J. Nonlinear Var. Anal. 8 (2024), No. 1, pp. 125-142
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.8.2024.1.07

A MODIFICATION PIECEWISE CONVEXIFICATION METHOD WITH A CLASSIFICATION STRATEGY FOR BOX-CONSTRAINED NON-CONVEX OPTIMIZATION PROGRAMS

QIAO ZHU[1,2], LIPING TANG[1,*], XINMIN YANG[1]

[1]*National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China*
[2]*College of Mathematics, Sichuan University, Chengdu 610065, China*

Abstract. This paper presents a piecewise convexification method with a box classification strategy to approximate the entire globally optimal solution set of non-convex optimization problems with box constraints. First, the box classification strategy is proposed based on the convexity of the objective function on the sub-boxes, which helps to reduce the number of box divisions and improve the computational efficiency. At the same time, we construct the piecewise convexification problem of the original non-convex optimization problem by applying the α-based Branch-and-Bound (αBB) method, and we define the (approximate) solution set of the piecewise convexification problem based on the result of classifying the sub-boxes. Then, it is deduced that the globally optimal solution set can be approximated by the (approximate) solution set of the piecewise convexification problem. Finally, a piecewise convexification algorithm is proposed that includes a new subset selection technique for division and two new termination tests. The results of our experiments demonstrate the effectiveness and general superiority of our approach over the competition.

Keywords. α-based Branch-and-Bound; Global optimization; Non-convex programming; Optimal solution set; Piecewise convexification.

1. INTRODUCTION

Non-convex optimization problems arise frequently in machine learning, such as the image recovery problem [1, 2] and the robust support vector regression problem [3, 4]. Meanwhile, one is generally only interested in globally (approximate) optimal solutions in applications, so how to effectively solve non-convex optimization problems has received much attention. Up to now, most global approximation algorithms have been developed such as convex relaxation-based method; see [5, 6, 7, 8, 9, 10, 11]. In particular, the αBB method, which is one of the global optimization methods and is based on the idea of the convex relaxation, plays a substantial role in the design of efficient and computationally tractable numerical algorithms for non-convex optimization problems; see, e.g., [12, 13, 14, 15]. It is worth remarking that these algorithms

*Corresponding author.
E-mail address: math_qiaozhu@163.com (Q. Zhu), tanglipings@163.com (L.P. Tang), xmyang@cqnu.edu.cn (X. Yang).

Received 21 March 2023; Accepted 24 May 2023; Published online 13 December 2023

generally aim at finding a single globally optimal solution, but most application problems may exist with many or even an infinite number of globally optimal solutions; see, e.g., [16, 17]. To the best of our knowledge, research on global algorithms to determine the set of globally optimal solutions is not abundant. In [18], Eichfelder et al. generalized the classical αBB method to find the globally optimal solution set with predefined quality for a non-convex optimization problem. However, as they pointed out, some additional variables and the additional while loop were required.

Motivated by [18], we develop a piecewise convexification method with a box classification strategy to approximate the entire globally optimal solution set of non-convex optimization problems. In order to reduce the number of box divisions and improve the computational efficiency, we first classify the sub-boxes by the convexity of the objective function on them and further partition only on some sub-boxes in the subsequent division. The αBB method is applied to some sub-boxes to construct a piecewise convexification problem of the original non-convex optimization problem. Then, we construct the (approximate) solution sets of the piecewise convexification problem using the results of the box classification, and we show that these constructed sets approximate the globally optimal solution set with a predefined quality. Finally, a new piecewise convexification algorithm is proposed that includes a new sub-box selection rule for partitioning and two new termination rules. Furthermore, several instances verify that these rules are conducive to improving the effectiveness of the algorithm.

This paper is organized as follows. Section 2 summarizes some basic definitions of the optimization problem. It also introduces the αBB method and the box division. In Section 3, incorporating the αBB method and the box division criterion, we first propose a piecewise convexification method for the non-convex optimization problem, and then analyze the solution set of this piecewise convexification optimization problem. More importantly, some relationships between the (approximate) optimal solution set of the convexification problem and the original optimization problem are also stated in detail. A new algorithm that generates the subset of approximate global solutions is presented in Section 4. Finally, some numerical experiments are reported and discussed in Section 5, the last section.

2. Preliminaries

Let $\mathbb{I}$ denote the set of all real nonempty closed boxes and $\mathbb{I}^n$ denote the set of all n-dimensional boxes. For a given box $X \in \mathbb{I}^n$, we set $X = [a,b] := \prod_{i=1}^{n} [a_i, b_i]$, where $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$. Thus, $x \in X$ denotes $x_i \in [a_i, b_i]$ for each $i \in \{1, \cdots, n\}$.

In this paper, we consider the following non-convex optimization problem (NCOP):

$$\text{(NCOP)} \ \min_{x \in X} f(x),$$

where $X = [a,b]$ is a box and $f : \mathbb{R}^n \to \mathbb{R}$ is a non-convex twice continuously differentiable function. We start with an overview of the (approximate) optimal solution of (NCOP).

Definition 2.1. ([18]) Let $X \subseteq \mathbb{R}^n$ be a nonempty set, $\widetilde{x} \in X$, and $f : X \to \mathbb{R}$.

(a) If $f(\widetilde{x}) \leq f(y)$ for any $y \in X$, then $\widetilde{x}$ is an optimal solution to f w.r.t. X. Therefore, the optimal solution set is denoted by $X_{op} := \{x \in X : f(x) \leq f(y), \ \forall y \in X\}$.

(b) For any given $\varepsilon > 0$, if $f(\widetilde{x}) \leq f(y) + \varepsilon$ for any $y \in X$, then $\widetilde{x}$ is a ε-minimal point of f w.r.t. X, and $X_{op}^{\varepsilon} := \{x \in X : f(x) \leq f(y) + \varepsilon, \forall y \in X\}$ is the ε-optimal solution set.

Before proceeding, we briefly describe the αBB method and the box division, as they play an important role in the devise of the piecewise convexification method.

2.1. **The αBB method.** The αBB method is a global optimization method that constructs a convex relaxation estimator of a non-convex objective function f w.r.t. X; see, e.g., [18, 19, 20, 21]. More precisely, let $f : X \to \mathbb{R}$ be a real-valued twice continuously differentiable function and X be a box, i.e., $X = [a,b] \in \mathbb{I}^n$. A convex lower relaxation function $F_X^\alpha : X \to \mathbb{R}$ of f by the idea of the αBB method was defined in [20] as follows:

$$F_X^\alpha(x) = f(x) + \sum_{i=1}^{n} \alpha_i (a_i - x_i)(b_i - x_i),$$

where parameter $\alpha := (\alpha_1, \cdots, \alpha_n)$ guarantees the convexity of F_X^α on X.

There are several methods for estimating the value of α; see, e.g., [20, 22, 23]. In this article, we directly adopt the following method from [20] to roughly compute the value of α_i for each $i \in \{1, \cdots, n\}$, as defined by

$$\alpha_i := \max\left\{0, -\frac{1}{2}\left(\min_{x\in X} \nabla^2 f(x)_{ii} - \sum(X,i,d)\right)\right\} \tag{2.1}$$

where $d := b - a \in \mathbb{R}^n$, Hessian matrix $\nabla^2 f(x) = \left(\nabla^2 f(x)_{ij}\right)$, and

$$\sum(X,i,d) := \sum_{i\neq j} \max\left\{\left|\min_{x\in X} \nabla^2 f(x)_{ij}\right|, \left|\max_{x\in X} \nabla^2 f(x)_{ij}\right|\right\} \frac{d_j}{d_i}. \tag{2.2}$$

A lower bound of the minimum eigenvalue of $\nabla^2 f(x)$ w.r.t. X was given in [20], i.e.,

$$\lambda_{\min}^X \geq \min_i \left(\min_{x\in X} \nabla^2 f(x)_{ii} - \sum(X,i,e)\right),$$

where $e = (1, \cdots, 1) \in \mathbb{R}^n$. Obviously, if the lower bound

$$\tilde{\lambda}_X := \min_i \left(\min_{x\in X} \nabla^2 f(x)_{ii} - \sum(X,i,e)\right) \geq 0, \tag{2.3}$$

then f is obviously convex on X, not vice versa. It is clear that $\tilde{\lambda}_{X_2} \leq \tilde{\lambda}_{X_1}$ with $X_1 \subseteq X_2$.

Moreover, in [22], the maximum separation distance between f and F_X^α over X is

$$D(X) = \max_{x\in X} \|f(x) - F_X^\alpha(x)\| = \sum_{i=1}^{n} \alpha_i \left(\frac{b_i - a_i}{2}\right)^2, \tag{2.4}$$

which shows that $D(X)$ is determined by the interval $[a,b]$ and α.

2.2. **The box division.** As shown in equation (2.4), a smaller box helps to generate a tighter underestimation of the original function. Thus, we divide the whole box into some sub-boxes to well approximate the original function.

Let μ^n be the Lebesgue measure on $\mathbb{R}^n$. If the box set $\mathbb{Y}^t := \{Y^1, Y^2, \cdots, Y^{M_t}\}$ satisfies

$$X = \bigcup_{k_t=1}^{M_t} Y^{k_t} \text{ and } \mu^n(Y^{k_t} \cap Y^{j_t}) = 0 \; \forall k_t, j_t \in \{1, \cdots, M_t\},$$

then $\mathbb{Y}^t$ is a subdivision of X. The sub-box Y^{k_t} abbreviates as $Y^{k_t}=[a^{k_t},b^{k_t}]=\prod_{i=1}^{n}[a_i^{k_t},b_i^{k_t}]$. It is worth noting that the subdivision $\mathbb{Y}^{t+1}$ of X w.r.t $t+1$ is based on $\mathbb{Y}^t$, i.e., selecting one or more sub-boxes from $\mathbb{Y}^t$ to divide. In the following, we introduce the box division way; see [18].

The division way: *For a given box $Y^{k_t}=[a^{k_t},b^{k_t}]\in\mathbb{Y}^t$, the branching index l is defined by*

$$l:=\min\{j\in\{1,\cdots,n\}:j\in\operatorname*{arg\,max}_{j\in\{1,\cdots,n\}}(b_j^{k_t}-a_j^{k_t})\},$$

and then Y^{k_t} splits into two subsets $Y^{k_t,1}$ and $Y^{k_t,2}$ based on direction l by

$$Y^{k_t,1}:=\prod_{i=1,i\neq l}^{n}[a_i^{k_t},b_i^{k_t}]\times\left[a_l^{k_t},\frac{a_l^{k_t}+b_l^{k_t}}{2}\right]\quad\textit{and}\quad Y^{k_t,2}:=\prod_{i=1,i\neq l}^{n}[a_i^{k_t},b_i^{k_t}]\times\left[\frac{a_l^{k_t}+b_l^{k_t}}{2},b_l^{k_t}\right].$$

Clearly, $Y^{k_t}=Y^{k_t,1}\cup Y^{k_t,2}$ and $Y^{k_t,1},Y^{k_t,2}\in\mathbb{Y}^{t+1}$. For simplicity, we define the splitting operator $\mathrm{Sp}(Y^{k_t}):=\{Y^{k_t,1},Y^{k_t,2}\}$ and the length of the subdivision $\mathbb{Y}^t$ of X by

$$|T(\mathbb{Y}^t)|:=\max_{k_t\in\{1,\cdots,M_t\}}\left\{\|a^{k_t}-b^{k_t}\|_2^2\right\}=\max_{k_t\in\{1,\cdots,M_t\}}\left\{\sum_{i=1}^{n}(b_i^{k_t}-a_i^{k_t})^2\right\}.$$

Remark 2.1. Let $x^*\in X_{op}$. For any given subdivision $\mathbb{Y}^t$ of X, there exists $Y^{k_t}\in\mathbb{Y}^t$ such that $x^*\in Y^{k_t}$ and x^* is an optimal solution to f w.r.t. Y^{k_t}.

3. Piecewise Convexification Method for (NCOP)

In this section, we first introduce the piecewise convexification problem for the non-convex optimization problem (PC-NCOP). The solution sets of the piecewise convexification optimization problem are constructed and discussed in detail. Finally, we analyze some relationships between the solution sets of the piecewise convexification optimization problem and the (approximate) globally optimal solution set of the original non-convex optimization problem.

3.1. Piecewise convexification problem.

In order to approximate the globally optimal solution set of a non-convex optimization problem, we use the box division technique to divide X into several sub-boxes and use the αBB method to relax this problem on each sub-box of X instead of on X itself. Thus it is referred to as the piecewise convexification method. In the following, we discuss this method in detail.

Let $\mathbb{Y}^t:=\{Y^1,Y^2,\cdots,Y^{M_t}\}$ be a subdivision of X. Then we consider the same convex relaxation subproblem on $Y^{k_t}=[a^{k_t},b^{k_t}]\in\mathbb{Y}^t$ as [20], i.e.,

$$\min_{x\in Y^{k_t}}F_{k_t}^{\alpha^{k_t}}(x):=f(x)+\sum_{i=1}^{n}\alpha_i^{k_t}(a_i^{k_t}-x_i)(b_i^{k_t}-x_i),\tag{3.1}$$

where $\alpha^{k_t}:=(\alpha_1^{k_t},\cdots,\alpha_n^{k_t})$. If $\tilde{\lambda}_{Y^{k_t}}\geq 0$ estimated by (2.3), then $\alpha_i^{k_t}=0$. Otherwise, $\alpha_i^{k_t}$ is computed by (2.1) for any $i\in\{1,\cdots,n\}$. Similarly, $\tilde{\lambda}_{Y^{k_t}}\geq 0$ implies that $F_{k_t}^{\alpha^{k_t}}$ is a convex lower bound estimation function of $f(x)$ on Y^{k_t} and $F_{k_t}^{\alpha^{k_t}}(x)=f(x)$ for any $x\in Y^{k_t}$.

Let $X_{ap}^{k_t}$ and $X_{ap}^{k_t,\varepsilon}$ denote the optimal solution set and the ε-solution set of (3.1), respectively,

$$X_{ap}^{k_t} := \{x \in Y^{k_t} : F_{k_t}^{\alpha^{k_t}}(x) \le F_{k_t}^{\alpha^{k_t}}(y) \text{ for any } y \in Y^{k_t}\},$$
$$X_{ap}^{k_t,\varepsilon} := \{x \in Y^{k_t} : F_{k_t}^{\alpha^{k_t}}(x) \le F_{k_t}^{\alpha^{k_t}}(y) + \varepsilon \text{ for any } y \in Y^{k_t}\}, \tag{3.2}$$

where $\varepsilon > 0$. Obviously, $X_{ap}^{k_t}$ and $X_{ap}^{k_t,\varepsilon}$ are not empty sets.

As mentioned above, for any $k_t \in \{1, \cdots, M_t\}$ and $Y^{k_t} \in \mathbb{Y}^t$, $F_{k_t}^{\alpha^{k_t}}$ is a convex relaxation sub-problem of (NCOP) w.r.t. Y^{k_t}. Then, we use all convex sub-problems to constitute the piecewise convexification optimization problem of (NCOP) with respect to X.

3.2. **The solution set of the piecewise convexification problem.** In this subsection, we construct the solution set of the piecewise convexification problem w.r.t. X for the subdivision $\mathbb{Y}^t$. The construction of this solution set is crucial because it relates to the approximation of the globally optimal solution set and directly affects the performance of the algorithm.

Let M_t be the index set of all sub-boxes of the subdivision $\mathbb{Y}^t$ and $X = \bigcup_{k_t \in M_t} Y^{k_t}$. As we all know, if f is convex on a current box, then it is also convex on any subset of that box. Thus, in this paper, we will check whether f is already convex on the current box before dividing it. Obviously, we can first define two auxiliary indicator sets to classify those sub-sets based on the convexity of f on its corresponding box, i.e,

$$M_1(t) := \left\{k_t \in M_t : \tilde{\lambda}_{Y^{k_t}} \ge 0\right\} \text{ and } M_2(t) := \left\{k_t \in M_t : \tilde{\lambda}_{Y^{k_t}} < 0\right\},$$

where $\tilde{\lambda}_{Y^{k_t}}$ is defined by (2.3). Clearly, $M_t = M_1(t) \cup M_2(t)$ and f is convex on the subset Y^{k_t} for any $k_t \in M_1(t)$. However, for any $k_t \in M_2(t)$ one cannot claim that f must be non-convex on Y^{k_t}, because we only obtain $\tilde{\lambda}_{Y^{k_t}} < 0$ and not $\lambda_{\min}^{Y^{k_t}} < 0$. That is, the convexity of f on Y^{k_t} is uncertain for any $k_t \in M_2(t)$. Thus, we need to be more concerned about the indices in $M_2(t)$ rather than in $M_1(t)$ for any t. Then a box division criterion is proposed.

Box division criterion: *No division is applied to the box if f is convex on it, and the box on which f is non-convex is selected for subdivision.*

Some notations that help us to clearly define the solution set of the piecewise convexification problem, corresponding to the above index sets, are presented.

$$X_{ap}^{C}(t) := \bigcup_{k_t \in M_1(t)} X_{ap}^{k_t} \text{ and } X_{ap}^{UC}(t) := \bigcup_{k_t \in M_2(t)} X_{ap}^{k_t}, \tag{3.3}$$

where $X_{ap}^{k_t}$ is an optimal solution set of the convex relaxation optimization problem (3.1).

Finally, using these auxiliary sets above, we in this paper, directly define the set of the piecewise convexification problem $X_{ap}(t)$ of the following form

$$X_{ap}(t) = \left\{x \in X_{ap}^{C \cup UC}(t) : f(x) \le f(y) \text{ for any } y \in X_{ap}^{C \cup UC}(t)\right\} \tag{3.4}$$

where $X_{ap}^{C \cup UC}(t) := X_{ap}^{C}(t) \cup X_{ap}^{UC}(t)$.

Remark 3.1. When f is convex on X, it is easy to check that $X_{ap}(t) = X_{op}$ for any t, i.e., this definition can be reduced to the classical way of defining a solution set. This shows that this definition of solution set of the piecewise convexification optimization problem is reasonable.

In what follows, we analyze the solution set $X_{ap}(t+1)$ we defined from the perspective of the division process, which helps us to better understand the advantages of this definition. As mentioned above, no subdivision is applied to this box on which f is convex, that is, from the subdivision $\mathbb{Y}^t$ to $\mathbb{Y}^{t+1}$, we only divide the boxes whose indicators belong to $M_2(t)$, instead of dividing all the boxes corresponding to M_t. Then, this division process yields to two new auxiliary index sets based on $M_2(t)$, as defined by

$$M_t^{t+1,C} := \left\{ j_t : \tilde{\lambda}_{Y^{j_t}} \geq 0 \text{ where } Y^{j_t} \in \mathrm{Sp}(Y^{k_t}) \text{ and } k_t \in M_2(t) \right\}, \tag{3.5}$$

$$M_t^{t+1,UC} := \left\{ j_t : \tilde{\lambda}_{Y^{j_t}} < 0 \text{ where } Y^{j_t} \in \mathrm{Sp}(Y^{k_t}) \text{ and } k_t \in M_2(t) \right\}, \tag{3.6}$$

which indicate that we classify the new sub-boxes, that is, we put the index of the new box that makes f convex in $M_t^{t+1,C}$, otherwise, put it in $M_t^{t+1,UC}$. Obviously, these definitions imply that

$$M_1(t+1) = M_1(t) \cup M_t^{t+1,C} \text{ and } M_2(t+1) = M_t^{t+1,UC}. \tag{3.7}$$

The union of solution sets around (3.5) and (3.6) are similarly represented by

$$X_{ap}^{newC}(t,t+1) := \bigcup_{k_t \in M_t^{t+1,C}} X_{ap}^{k_t} \text{ and } X_{ap}^{newUC}(t,t+1) := \bigcup_{k_t \in M_t^{t+1,UC}} X_{ap}^{k_t}.$$

Combining this with (3.7), one can conduct that

$$X_{ap}^{C}(t+1) = X_{ap}^{C}(t) \cup X_{ap}^{newC}(t,t+1) \text{ and } X_{ap}^{UC}(t+1) = X_{ap}^{newUC}(t,t+1).$$

Therefore, $X_{ap}(t+1)$ can be equivalently expressed in the following form:

$$X_{ap}(t+1) := \left\{ x \in X_{ap}(t,t+1) : f(x) \leq f(y) \text{ for any } y \in X_{ap}(t,t+1) \right\},$$

where $X_{ap}(t,t+1) := X_{ap}^{C}(t) \cup X_{ap}^{newC}(t,t+1) \cup X_{ap}^{newUC}(t,t+1)$.

This equivalent form demonstrates the rationality and advantage of this way of defining $X_{ap}(t+1)$. These are summarized in the following remark.

Remark 3.2. (i) There is a significant relationship between $X_{ap}(t)$ and $X_{ap}(t+1)$ since $X_{ap}(t+1)$ uses the information $X_{ap}^{C}(t)$ from the subdivision $\mathbb{Y}^t$. In fact, the subdivision $\mathbb{Y}^{t+1}$ is always based on the result of the subdivision $\mathbb{Y}^t$, and it follows that this relation is reasonable.

(ii) From the subdivision $\mathbb{Y}^t$ to $\mathbb{Y}^{t+1}$, we do not consider the boxes that make f convex in the subdivision $\mathbb{Y}^t$. Moreover, we use $X_{ap}^{C}(t)$ directly from the result of the subdivision t to construct $X_{ap}(t+1)$, instead of solving these sub-problems repeatedly. These techniques can reduce the number of sub-problems to solve in the piecewise convexification method.

Next, we discuss the relationship between $X_{ap}(t)$ and X_{op} when f has the piecewise convex property on X, i.e., f is non-convex on X and f is convex on each sub-box of X for some subdivision $\mathbb{Y}^{t_0}$.

Theorem 3.1. *If there exists a subdivision $\mathbb{Y}^{t_0}$ of X such that $M_2(t_0) = \emptyset$, then $X_{op} = X_{ap}(t_0)$.*

Proof. Obviously, $M_2(t_0) = \emptyset$ implies that $X_{ap}^{UC}(t_0) = \emptyset$, $M_{t_0} = M_1(t_0)$, and

$$X_{ap}(t_0) = \{x \in X_{ap}^{C}(t_0) : f(x) \leq f(y) \text{ for any } y \in X_{ap}^{C}(t_0)\}. \tag{3.8}$$

Since $\mathbb{Y}^{t_0}$ is a subdivision of X, then $X = \bigcup_{k_t \in M_{t_0}} Y^{k_t}$. In what follows, we prove that $X_{ap}(t_0) = X_{op}$, that is, $X_{op} \subseteq X_{ap}(t_0)$ and $X_{ap}(t_0) \subseteq X_{op}$. If $X_{op} \nsubseteq X_{ap}(t_0)$, then there exists $\widehat{x} \in X_{op}$ such that $\widehat{x} \notin X_{ap}(t_0)$. Based on the definition of $X_{ap}(t_0)$, shown in (3.8), it is easy to verify that $\widehat{x} \notin X_{op}$. Therefore, we conduct that $X_{op} \subseteq X_{ap}(t_0)$.

Next, we assume that $X_{ap}(t_0) \nsubseteq X_{op}$, i.e., there exists $\widehat{x} \in X_{ap}(t_0)$ with $\widehat{x} \notin X_{op}$. Thus, one can find $\widehat{y} \in X$ satisfying $f(\widehat{y}) < f(\widehat{x})$. Since $\mathbb{Y}^{t_0}$ is a subdivision, then let $\widehat{y} \in Y^{j_{k_0}}$, where $k_{t_0} \in M_{t_0} = M_1(t_0)$. Due to $X_{ap}^{k_{t_0}} \neq \emptyset$, there must exist $\widehat{z}$ such that $\widehat{z} \in X_{ap}^{k_{t_0}}$. It follows that $f(\widehat{z}) \leq f(\widehat{y})$, which yields $f(\widehat{z}) < f(\widehat{x})$ with $\widehat{z} \in X_{ap}^{k_{t_0}} \subseteq X_{ap}^{C}(t_0)$. This seems to contradict the fact that $\widehat{x} \in X_{ap}(t_0)$. Thus $X_{ap}(t_0) \subseteq X_{op}$.

Obviously, from the above analysis, one can conclude that $X_{ap}(t_0) = X_{op}$. □

This theorem demonstrates that the proposed piecewise convexification method can explore all globally optimal solutions of the non-convex problem with the piecewise convex properties.

3.3. **Approximate the globally optimal solution set.**

In this subsection, we show that $X_{ap}(t)$ is a lower bound set of X_{op}^{ε}, and a new approximate solution set of the piecewise convexification optimization problem is presented to obtain the upper bound set of X_{op}^{ε}.

Theorem 3.2. *For any $\varepsilon > 0$, there exist $t \in \mathbb{N}$ and the subdivision $\mathbb{Y}^t$ of X such that*

$$X_{ap}(t) \subseteq X_{op}^{\varepsilon},$$

where the subdivision $\mathbb{Y}^t$ satisfies $\max\limits_{k_t \in M_2(t)} \sum\limits_{i=1}^{n} \alpha_i^{k_t} \left(\frac{a_i^{k_t} - b_i^{k_t}}{2}\right)^2 \leq \varepsilon$

Proof. From the Lemma 5 in [18], there exists the subdivision $\mathbb{Y}^t$ such that $|T(\mathbb{Y}^t)| \to 0$ as $t \to \infty$. Since α_i is a finite value for any i, then there exist t_0 and a subdivision $\mathbb{Y}^{t_0}$ such that

$$\max_{k_{t_0} \in M_2(t_0)} \sum_{i=1}^{n} \alpha_i^{k_{t_0}} \left(\frac{a_i^{k_{t_0}} - b_i^{k_{t_0}}}{2}\right)^2 \leq \varepsilon. \tag{3.9}$$

Next, we just need to prove that $X_{ap}(t_0) \subseteq X_{op}^{\varepsilon}$. Suppose that $X_{ap}(t_0) \nsubseteq X_{op}^{\varepsilon}$, that is, there exists $\widehat{x} \in X_{ap}(t_0)$ such that $\widehat{x} \notin X_{op}^{\varepsilon}$. Obviously, $\widehat{x} \in X_{ap}^{C \cup NC}(t_0)$ and then one can find $\widehat{y} \in X$ satisfying

$$f(\widehat{y}) + \varepsilon < f(\widehat{x}). \tag{3.10}$$

Let $\widehat{y}$ belong to the sub-box $Y^{k_{t_0}}$ of X, i.e., $\widehat{y} \in Y^{k_{t_0}} \in \mathbb{Y}^{t_0}$ where $k_{t_0} \in M_1(t_0) \cup M_2(t_0)$.

If $\widehat{y} \in X_{ap}^{k_{t_0}} \subseteq Y^{k_{t_0}}$, then, according to (3.10) and $X_{ap}^{k_{t_0}} \subseteq X_{ap}^{C \cup NC}(t_0)$, it contradicts $\widehat{x} \in X_{ap}(t_0)$. Thus $\widehat{y} \in Y^{k_{t_0}} \backslash X_{ap}^{k_{t_0}}$. Then, there exists $\widehat{z} \in X_{ap}^{k_{t_0}}$ satisfying $F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{z}) < F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{y})$, where $k_{t_0} \in M_{t_0}$. Note that $M_{t_0} := M_1(t_0) \cup M_2(t_0)$. Obviously, if $k_{t_0} \in M_1(t_0)$, i.e., f is convex on $Y^{k_{t_0}}$, then $f(\widehat{z}) < f(\widehat{y})$ and $f(\widehat{z}) < f(\widehat{y}) + \varepsilon < f(\widehat{x})$ from (3.10). This is contrary to $\widehat{x} \in X_{ap}(t_0)$ by $\widehat{z} \in X_{ap}^{k_{t_0}} \subset X_{ap}^{C}(t_0)$. However, if $k_{t_0} \in M_2(t_0)$, then $\widehat{y} \notin X_{ap}^{k_{t_0}}$ implies that there exists $\widehat{z} \in X_{ap}^{k_{t_0}} \subset X_{ap}^{UC}(t_0)$ such that $F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{z}) < F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{y})$. On account of (3.10), one can obtain $F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{z}) < f(\widehat{x}) - \varepsilon$, that is,

$$f(\widehat{z}) + \sum_{i=1}^{n} \alpha_i^{k_{t_0}} (a_i^{k_{t_0}} - \widehat{z}_i)(b_i^{k_{t_0}} - \widehat{z}_i) < f(\widehat{x}) - \varepsilon,$$

which implies that $f(\widehat{z}) < f(\widehat{x})$ by (3.9). This contradicts $\widehat{x} \in X_{ap}(t_0)$ by $\widehat{z} \in X_{ap}^{k_{t_0}} \subset X_{ap}^{UC}(t_0)$. Consequently, we infer that $\widehat{y} \notin Y^{k_{t_0}} \backslash X_{ap}^{k_{t_0}}$.

In summary, the above analysis violates the assumption $\widehat{y} \in Y^{k_{t_0}}$. Therefore, this theorem has now been proven. $\square$

The above theorems demonstrate that the solution set of the piecewise convexification optimization problem is a lower bound set of X_{op}^{ε}. To construct the upper bound set of X_{op}^{ε}, the approximate solution set of the piecewise convexification optimization problem is introduced as

$$X_{ap}^{\varepsilon}(t) := \left\{ x \in X_{ap}^{C\cup UC,\varepsilon}(t) : f(x) \leq f(y) + \varepsilon \text{ for any } y \in X_{ap}^{C\cup UC,\varepsilon}(t) \right\},$$

where $X_{ap}^{k_t,\varepsilon}$ is defined by (3.2), $X_{ap}^{C\cup UC,\varepsilon}(t) := X_{ap}^{C,\varepsilon}(t) \cup X_{ap}^{UC,\varepsilon}(t)$, and

$$X_{ap}^{C,\varepsilon}(t) := \bigcup_{k_t \in M_1(t)} X_{ap}^{k_t,\varepsilon},\; X_{ap}^{UC,\varepsilon}(t) := \bigcup_{k_t \in M_2(t)} X_{ap}^{k_t,2\varepsilon}.$$

In what follows, we prove that a new set $X_{ap}^{\varepsilon}(t)$ is an upper bound set of X_{op}^{ε}.

Theorem 3.3. *For any $\varepsilon > 0$, there exist $t \in \mathbb{N}$ and the subdivision $\mathbb{Y}^t$ of X such that*

$$X_{op}^{\varepsilon} \subseteq X_{ap}^{\varepsilon}(t),$$

where the subdivision $\mathbb{Y}^t$ satisfies $\max\limits_{k_t \in M_2(t)} \sum\limits_{i=1}^{n} \alpha_i^{k_t} \left(\frac{b_i^{k_t} - a_i^{k_t}}{2} \right)^2 \leq \varepsilon$,

Proof. Similarly, there exist t_0 and a subdivision $\mathbb{Y}^{t_0}$ of X such that (3.9) holds. It remains to show that $X_{op}^{\varepsilon} \subseteq X_{ap}^{\varepsilon}(t_0)$. Assume that $X_{op}^{\varepsilon} \nsubseteq X_{ap}^{\varepsilon}(t_0)$ for some t_0. Then one can find $\widehat{x} \in X_{op}^{\varepsilon}$ and $\widehat{x} \notin X_{ap}^{\varepsilon}(t_0)$.

Now, we can distinguish two cases for $\widehat{x} \notin X_{ap}^{\varepsilon}(t_0)$ from the definition of $X_{ap}^{\varepsilon}(t_0)$, and the first of which is $\widehat{x} \in X_{ap}^{C\cup UC,\varepsilon}(t_0)$. It implies that there exists $\widehat{y} \in X_{ap}^{C\cup UC,\varepsilon}(t_0) \subseteq X$ satisfying $f(\widehat{y}) + \varepsilon < f(\widehat{x})$. This contradicts the fact that $\widehat{x} \in X_{op}^{\varepsilon}$, that is, the first case is not true.

The second case is $\widehat{x} \notin X_{ap}^{C\cup UC,\varepsilon}(t_0)$. For this subdivision $\mathbb{Y}^{t_0}$, without loss of generality, let $\widehat{x} \in Y^{k_{t_0}}$ for some $k_{t_0} \in M_{t_0} = M_1(t_0) \cup M_2(t_0)$. If $k_{t_0} \in M_1(t_0)$, then $X_{ap}^{k_{t_0},\varepsilon} \subseteq X_{ap}^{C,\varepsilon}(t_0)$. Moreover, $\widehat{x} \notin X_{ap}^{C\cup UC,\varepsilon}(t_0)$ indicates that $\widehat{x} \notin X_{ap}^{k_{t_0},\varepsilon}$. Then, there exists $\widehat{y} \in Y^{k_{t_0}}$ with $F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{y}) + \varepsilon < F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{x})$. Obviously, $F_{k_{t_0}}^{\alpha^{k_{t_0}}}(x) = f(x)$ for any $x \in Y^{k_{t_0}}$ as $k_{t_0} \in M_1(t_0)$. Then it conducts that

$$f(\widehat{y}) + \varepsilon = F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{y}) + \varepsilon < F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{x}) = f(\widehat{x}),$$

which contradicts $\widehat{x} \in X_{op}^{\varepsilon}$. This yields to $k_{t_0} \notin M_1(t_0)$, that is, $k_{t_0} \in M_2(t_0)$. This indicates that $X_{ap}^{k_{t_0},2\varepsilon} \subseteq X_{ap}^{UC,\varepsilon}(t_0)$ and $\widehat{x} \notin X_{ap}^{k_{t_0},2\varepsilon}$ by $\widehat{x} \notin X_{ap}^{C\cup UC,\varepsilon}(t_0)$. Thus there exists $\widehat{y} \in Y^{k_t}$ such that $F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{y}) + 2\varepsilon < F_{k_{t_0}}^{\alpha^{k_{t_0}}}(\widehat{x})$, which implies from (3.9) that $f(\widehat{y}) + \varepsilon < f(\widehat{x})$. This contradicts $\widehat{x} \in X_{op}^{\varepsilon}$, which means that $k_{t_0} \in M_2(t_0)$ is also false. Thus, the second case would not hold.

Consequently, the above assumption is not true, that is, $X_{op}^{\varepsilon} \subseteq X_{ap}^{\varepsilon}(t_0)$ holds. $\square$

Remark 3.3. Combining Theorems 3.2 and 3.3, the lower and upper bound sets of the set $X_{ap}(t)$ are obtained by constructing solution sets of the piecewise convexification optimization problem when the subdivision, that is,

$$X_{ap}(t) \subseteq X_{op}^{\varepsilon} \subseteq X_{ap}^{\varepsilon}(t), \forall \varepsilon > 0,$$

where the subdivision $\mathbb{Y}^t$ satisfies $\max_{k_t \in M_2(t)} \sum_{i=1}^{n} \alpha_i^{k_t} \left(\frac{a_i^{k_t} - b_i^{k_t}}{2} \right)^2 \leq \varepsilon$. However, Eichfelder, Gerlach, and Sumi only obtained a lower bound set of $X_{ap}(t)$ in [18].

4. A Piecewise Convexification Algorithm

In this section, the piecewise convexification algorithm for the non-convex optimization problem is devised. Furthermore, the theoretical property of this algorithm is obtained. It is necessary to introduce some notations in Table 1 that are used in the algorithm.

Table 1. Notations in algorithm

Abbreviation	Denotation
$\widehat{X}$	The sub-box of X
$\widehat{\mu}$	The function values of $F_{\widehat{X}}^{\widehat{\alpha}}$ on $\widehat{X}$
$X_{ap}^{\widehat{X}}$	The optimal solution set of $F_{\widehat{X}}^{\widehat{\alpha}}$ on $\widehat{X}$
v_{glob}	The smallest objective function value found for all current sub-boxes
$w(\widehat{X}, \widehat{\alpha})$	The modified width of $\widehat{X}$ and $w(\widehat{X}, \widehat{\alpha}) := \sum_{i=1}^{n} \widehat{\alpha}_i \left(\frac{\widehat{a}_i - \widehat{b}_i}{2} \right)^2$
$\tilde{\lambda}_{\widehat{X}}$	A lower bound of $\lambda_{\min}(\widehat{X})$ computed by (2.3)

Next, we present the piecewise convexification algorithm with a new box selection rule, a box discard and two new termination strategies, which is used to obtain a subset of the approximate globally optimal solution set, as shown in Algorithm 1.

Remark 4.1. The algorithm is explained in more detail below.

(1) It terminates after a finite number of iterations because there exists the subdivision $\mathbb{Y}^t$ of X such that $\max_{(\widetilde{X}, \tilde{x}, \tilde{\mu}, \widetilde{\alpha}) \in L_{UC}} w(\widetilde{X}, \widetilde{\alpha}) \leq \varepsilon$.

(2) At the end of the algorithm, the union of sub-boxes $\bigcup_{(\widehat{X}, \widehat{x}, \widehat{\mu}, \widehat{\alpha}) \in M_1} \widehat{X}$ is constructed a subdivision of X, where $M_1 = M \cup L_{UC}$.

(3) The same box discarding technique in [18], i.e., $\widehat{\mu} > v_{glob}$, is applied in our algorithm. Thus, it also holds that $\widehat{X} \cap X_{op} = \emptyset$.

(4) Noticeable point is that there is only one termination condition in [18], i.e., $L = \emptyset$. However, our algorithm contains two termination conditions, that is,

$$\max_{(\widetilde{X}, \tilde{x}, \tilde{\mu}, \widetilde{\alpha}) \in L_{UC}} w(\widetilde{X}, \widetilde{\alpha}) \leq \varepsilon \text{ or } L_{UC} = \emptyset.$$

Clearly, if $L_{UC} = \emptyset$, then $\max_{(\widetilde{X}, \tilde{x}, \tilde{\mu}, \widetilde{\alpha}) \in L_{UC}} w(\widetilde{X}, \widetilde{\alpha}) = 0$, and not vice versa, i.e., $L_{UC} \neq \emptyset$ may be hold when $\max_{(\widetilde{X}, \tilde{x}, \tilde{\mu}, \widetilde{\alpha}) \in L_{UC}} w(\widetilde{X}, \widetilde{\alpha}) \leq \varepsilon$. The following numerical experiments, in Section 5, will show that, for some complex problems, these two termination conditions are more conducive to speeding up the algorithm than having only one termination condition in [18].

(5) Another noticeable point is that, in Algorithm 1, we set a new selection rule of sub-box for the division, that is, the box with the maximum modified width in L_{UC} is selected to

Algorithm 1 The Piecewise Convexification Algorithm for (NCOP) (PCA–αBB).

Require: $X^0 = [a,b] \in \mathbb{I}^n$, $f \in \mathbb{C}^2(\mathbb{R}^n,\mathbb{R}), \varepsilon > 0$;
Ensure: X_{ap}^{new};
1: Compute an α^0 of f on X^0 according to (2.1)-(2.2);
2: Set $X^* := X^0$, $x^* := \frac{a+b}{2}$, $\mu^* := -\infty$, $\alpha^* := \alpha_0$, $x_{act} := x^*$, $v_{glob} = v_{act} = +\infty$, $L_{UC} := \{(X^*,x^*,\mu^*,\alpha^*)\}$, $M_C = M_D = \emptyset$ and $k := 0$.
3: **while** $L_{UC} \neq \emptyset$ and $\max\limits_{(\widetilde{X},\widetilde{x},\widetilde{\mu},\widetilde{\alpha})\in L_{UC}} w(\widetilde{X},\widetilde{\alpha}) > \varepsilon$ **do**
4: $k := k+1$.
5: Define (X^*,x^*,μ^*,α^*) at the first element of L_{UC} with $\max\limits_{(\widetilde{X},\widetilde{x},\widetilde{\mu},\widetilde{\alpha})\in L_{UC}} w(\widetilde{X},\widetilde{\alpha})$.
6: Delete (X^*,x^*,μ^*,α^*) from L_{UC}.
7: **for** all $\widehat{X} \in \mathrm{Sp}(X^*)$ **do**
8: Compute $\widetilde{\lambda}_{\widehat{X}}$ by (2.3) and $\widehat{\alpha}$ by (2.1), respectively.
9: Compute $\widehat{x} \in \arg\min\limits_{x\in\widehat{X}} F_{\widehat{X}}^{\widehat{\alpha}}(x)$. Let $\widehat{\mu} = \min\limits_{x\in\widehat{X}} F_{\widehat{X}}^{\widehat{\alpha}}(x)$.
10: **if** $\widehat{\mu} \leq v_{glob}$ **then**
11: **if** $\widetilde{\lambda}_{\widehat{X}} \geq 0$ **then**
12: Add $(\widehat{X},\widehat{x},\widehat{\mu},\widehat{\alpha})$ as the last element to M_C and add $\widehat{x}$ to $X_{ap}^{\widehat{X}}$.
13: **else**
14: Add $(\widehat{X},\widehat{x},\widehat{\mu},\widehat{\alpha})$ as the last element to L_{UC} and add $\widehat{x}$ to $X_{ap}^{\widehat{X}}$.
15: **end if**
16: **if** $f(\widehat{x}) \leq v_{act}$ **then**
17: Set $x_{act} = \widehat{x}, v_{act} = f(x_{act}), v_{glob} = \min\{v_{act}, v_{glob}\}$
18: Delete $(X,x,\mu,\alpha) \in L_{UC}$ with $\mu > v_{glob}$ from L_{UC}, and
19: add $(X,x,\mu,\alpha) \in L_{UC}$ with $\mu > v_{glob}$ to M_D.
20: **end if**
21: **else**
22: Add $(\widehat{X},\widehat{x},\widehat{\mu},\widehat{\alpha})$ as the last element to M_D.
23: **end if**
24: **end for**
25: **end while**
26: $M := \bigcup\{\widetilde{X} : (\widetilde{X},\widetilde{x},\widetilde{\mu},\widetilde{\alpha}) \in L_{UC}\} \cup M_C$.
27: $X_{ap}^{new} := \left\{ x \in \bigcup\limits_{\widehat{X}\in M} X_{ap}^{\widehat{X}} : \ f(x) \leq f(y),\ \forall\, y \in \bigcup\limits_{\widehat{X}\in M} X_{ap}^{\widehat{X}} \right\}$.

divide into two sub-boxes as shown in line 5. This selection approach is different from the one proposed in [18].

(6) More importantly, Algorithm 1 requires only one while loop and no additional parameters, in contrast to the modified αBB method [18] with two while loops and some additional parameters. parameters.

The following theorem shows the theoretical result of this algorithm.

Theorem 4.1. *At the end of Algorithm 1, the set X_{ap}^{new} output of Algorithm 1 is constructed a subset of the approximate global optimal solution set X_{op}^{ε}.*

Proof. The proof process is similar to Theorem 3.2 and is omitted here. □

5. Numerical Experiments

In this section, we show the results of computational experiments, focusing on the comparisons between PCA–αBB, i.e., Algorithm 1, and M-αBB(the mod $\alpha_{i,d=u-l}^{loc}$BB, [18]). All computations were performed on a computer with Iter(R)Core(TM)i5-8250U CPU and 8 Gbytes RAM. The code for two algorithms was written in Python 3.

As stated in [18], the authors used INTLAB ToolBox to automatically compute the elements $\nabla^2 f(X)_{ij}$. However, due to the software copyright, we cannot directly use INTLAB ToolBox, so we solve the optimization problem $\min_{x\in\widetilde{X}} \nabla^2 f(\widetilde{X})_{ij}$, to estimate $\alpha_i^{k_t}$ on each sub-box, and we rewrite the code of the modified αBB method in [18], without INTLAB.

For the sake of brevity, we list in Table 2 some of the notations used to record the numerical results. Note that if $flag = 1$, then this algorithm satisfies $L_{UC} = \emptyset$. Otherwise,

Table 2. Notations in numerical experiments

Abbreviation	Denotation
$iter$	Number of iterations required
t	Required CPU time in seconds
N_ε	Number of ε-optimal solutions of the algorithm, let $\varepsilon = 10^{-3}$
$flag$	The algorithm termination condition indicator, $flag = 1$ or 0
-	The algorithm does not record a certain value

$\max_{(\widetilde{X},\widetilde{x},\widetilde{\mu},\widetilde{\alpha})\in L_{UC}} w(\widetilde{X},\widetilde{\alpha}) \le \varepsilon$ when $flag = 0$. The $flag$ is introduced to mark the termination condition of the algorithm and to illustrate the validity of two termination conditions in Algorithm 1. In addition, we replace $\bar{\mu} \le v_{glob}$ in line 11 by $\bar{\mu} \le v_{glob} + 10^{-6}$ for all instances.

First, we demonstrate the performance of two approaches on eight test instances listed in Table 3, which includes the objective functions f, the feasible sets X^0, the number of globally optimal solutions $|\arg\min_{x\in X} f(x)|$, and the globally optimal values $\min_{x\in X} f(x)$. Most of the test instances in Table 3 have multiple optimal solutions and all examples are two-dimensional.

Numerical results of these two algorithms are presented in Table 4. It is easy to see that for all the test examples, the $iter$ and t values of PCA-αBB are significantly smaller than those of M-αBB[18]. As for the $|N_\varepsilon|$ values, there is only a slight difference for ***Branin***, i.e., two algorithms can find only two globally optimal solutions, not three. These results demonstrate that PCA-αBB can find almost all optimal solutions of the original non-convex problem and performs better than M-αBB[18]. In addition, for ***Branin***, ***Himmelblau*** and ***Shubert*** the termination condition of PCA-αBB is the same, that is, $\max_{(\widetilde{X},\widetilde{x},\widetilde{\mu},\widetilde{\alpha})\in L_{UC}} w(\widetilde{X},\widetilde{\alpha}) \le \varepsilon$. The termination condition of PCA-αBB for other remaining instances is $L_{UC} = \emptyset$.

The partitioning results of X are clearly shown in Fig.1, where the first and third columns show the results obtained by PCA-αBB, while the second and fourth columns show the results of M-αBB[18]. The red star indicates the ε-optimal solutions. Obviously, Fig.1 indicates that the way of selecting the boxes to divide can effectively reduce the number of iterations. In fact, M-αBB[18] has numerous subdivisions of the box near the optimal solution, while our algorithm has only a few subdivisions. The partial graph intuitively reflects this claim, see Fig.2.

TABLE 3. Test instances with finite number of optimal solutions

	$f:\mathbb{R}^2\to\mathbb{R}$ with $f(x)$	X	$\lvert\arg\min_{x\in X} f(x)\rvert$	$\min_{x\in X} f(x)$
Rastrigin[18]	$20+x_1^2+x_2^2-10(\cos(2\pi x_1)+\cos(2\pi x_2))$	$\left[\begin{pmatrix}-5.12\\-5.12\end{pmatrix},\begin{pmatrix}5.12\\5.12\end{pmatrix}\right]$	1	0
6-Hump[24]	$(4-2.1x_1^2+\frac{1}{3}x_1^4)x_1^2+x_1x_2-(4-4x_2^2)x_2^2$	$\left[\begin{pmatrix}-1.9\\-1.1\end{pmatrix},\begin{pmatrix}1.9\\1.1\end{pmatrix}\right]$	2	≈ -1.031629
Branin[24]	$(x_2-\frac{5.1}{4\pi^2}x_1^2+\frac{5}{\pi}x_1-6)^2+10(1-\frac{1}{8\pi})\cos(x_1)+10$	$\left[\begin{pmatrix}-5\\0\end{pmatrix},\begin{pmatrix}10\\15\end{pmatrix}\right]$	3	≈ 0.397886
Himmelblau[24]	$(x_1^2+x_2-11)^2+(x_1+x_2^2-7)^2$	$\left[\begin{pmatrix}-6\\-6\end{pmatrix},\begin{pmatrix}6\\6\end{pmatrix}\right]$	4	0
Rastrigin mod[24]	$20+x_1^2+x_2^2+10(\cos(2\pi x_1)+\cos(2\pi x_2))$	$\left[\begin{pmatrix}-5.12\\-5.12\end{pmatrix},\begin{pmatrix}5.12\\5.12\end{pmatrix}\right]$	4	≈ 0.497480
Shubert[24]	$\sum_{i=1}^{5}[i\cos((i+1)x_1+1)]\sum_{j=1}^{5}[j\cos((j+1)x_2+j)]$	$\left[\begin{pmatrix}-10\\-10\end{pmatrix},\begin{pmatrix}10\\10\end{pmatrix}\right]$	18	≈ -186.730909
Deb 1[24]	$-\frac{1}{2}(\sin^6(5\pi x_1)+\sin^6(5\pi x_2))$	$\left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right]$	25	-1
Vincent[24]	$-\frac{1}{2}(\sin(10\ln(x_1))+\sin(10\ln(x_2)))$	$\left[\begin{pmatrix}0.25\\0.25\end{pmatrix},\begin{pmatrix}10\\10\end{pmatrix}\right]$	36	-1

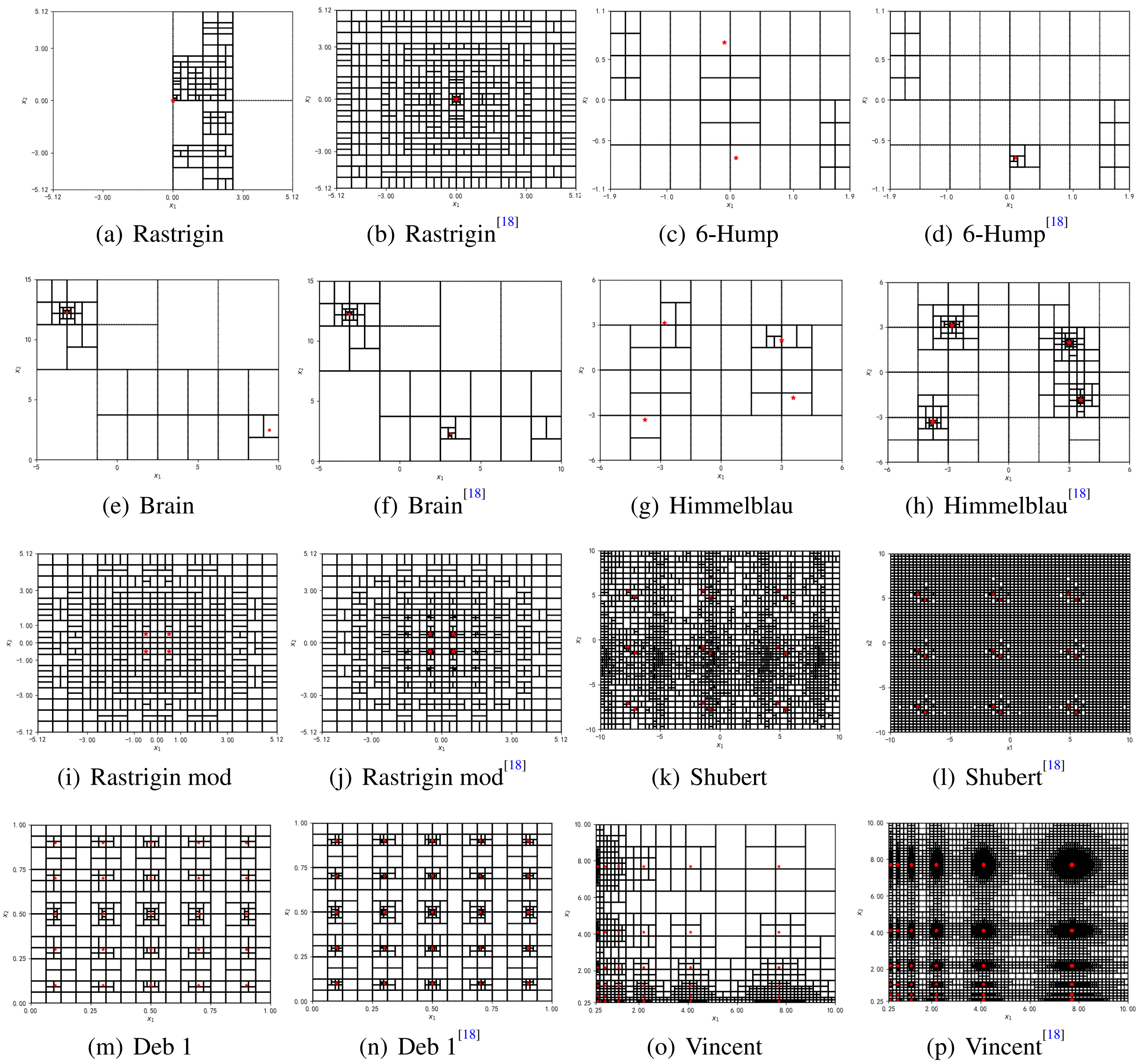
TABLE 4. Numerical results for test instances in Table 3.

Problem	PCA-αBB				M-αBB[18]			
	iter	t	N_ε	*flag*	*iter*	t	N_ε	*flag*
Rastrigin	**104**	**1.045**	**1**	1	551	30.959	**1**	1
6-Hump	**47**	**0.402**	**2**	1	49	0.418	1	1
Branin	**52**	**0.576**	**2**	0	67	0.726	**2**	1
Himmelblau	**43**	**0.558**	**4**	0	382	4.345	**4**	1
Rastrigin mod	**571**	**6.734**	**4**	1	859	11.046	**4**	1
Shubert	**3091**	**56.073**	**18**	0	5056	136.460	**18**	1
Deb 1	**391**	**5.390**	**25**	1	863	17.009	**25**	1
Vincent	**1169**	**10.820**	**36**	1	11705	166.680	**36**	1

In what follows, we consider four numerical tests with an infinite number of globally optimal solutions listed in [18], as defined by Table 5. Furthermore, the numerical results of the instance tests in Table 5 are shown in Table 6. The *iter* and t values of PCA-αBB are significantly better than those of M-αBB[18]. Except for ***Test03***, the number of solutions of PCA-αBB is also higher than that of M-αBB[18]. In addition, for PCA-αBB the termination condition

$$\max_{(\widehat{X},\widehat{x},\widehat{\mu},\widehat{\alpha})\in L_{UC}} w(\widehat{X},\widehat{\alpha})\le\varepsilon$$

is satisfied for these test problems as $flag=0$. In other words, this termination condition is meaningful in PCA-αBB and could be helpful to reduce the number of iterations. Moreover, the results of the interval partiting and solutions in set X_{ap}^{new} for PCA-αBB are shown in Fig.3. This figure shows that the distribution of these optimal solutions obtained from the PCA-αBB can be used to describe the distribution of the optimal solutions of the original problem. Meanwhile,



(a) Rastrigin (b) Rastrigin[18] (c) 6-Hump (d) 6-Hump[18]

(e) Brain (f) Brain[18] (g) Himmelblau (h) Himmelblau[18]

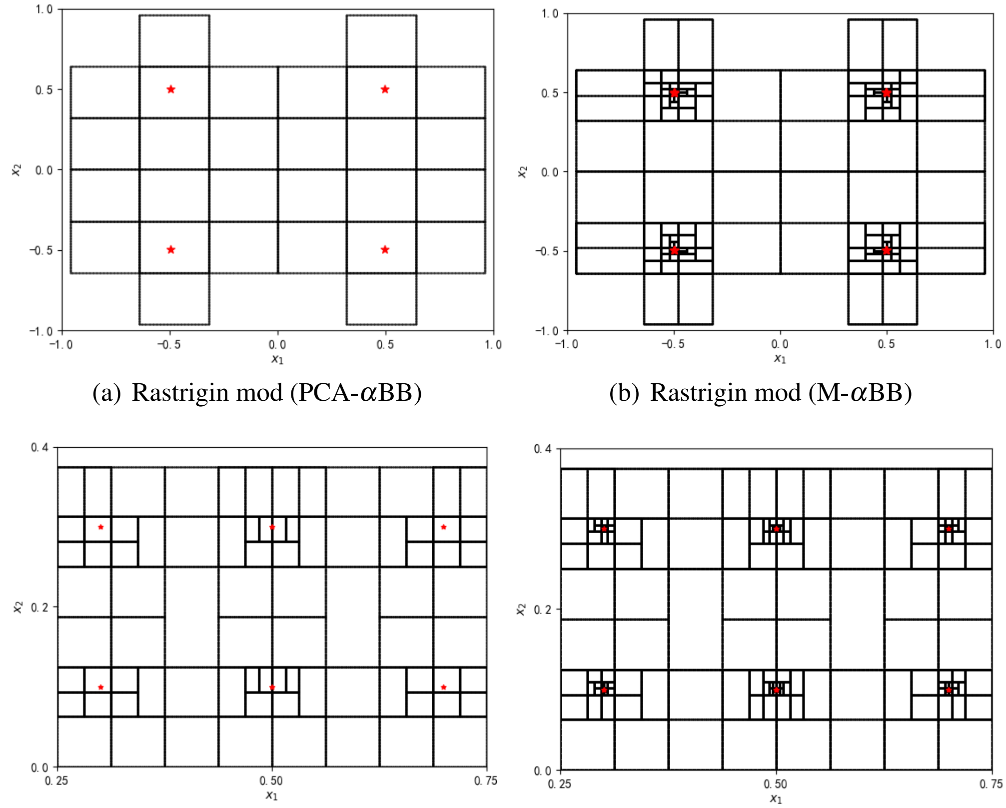
(i) Rastrigin mod (j) Rastrigin mod[18] (k) Shubert (l) Shubert[18]

(m) Deb 1 (n) Deb 1[18] (o) Vincent (p) Vincent[18]

FIGURE 1. Box division results of test instances in Table 3

TABLE 5. Test instances with an infinite number of optimal solutions

	$f:\mathbb{R}^2\to\mathbb{R}$	X	$\arg\min\limits_{x\in X} f(x)$
Test01	$\left(\frac{x_1^2}{4^2}+\frac{x_2^2}{2^2}-1\right)^2$	$\left[\begin{pmatrix}-5\\-5\end{pmatrix},\begin{pmatrix}5\\5\end{pmatrix}\right]$	$\left\{\begin{pmatrix}x_1\\x_2\end{pmatrix}\middle\vert \frac{x_1^2}{4^2}+\frac{x_2^2}{2^2}=1\right\}$
Test02	$\frac{1}{10}(x_1(1-x_2)+x_2(1-x_1))^2$	$\left[\begin{pmatrix}-5\\-5\end{pmatrix},\begin{pmatrix}5\\5\end{pmatrix}\right]$	$\left\{\begin{pmatrix}x_1\\x_2\end{pmatrix}\middle\vert x_1\in[-5,5]\backslash\{\frac{1}{2}\}, x_2=-\frac{x_1}{1-2x_1}\right\}$
Test03	$\sin^2(\frac{5}{4}x_1+x_2-3)$	$\left[\begin{pmatrix}0\\-2\end{pmatrix},\begin{pmatrix}4\\3\end{pmatrix}\right]$	$\left\{\begin{pmatrix}x_1\\x_2\end{pmatrix}\middle\vert \frac{5}{4}x_1+x_2=3+a, a\in\{0,\pm\pi\}\right\}\bigcap X$
Test04	$(x_1+\sin^2(x_1))\cos^2(x_2)$	$\left[\begin{pmatrix}0\\-2\end{pmatrix},\begin{pmatrix}4\\3\end{pmatrix}\right]$	$\left\{\begin{pmatrix}0\\x_2\end{pmatrix}\middle\vert x_2\in[-2,3]\right\}\bigcup\left\{\begin{pmatrix}x_1\\x_2\end{pmatrix}\middle\vert x_1\in[0,4], x_2\in\{\pm\frac{\pi}{2}\}\right\}$



(a) Rastrigin mod (PCA-αBB) (b) Rastrigin mod (M-αBB)

(c) Deb 1 (PCA-αBB) (d) Deb 1(M-αBB)

FIGURE 2. Subdivisions on sub-box for PCA-αBB (Algorithm 1)

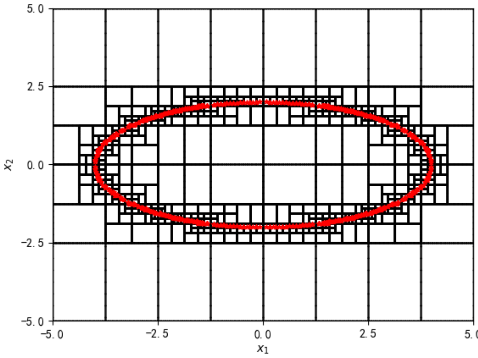
TABLE 6. Numerical results for test instance in Table 5.

Problems	PCA-αBB				M-αBB[18]			
	$iter$	t	N_ε	$flag$	$iter$	t	N_ε	$flag$
Test01	**559**	**6.717**	**592**	0	1355	18.4102	588	1
Test02	**672**	**6.891**	**649**	0	1156	17.511	433	1
Test03	**1189**	**11.511**	1237	0	3019	52.353	**1336**	1
Test04	**2343**	**31.722**	**3226**	0	4863	121.239	2123	1

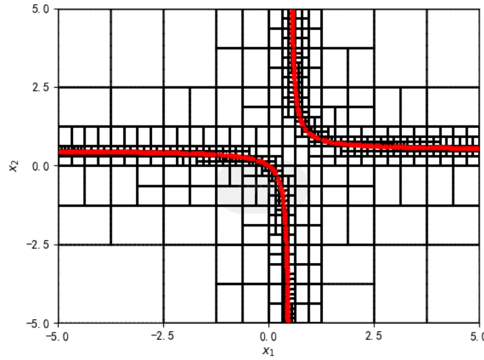
Fig. 4 shows the partition of two algorithms on the sub-box. Obviously, compared to PCA-αBB, M-αBB[18] has many redundant partitions close to the effective solution. Note that M-αBB[18] obtained numerous locally optimal solutions for ***Test04***, see Fig.4 (h).

Finally, in order to verify the efficiency of the proposed algorithm for the high-dimensional instances, Table 7 shows a high-dimensional test problem, which is selected from [18].

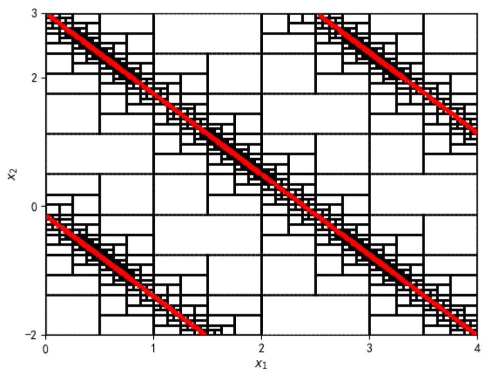
Table 8 shows the results of applying the PCA-αBB and M-αBB[18] to $\mathbf{TestDim}_d$. From Table 8, the experimental results demonstrate that both the proposed algorithm and M-αBB[18] lead to the same values of N_ε. However, compared with PCA-αBB, M-αBB[18] requires more $iter$ value and t value, and the advantage of PCA-αBB is more prominent as the dimension increases. For these high-dimensional problems, $flag = 0$ indicates that the stopping condition



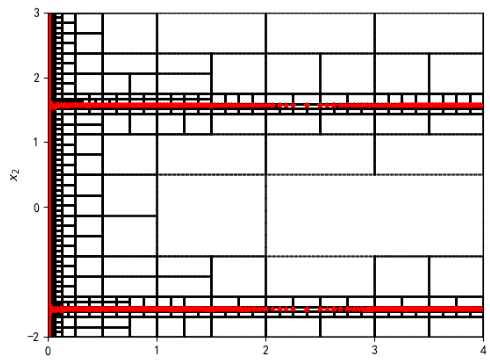
(a) ***Test01***



(b) ***Test02***

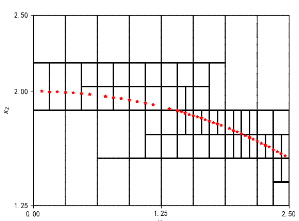


(c) ***Test03***

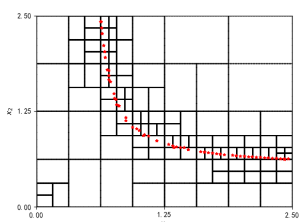


(d) ***Test04***

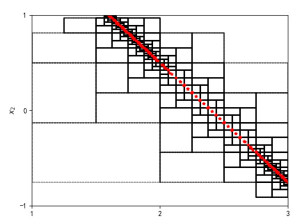
FIGURE 3. The box division and solutions in set X_{ap}^{new} for Algorithm 1 and ***Test01-04***



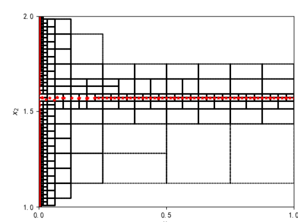
(a) ***Test01***



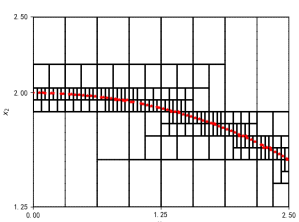
(b) ***Test02***



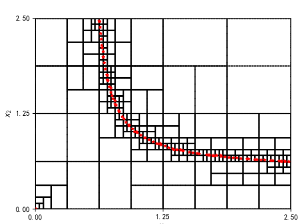
(c) ***Test03***



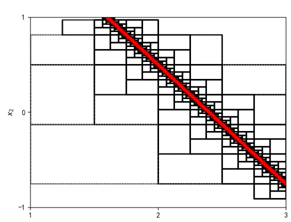
(d) ***Test04***



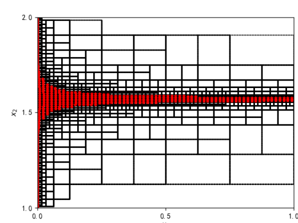
(e) ***Test01***(M-αBB)



(f) ***Test02***(M-αBB)



(g) ***Test03***(M-αBB)



(h) ***Test04***(M-αBB)

FIGURE 4. Subdivision on subbox for two algorithms.

TABLE 7. Test instances with high dimensional

	$f:\mathbb{R}^d\to\mathbb{R}$	X	$\arg\min_{x\in X} f(x)$
TestDim$_d$[18] $(d\in\mathbb{N})$	$\sum_{i=1}^{d}(\cos(2\pi x_i))^2$	$\left[\begin{pmatrix}-\frac{1}{4}\\ \cdots\\ -\frac{1}{4}\end{pmatrix},\begin{pmatrix}\frac{1}{4}\\ \cdots\\ \frac{1}{4}\end{pmatrix}\right]$	$\left\{x\in\mathbb{R}^d \middle\vert x_i\in\{-\frac{1}{4},\frac{1}{4}\}, i\in\{1,\cdots,d\}\right\}$

$\max_{(\widehat{X},\widehat{x},\widehat{\mu},\widehat{\alpha})\in L_{UC}} w(\widehat{X},\widehat{\alpha})\le\varepsilon$ is satisfied and $L_{UC}\neq\emptyset$ holds. These help to reduce the number of iterations and the CPU time.

TABLE 8. Numerical results for test instance in Table 7.

TestDim$_d$	PCA-αBB				M-αBB[18]			
	iter	t	N_ε	*flag*	*iter*	t	N_ε	*flag*
$d=2$	**11**	**0.061**	2	0	51	0.416	2	1
$d=3$	**47**	**0.491**	4	0	255	4.502	4	1
$d=4$	**175**	**2.915**	16	0	655	12.994	16	1
$d=5$	**607**	**17.107**	32	0	2076	84.313	32	1
$d=6$	**2047**	**78.059**	64	0	17919	267.484	64	1
$d=7$	**6783**	**420.671**	128	0	28353	1223.365	128	1
$d=8$	**22272**	**1103.478**	256	0	89871	5391.845	256	1
$d=9$	**72704**	**4940.783**	512	0	282072	32916.650	512	1

From the above numerical experiments, for these instances with infinite number of optimal solutions or with high dimension, the termination condition $\max_{(\widehat{X},\widehat{x},\widehat{\mu},\widehat{\alpha})\in L_{UC}} w(\widehat{X},\widehat{\alpha})\le\varepsilon$ is easier to satisfy than $L_{UC}=\emptyset$. This means that two termination conditions in the proposed algorithm can be used to reduce the number of iterations of the algorithm, thereby improving the efficiency of the algorithm. At the same time, this box classification strategy is beneficial to reduce the number of box divisions.

6. CONCLUSIONS

A piecewise convexification method based on the αBB method with box classification strategy was studied for non-convex single-objective optimization problems. The box classification strategy was proposed according to the convexity of the objective function on the sub-boxes when dividing the boxes, which helps to reduce the number of box divisions and improve the computational efficiency. We applied the αBB method to construct the piecewise convexification problem, and used the solution set of the piecewise convexification problem to approximate the globally optimal solution set from two different directions, which is obviously different from [18]. Furthermore, based on the theoretical results, a piecewise convexification algorithm with two termination conditions was proposed, and numerical experiments demonstated that this algorithm can obtain a large number of globally optimal solutions more faster than another algorithm in [18].

Acknowledgments
This research was supported by the Major Program of the National Natural Science Foundation of China (No. 11991020 and No. 11991024), the NSFC-RGC (Hong Kong) Joint Research Program (No. 12261160365), the National Natural Science Foundation of China (No. 11971084, and No. 12171060), the Team Project of Innovation Leading Talent in Chongqing (No. CQYC20210309536), the Natural Science Foundation of Chongqing (No. ncamc2022-msxm01), the Postdoctoral Science Foundation of Chongqing (No. CSTB2023NSCQ-BHX0017), the Science and Technology Research Program of Chongqing Education Commission of China (No. KJQN202300552), and the Foundation of Chongqing Normal University (No. 22XLB005 and No. 22XLB006)

REFERENCES

[1] L. Pan, X. Chen, Group sparse optimization for images recovery using capped folded concave functions, SIAM J. Imaging Sci. 14 (2021), 1-25.
[2] X. Li, Z. Zhu, A.M.C. So, R. Vidal, Nonconvex robust low-rank mtrix recovery, SIAM J. Optim. 30 (2020), 660-686.
[3] P. Zhong, Training robust support vector regression with smooth non-convex loss function, Optim. Methods Softw. 27 (2012), 1039-1058.
[4] S. Manisha, G. Debdas, K.K. Shukla, P. Witold, Robust twin support vector regression based on rescaled Hinge loss, Pattern Recog. 105 (2020), 107395.
[5] M. Locatelli, F. Schoen, Global optimization: Historical notes and recent developments, Euro. J. Comput. Optim. 9 (2021), 100012.
[6] A. Marmin, M. Castella, J.C. Pesquet, How to globally solve non-convex optimization problems involving an approximate l_0 penalization, In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5601-5605, IEEE, 2019.
[7] S. Qu, H. Yin, K. Zhang, A global optimization algorithm using linear relaxation, Appl. Math. Comput. 178 (2006), 510-518.
[8] F. Schoen, L. Tigli, Efficient large scale global optimization through clustering-based population methods, Comput. Oper. Res. 127 (2021), 105-165.
[9] Y. Xia, A survey of hidden convex optimization, J. Oper. Res. Soc. Chin. 8 (2020), 1-28.
[10] Y. Yang, M. Pesavento, S. Chatzinotas, B. Ottersten, Successive convex approximation algorithms for sparse signal estimation with nonconvex regularizations, IEEE J. Sel. Topics Signal Process, 12 (2018), 1286-1302.
[11] M. Jiang, R. Shen, Z. Meng, C. Dang, Optimization conditions and decomposable algorithms for convertible nonconvex optimization, J. Nonlinear Var. Anal. 7 (2023), 103-128.
[12] M. Hladík, An extension of the αBB-type underestimation to linear parametric Hessian matrices, J. Global. Optim. 64 (2016), 217-231.
[13] N. Kazazakis, C.S. Adjiman, Arbitrarily tight αBB underestimators of general non-linear functions over sub-optimal domains, J. Global. Optim. 71 (2018), 815-844.
[14] Y. Matanga, Y.X. Sun, Z.H.Wang, Hybrid PSO-αBB global optimisation for C^2 box-constrained multimodal NLPs, IEEE Access 10 (2021), 805-818.
[15] H. Milan, On the efficient gerschgorin inclusion usage in the global optimization αBB method, J. Global. Optim. 61 (2014), 235-253.
[16] E. Algaba, V. Fragnelli, N. Llorca, J.Sánchez-Soriano, Horizontal cooperation in a multimodal public transport system: The profit allocation problem, Eur. J. Oper. Res. 275 (2019), 659-665.
[17] C.H. Yoo, D.K. Lim, H.K. Jung, A novel multimodal optimization algorithm for the design of electromagnetic machines, IEEE T. Magn. 52 (2016), 1-4.
[18] G. Eichfelder, T. Gerlach, S. Sumi, A modification of the αBB method for box-constrained optimization and an application to inverse kinematics, Eur. J. Oper.Res. 4 (2016), 93-121.

[19] C.S. Adjiman, I.P. Androulakis, C.A. Floudas, A global optimization method, αBB, for general twice-differentiable constrained NLPs-II. Implementation and computational results, Comput. Chem. Eng. 22 (1998), 1159-1179.

[20] C.S. Adjiman, S. Dallwig, C.A. Floudas, A. Neumaier, A global optimization method, αBB, for general twice-differentiable constrained NLPs-I. Theoretical advances, Comput. Chem. Eng. 22 (1998), 1137-1158.

[21] I.P. Androulakis, C.D. Maranas, C.A. Floudas, αBB: A global optimization method for general constrained nonconvex problems, J. Global. Optim. 7 (1995), 337-363.

[22] D. Nerantzis, C.S. Adjiman, Tighter αBB relaxations through a refinement scheme for the scaled gerschgorin theorem, J. Global. Optim. 73 (2019), 467-483.

[23] A. Skjäl, T. Westerlund, New methods for calculating αBB-type underestimators, J. Global. Optim. 58 (2014), 411-427.

[24] M.G. Epitropakis, V.P. Plagianakos, M.N. Vrahatis, Finding multiple global optima exploiting differential evolution's niching capability, In: Proceedings of IEEE SDE, pp. 0-87, Paris, France, 2011.