. Each central observation vector $V^{(k)}$ is associated with a conical sub-region in which the observation vector of any objective vector is closer to $V^{(k)}$ than to the other central observation vectors. Specifically, the index of the sub-region in which a solution $\mathbf{x} \in \Omega$ actually lies is $k = \lfloor \frac{(N-1)(f_1(x) - f_1^{\triangle}(\Omega))}{f_1(x) - f_1^{\triangle}(\Omega) + f_2(x) - f_2^{\triangle}(\Omega)} + \frac{1}{2} \rfloor$ and $\lfloor . \rfloor$ is the bottom integral function.

Although the hypervolume indicator processes the highly desirable property in measuring the quality of the Pareto solution set, it is of high computational complexity [10]. Based on the cone decomposition strategy and the idea of hypervolume, an efficient conical area indicator [9] has been designed for bi-objective optimization. Let $\mathbf{y}' \in C^{(k)}, 0 \leq k \leq N-1$, the conical area for $\mathbf{y}'$ in $C^{(k)}$ is the area of the portion $\tilde{C}(\mathbf{y}') = \{\mathbf{y} \in C^{(k)}|\neg(\mathbf{y}' \prec \mathbf{y}) \wedge \mathbf{y} \prec \mathbf{y}^r\}$, written as $s_k(\mathbf{y}')$, where $\mathbf{y}^r$ is the reference point, which should be an approximate infinity point dominated by all feasible solutions. As a substitute for hypervolume, the conical area can be figured out by only one individual, which avoids the high computational complexity of hypervolume of the whole population. Combining the cone decomposition strategy with the conical area indicator, the kth conical subproblem $g^{cone,k}(\mathbf{x})$ can be described in the following form:

$$\begin{aligned} minimize \quad & g^{cone,k}(\mathbf{x}) = s_k(\mathbf{f}(\mathbf{x}) - \mathbf{f}^{\triangle}(\Omega)) \\ subject \quad to \quad & \mathbf{x} \in \Omega^{(k)}. \end{aligned} \tag{2}$$

B. Local Pareto Optimality of Cone Decomposition

Two desirable properties of cone decomposition have been proved in [9]. Firstly, a solution which is dominated by another solution in the same conical subregion, must have a larger conical area value, as described in Lemma 1. Afterwards, Theorem 1 concludes the local Pareto optimality of cone decomposition. The local Pareto optimality guarantees that the optimal solution of each conical subproblem, which is the one with the minimal conical area, must be Pareto non-dominated in the associated sub-region.

Lemma 1. Let $\mathbf{y}' = ({y_1}', {y_2}') \in C^{(k)}$ and $\mathbf{y}'' = ({y_1}'', {y_2}'') \in C^{(k)}$, $0 \le k \le N-1$, if $\mathbf{y}' \prec \mathbf{y}''$, then $s_k(\mathbf{y}') < s_k(\mathbf{y}'')$.

Theorem 1. *If* $\boldsymbol{x}^* \in \Omega^{(k)}$ *and* $\forall \boldsymbol{x} \in \Omega^{(k)} : s_k(\mathbf{f}(\boldsymbol{x}^*) - \mathbf{f}^{\triangle}(\Omega)) \le s_k(\mathbf{f}(\boldsymbol{x}) - \mathbf{f}^{\triangle}(\Omega))$, *then* $\neg\exists \boldsymbol{x} \in \Omega^{(k)} : \boldsymbol{x} \prec \boldsymbol{x}^*$.

III. Global Pareto Optimality

The local Pareto optimality of cone decomposition doesn't guarantee that the optimal solution of a conical subproblem remains Pareto non-dominated in the entire bi-objective space. Therefore, whether the cone decomposition approach possesses the better property of global Pareto optimality is critical to the quality of frontiers obtained by this approach. This section firstly gives the function describing the Pareto frontier and then proves the global Pareto optimality of the cone decomposition approach under certain continuity assumptions.

A. Function Representation of the Pareto Frontier

In the bi-objective case, the true Pareto frontier can be described in terms of a function $\mathcal{F}$ mapping the image of the Pareto set under f_1 onto the image of the Pareto set under f_2. At first, let a real set $Y_1^\star \in \mathbb{R}$ denote the image of the Pareto set under f_1. Then, the true Pareto frontier can be represented in terms of the function:

$$\mathcal{F} : y_1^\star \in Y_1^\star \mapsto y_2^\star. \tag{3}$$

It is worth noting that, for any objective vector $\mathbf{y}^\star$, $\mathbf{y}^\star \in \mathcal{F}$ implies that

$$\exists \mathbf{x}^\star \in \Omega : \mathbf{y}^\star = \mathbf{f}(\mathbf{x}^\star) \wedge \neg\exists \mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) \prec \mathbf{y}^\star. \tag{4}$$

Furthermore, it can be concluded that, since $\mathcal{F}$ represents the shape of the Pareto frontier, the function $\mathcal{F}$ must be strictly monotonically decreasing in $Y_1^\star$ for BOPs with minimization objectives. That is, if $y_1^{\star,a} \in Y_1^\star$, $y_1^{\star,b} \in Y_1^\star$, and $y_1^{\star,a} < y_1^{\star,b}$, then it holds that $y_2^{\star,a} = \mathcal{F}(y_1^{\star,a}) > \mathcal{F}(y_1^{\star,b}) = y_2^{\star,b}$. Otherwise, it can be reached from $y_1^{\star,a} < y_1^{\star,b}$ and $y_2^{\star,a} \le y_2^{\star,b}$ that $\exists \mathbf{x}^{\star,a} \in \Omega : \mathbf{y}^{\star,a} = \mathbf{f}(\mathbf{x}^{\star,a}) \prec \mathbf{y}^{\star,b}$, which contradicts $\neg\exists \mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) \prec \mathbf{y}^{\star,b}$.

Specifically, both the left extremal point $\mathbf{y}^{\star,\triangleright} = (y_1^{\star,\triangleright}, \mathcal{F}(y_1^{\star,\triangleright}))$ and the right one $\mathbf{y}^{\star,\triangleleft} = (y_1^{\star,\triangleleft}, \mathcal{F}(y_1^{\star,\triangleleft}))$ of the true Pareto frontier $\mathcal{F} = \{\mathbf{y}^\star = (y_1^\star, y_2^\star) \in \mathbb{R}^2 | y_2^\star = \mathcal{F}(y_1^\star) \wedge y_1^\star \in Y_1^\star\}$ are the anchor points, where $y_1^{\star,\triangleright} = \min\{Y_1^\star\} = f_1^{\triangle}(\Omega)$, $y_1^{\star,\triangleleft} = \max\{Y_1^\star\}$, and $\mathcal{F}(y_1^{\star,\triangleleft}) = f_2^{\triangle}(\Omega)$.

Since the utopian point $(y_1^{\star,\triangleright}, \mathcal{F}(y_1^{\star,\triangleleft}))$ of an arbitrary front shape $\mathcal{F}$ can be transformed into the origin $(0,0)$ by a new coordinate system as mentioned above, we may thereby assume without loss of generality that any Pareto frontier shape is described by a more concise function:

$$\mathcal{F} : y_1^\star \in [0, y_1^{\star,\triangleleft}] \mapsto y_2^\star, where \mathcal{F}(y_1^{\star,\triangleleft}) = 0. \tag{5}$$

B. Proof of Global Pareto Optimality

Lemma 2. *Let* $\mathcal{F} : y_1^\star \in [0, y_1^{\star,\triangleleft}] \mapsto y_2^\star$ *with* $\mathcal{F}(y_1^{\star,\triangleleft}) = 0$ *be the Pareto frontier of a BOP bo,* $\boldsymbol{y}^{k,\sharp} = \boldsymbol{f}(\boldsymbol{x}^{k,\sharp}) \in C^{(k)}$ *be the objective vector of the optimal solution* $\boldsymbol{x}^{k,\sharp} \in \Omega^{(k)}$ *of the k-th conical subproblem* $g^{cone,k}(\boldsymbol{x})$ *of cone decomposition for bo, i.e.* $\forall \boldsymbol{x} \in \Omega^{(k)} : s_k(\boldsymbol{y}^{k,\sharp}) \le s_k(\boldsymbol{f}(\boldsymbol{x}))$, $\boldsymbol{y}^{k,\triangleright}$ *and* $\boldsymbol{y}^{k,\triangleleft}$ *are the intersections of the frontier* $\mathcal{F}$*, respectively, with the upper boundary* $L^{k,\triangleright}$ *and the lower boundary* $L^{k,\triangleleft}$ *of the k-th conical sub-region* $C^{(k)}$*, i.e.* $\boldsymbol{y}^{k,\triangleright} = \mathcal{F} \cap L^{k,\triangleright}$ *and* $\boldsymbol{y}^{k,\triangleleft} = \mathcal{F} \cap L^{k,\triangleleft}$*, similarly hereinafter. If the Pareto frontier* $\mathcal{F}$ *is continuous in the conical sub-region* $C^{(k)}$ *as well as its lower boundary, i.e.* $\mathcal{F}(y_1^\star)$ *is continuous in the closed interval* $[y_1^{k,\triangleright}, y_1^{k,\triangleleft}]$*, then*

$$\exists \boldsymbol{x}^{k,\star} \in \Omega : \boldsymbol{y}^{k,\star} = \boldsymbol{f}(\boldsymbol{x}^{k,\star}) \in \mathcal{F} \cap C^{(k)} \wedge v_1(\boldsymbol{y}^{k,\star}) = v_1(\boldsymbol{y}^{k,\sharp}). \tag{6}$$

Proof: Given any objective vector $\mathbf{y}^\star \in \mathcal{F}$ in the Pareto frontier, because $y_2^\star = \mathcal{F}(y_1^\star)$, the first component of the observation vector of $\mathbf{y}^\star \in \mathcal{F}$ can be represented as a composite function $v_1'(y_1^\star)$ of the first objective $y_1^\star$:

$$v_1'(y_1^\star) = v_1(\mathbf{y}^\star) = v_1((y_1^\star, \mathcal{F}(y_1^\star))) = \frac{y_1^\star}{y_1^\star + \mathcal{F}(y_1^\star)}. \tag{7}$$

Since $\mathcal{F}(y_1^\star)$ is continuous in the closed interval $[y_1^{k,\triangleright}, y_1^{k,\triangleleft}]$, the composite function $v_1'(y_1)$ is also continuous in $[y_1^{k,\triangleright}, y_1^{k,\triangleleft}]$ by the continuity of composite functions.

From $\mathbf{y}^{k,\triangleright} = \mathcal{F} \cap L^{k,\triangleright}$ and the definition of $L^{k,\triangleright}$, we know $v_1(\mathbf{y}^{k,\triangleright}) = v_1^{k,\triangleright}$. In a similar way, $v_1(\mathbf{y}^{k,\triangleleft}) = v_1^{k,\triangleleft}$. From $\mathbf{y}^{k,\sharp} \in C^{(k)}$ and the definition of subregion $C^{(k)}$, we also obtain $v_1^{k,\triangleright} \le v_1(\mathbf{y}^{k,\sharp}) < v_1^{k,\triangleleft}$. Afterwards, applying the intermediate value theorem for the continuous function $v_1'(y_1^\star)$ in $[y_1^{k,\triangleright}, y_1^{k,\triangleleft}]$, we get that there exists $y_1^{k,\star} \in [y_1^{k,\triangleright}, y_1^{k,\triangleleft})$ such that

$$v_1'(y_1^{k,\star}) = v_1(\mathbf{y}^{k,\sharp}) \tag{8}$$

from $v_1'(y_1^{k,\triangleright}) = v_1(\mathbf{y}^{k,\triangleright}) = v_1^{k,\triangleright}$, $v_1'(y_1^{k,\triangleleft}) = v_1(\mathbf{y}^{k,\triangleleft}) = v_1^{k,\triangleleft}$, and $v_1^{k,\triangleright} \leq v_1(\mathbf{y}^{k,\sharp}) < v_1^{k,\triangleleft}$. Let $y_2^{k,\star} = \mathcal{F}(y_1^{k,\star})$. Then, we have $\mathbf{y}^{k,\star} = (y_1^{k,\star}, y_2^{k,\star}) \in \mathcal{F}$, which implies that

$$\exists \mathbf{x}^{k,\star} \in \Omega : \mathbf{y}^{k,\star} = \mathbf{f}(\mathbf{x}^{k,\star}) \in \mathcal{F}. \tag{9}$$

Hence, by integrating Eq. 8, Eq. 9, and $v_1'(y_1^{k,\star}) = v_1(\mathbf{y}^{k,\star})$, we obtain Eq. 6, which concludes Lemma 2. ■

Theorem 2. *Let $\mathcal{F} : y_1^\star \in [0, y_1^{\star,\triangleleft}] \mapsto y_2^\star$ with $\mathcal{F}(y_1^{\star,\triangleleft}) = 0$ be the Pareto frontier of a BOP bo, $\mathbf{y}^{k,\sharp} = \boldsymbol{f}(\mathbf{x}^{k,\sharp}) \in C^{(k)}$ be the objective vector of the optimal solution $\boldsymbol{x}^{k,\sharp} \in \Omega^{(k)}$ of the k-th conical subproblem $g^{cone,k}(\mathbf{x})$ of cone decomposition for bo. If the Pareto frontier $\mathcal{F}$ is continuous in the conical sub-region $C^{(k)}$ as well as its lower boundary, then $\mathbf{y}^{k,\sharp}$ is Pareto optimal, i.e.*

$$\neg \exists \boldsymbol{x} \in \Omega : \boldsymbol{f}(\boldsymbol{x}) \prec \mathbf{y}^{k,\sharp}. \tag{10}$$

Proof: We first assume $\mathbf{y}^{k,\sharp} \notin \mathcal{F}$ and derive a contradiction.

By using Lemma 2, we have Eq. 6. In this situation, we are going to prove that $\mathbf{y}^{k,\star} \prec \mathbf{y}^{k,\sharp}$ for $\mathbf{y}^{k,\star}$ in Eq. 6. From Eq. 6, we know $v_1(\mathbf{y}^{k,\star}) = v_1(\mathbf{y}^{k,\sharp})$, thereby

$$\frac{y_1^{k,\star}}{y_1^{k,\star} + y_2^{k,\star}} = \frac{y_1^{k,\sharp}}{y_1^{k,\sharp} + y_2^{k,\sharp}} \tag{11}$$

where $y_1^{k,\sharp} \geq 0$. Furthermore, two different cases, $y_1^{k,\sharp} = 0$ and $y_1^{k,\sharp} > 0$, are respectively considered. In case $y_1^{k,\sharp} = 0$, $y_1^{k,\star} = 0$. Moreover, it can be inferred that $y_2^{k,\star} < y_2^{k,\sharp}$. Otherwise if $y_2^{k,\star} \geq y_2^{k,\sharp}$, then either $\mathbf{y}^{k,\sharp} = \mathbf{y}^{k,\star}$ which contradicts $\mathbf{y}^{k,\star} \in \mathcal{F} \wedge \mathbf{y}^{k,\sharp} \notin \mathcal{F}$, or $\mathbf{y}^{k,\sharp} = \mathbf{f}(\mathbf{x}^{k,\sharp}) \prec \mathbf{y}^{k,\star}$ which contradicts $\neg \exists \mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) \prec \mathbf{y}^{k,\star}$. Thus, we obtain $\mathbf{y}^{k,\star} \prec \mathbf{y}^{k,\sharp}$ in case $y_1^{k,\sharp} = 0$. In case $y_1^{k,\sharp} > 0$, we have $\frac{y_1^{k,\star}}{y_1^{k,\sharp}} = \frac{y_1^{k,\star} + y_2^{k,\star}}{y_1^{k,\sharp} + y_2^{k,\sharp}} = t$ from Eq. 11, thereby

$$y_1^{k,\star} = t y_1^{k,\sharp} \wedge y_2^{k,\star} = t y_2^{k,\sharp} \tag{12}$$

where $t > 0$. Moreover, it can be inferred that $t < 1$. Otherwise if $t \geq 1$, then either $t = 1$ and $\mathbf{y}^{k,\sharp} = \mathbf{y}^{k,\star}$ which contradicts $\mathbf{y}^{k,\star} \in \mathcal{F} \wedge \mathbf{y}^{k,\sharp} \notin \mathcal{F}$, or $t > 1$ and $\mathbf{y}^{k,\sharp} = \mathbf{f}(\mathbf{x}^{k,\sharp}) \prec \mathbf{y}^{k,\star}$ which contradicts $\neg \exists \mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) \prec \mathbf{y}^{k,\star}$. Thus, from Eq. 12 and $t < 1$, it can be reached that $y_1^{k,\star} < y_1^{k,\sharp}$ and $y_2^{k,\star} < y_2^{k,\sharp}$, thereby $\mathbf{y}^{k,\star} \prec \mathbf{y}^{k,\sharp}$ in case $y_1^{k,\sharp} > 0$. To sum up both cases, we have $\mathbf{y}^{k,\star} \prec \mathbf{y}^{k,\sharp}$.

From $v_1(\mathbf{y}^{k,\star}) = v_1(\mathbf{y}^{k,\sharp})$ and $\mathbf{y}^{k,\sharp} \in C^{(k)}$, it can be inferred that $\mathbf{y}^{k,\star} = \mathbf{f}(\mathbf{x}^{k,\star}) \in C^{(k)}$ and $\mathbf{x}^{k,\star} \in \Omega^{(k)}$. Applying Lemma 1, we obtain from $\mathbf{f}(\mathbf{x}^{k,\star}) \in C^{(k)}$, $\mathbf{y}^{k,\sharp} \in C^{(k)}$ and $\mathbf{f}(\mathbf{x}^{k,\star}) \prec \mathbf{y}^{k,\sharp}$ that $\exists \mathbf{x}^{k,\star} \in \Omega^{(k)} : s_k(f(\mathbf{x}^{k,\star})) < s_k(\mathbf{y}^{k,\sharp})$, which contradicts the premise $\forall \mathbf{x} \in \Omega^{(k)} : s_k(\mathbf{y}^{k,\sharp}) \leq s_k(f(\mathbf{x}))$. Therefore, the assumption $\mathbf{y}^{k,\sharp} \notin \mathcal{F}$ must be false and the opposite $\mathbf{y}^{k,\sharp} \in \mathcal{F}$ holds. According to the definition of the Pareto frontier $\mathcal{F}$, Eq. 10 is obtained and Theorem 1 is concluded. ■

According to Theorem 2, under the assumption that the Pareto frontier $\mathcal{F}$ is continuous in the k-th conical subregion $C^{(k)}$ as well as its lower boundary, the conical subproblem $g^{cone,k}(\mathbf{x})$ not only has the local Pareto optimality, but also possesses the better property of global Pareto optimality, which guarantees that the optimal solution of the conical subproblem remains Pareto non-dominated in the entire bi-objective space. In most situations, the Pareto frontier of a continuous BOP consists of piecewise continuous segments, which guarantees that the continuity assumption in the majority of subregions are true.

IV. Empirical results and discussion

A. Experimental Setup

In order to empirically validate the global Pareto optimality of the cone decomposition approach, the optimal solution of each conical subproblem should be approximately found. CAEA constructs the Pareto frontier by solving multiple conical sub-problems simultaneously and collaboratively. In view of this consideration, CAEA is natively capable of being viewed as a method to solve all conical sub-problems. For the sub-problem $g^{cone,k}(\mathbf{x})$, CAEA maintains only one best individual in subregion $C^{(k)}$ responsible for the optimization of $g^{cone,k}(\mathbf{x})$ and produces only one offspring at each generation. Since each offspring lies in the conical sub-region of only one sub-problem, there is only one individual associated with this sub-problem need to be updated. Importantly, CAEA both adopts the cone decomposition strategy to maintain the diversity of distribution and keeps the Pareto optimality of individuals in the population through the minimization of conical area.

In this section, CAEA is compared against the other four existing popular MOEAs, namely, NSGA-II, NSGA-III, MOEA/D and MOEAD-DE, on 5 bi-objective instances used in [11]. The population size N is set as 100, and the stopping criterion is after 3×10^5 function evaluations. The distribution indexes in simulated binary crossover (SBX) and polynomial mutation are respectively set to be 15 and 20. The crossover rate and mutation rate are, respectively, 1.00 and *1/n*, where *n* denotes the number of decision variables. Besides, the scaling factor (F) and the crossover control parameter (C_r) of differential evolution (DE) are 0.5 and 1. The size T of the neighborhood of each weight vector in MOEA/D or MOEAD-DE is set to 20, which is the same as in [5], [7]. To assess the overall performance, a total of 30 statistically independent runs of each algorithm have been conducted for each test instance.

B. Performance Metrics

At first, the *set coverage (C-metric)* [12] is adopted to evaluate the Pareto optimality of obtained optimal solutions of conical subproblems by comparative analysis in our experiments. The C-metric is defined as

$$C(X, X') = \frac{|\, \{a' \in X' | \exists a \in X : a \prec a'\} \,|}{|\, X' \,|},$$

Table I
C-METRIC VALUES ACHIEVED BY CAEA IN COMPARISON WITH THE OTHER FOUR MOEAS ON 5 BENCHMARK TEST INSTANCES

C	MOP1	MOP2	MOP3	MOP4	MOP5
C(CAEA,NSGA-II)	30.0%	42.7%	35.1%	36.9%	48.7%
C(NSGA-II,CAEA)	5.9%	4.0%	3.9%	3.2%	2.8%
C(CAEA,NSGA-III)	42.1%	21.1%	28.0%	39.6%	37.6%
C(NSGA-III,CAEA)	6.2%	3.9%	3.0%	3.5%	2.9%
C(CAEA,MOEA/D)	38.0%	29.5%	33.0%	35.8%	43.2%
C(MOEA/D,CAEA)	0%	10.4%	9.6%	8.3%	6.7%
C(CAEA,MOEAD-DE)	46.9%	23.5%	37.1%	47.1%	51.4%
C(MOEAD-DE,CAEA)	0%	0.8%	1.6%	4.2%	3.6%

where X and X' are two solution sets, and $C(X, X')$ means the ratio of the individuals in X' dominated by X. To be specific, the C-metric appraises the relative extent of the solutions approaching the PF. Since $C(X, X')$ is not necessarily equal to $1 - C(X', X)$, both $C(X, X')$ and $C(X', X)$ have to be considered.The C-metric is Pareto compliant and provides a relative comparison based upon dominance numbers between MOEAs. If $C(X, X')$ is found to be significantly higher than $C(X', X)$ over many trials, we may argue that X has the better Pareto optimality than the solution set X'.

As a perform indicator, the inverted generational distance(IGD) has frequently adopted to measure how far the elements in the PF are from those in the front obtained by the algorithms. Considering the Pareto dominance relation between a solution and a reference point when their distance is calculated, IGD$^+$[13] was proposed and it can be calculated as (13):

$$IGD^+(A) = \frac{1}{|Z|}\sum_{j=1}^{|Z|}\min_{\mathbf{x}_i \in A} d^+(\mathbf{F}(\mathbf{x}_i), \mathbf{z}_j)$$

$$d^+(\mathbf{F}(\mathbf{x}), \mathbf{z}) = \sqrt{\sum_{k=1}^{m}(\max\{f_k(\mathbf{x}) - z_k, 0\})^2} \tag{13}$$

in which $A = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{|A|}\}$ is the non-dominated solution sets for evaluating and Z is the set of uniformly distributed points along the PF.

C. Performance Comparisons in Terms of Pareto Dominance

In order to empirically validate the global Pareto optimality of cone decomposition, the set of solutions found by CAEA is also regarded as the approximate optima of all conical subproblems. Table I lists the C-metric values achieved by CAEA in comparison with the other four MOEAs on 5 benchmark test instances. All the C-metric values in Table I are averaged over 30 trials. Table I clearly indicates that the outcomes of CAEA dominates $21.1\% \sim 51.4\%$ of those of the other four MOEAs, while the other four MOEAs achieves only $0\% \sim 10.4\%$ coverage regarding CAEA. It implies the solution sets found by CAEA remarkably dominate more individuals in the solution sets by all the other four algorithms on all 5 benchmark test instances.

In addition, it is worth noting that MOP_3 has a disconnected Pareto frontier. Nevertheless, CAEA still achieves obviously higher C-metric values ($28.0\% \sim 37.1\%$ versus $1.6\% \sim 9.6\%$) in comparison with the other four MOEAs. This is due to the present of continuous frontier segment within most subregions caused by the disconnected and piecewise continuous frontier. In conclusion, the optimal solutions accquired by CAEA have a higher coverage based on Pareto dominance compared with the other four algorithms, which reveals that the cone decomposition approach used by CAEA has a good character of global Pareto optimality.

D. Performance Comparisons in Terms of IGD$^+$

Considering that the above set coverage metric only evaluates the convergence of obtained frontiers based on the Pareto dominance, it is necessary to further assess the overall performance on both convergence and spread of obtained frontiers according to the IGD$^+$ metric. Fig. 2 presents the convergence curves of average IGD$^+$ values over 30 runs of each algorithm versus the number of function evaluations on each of 5 benchmark test instances.

It is evident from Fig. 2 that CAEA achieves both significantly lower and more stable IGD$^+$ values at the maximal evolutionary generation on all 5 test instances. The other four MOEAs obtain very similar IGD$^+$ scores on MOP_{1-5}. Not only the global Pareto optimality of the cone decomposition approach guarantees the convergence ability of CAEA, but also the uniformly distributed subregions effectively maintain the diversity of population and the spread along the frontier. As a consequence, although CAEA maybe converges slower at the very early stage of about 50000 function evalutions, it always converges fastest and gains the lowest IGD$^+$ values at the middle and terminal stages in comparison with NSGA-II, NSGA-III, MOEA/D and MOEA/D-DE in our experiments. To sum up, CAEA acquires both the significantly better Pareto optimality and obviously higher qualities of frontiers than the other four state-of-the-art algorithms on solving bi-objective problems.

V. Conclusion

In this paper, we prove the global Pareto optimality of conical subproblems in such a case that the front segments are continuous in the corresponding conical sub-regions. Moreover, the global optimality of cone decomposition is validated and the performance of CAEA is evaluated on 5 benchmark bi-objective problems. Experimental results show that CAEA acquires better Pareto optimality and higher qualities of frontiers in terms of IGD+ than the other four comparing algorithms on solving bi-objective problems by

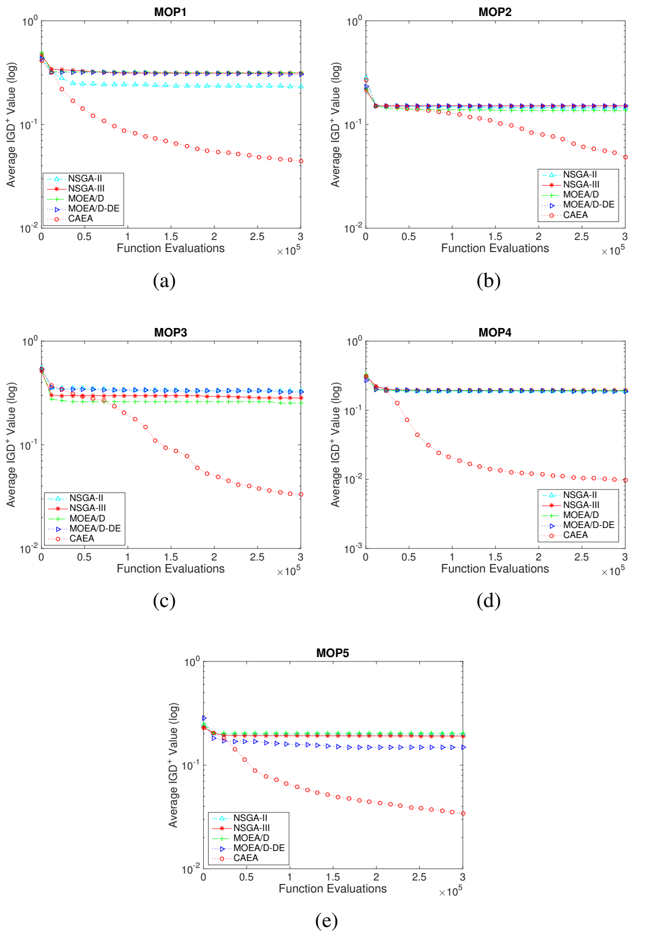

Figure 2. Convergence curves of average IGD+ values over 30 runs of each algorithm versus the number of function evaluations on each of 5 benchmark test instances. (a) MOP_1. (b) MOP_2. (c) MOP_3. (d) MOP_4. (e) MOP_5.

the use of the cone decomposition approach. Our future research will focus on designing and analyzing the cone decomposition approaches in three or higher objective spaces.

ACKNOWLEDGMENT

This work was supported partly by the Natural Science Foundation of Guangdong Province, China (No.2015A030313204), the Pearl River S&T Nova Program of Guangzhou (No.2014J2200052), the National Natural Science Foundation of China (No.61203310, 61503087), the Fundamental Research Funds for the Central Universities, SCUT (No.2017MS043), and the China Scholarship Council (CSC) (No.201406155076, 201408440193).

REFERENCES

[1] C. Liu, J. Liu, and Z. Jiang, "A multiobjective evolutionary algorithm based on similarity for community detection from signed social networks," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2274–2287, 2014.

[2] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, 2014.

[3] K. Deb, "Multiobjective optimization using evolutionary algorithms," *Computational Optimization & Applications*, vol. 39, no. 1, pp. 75–96, 2016.

[4] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002.

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

[6] L. Wang, Q. Zhang, A. Zhou, M. Gong, and L. Jiao, "Constrained subproblems in a decomposition-based multiobjective evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 475–480, 2016.

[7] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 284–302, 2009.

[8] Z. Wang, Q. Zhang, A. Zhou, M. Gong, and L. Jiao, "Adaptive replacement strategies for MOEA/D," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 474–486, 2016.

[9] W. Ying, X. Xu, Y. Feng, and Y. Wu, "An efficient conical area evolutionary algorithm for bi-objective optimization," *IEICE Trans. Fund. Elec. Comm. Comp. Sci.*, vol. E95A, no. 8, pp. 1420–1425, 2012.

[10] M. Emmerich, N. Beume, and B. Naujoks, "An EMO algorithm using the hypervolume measure as selection criterion," in *Proc. Evol. Multi-Criterion Opt.*, 2005, pp. 62–76.

[11] H. L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 450–455, 2014.

[12] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, 1999.

[13] H. Ishibuchi, H. Masuda, and Y. Nojima, "A study on performance evaluation ability of a modified inverted generational distance indicator," in *Proc. Annual Conference on Genetic and Evolutionary Computation*. ACM, 2015, pp. 695–702.