

A neurodynamic approach for a class of pseudoconvex semivectorial bilevel optimization problems

Tran Ngoc Thang, Dao Minh Hoang & Nguyen Viet Dung

To cite this article: Tran Ngoc Thang, Dao Minh Hoang & Nguyen Viet Dung (14 Aug 2024): A neurodynamic approach for a class of pseudoconvex semivectorial bilevel optimization problems, Optimization Methods and Software, DOI: [10.1080/10556788.2024.2380688](https://doi.org/10.1080/10556788.2024.2380688)

To link to this article: <https://doi.org/10.1080/10556788.2024.2380688>



Published online: 14 Aug 2024.



Submit your article to this journal [↗](#)



Article views: 40



View related articles [↗](#)



View Crossmark data [↗](#)



A neurodynamic approach for a class of pseudoconvex semivectorial bilevel optimization problems

Tran Ngoc Thang^{a,b}, Dao Minh Hoang^{c,d} and Nguyen Viet Dung^{a,b}

^aFaculty of Mathematics and Informatics, Hanoi University of Science and Technology, Hanoi, Vietnam;

^bCenter for Digital Technology and Economy (BK Fintech), Hanoi University of Science and Technology, Hanoi, Vietnam; ^cFaculty of Computer Science, Technische Universität Dresden, Dresden, Germany; ^dCenter for Systems Biology Dresden, Dresden, Germany

ABSTRACT

The article proposes an exact approach to finding the global solution of a nonconvex semivectorial bilevel optimization problem, where the objective functions at each level are pseudoconvex, and the constraints are quasiconvex. Due to its non-convexity, this problem is challenging, but it attracts more and more interest because of its practical applications. The algorithm is developed based on monotonic optimization combined with a recent neurodynamic approach, where the solution set of the lower-level problem is inner approximated by copolyblocks in outcome space. From that, the upper-level problem is solved using the branch-and-bound method. Finding the bounds is converted to pseudoconvex programming problems, which are solved using the neurodynamic method. The algorithm's convergence is proved, and computational experiments are implemented to demonstrate the accuracy of the proposed approach.

ARTICLE HISTORY

Received 24 April 2023

Accepted 11 June 2024

KEYWORDS

Semivectorial bilevel optimization; pseudoconvex functions; neurodynamic method; monotonic optimization; branch-and-bound method

2020 MATHEMATICS

SUBJECT

CLASSIFICATIONS

90C26; 58E17

1. Introduction

The bilevel optimization problem was first examined with mathematical models by Bracken and McGill [7], following a formulation in the form of the Stackelberg game [13]. The focal difference between them and the prevalent single-level optimization problem is the hierarchical structure of the problem, where constraints for an optimization problem are affected by another problem. The affected problem is the lower-level problem, while the other is the upper-level problem. Despite getting more attention lately, it has been crucial because of its application to practical problems.

Unlike single-level optimization problems, which can be imagined as one agent trying to achieve objectives alone, bilevel ones are more like two agents interacting with each other by their former decisions while achieving objectives. Furthermore, since the majority of realistic problems are affected by many agents, the number of bilevel optimization research has increased quickly in the last decades, and they have spread in all fields of application. Some practical ones such as the Nash equilibrium problem in economics,

optimal operation control problem [18], management problem in supply chain systems [12,24], traffic and transportation network design problem [2,40], machine learning problem [11,25],... There are two main ways to classify bilevel optimization problems. One considers their type of functions, such as linear, convex, nonconvex, or discrete ones. Others consider the context of the problem, like cooperating with the optimistic problem in contrast to the pessimistic problem, optimization over the efficient set, multiobjective bilevel optimization problem, multi leaders/followers problem, bilevel with fuzzy and stochastic extension. Although in different contexts, they still share the same bilevel optimization formation.

In the method aspect, exact algorithms with classic approaches were researched for decades, in contrast to the heuristic group with later approaches. The exact methods that should be mentioned are extreme-point approaches, branch-and-bound, complementary pivoting, descent methods, penalty function methods, or trust-region methods. The heuristic groups contain evolutionary, local search-based, or neural network-based approaches. Lu et al. wrote a conscientious survey about these approaches (see [23]).

This paper considers a bilevel optimization problem in the form of optimization over an efficient set scenario. Also, in the lower-level problem, we examine vector function, a multiobjective optimization problem that makes an overall formulation called semivectorial bilevel optimization. This topic has been researched for years, with the main target being linear functions or convex ones in lower level or upper level [4,15–17,20,28–30]. Popular approaches were the penalty method, which converts the problem into a single-level optimization problem [10,20,41], and the branch-and-bound method, which solves iterated relaxed subproblems for boundings [4,15,28]. Recently, there have been some nonconvex ones, such as [27,32], but the algorithms are all developed using heuristic methods. Our research considers pseudoconvex objective functions on both levels and proposes a novel algorithm. By utilizing the nice properties of functions, the algorithm is developed based on monotonic optimization combined with a recent neurodynamic approach, where the solution set of the lower-level problem is inner approximated by copolyblocks in outcome space. Afterwards, the branch-and-bound technique is used to resolve the upper-level problem. The neurodynamic approach is then used to solve the resulting pseudoconvex programming problem for finding the bounds.

In the next section, we present the equivalent outcome space problem of the considered one. Section 3 gives some mathematics bases for building the proposed algorithm. Section 4 presents the proposed algorithm. We prove the algorithm's convergence in Section 5 before showing some special cases of the upper-level objective function. Computational experiments are shown in Section 6, and the conclusion is ended in the last section.

2. Preliminaries

2.1. Notations and definitions

The positive orthant cone in p -dimension space and its interior are denoted as $\mathbb{R}_+^m := \{x \in \mathbb{R}^p \mid x \geq 0\}$ and $\text{int } \mathbb{R}_+^m$, respectively. Let's recall the definition of *nondominated point* by considering two vectors $a, b \in \mathbb{R}^p$, where $a \leq b$ if all components $a_i \leq b_i$, $i = 1, 2, \dots, p$ and similarly $a < b$ if $a_i < b_i$ for all $i = 1, 2, \dots, p$. Let $Q \subset \mathbb{R}^p$ be some nonempty set. A point q in Q is considered an *nondominated point* if there is no other point q' in Q such that

q' is less than or equal to q . Similarly, a point q in Q is considered a *weakly nondominated point* if there is no other point q' in Q such that q' is less than q . The set of all nondominated points and weakly nondominated points of Q are denoted $\text{Min}Q$ and $\text{WMin}Q$, respectively.

Definition 2.1 (Clarke [9]): Suppose φ is a function that maps from \mathbb{R}^n to \mathbb{R} and is locally Lipschitz near a point x in \mathbb{R}^n . We can define the generalized directional derivative of φ at x , in the direction of a vector v in \mathbb{R}^n , as follows

$$\varphi^\circ(x; v) = \limsup_{y \rightarrow x, t \downarrow 0} \frac{\varphi(y + tv) - \varphi(y)}{t}.$$

The Clarke's generalized subgradient of φ at x is given by $\partial\varphi(x) = \{\zeta \in \mathbb{R}^n : \varphi^\circ(x; v) \geq \zeta^T v, \text{ for all } v \text{ in } \mathbb{R}^n\}$.

Definition 2.2 (Clarke [9]): Suppose that $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ is locally Lipschitz near $x \in \mathbb{R}^n$. The one-side directional derivative for any direction $v \in \mathbb{R}^n$ is defined by

$$\varphi'(x; v) = \lim_{t \downarrow 0} \frac{\varphi(x + tv) - \varphi(x)}{t},$$

and we say φ is regular at x , if the one-side directional derivative exists and for all $v \in \mathbb{R}^n$, $\varphi^\circ(x; v) = \varphi'(x; v)$. Moreover, φ is considered regular on \mathbb{R}^n provided φ is regular at any $x \in \mathbb{R}^n$.

Definition 2.3 (Clarke [9]): Let $S \subseteq \mathbb{R}^n$ be a convex set. A function $\varphi : S \rightarrow \mathbb{R}$ is said to be a real-valued convex function provided that, for all $x, x' \in S$ and $\lambda \in [0, 1]$, one has $\varphi(\lambda x + (1 - \lambda)x') \leq \lambda\varphi(x) + (1 - \lambda)\varphi(x')$. If $\varphi : S \rightarrow \mathbb{R}$ is convex, then $\partial\varphi(x) = \{\zeta \in \mathbb{R}^n : \varphi(x) - \varphi(x') \leq \zeta^T(x - x'), \text{ for all } x' \in \mathbb{R}^n\}$.

It's important to know that convex functions are regular, as stated in Proposition 2.3.6 in Clarke's work [9].

Definition 2.4 (Cambini & Martein [1]): Let $G : \mathbb{R}^n \rightarrow \mathbb{R}$ be locally Lipschitz, then

- (1) G is pseudoconvex if for any $x, \tilde{x} \in \mathbb{R}^n$, $\exists \zeta \in \partial G(x) : \zeta^T(\tilde{x} - x) \geq 0 \Rightarrow G(\tilde{x}) \geq G(x)$;
- (2) G is quasiconvex if for any $x, \tilde{x} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$, $G(\lambda x + (1 - \lambda)\tilde{x}) \leq \max\{G(x), G(\tilde{x})\}$.

Theorem 2.5 (Cambini & Martein [1]): Let $z(x) = \frac{f(x)}{g(x)}$ be the ratio of two differentiable functions f and g defined on an open convex set $S \subset \mathbb{R}^n$.

- (1) If f is a convex function and g is both positive and affine, then z is a pseudoconvex function;
- (2) If f is non-negative and convex, and g is positive and concave, then z is pseudoconvex;

Theorem 2.6 (Carosi & Martein [8]): Consider the ratio function $h(x) = \frac{1}{2} \frac{x^T A x}{(b^T x + b_0)^3}$ where A is a non-zero $n \times n$ symmetric matrix, $b \in \mathbb{R}^n$ and $b_0 \in \mathbb{R}$. The function h is pseudoconvex on $S = \{x \in \mathbb{R}^n : b^T x + b_0 > 0\}$ if and only if h is of the following form:

$$h(x) = \frac{1}{2} \frac{\mu (b^T x)^2}{(b^T x + b_0)^3} \quad \text{where } b_0 < 0.$$

2.2. Problem formulation

The pseudoconvex semivectorial bilevel optimization problem under consideration can be expressed in the following manner

$$\begin{aligned} \min \quad & h(x, y) \\ \text{s.t.} \quad & g(x, y) \leq 0, y \in \mathbb{R}_+^m, \\ & x \in \text{Argmin} \{f(x) \mid x \in X\}, \end{aligned} \tag{BP}$$

where it will be assumed that

- (A1) $X = \{x \in \mathbb{R}^n \mid s(x) \leq 0\}$ is a nonempty bounded and convex set;
- (A2) The real-valued function $h : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous and the vector-valued functions $g : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^\ell, f : \mathbb{R}^n \rightarrow \mathbb{R}^p, s : \mathbb{R}^n \rightarrow \mathbb{R}^q$ are continuously differentiable, where the integer numbers $m, n, \ell, p, q \geq 2$;
- (A3) The objective functions h, f are pseudoconvex and regular;
- (A4) The constraint functions g, s are quasiconvex.

In the above formulation, the scalar function h is called the upper-level objective function, and the vectorial function $f = (f_1, f_2, \dots, f_p)$ is called the lower-level objective function. Recall that for the vector optimization problem

$$\begin{aligned} \text{Min} \quad & f(x) \\ \text{s.t.} \quad & x \in X, \end{aligned}$$

a feasible solution \bar{x} is said to be an efficient solution (resp., weakly efficient solution) of this problem if there is no solution $x \in X$ such that $f(x) \leq f(\bar{x})$ and $f(\bar{x}) \neq f(x)$ (resp., $f(x) < f(\bar{x})$) (see [36]). For convenience, as usual, we denote X_{WE} as the weakly efficient solution set of the vector optimization problem $\text{Min}\{f(x) \mid x \in X\}$. Therefore, the constraint $x \in \text{Argmin}\{f(x) \mid x \in X\}$ can be replaced by $x \in X_{WE}$.

Problem (BP) covers a large class of complex nonconvex optimization problems, such as nonconvex multiplicative programming, optimization over the efficient set, and meta-learning problems in machine learning [11,14,25,38]. The restricted form of (BP) for optimizing over the efficient set has been studied in some previous work such as [19,33–35,39] with the effective algorithms.

2.3. Neurodynamic method for solving nonsmooth pseudoconvex programming problem

We now consider a pseudoconvex programming problem as follows:

$$\begin{aligned} \min \quad & r(x) \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{SQ}$$

where the objective function $r(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ is nonsmooth pseudoconvex, the constraint set $X = \{x \in \mathbb{R}^n \mid s(x) \leq 0\}$, $s(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $s_i (i = 1, \dots, m) : \mathbb{R}^n \rightarrow \mathbb{R}$ are quasi-convex and differentiable.

To tackle this problem, we employ the neurodynamic approach introduced by Liu et al. [22]. This method differs from others such as [6,21,26,42] because it permits using quasiconvex functions for inequality constraints. Even though the zero-level sets of both quasiconvex and convex functions are convex and a quasiconvex constraint can be substituted with an equivalent convex one, identifying such a replacement is typically not straightforward.

To solve Problem (SQ), Liu initially defines the set-valued function $\Psi : \mathbb{R} \rightrightarrows [0, 1]$, which is characterized by:

$$\Psi(\xi) = \begin{cases} 1, & \xi > 0; \\ [0, 1], & \xi = 0; \\ 0, & \xi < 0. \end{cases} \tag{1}$$

In fact, $\Psi(\xi) = \partial \max\{0, \xi\}$, and Ψ is upper semicontinuous on \mathbb{R} . Define the following function:

$$S_m(x) = \sum_{i=1}^m \max\{0, s_i(x)\}, \tag{2}$$

where $s_i(x)$, $i = 1, 2, \dots, m$ are given in (SQ). The closed-form for $\partial S_m(x)$ is derived as

$$\partial S_m(x) = \begin{cases} 0, & x \in \text{int}(X); \\ \sum_{i \in I_0(x)} [0, 1] \partial s_i(x), & x \in \text{bd}(X); \\ \sum_{i \in I_+(x)} \partial s_i(x) + \sum_{i \in I_0(x)} [0, 1] \partial s_i(x), & x \notin X, \end{cases}$$

where $I_+(x) = \{i \in \{1, 2, \dots, m\} : s_i(x) > 0\}$, $I_0(x) = \{i \in \{1, 2, \dots, m\} : s_i(x) = 0\}$, $\text{int}(X)$ and $\text{bd}(X)$ denote the interior and boundary of the feasible set X , respectively.

The neurodynamic model for solving Problem (SQ) is described in the form of a dynamic system as follows:

$$\begin{cases} x(0) & \in X; \\ \frac{d}{dt}x(t) & \in -c(x(t))\partial r(x(t)) - \partial S(x(t)), \end{cases} \tag{3}$$

where

$$c(x(t)) = \left\{ \prod_{i=1}^m c_i(t) \mid c_i(t) \in 1 - \Psi(s_i(x(t))), i = 1, 2, \dots, m \right\}. \tag{4}$$

In the neurodynamic model (3), the state $x(0)$ is any initial state within X . The second expression is specified from the gradient descent method where the $-\partial S(x(t))$ term serves to steer the state towards the feasible region X , while the term $-\partial r(x(t))$ causes the state to follow a descent direction of the objective function r . The $c(x(t))$ term adjusts the step-size of the state in the model (3). If the state is outside of X , the model will direct it towards X . However, if the state is within X , the model will direct it towards a feasible descent direction of r .

Theorem 2.7 ([22, Theorem 4.3]): *The state $x(t)$ of neural network (3) with any $x(0) \in X$, converges to an optimal solution of Problem (SQ).*

3. The equivalent problem in outcome space

We denote the outcome set $\mathcal{Z} := \{z \in \mathbb{R}^p \mid z = f(x), x \in X\}$ and $\mathcal{G} = \{(x, y) \mid x \in X, y \in \mathbb{R}_+^m, g(x, y) \leq 0\}$. Since X is bounded and f consists of continuous functions, \mathcal{Z} is bounded. We thus can find a box $[\omega, \Omega]$ which contains \mathcal{Z} , i.e. $\omega \leq \mathcal{Z} \leq \Omega$. To determine ω , it is sufficient to solve the following problems for each component $\omega_i, i = 1, 2, \dots, p$,

$$\min f_i(x), \quad \text{s.t. } x \in X. \quad (P_i^\omega)$$

Since $f_i(x)$ is a pseudoconvex function, any local minimum of the above equation is also a global optimum [3]. It is straightforward to apply the neurodynamic model presented in Section 2.3 to (P_i^ω) as follows:

$$\begin{cases} x(0) & \in X; \\ \frac{d}{dt}x(t) & \in -c(x(t))\partial f_i(x(t)) - \partial S_m(x(t)), \end{cases} \quad \text{Solve } (P_i^\omega)$$

where $S_m(x)$, $c(x(t))$ are defined from (2) and (4). However, a similar reasoning cannot be applied to find Ω because the problem $\max\{f_i(x), \text{ s.t. } x \in X\}$ maximizing a pseudoconvex function over a convex set that is nonconvex. For that purpose, we bound X in a simplex Δ with vertex set $V(\Delta) = \{\Delta^0, \Delta^1, \dots, \Delta^n\}$, where $\Delta^0 = (\Delta_1^0, \Delta_2^0, \dots, \Delta_n^0) \in \mathbb{R}^n$ with $\Delta_i^0 = \min_{x \in X} x_i, i = 1, 2, \dots, n$ and $\Delta^i = (\Delta_1^i, \Delta_2^i, \dots, \Delta_n^i), i = 1, 2, \dots, n$ are defined by

$$\Delta_k^i = \begin{cases} \Delta_k^0, & \text{if } k \neq i; \\ U - \sum_{j \neq k} \Delta_j^0, & \text{if } k = i, \end{cases}$$

where U is the optimal value of the problem: $\max\{\langle e, x \rangle \mid x \in X\}, e \in \mathbb{R}^n, e = (1, 1, \dots, 1)^T$. This definition, combined with the fact that f is a pseudoconvex function, leads to $X \subset \Delta$ and, therefore, a way to specify Ω , such as

$$\begin{aligned} \Omega_i &:= \max\{f_i(x) \mid x \in V(\Delta)\} \\ &= \max\{f_i(x) \mid x \in \Delta\} \\ &\geq \max\{f_i(x) \mid x \in X\}, \quad i = 1, 2, \dots, p. \end{aligned} \quad (P_i^\Omega)$$

We can see that \mathcal{Z} is not full-dimensional and nonconvex. We therefore define $\mathcal{Z}^+ := \mathcal{Z} + \mathbb{R}_+^p = \{z \in \mathbb{R}^p \mid \exists z^0 \in \mathcal{Z}, z^0 \leq z\}$ which is a full-dimensional convex set. One problem of

\mathcal{Z}^+ is that it is not bounded. We then bound the set \mathcal{Z} by another equivalently efficient set denoted as \mathcal{Z}^\diamond ,

$$\mathcal{Z}^\diamond := \mathcal{Z}^+ \cap (\Omega - \mathbb{R}_+^m).$$

The efficient equivalence of \mathcal{Z} , \mathcal{Z}^+ and \mathcal{Z}^\diamond is shown in Proposition 3.1 where $\text{Min } \mathcal{Z}$, $\text{Min } \mathcal{Z}^+$ and $\text{Min } \mathcal{Z}^\diamond$ denote the efficient sets of \mathcal{Z} , \mathcal{Z}^+ and \mathcal{Z}^\diamond , respectively.

Proposition 3.1 ([36]): *We have*

- (i) $\text{Min } \mathcal{Z} = \text{Min } \mathcal{Z}^+ = \text{Min } \mathcal{Z}^\diamond$;
- (ii) $\text{WMin } \mathcal{Z} = \text{WMin } \mathcal{Z}^+ \cap \mathcal{Z} = \text{WMin } \mathcal{Z}^\diamond \cap \mathcal{Z}$.

We will use some concepts of monotonic optimization presented in [36]. A real-valued function d is said to be increasing (decreasing) over $S \subseteq \mathbb{R}^n$ if and only if $d(x) \geq d(y)$ ($d(x) \leq d(y)$), $\forall x \in y + \mathbb{R}_+^n$. A set $Q \subset [\omega, \Omega]$ is called *normal* if $[\omega, z] \subset Q, \forall z \in Q$ and *conormal* if $[z, \Omega] \subset Q, \forall z \in Q$. The concept of normal sets and conormal sets in 2-dimensional space is illustrated in Figure 1.

It is obvious that \mathcal{Z}^\diamond is a conormal set. We then convert Problem (BP) into a monotonic problem by constructing a function $\varphi : \mathcal{Z}^\diamond \rightarrow \mathbb{R}$ as follows

$$\varphi(z) = \min\{h(x, y) \mid (x, y) \in \mathcal{G}, f(x) \leq z\}. \quad (\text{MP}(z))$$

As we can see, when z increases, $\varphi(z)$ decreases due to the expansion of decision space X , so φ is a decreasing function over the conormal set \mathcal{Z}^\diamond .

Proposition 3.2: *The monotonic Problem (BP) is equivalent to the outcome-space problem*

$$\min\{\varphi(z) \mid z \in \text{WMin } \mathcal{Z}^\diamond\}, \quad (\text{OP})$$

which means that (BP) is solved if and only if (OP) is solved.

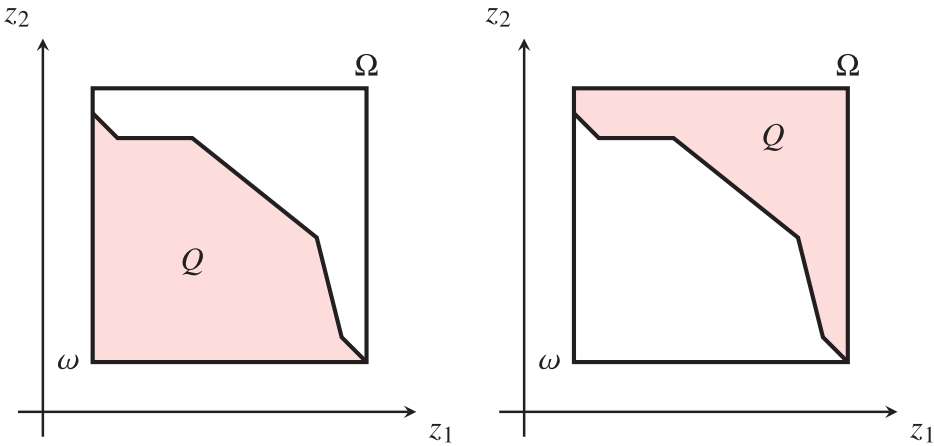


Figure 1. A normal set (left) and a conormal set (right) in 2D space.

Proof: As presented in Section 2.2, the problem (BP) is equivalent to

$$\min\{h(x, y) \mid g(x, y) \leq 0, x \in X_{WE}, y \in \mathbb{R}_+^m\},$$

whose constraints can be rewritten as,

$$\{x \in X, y \in \mathbb{R}_+^m \mid g(x, y) \leq 0, f(x) \in \text{WMin } \mathcal{Z}\}$$

which can be reduced as follows,

$$\begin{aligned} & \{(x, y) \in \mathcal{G} \mid f(x) \in \text{WMin } \mathcal{Z}\} \\ &= \{(x, y) \in \mathcal{G} \mid f(x) \in \text{WMin } \mathcal{Z}^\diamond \cap \mathcal{Z}\} \text{ (by Proposition 3.1(ii))} \\ &= \{(x, y) \in \mathcal{G} \mid f(x) \in \text{WMin } \mathcal{Z}^\diamond \text{ and } f(x) \in \mathcal{Z}\}, \\ &= \{(x, y) \in \mathcal{G} \mid f(x) \in \text{WMin } \mathcal{Z}^\diamond\} \end{aligned} \quad (*)$$

On the one hand, suppose \bar{z} is optimal solution of $\min\{\varphi(z) \mid z \in \text{WMin } \mathcal{Z}^\diamond\}$ and (\bar{x}, \bar{y}) is optimal solution of $MP(\bar{z})$. Since $\varphi(\bar{z})$ reaches its optimal value at (\bar{x}, \bar{y}) , we get

$$\begin{cases} \varphi(\bar{z}) = h(\bar{x}, \bar{y}); \\ f(\bar{x}) \leq \bar{z}, \quad \bar{z} \in \text{WMin } \mathcal{Z}^\diamond; \\ h(\bar{x}, \bar{y}) \leq h(x, y) \quad \forall (x, y) \in \mathcal{G}, f(x) \leq \bar{z}. \end{cases} \quad (5)$$

From the fact that $\bar{z} \in \text{WMin } \mathcal{Z}^\diamond$ and $f(x) \leq \bar{z}$, it infers $f(x) \in \text{Wmin } \mathcal{Z}^\diamond$. Thus, we have

$$h(\bar{x}, \bar{y}) \leq h(x, y) \quad \forall (x, y) \in \mathcal{G}, f(x) \in \text{Wmin } \mathcal{Z}^\diamond.$$

Therefore, by (*),

$$h(\bar{x}, \bar{y}) \leq h(x, y) \quad \forall (x, y) \in \mathcal{G}, f(x) \in \text{Wmin } \mathcal{Z},$$

which means that (\bar{x}, \bar{y}) is also optimal solution of (BP).

On the other hand, suppose we have a solution (\bar{x}, \bar{y}) of Problem (BP), we then have $f(\bar{x}) \in \text{Wmin } \mathcal{Z}^\diamond$ and $\varphi(z)$ reaches its optimal value if and only if $z = f(\bar{x})$ implying $\min \varphi(z) = \varphi(f(\bar{x}))$ implying $z = f(\bar{x})$ is an optimal solution to (OP). ■

4. The branch-and-bound scheme for solving Problem (OP)

4.1. Cutting cones and inner approximations of \mathcal{Z}^\diamond

In our previous work [36], it was established that the union of an arbitrary union of normal (respectively, conormal) sets is itself a normal (respectively, conormal) set. The union of all normal (respectively, conormal) sets contained in Q is denoted as the normal (respectively, conormal) hull of Q , represented by $\mathcal{N}(Q)$ (respectively, $\mathcal{M}(Q)$), which is also the minimal conormal set containing Q . From the definition, we can see that $\mathcal{N}(Q) := (Q - \mathbb{R}_+^p) \cap [\omega, \Omega]$ and $\mathcal{M}(Q) = (Q + \mathbb{R}_+^p) \cap [\omega, \Omega]$.

A polyblock \mathcal{P} is defined as the normal hull of a finite set of vertices V within the interval $[\omega, \Omega]$, such that $\mathcal{P} = \bigcup_{v \in V} [\omega, v]$ or equivalently $\mathcal{P} = \mathcal{N}(V)$. A copolyblock \mathcal{P} is defined as the conormal hull of a finite set of vertices V within the interval $[\omega, \Omega]$, such that $\mathcal{P} =$

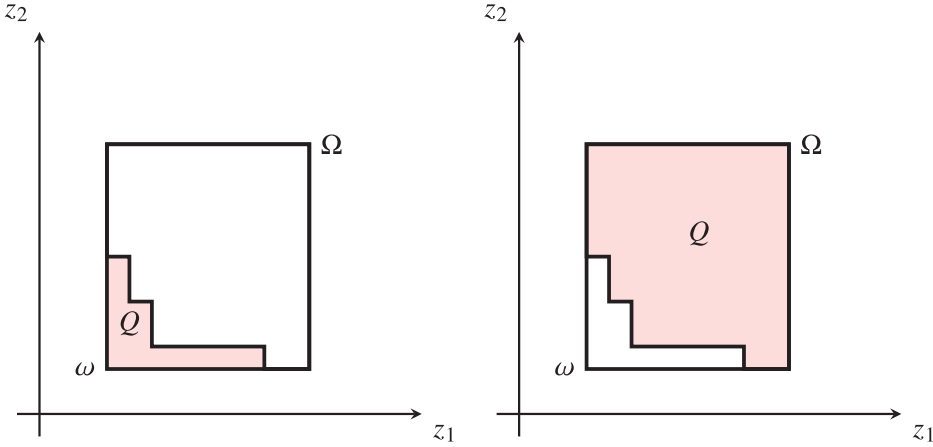


Figure 2. A polyblock (left) and a copolyblock (right) in 2D space.

$\bigcup_{v \in V} [v, \Omega]$ or equivalently $\mathcal{P} = \mathcal{M}(V)$. The concept of polyblocks and copolyblocks in 2-dimensional space is illustrated in Figure 2.

Let v be a vertex in $\mathcal{N}(Q)$, v is *proper* if $\nexists v' \in \mathcal{N}(Q)$ such that $v' \neq v$ and $v' \geq v$. An *improper* vertex in $\mathcal{N}(Q)$ is not proper. A polyblock is determined by its set of proper vertices as it is the normal hull of its proper vertices. Similarly, a copolyblock is the conormal hull of its proper vertices where a proper vertex v satisfies $\nexists v' \in \mathcal{M}(Q)$ such that $v' \neq v$ and $v' \leq v$.

In our research, we developed an inner approximation algorithm, so instead of determining an outer copolyblock by its proper vertices V , we regard the proper vertices as ones of the opposite polyblock and construct the inner copolyblock as follows

$$\mathcal{L}(V) := [\omega, \Omega] \setminus \text{int}(\mathcal{N}(V) - \mathbb{R}_+^p).$$

For convenience, throughout our paper, copolyblock is understood as one built in this manner, the inner copolyblock $\mathcal{L}(V)$ constructed above generates new proper vertices, and initial proper vertices V are now called *co-proper* vertices. Now, we recall some main properties of copolyblocks in the following proposition.

Proposition 4.1 ([36]): (i) *A finite union of copolyblocks is a copolyblock.*
(ii) *The union of a finite number of conormal sets forms a compact conormal set.*
(iii) *The intersection of a finite number of copolyblocks forms a compact conormal set.*

Proposition 4.2: *Given $v \in \mathcal{Z}^\diamond$ and w_v determined by Remark 4.1, the new copolyblock \mathcal{P}' obtained by applying the cutting cone of \mathcal{Z}^\diamond on \mathcal{P} at w_v has vertex set V' , where \mathcal{P} is a copolyblock in the box $[\omega, \Omega]$ with co-proper vertex set V such that $\mathcal{P} \subseteq \mathcal{Z}^\diamond$, and*

$$V' = (V \setminus \{v\}) \cup \{v - (v_i - w_i)e^i\}, \quad i = 1, \dots, p.$$

According to Proposition 4.1(iii), any compact conormal set can be approximated as closely as desired by a copolyblock. As a result, a family of copolyblocks can be used to approximate the compact conormal set \mathcal{Z}^\diamond . In particular, Proposition 4.2 generates a

nested sequence of copolyblocks that inner-approximates the outcome set \mathcal{Z}^\diamond , such that $\mathcal{P}^0 \subset \mathcal{P}^1 \subset \mathcal{P}^2 \subset \dots \subset \mathcal{P}^k \subset \mathcal{P}^{k+1} \subset \dots \subset \mathcal{Z}^\diamond$, where the initial copolyblock $\mathcal{P}^0 = [\Omega, \Omega]$ is constructed by calculating Ω as in Section 3.

In each iteration, the process where copolyblock \mathcal{P}^{k+1} is generated from \mathcal{P}^k and the resulting inner approximation is described in Procedure CopolyblockCut,

Procedure 1: *CopolyblockCut*

Input: A copolyblock V^k or a set of its vertices, the current considering vertex v^k and the intersection w^k of the ray starting from v^k along direction \hat{d} and $\partial \mathcal{Z}^+$

Output: The new inner approximate outcome set V^{k+1}

```

1 Set  $V^{k+1} \leftarrow V^k \setminus \{v^k\}$ .
2 for  $i \leftarrow 1$  to  $p$  do
3    $z^i = v^k - (v_i^k - w_i^k)e^i$ ;
4   if  $z_i^i \neq \omega_i$  then
5      $V^{k+1} \leftarrow (V^{k+1} \cup \{z^i\})$ .
```

4.2. Brand-and-bound algorithm scheme

Utilizing the outcome space approach, the solution of (BP) is attained by enhancing upper and lower bounds for the objective function subsequent to each iteration. Furthermore, the outcome space is approximated recurrently via the cutting cone methodology on inner copolyblocks. Commencing with copolyblock $\mathcal{P}^0 = [\Omega, \Omega]$, we construct a sequence of copolyblocks \mathcal{P}^k iteratively such that

$$\mathcal{P}^0 \subset \mathcal{P}^1 \subset \mathcal{P}^2 \subset \dots \subset \mathcal{P}^k \subset \mathcal{P}^{k+1} \subset \dots \subset \mathcal{Z}^\diamond.$$

The subsequent notations shall be employed:

- The set of all co-proper vertices is denoted by V^k which defines the copolyblock $\mathcal{P}^k = \mathcal{L}(V^k)$.
- The upper and lower bounds are denoted by α_k and β_k , respectively.

At the initial step with $k = 0$, we have $V^0 = \{\Omega\}$, $\mathcal{P}^0 = [\Omega, \Omega]$, $\alpha_0 = +\infty$. In a typical iteration k , by solving $\min\{\varphi(v) \mid v \in V^k\}$ the lower bound β_k is assigned, then we determine v^k such that $\beta_k = \varphi(v^k)$. By solving $(P^2(v^k))$, we find a new weakly efficient point $z^k = f(x^k)$ of the lower problem, after that we find solution y^k for problem (BP) with the x^k found, if the problem is feasible, we will obtain a solution (x^k, y^k) . After that we compare $h(x^k, y^k)$ to α_k to update upper bound. If α_k and β_k satisfy the terminate condition, meaning that the upper bound is close enough to the lower bound, the algorithm stops and returns $((x^k, y^k), h(x^k, y^k))$. Otherwise, a new co-proper vertices set V^{k+1} is created by procedure CopolyblockCut, before new inner approximate outcome set $\mathcal{P}^{k+1} = \mathcal{L}(V^{k+1})$ is determined, and the next iteration's upper bound is assigned.

4.3. Determining the upper bounds

A weakly nondominated point of \mathcal{Z}^+ can be easily determined by the following remark.

Remark 4.1: Let fix a vector $\hat{d} > 0$ in \mathbb{R}^n and v an arbitrary point in \mathbb{R}^p . Then the intersection w_v of the line through v along direction \hat{d} can be determined by

$$w_v = v + t_v \hat{d} \quad (6)$$

where t_v is the optimal value of the following problem

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & v + t\hat{d} \in \mathcal{Z}^+, \quad t \in \mathbb{R}. \end{aligned} \quad (P^0(v))$$

Lemma 4.3 shows that $w_v := v + t_v \hat{d}$ is a weakly nondominated point of \mathcal{Z}^+ .

Lemma 4.3 ([36, Lemma 2.1]): *For any point v in \mathbb{R}^p . Then there exists the unique point w_v determined by (6) that is a weakly nondominated point of \mathcal{Z}^+ .*

Lemma 4.4 (Clarke [9], Chain rule): *Let $V(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ be a regular function and $x(t) : \mathbb{R} \rightarrow \mathbb{R}^n$ be Lipschitz near t and differentiable at t . Then, for almost every $t \in [0, +\infty)$, it holds that $\dot{V}(x(t)) = \zeta^T \dot{x}(t)$ for all $\zeta \in \partial V(x(t))$.*

Lemma 4.5 (Clarke [9]): *Let $\{f_i, i = 1, 2, 3, \dots, n\}$ be a finite set of functions that are Lipschitz near x and regular. Define $\psi(x) = \max\{f_i(x), i = 1, 2, 3, \dots, n\}$. Then the subdifferential of $\psi(x)$ is given by $\partial\psi(x) = \text{conv}\{\partial f_i(x), i \in I(x)\}$ where $I(x)$ is the set of indices i such that $f_i(x) = \psi(x)$ and ‘conv’ denotes the convex hull of a set.*

Lemma 4.6: *Given $F(x) = \max\{f_i(x) \mid i = 1, \dots, n\}$ with $f_i(x) : S \rightarrow \mathbb{R}$ being pseudoconvex, continuous differential functions. Then $F(x)$ is also a pseudoconvex function.*

Proof: Since $f_i(x), i = \overline{1, n}$ are locally Lipschitz so their subgradient can be expressed as $\partial f_i(x) = \{\nabla f_i(x)\}$. By Lemma 4.5, we have

$$\partial F(x) = \text{conv}\{\partial f_i(x), f_i(x) = F(x)\}.$$

By Definition 2.4, for any $x, \tilde{x} \in S$, $\exists \xi_i \in \partial f_i(x) : \xi_i^T (\tilde{x} - x) \geq 0 \Rightarrow f_i(\tilde{x}) \geq f_i(x)$. Set $I(x) = \{i \in \{1, \dots, n\} \mid f_i(x) = F(x)\}$, we have

$$\sum_{i \in I(x)} \mu_i \xi_i^T (\tilde{x} - x) \geq 0, \quad \sum_{i \in I(x)} \mu_i = 1,$$

and

$$F(\tilde{x}) \geq F(x).$$

Therefore, we conclude that $F(x)$ is also a pseudoconvex function by Definition 2.4. ■

We can rewrite $(P^0(v))$ in its explicit form

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & f(x) - t\hat{d} - v \leq 0, \\ & x \in X, t \in \mathbb{R}. \end{aligned} \tag{P^1(v)}$$

It can be seen that $(P^1(v))$ is, in general, nonconvex. However, it is equivalent to the following problem

$$\begin{aligned} \min \quad & \max \left\{ \frac{f_j(x) - v_j}{\hat{d}_j} \mid j = 1, \dots, p \right\} \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{P^2(v)}$$

which is solvable as the objective function is pseudoconvex according to Lemma 4.6, and the feasible set is convex.

Lemma 4.7 ([36, Lemma 2.2]): *It can be inferred that problems $(P^0(v))$ and $(P^2(v))$ are equivalent. In other words, if Problem $(P^0(v))$ has an optimal solution (x^*, t^*) , then x^* is the optimal solution of Problem $(P^2(v))$. Conversely, $(P^2(v))$ has an optimal solution x^* with the corresponding optimal value t^* , then (x^*, t^*) is the optimal solution of $(P^0(v))$. Additionally, it should be noted that Problem $(P^2(v))$ is a pseudoconvex programming problem.*

By Lemma 4.7, finding a Pareto solution to Problem (BP) is transformed into finding the optimal solution to a pseudoconvex programming problem. The gradient descent technique and the neurodynamic method are two ways to solve quasiconvex programming problems. Since the objective function of $(P^2(v))$ is nonsmooth even though the function f is smooth, the neurodynamic approach is used to tackle this problem.

For that, we calculate the subgradient of $\max \left\{ \frac{f_j(x) - v_j}{\hat{d}_j} \mid j = 1, \dots, p \right\}$ with the help of Lemmas 4.4 and 4.5.

The dynamic system of the problem can then be formulated analogously to the problem in Section 2.3 by applying the theorems directly to $(P^2(v))$ as follows,

$$\begin{aligned} \partial \max \left\{ \frac{f_j(x) - v_j}{\hat{d}_j} \mid j = 1, \dots, p \right\} &= \text{conv} \left\{ \frac{\partial f_j(x) - v_j}{\hat{d}_j} \mid j \in I(x) \right\} \\ &= \text{conv} \left\{ \frac{\partial f_j(x(t))^T \dot{x}(t)}{\hat{d}_j} \mid j \in I(x) \right\}. \end{aligned}$$

The second equal sign can be derived directly from Lemma 4.4 and the fact that $\partial \frac{v_j}{\hat{d}_j} = \{0\} \forall j \in I(x)$. We then attain the dynamic model for solving $(P^2(v))$,

$$\begin{cases} x(0) \in X; \\ \frac{d}{dt}x(t) \in -c(x(t)) \times \text{conv} \left\{ \frac{\partial f_j(x(t))^T \dot{x}(t)}{\hat{d}_j} \mid j \in I(x) \right\} - \partial S_m(x(t)), \end{cases} \tag{NDP^2(v)}$$

where $c(x(t))$, Ψ and $S_m(x(t))$ is defined in (1), (2) and (4), respectively.

4.4. Determining the lower bounds

Since $h(x, y)$ is a pseudoconvex function, the subproblem of finding

$$\beta_k = \min \left\{ \varphi(v) \mid v \in V^k \right\}, \quad (LB(V^k))$$

in the proposed algorithm requires solving $(MP(z))$ where f_i are quasiconvex functions (by pseudoconvexity) and \mathcal{G} is a nonempty compact convex set with the quasiconvexity of g in it. These problems can be handled by the neurodynamic model mentioned in Section 2.3. We first define function $\tilde{G}(u)$ as,

$$\tilde{G}(u) = (s_1(x), \dots, s_m(x), -y, f_1(x) - z_1, \dots, f_p(x) - z_p, g_1(x, y), \dots, g_l(x, y))^T, \quad (7)$$

where $u = (x, y)^T \in \mathbb{R}^{n+1}$ and $u \in \tilde{\mathcal{G}} := \mathcal{G} \cup \{x \in X \mid f(x) \leq z\}$. Analogously to Section 2.3, we then introduce a function \tilde{G}_m ,

$$\tilde{G}_m(u) = \sum_{i=1}^{m+p+l+1} \max\{0, \tilde{G}_i(u)\}. \quad (8)$$

We can calculate subgradient of $\tilde{G}_m(u)$ as

$$\partial \tilde{G}_m(u) = \begin{cases} 0, & u \in \text{int}(\tilde{\mathcal{G}}); \\ \sum_{i \in I_0(u)} [0, 1] \partial \tilde{G}_i(u), & u \in \text{bd}(\tilde{\mathcal{G}}); \\ \sum_{i \in I_+(u)} \partial s_i(u) + \sum_{i \in I_0(u)} [0, 1] \partial \tilde{G}_i(u), & u \notin \tilde{\mathcal{G}}, \end{cases}$$

where $I_+(u) = \{i \in \{1, 2, \dots, m\} : \tilde{G}_i(u) > 0\}$, $I_0(u) = \{i \in \{1, 2, \dots, m\} : \tilde{G}_i(u) = 0\}$, $\text{int}(\tilde{\mathcal{G}})$ and $\text{bd}(\tilde{\mathcal{G}})$ denote the interior and boundary of the feasible set $\tilde{\mathcal{G}}$, respectively.

The neurodynamic model for solving Problem $(MP(z))$ are constructed as,

$$\begin{cases} u(0) \in \tilde{\mathcal{G}}; \\ \frac{d}{dt} u(t) \in -c(u(t)) \partial h(u) - \partial \tilde{G}_m(u(t)), \end{cases} \quad (NDMP(z))$$

where

$$c(u(t)) = \left\{ \prod_{i=1}^m c_i(t) \mid c_i(t) \in 1 - \Psi(\tilde{G}_i(u(t))), i = 1, 2, \dots, m \right\}, \quad (9)$$

where Ψ is defined in (1).

4.5. The description of the proposed algorithm

We aim to find the approximate solutions to (BP) and (OP). Let a small tolerance $\varepsilon > 0$, a point $z^* \in W \min \mathcal{Z}^\diamond$ is called an ε -optimal solution to Problem (OP) if there exists an upper bound α^* for Problem (OP) such that $\alpha^* - \varphi(z^*) < \varepsilon(1 + |\varphi(z^*)|)$. Any $(x^*, y^*) \in X_{WE} \times \mathbb{R}_+^m$ which is an optimal solution of $(MP(z^*))$ is called an approximate optimal solution to Problem (BP). Below is a detailed description of Algorithm *Solve* (BP) to find the approximate optimal solution (x^*, y^*) to Problem (BP).

Algorithm: Solve (BP)**Input:** A pseudoconvex problem in form of (BP)**Output:** The approximate optimal solution with arbitrary error threshold

```

1 Choose a sufficient small tolerance level  $\varepsilon > 0$ . Solve problems  $(P_i^\omega)$  and  $(P_i^{\Omega})$ ,
   $i = 1, \dots, p$  to determine the box  $[\omega, \Omega]$ . Set  $\mathcal{P}^0 \leftarrow [\Omega, \Omega]$ ,  $V^0 \leftarrow \{\Omega\}$  and
  choose a direction  $\hat{d} \in \mathbb{R}_+^p$  (e.g.  $\hat{d} = e$ ). for  $i \leftarrow 1$  to  $p$  do
2    $z^i = \Omega - (\Omega_i - \omega_i)e^i$ , solve (NDMP( $z$ )) with  $z = z^i$ ;
3   if (MP( $z^i$ )) has a solution  $(x^i, y^i)$  then
4      $\alpha^i = h(x^i, y^i)$ .
5   else
6      $\alpha^i = \infty$ .
7 Set initial upper bound  $\alpha_0 \leftarrow \min\{\alpha^i \mid i = 1, 2, \dots, p\}$ ,  $k \leftarrow 0$  and boolean
   $update \leftarrow False$ .
8 Set current best solution  $(x^*, y^*)$  corresponding to problem (MP( $z^i$ )) which
  satisfies  $\varphi(z^i) = \alpha_0$ .
9 for  $k \leftarrow 1$  to  $\infty$  do
10   foreach  $v \in V^k$  do
11     Solve (NDMP( $z$ )) with  $z = v$ ;
12     if (MP( $v$ )) has a solution  $(x^*, y^*)$  then
13       else
14         if  $\nexists i : f_i(x^*) = \omega_i$  then
15            $\beta_v \leftarrow \varphi(v)$ .
16         else
17            $V^k \leftarrow V^k \setminus \{v\}$ .
18          $V^k \leftarrow V^k \setminus \{v\}$ .
19   Solve (LB( $V^k$ )) to get a new lower bound  $\beta_k$  and  $v^k \in V^k$  such that
     $\varphi(v^k) = \beta_k$ ; Solve problem  $(P^2(v^k))$  to find an optimal solution  $(x^k, t_k)$  and
    set  $w^k \leftarrow v^k + t_k \hat{d}$ ;  $z^k \leftarrow f(x^k)$ ;
20   if  $w^k = z^k$  and  $w_i^k > \omega_i, \forall i$  then
21     Find a feasible  $y^k$  satisfying  $(x^k, y^k) \in \mathcal{G}$ ;
22     if  $y^k$  is found then
23        $update \leftarrow True$ .
24     else
25        $update \leftarrow False$ .
26   if  $update$  and  $h(x^k, y^k) < \alpha_k$  then
27     Update the upper bound  $\alpha_k \leftarrow h(x^k, y^k)$ ,  $x^* \leftarrow x^k$ ,  $y^* \leftarrow y^k$ .
28   if  $\alpha_k - \beta_k \leq \varepsilon(1 + |\beta_k|)$  then
29     Terminate.
30   else
31     Determine the new set  $V^{k+1}$  by using Procedure 1.
32   Determine the new inner approximate outcome set  $\mathcal{P}^{k+1} \leftarrow \mathcal{L}(V^{k+1})$ ;
     $\alpha_{k+1} \leftarrow \alpha_k$ .

```

5. The convergence of the proposed algorithm

The convergence of the proposed algorithm, when k is sufficiently large, is proven through the following lemmas.

Lemma 5.1: *The number k tends to infinity and*

$$\lim_{k \rightarrow \infty} \max_{v \in V^k} \|w_v - v\| = 0,$$

where V^k denotes the set of all proper vertices of \mathcal{P}^k and w_v denotes the corresponding weakly nondominated point of \mathcal{Z}^+ obtained by solving Problem $(P^2(v))$.

Proof: Consider a vertex $v^k \in \mathcal{P}^k$ chosen at the k th iteration and the optimal value t_k of $(P^2(v^k))$. As in (6), let $w_{v^k} = v^k + t_k \hat{d}$, we have

$$\text{Vol}[v^k, w_{v^k}] = (t_k)^p \text{Vol}[0, \hat{d}]. \quad (10)$$

The lemma stays valid if $\max_{v \in V^k} \|w_v - v\| = 0$ at some $k \geq 0$. Otherwise, there exists $v^k \in V^k$ such that $\|w_{v^k} - v^k\| = \max_{v \in V^k} \|w_v - v\| > 0$. We also have $\mathcal{P}^k \subseteq \mathcal{P}^{k+1} \setminus (v^k - \text{int } \mathbb{R}_+^m)$, since $[v^k, w_{v^k}] \subseteq \mathcal{P}^k$ deduced from the definition of w_{v^k} , the volume of \mathcal{P}^k satisfies

$$\text{Vol } \mathcal{P}^{k+1} - \text{Vol } \mathcal{P}^k \geq \text{Vol}[v^k, w_{v^k}]. \quad (11)$$

Combining (10) with (11), we obtain

$$\text{Vol } \mathcal{P}^{k+1} - \text{Vol } \mathcal{P}^k \geq (t_k)^p \text{Vol}[0, \hat{d}].$$

Therefore,

$$\sum_{i=0}^k (\text{Vol } \mathcal{P}^{i+1} - \text{Vol } \mathcal{P}^i) \geq \left(\sum_{i=0}^k (t_i)^p \right) \text{Vol}[0, \hat{d}].$$

We deduce

$$\text{Vol } \mathcal{Z}^\diamond \geq \text{Vol } \mathcal{P}^{k+1} \geq \text{Vol } \mathcal{P}^{k+1} - \text{Vol } \mathcal{P}^0 \geq \left(\sum_{i=0}^k (t_i)^p \right) \text{Vol}[0, \hat{d}],$$

for all $k \geq 1$. As k approaches infinity, the positive series $\sum_{i=0}^k (t_i)^p$ is upper bounded by $\text{Vol } \mathcal{Z}^\diamond / \text{Vol}[0, \hat{d}]$, which implies its convergence and $\lim_{i \rightarrow \infty} t_i = 0$. Since \hat{d} is bounded, for any $i \geq 1$, we have

$$\lim_{i \rightarrow \infty} \max_{v \in V^i} \|w_v - v\| = \lim_{i \rightarrow \infty} \|w_{v^i} - v^i\| = \lim_{i \rightarrow \infty} t_i \|\hat{d}\| = 0. \quad \blacksquare$$

Lemma 5.2: *If the solution (\bar{x}, \bar{y}) of Problem (BP) satisfies $\exists i : f_i(\bar{x}) = \omega_i$, then \bar{x} can be obtained by solving Problem $\varphi(\bar{z})$ with \bar{z} is the solution of the following problem $\min\{\varphi(z) \mid z \in \{\Omega - (\Omega_j - \omega_j)e^j \mid j = 1, 2, \dots, p\}\}$.*

Proof: Since (\bar{x}, \bar{y}) is the solution to the problem (BP), we have

$$h(\bar{x}, \bar{y}) \leq \varphi(z), \quad \forall z \in \{\Omega - (\Omega_j - \omega_j)e^j \mid j = 1, 2, \dots, p\}. \quad (12)$$

Consider $z^i = \Omega - (\Omega_i - \omega_i)e^i$, because $f_i(\bar{x}) = \omega_i$ and $f(\bar{x}) \leq \Omega$, we have $f(\bar{x}) \leq z^i$ and therefore \bar{x}, \bar{y} is a feasible solution of the problem

$$\varphi(z^i) = \min\{h(x, y) \mid (x, y) \in \mathcal{G}, f(x) \leq z^i\},$$

thus

$$\varphi(z^i) \leq h(\bar{x}, \bar{y}). \quad (13)$$

From (12) and (13), we must have

$$h(\bar{x}, \bar{y}) = \varphi(z^i).$$

Because of that, (\bar{x}, \bar{y}) can be obtained by selecting best solution after solving p problems $MP(z^j)$, $j = 1, 2, \dots, p$. ■

Lemma 5.3: For any $z \in \mathcal{Z}^\diamond$ such that $(MP(z))$ has a solution, then if we continuously solve $(P^2(v^k))$ with initial $v^0 \equiv z$ to obtain p new $v^k - (v_i^k - w_i^k)e^i$ points adding to V^{k+1} and remove any $v \in V^{k+1}$ from V^{k+1} if $MP(v)$ has no solution, two following statements will be true,

- (i) If $\nexists (\bar{x}, \bar{y}) \in \mathcal{G}$ such that $f(\bar{x}) \in \text{Min } \mathcal{Z} \cap (z - \mathbb{R}_+^m)$, we will obtain $V^k = \emptyset$ when k tends to infinity.
- (ii) If $\exists (\bar{x}, \bar{y}) \in \mathcal{G}$ such that $f(\bar{x}) \in \text{Min } \mathcal{Z} \cap (z - \mathbb{R}_+^m)$, then we can extract a sequence $\{u^k\}_{k=0}^\infty$, $u^k \in V^k$ with $u^0 \equiv z$ such that if w_{u^k} is the weakly nondominated point of \mathcal{Z}^+ induced by the solution (x^k, t_k) of problem $(P^2(u^k))$, when k tends to infinity

$$\lim_{k \rightarrow \infty} \|w_{u^k} - f(\bar{x})\| = 0.$$

Proof: (i) Since $\nexists (\bar{x}, \bar{y}) \in \mathcal{G}$ such that $f(\bar{x}) \in \text{Min } \mathcal{Z} \cap (z - \mathbb{R}_+^m)$, by denoting

$$\begin{aligned} A &= \text{Min } \mathcal{Z} \cap (z - \mathbb{R}_+^m), \\ B &= (z - \mathbb{R}_+^m) \cap \{f(x) \mid (x, y) \in \mathcal{G}\}, \end{aligned}$$

we have

$$A \cap B = \emptyset.$$

Therefore, we can assume that the infimum distance between these two set, which is Hausdorff distance, is a number $\varepsilon > 0$.

Also, because of Lemma 5.1, we can chose a number k such that

$$\lim_{k \rightarrow \infty} \max_{v \in V^k} \|w_v - v\| < \varepsilon.$$

Thus, since $w_v \in A$, $\forall v \in V^k$, it does not exist any $v \in V^k$ such that $\exists b \in B : b \leq v$ due to the Hausdorff distance $\varepsilon > 0$.

As a consequence, $\nexists (x, y) \in \mathcal{G}$ such that $f(x) \leq v$, $\forall v \in V^k$, hence $MP(v)$ attains no solution for any $v \in V^k$ and so $\forall v \in V^k$ is removed from V^k .

(ii) We assume some $v^k \in V^k$ satisfying $v^k > f(\bar{x})$ and consider p points

$$v^{k,i} = v^k - \left(v_i^k - w_{v^k,i} \right) e^i, \quad i = 1, 2, \dots, p.$$

We will prove that $\exists i : v_i^{k,i} \geq f_i(\bar{x})$, if not, we assume $v_i^{k,i} < f_i(\bar{x})$, $\forall i$. Besides, we also have $w_{v^k,i} \leq v_i^{k,i}$, $\forall i$, therefore $w_{v^k,i} < f_i(\bar{x})$, $\forall i$.

We can conclude that $w_{v^k} < f(\bar{x})$. However, this results in a contradiction since both w_{v^k} and $f(\bar{x})$ belong to $\text{WMin} \mathcal{Z}^\diamond$. Therefore, the assumption is false and we have

$$\exists i : v_i^{k,i} \geq f_i(\bar{x}).$$

Also, since $v^k > f(\bar{x})$ and $v_j^k = v_j^{k,i}$, $\forall j \neq i$, we have

$$\exists i : v^{k,i} \geq f(\bar{x}).$$

Given that $v^{k,i}$ is an element of V^{k+1} , we can use mathematical induction to derive a sequence $\{u^k\}_{k=0}^\infty$ where u^k is an element of V^k and u^0 is equal to z . This sequence satisfies the condition that $u^k > f(\bar{x})$ because $u^0 = z > f(\bar{x})$.

Now, we will prove that when k tends to infinity.

$$\lim_{k \rightarrow \infty} \|w_{u^k} - f(\bar{x})\| = 0.$$

Since $u^{k+1} = u^k - (u_i^k - w_{u^k,i})e^i$ with some i , we have $u^{k+1} < u^k$, thus sequence $\{u^k\}_{k=0}^\infty$ is a decreasing one, and therefore, it converges because of having a lower bound $f(\bar{x})$.

If $\{u^k\}_{k=0}^\infty$ converges at a point $u \in \text{Min} \mathcal{Z} \cap (z - \mathbb{R}_+^m)$ such that $u > f(\bar{x})$, we have a contradiction since $u, f(\bar{x}) \in \text{Min} \mathcal{Z} \cap (z - \mathbb{R}_+^m)$, hence

$$\lim_{k \rightarrow \infty} \|u^k - f(\bar{x})\| = 0.$$

Combining with Lemma 5.1, for any $\varepsilon > 0$, we can find a $u^t \in \{u^k\}_{k=0}^\infty$ such that $\|w_{u^t} - u^t\| < \frac{\varepsilon}{2}$ and $\|u^t - f(\bar{x})\| < \frac{\varepsilon}{2}$.

We have triangle inequality,

$$\|w_{u^t} - f(\bar{x})\| \leq \|w_{u^t} - u^t\| + \|u^t - f(\bar{x})\| < \varepsilon.$$

And therefore, we must have

$$\lim_{k \rightarrow \infty} \|w_{u^k} - f(\bar{x})\| = 0. \quad \blacksquare$$

Lemma 5.4: At the k th iteration, let w_{v^k} be the weakly nondominated point of \mathcal{Z}^+ induced by the solution (x^k, t_k) of Problem $(P^2(v^k))$, we consider the situation when $\nexists i : w_{v^k,i} = \omega_i$, when k tends to infinity

$$\lim_{k \rightarrow \infty} \|\varphi(f(x^k)) - \varphi(v^k)\| = 0.$$

Proof: Given that f, g and h are continuous functions with finite values and X is a nonempty compact convex set, it follows that φ is also a continuous function with finite values.

Besides, as a result of Lemma 5.1

$$\lim_{k \rightarrow \infty} \|w_{v^k} - v^k\| \leq \lim_{k \rightarrow \infty} \max_{v \in V^k} \|w_v - v\| = 0.$$

Since $\nexists i : w_{v^k, i} = \omega_i$, thus $w_{v^k} \in \text{Min } \mathcal{Z}$ and we have $w_{v^k} = v^k + t_k \hat{d} = f(x^k)$.

Because $\lim_{k \rightarrow \infty} \|w_{v^k} - v^k\| = 0$ and φ is a continuous function, we have

$$\lim_{k \rightarrow \infty} \|\varphi(w_{v^k}) - \varphi(v^k)\| = 0.$$

Therefore

$$\lim_{k \rightarrow \infty} \|\varphi(f(x^k)) - \varphi(v^k)\| = \lim_{k \rightarrow \infty} \|\varphi(w_{v^k}) - \varphi(v^k)\| = 0. \quad \blacksquare$$

Theorem 5.5: *If problem (OP) has an optimal solution, then for any given $\varepsilon > 0$, the algorithm will terminate after a finite number of iterations and return an ε -optimal solution to Problem (OP).*

Proof: According to Lemma 5.2, if the global solution (\bar{x}, \bar{y}) of Problem (BP) satisfies $\exists i : f_i(\bar{x}) = \omega_i$ which means $f(\bar{x}) \in \text{WMin } \mathcal{Z} \setminus \text{Min } \mathcal{Z}$, we will have \bar{x} obtained by the proposed algorithm.

On the other hand, when $f(\bar{x})$ is on $\text{Min } \mathcal{Z}$, since V^k is the set of vertices of the inner approximate outcome set, we must have some $v^k \in V^k$ such that $f(\bar{x}) \in \text{Min } \mathcal{Z} \cap (v^k - \mathbb{R}_+^m)$. Applying Lemma 5.3(ii), it exists a sequence pair (u^t, w_{u^t}) such that

$$\lim_{k \rightarrow \infty} \|w_{u^k} - f(\bar{x})\| = 0,$$

thus

$$\lim_{k \rightarrow \infty} \|\varphi(w_{u^k}) - \varphi(f(\bar{x}))\| = 0,$$

so we can find $k > 0$ such that

$$\|\varphi(w_{u^k}) - \varphi(f(\bar{x}))\| < \frac{\varepsilon}{2}.$$

Also, thanks to Lemma 5.4, we have

$$\lim_{k \rightarrow \infty} \|\varphi(w_{u^k}) - \varphi(u^k)\| = 0.$$

According to the construction of the bound α_k and β_k , we have

$$0 \leq \alpha_k - \beta_k = h(x^k, y^k) - \varphi(v^k) = \varphi(f(x^k)) - \varphi(v^k).$$

In case we chose $v^k = u^k$, such that

$$\|\varphi(f(\bar{x})) - \varphi(u^k)\| < \frac{\varepsilon}{2},$$

we will obtain

$$0 \leq \alpha_k - \beta_k$$

$$\begin{aligned}
&= \varphi(f(x^k)) - \varphi(f(u^k)) \\
&= \varphi(w_{u^k}) - \varphi(f(u^k)) \\
&\leq \|\varphi(w_{u^k}) - \varphi(f(\bar{x}))\| + \|\varphi(f(\bar{x})) - \varphi(u^k)\| \\
&< \varepsilon.
\end{aligned}$$

Furthermore, Lemma 5.3(i) proves that if we select a vertex $v^k \in V^k$ which causes $MP(v^k)$ to have no feasible solution in $\text{Min } \mathcal{Z} \cap (v^k - \mathbb{R}_+^m)$, the vertex set V^k will become empty in a finite number of iterations and we can continue for other $v^k \in V^k$. Moreover, if we choose a vertex $v^k \in V^k$ such that

$$\nexists (x, y) \in \mathcal{G} : f(x) \in (v^k - \mathbb{R}_+^m),$$

problem $MP(v^k)$ will have no solution resulting in the elimination of v^k from V^k and the algorithm goes on for other another vertex.

As a result, we can conclude that the algorithm terminates in a finite number of iterations and $f(\bar{x})$ is an ε -optimal solution to Problem (OP). ■

6. Computational experiments

6.1. Application to portfolio selection problems

The bilevel optimization problem (BP) can be modelled as a class of portfolio optimization problems. Indeed, as mentioned in previous works, the portfolio optimization problem can be modelled as a multi-objective optimization problem $\text{Min } \{f(x) | x \in X\}$. For instance, in Markowitz's model, the two objectives are maximizing profit and minimizing risk. From the efficient portfolio set $\text{Argmin } \{f(x) | x \in X\}$, we can add other constraints, such as limiting profits and risks within a given range, given as the form $g(x, y) \leq 0, y \in \mathbb{R}_+^m$. Finally, to find the best portfolio, the Decision Maker will provide an extra-criterion function $h(x, y)$, which is the upper-level objective function of the bilevel problem (BP). Below is a detailed explanation of how to model the problem.

Markowitz's portfolio theory assumes that investors seek the maximum feasible projected return while being risk averse. It also implies that investors pick portfolios based on the desired return and the variance of return. The random vector $\mathbf{R} = (R_1, R_2, \dots, R_n)^T$ in \mathbb{R}^n represents the random returns of n assets, with $p(\mathbf{R})$ being its probability distribution. The mean vector of \mathbf{R} is denoted as $\mathbf{L} = (L_1, \dots, L_n)^T$, and its covariance matrix is $\mathbf{Q} = (\sigma_{ij})_{n \times n}$, where σ_{jj}^2 is the variance of R_j and σ_{ij}^2 is the correlation coefficient between R_i and R_j for $i, j = 1, \dots, n$. Considering a portfolio represented by $\mathbf{x} = (x_1, \dots, x_n)^T$ where $\sum_{j=1}^n x_j = 1$ and $x_j \geq 0$ for all j , the expected return is calculated as $\mathcal{E}(\mathbf{x}) = E[\mathbf{R}^T \mathbf{x}] = \sum_{j=1}^n L_j x_j$. The variance of profit is given by $\mathcal{V}(\mathbf{x}) = \text{Var}(\mathbf{R}^T \mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_j x_i = \mathbf{x}^T \mathbf{Q} \mathbf{x}$. Set $M = \{\mathbf{x} \in \mathbb{R}_+^n \mid \sum_{j=1}^n x_j = 1\}$. Then the portfolio optimization problem is therefore presented as follows.

$$\max \quad \mathcal{E}(\mathbf{x}) = E[\mathbf{R}^T \mathbf{x}] = \sum_{j=1}^n L_j x_j$$

$$\begin{aligned} \min \quad & \mathcal{V}(\mathbf{x}) = \text{Var}(\mathbf{R}^T \mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_j x_i = \mathbf{x}^T \mathbf{Q} \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x} \in M. \end{aligned}$$

Matching with the problem (BP), we choose $f_1(\mathbf{x}) = -\mathcal{E}(\mathbf{x})$, $f_2(\mathbf{x}) = \mathcal{V}(\mathbf{x})$ and $X = M$.

The Sharpe Ratio, denoted as $\mathcal{S}(\mathbf{x})$, is a metric that evaluates a portfolio's returns adjusted for risk. It is calculated using the formula $\mathcal{S}(\mathbf{x}) = \frac{\mathcal{E}(\mathbf{x}) - r_f}{\sqrt{\mathcal{V}(\mathbf{x})}}$, where r_f is the risk-free rate. Essentially, a higher Sharpe Ratio indicates that a portfolio is yielding greater returns for the level of risk undertaken by the investor. This index can be chosen as an extra-criterion function to select the best portfolio with the objective of maximizing the Sharpe Ratio function. Matching with the problem (BP), we choose $h(\mathbf{x}) = -\mathcal{S}(\mathbf{x})$. As proved in [37], the function $h(\mathbf{x})$ is pseudoconvex.

We can add additional constraints on the limits of profit and risk. Specifically, the profit needs to be at some minimum level \mathcal{E}_0 , and the risk is less than the maximum \mathcal{V}_0 that can be tolerated. Matching with the constraint $g(\mathbf{x}) \leq 0$ of problem (BP), we choose $g(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}))$, $g_1(\mathbf{x}) = \mathcal{V}(\mathbf{x}) - \mathcal{V}_0$ and $g_2(\mathbf{x}) = \mathcal{E}_0 - \mathcal{E}(\mathbf{x})$. It is easy to see that the function $g(\mathbf{x})$ is convex because $\mathcal{E}(\mathbf{x})$ is linear and $\mathcal{V}(\mathbf{x})$ is convex.

The portfolio selection modelled above is a pseudoconvex semivectorial bilevel optimization problem. A numerical example of this problem will be implemented in Example 6.7.

6.2. Numerical examples

We demonstrate the efficiency of our algorithm through numerical results in this section. The experiments were performed on 2.6 GHz Intel Core i7 (four logical cores), 8 Gb RAM. The code is implemented on Matlab R2018a. In the following examples, 'Gap' is defined as:

$$\text{Gap} = |\alpha - \beta|,$$

with α and β being the upper bound and lower bound to Problem (BP), respectively.

Example 6.1: Consider problem (BP) where functions are defined by

$$\begin{aligned} h(\mathbf{x}) &= x_1 + x_2^2 \\ f_1(\mathbf{x}) &= x_1^2 + x_2^2 + 0.4x_1 - 4x_2 \\ f_2(\mathbf{x}) &= \max\{-0.5x_1 - 0.25x_2 - 0.2, -2x_1 + 4.6x_2 - 5.8\} \end{aligned}$$

and

$$X = \{\mathbf{x} \in \mathbb{R}^2 \mid A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq 0, c(\mathbf{x}) \leq 0\},$$

where

$$A = \begin{bmatrix} 1.0 & -2.0 \\ -1.0 & 1.0 \\ 2.0 & 1.0 \\ 2.0 & 5.0 \\ -1.0 & -1.0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1.0 \\ 1.0 \\ 4.0 \\ 10.0 \\ -1.5 \end{bmatrix},$$

Table 1. The computational result of Example 6.1.

k	v^k	α^k	β^k	Gap
1	(0.432442, 1.067557)	1.572121	1.250000	0.322121
2	(0.573591, 0.926408)	1.431824	1.250000	0.181824
3	(0.688942, 0.811057)	1.346757	1.250000	0.096757
		...		
31	(0.996542, 0.503457)	1.250012	1.250000	0.000012
32	(0.997561, 0.502438)	1.250006	1.250000	0.000006

and

$$c(x) = 0.5(x_1 - 1)^2 + 1.4(x_2 - 0.5)^2 - 1.1.$$

Observe that $h(x)$, $f_1(x)$ are convex functions, and hence pseudoconvex, $f_2(x)$ is the maximum of two affine functions so that it is also a pseudoconvex function. The restraint functions are linear to satisfy assumption (A4). At the initialization step, the box $[\omega, \Omega] = [-2.52, 0.49, 3.29, 1.20]$, the initial upper bound $\alpha^0 = \infty$ and the initial lower bound $\beta^0 = \varphi(\Omega) = 1.25$ where φ is the constructed function. It is natural to choose $\hat{d} = (1, 1)^T$ and $\varepsilon = 0.00001$. The algorithm terminates in 1.2682 seconds after 32 iterations and returns the optimal solution $x^* = (0.997561, 0.502439)$ with objective value $h(x^*) = 1.250006$. This result is better than those reported in [5], which are $x = (0.2585, 1.2415)$ and $h(x) = 1.7989$. The computation details are given in Table 1, and more close to the exact optimal results which are $x^* = (1.0, 0.5)$ with $h(x^*) = 1.25$.

Example 6.2: Consider the following fractional programming problem where $h(x)$ and $f_2(x)$ are pseudoconvex due to Theorems 2.5 and 2.6, respectively.

$$\min h(x) = \frac{2x_1 + 3x_2}{4x_1 + 5x_2 + 10}$$

where

$$f_1(x) = \frac{(3x_1 + x_2)^2}{(3x_1 + x_2 - 1)^3}$$

$$f_2(x) = \frac{x_1^2 - 2x_1 + x_2^2 - 8x_2}{x_2 + 1}$$

and the lower-level problem is $\min(f_1(x), f_2(x))$ with the feasible set

$$X = \{x_1, x_2 \in \mathbb{R} | 2x_1 + x_2 \leq 6, 3x_1 + x_2 \leq 8, x_1 - x_2 \leq 1, x_1, x_2 \geq 1\}$$

At the initialization step, the box $[\omega, \Omega] = [0.19, -4.34, 0.42, -3.67]$, the initial upper bound $\alpha^0 = \infty$ and the initial lower bound $\beta^0 = \varphi(\Omega) = 0.276170$ where φ is the constructed function. We choose $\hat{d} = (1, 1)^T$ and $\varepsilon = 0.01$. The algorithm terminates in 0.8376 seconds after 6 iterations and returns the optimal solution $x^* = (1.020247, 1.835256)$ with objective value $h(x^*) = 0.302334$. The computation details are given in Table 2.

Table 2. The computational result of Example 6.2.

k	v^k	α^k	β^k	Gap
1	(2.095784, 1.712646)	0.346225	0.276170	0.070055
2	(1.781338, 2.206333)	0.346225	0.285986	0.070055
3	(1.454576, 2.020253)	0.346067	0.285986	0.060081
4	(1.161954, 1.887563)	0.331592	0.311198	0.030346
5	(1.300950, 1.946424)	0.331592	0.315449	0.019604
6	(1.020247, 1.835256)	0.324469	0.315449	0.009020

Example 6.3: We consider the following problem where the lower problem has four objective functions

$$\min \quad h(x) = \frac{3x_1 + 2x_2 + 10x_3 + 11}{x_1 + x_2 + x_3 + 10}$$

where

$$f_1(x) = \frac{2x_1 + 5x_2 + 3x_3 + 10}{3x_2 + 3x_3 + 10}$$

$$f_2(x) = \frac{2x_1 + 4x_2 + 10}{4x_1 + 4x_2 + 5x_3 + 10}$$

$$f_3(x) = \frac{x_1 + 2x_2 + 5x_3 + 10}{x_1 + 5x_2 + 5x_3 + 10}$$

$$f_4(x) = \frac{x_1 + 2x_2 + 4x_3 + 10}{5x_2 + 4x_3 + 10}$$

The feasible set for the lower-level problem is the set of positive vectors in \mathbb{R}_+^3 such that

$$2x_1 + x_2 + 5x_3 \leq 10,$$

$$x_1 + 6x_2 + 3x_3 \leq 10,$$

$$5x_1 + 9x_2 + 2x_3 \leq 10,$$

$$9x_1 + 7x_2 + 3x_3 \leq 10$$

All objective functions are pseudoconvex by Theorem 2.5. At the initialization step, the box $[\omega, \Omega] = [1.00, 0.50, 1.65, 0.79, 1.17, 1.00, 3.00, 1.00]$, $\hat{d} = (1, 1, 1, 1)^T$ and $\varepsilon = 0.01$. The initial upper bound $\alpha^0 = \infty$, and the initial lower bound $\beta^0 = \varphi(\Omega) = 0.604$. The algorithm terminates in 1.5260 seconds after 7 iterations and returns the optimal solution $x^* = (0.000000, 0.775596, 0.007336)$ with objective value $h(x^*) = 0.607448$. The computation details are given in Table 3.

Example 6.4: We consider an example in [31] with linear objective in the upper problem

$$h(x) = -x_1 - 0.9$$

$$f_1(x) = x_1$$

$$f_2(x) = x_2$$

$$g(x) = x_1^2 + x_2^2 - 0.81$$

Table 3. The computational result of Example 6.3.

k	v^k	α^k	β^k	Gap
1	(0.000000, 0.483159, 0.697862)	0.894430	0.604000	0.290430
2	(0.000000, 0.783986, 0.483034)	0.818089	0.603704	0.214385
3	(0.000000, 0.877863, 0.084834)	0.648966	0.603704	0.045262
4	(0.000000, 0.822004, 0.292454)	0.648966	0.603704	0.045262
5	(0.000000, 0.825505, 0.127032)	0.648966	0.603704	0.045262
6	(0.000000, 0.823124, 0.223619)	0.648966	0.603704	0.045262
7	(0.000000, 0.775596, 0.007336)	0.607448	0.603704	0.003745

Table 4. The computational result of Example 6.4.

k	v^k	α^k	β^k	Gap
1	(-0.526290, -0.473709)	-1.426291	-1.800000	0.373709
2	(-0.775599, -0.224400)	-1.675600	-1.800000	0.124400
3	(-0.893699, -0.106300)	-1.793699	-1.800000	0.006301

and the feasible set for the lower-level problem is

$$X = \{x \in [-1, 1] | x_1 + x_2 + 1 \geq 0\}$$

At the initialization step, the box $[\omega, \Omega] = [-0.90, -0.90, 0.00, -0.00]$, $\hat{d} = (1, 1)^T$ and $\varepsilon = 0.01$. The initial upper bound $\alpha^0 = \infty$, and the initial lower bound $\beta^0 = \varphi(\Omega) = -1.8$ where φ is the constructed function. The algorithm terminates in 1.4638 seconds after 17 iterations and returns the optimal solution $x^* = (-0.893699, -0.106301)$ with objective value $h(x^*) = -1.793699$, which is significantly close to the exact optimal value reported in [31]. The computation details are given in Table 4.

Example 6.5: We consider another example in [31]

$$h_1(x) = (x_1 - 1)^2 + \sum_{i=2}^{14} x_i^2 + 0.25$$

$$f_1(x) = \sum_{i=1}^{14} x_i^2$$

$$f_2(x) = (x_1 - 0.5)^2 + \sum_{i=2}^{14} x_i^2$$

$$X = \{x \in \mathbb{R}^{14} | -1 \leq x_1, x_2, \dots, x_{14} \leq 2\}$$

At the initialization step, the box $[\omega, \Omega] = [0.00, 0.00, 0.25, 0.25]$, $\hat{d} = (1, 1)^T$ and $\varepsilon = 0.01$. The initial upper bound $\alpha^0 = \infty$, and the initial lower bound $\beta^0 = \varphi(\Omega) = 0.5$ where φ is the constructed function. The algorithm terminates in 1.4638 seconds after 17 iterations and returns the optimal solution $x^* = (0.5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ with objective value $h(x^*) = 0.5$ which is also the exact optimal value reported in [31]. The computation details are given in Table 5.

Table 5. The computational result of Example 6.5.

k	α^k	β^k	Gap
1	0.812500	0.500000	0.312500
2	0.566406	0.500000	0.066406
3	0.503922	0.500000	0.003922
4	0.500015	0.500000	0.000015
5	0.500000	0.500000	0.000000

Example 6.6: We now consider the following pseudoconvex optimization problem

$$\begin{aligned}
 \min \quad & h(x, y) = \frac{x_1^2 + 2x_2^2 + 10y_1^2 + y_2^2 + 11}{x_1 + x_3 + y_1 + 20} \\
 & f_1(x) = \frac{2x_1 + 5x_2 + 3x_3 + 10}{3x_2 + 3x_3 + 10} \\
 & f_2(x) = \frac{2x_1 + 4x_2 + 10}{4x_1 + 4x_2 + 5x_3 + 10} \\
 & f_3(x) = \frac{x_1 + 2x_2 + 5x_3 + 10}{x_1 + 5x_2 + 5x_3 + 10} \\
 & f_4(x) = \frac{x_1 + 2x_2 + 4x_3 + 10}{5x_2 + 4x_3 + 10} \\
 & X = \{x \in \mathbb{R}_+^3 \mid 2x_1 + x_2 + 5x_3 \leq 10, \\
 & \quad x_1 + 6x_2 + 3x_3 \leq 10, 5x_1 + 9x_2 + 2x_3 \leq 10, 9x_1 + 7x_2 + 3x_3 \leq 10\} \\
 & g_1(x) = -x_2 - x_3 - 2y_1 - y_2 + 2 \\
 & g_2(x) = x_2 + x_3 - 5y_1 + 2y_2 - 1 \\
 & y \in \mathbb{R}_+^2
 \end{aligned}$$

At the initialization step, the box $[\omega, \Omega] = [1.14, 0.68, 0.99, 0.98, 1.27, 0.92, 1.14, 1.17]$, $\hat{d} = (1, 1, 1, 1)^T$ and $\varepsilon = 0.01$. The initial bounds $\alpha^0 = \infty$, $\beta^0 = \varphi(\Omega) = 0.009170$. The algorithm terminates after 5 iterations, and returns the optimal solutions $x^* = (0.130662, 0.156198, 1.558087)$ and $y^* = (0.142857, 0.000000)$ with objective value $h(x^*, y^*) = 0.012365$. The computation details are given in Table 6.

In the following example, we consider an example of the portfolio selection problem presented in Section 6.1.

Example 6.7: In this example, we consider a portfolio selection optimization problem proposed in [37], which is reformulated into the form of Problem (BP) with: $h(x) =$

Table 6. The computational result of Example 6.6.

k	v^k	α^k	β^k	Gap
1	(0.000000, 0.422286, 1.291999, 0.142857, 0)	0.026160	0.009170	0.190285
2	(0.001831, 0.421485, 1.292799, 0.142857, 0)	0.026094	0.009170	0.016923
3	(0.329682, 0.276641, 1.437644, 0.142857, 0)	0.021261	0.009170	0.012091
4	(0.464042, 0.170190, 1.544095, 0.142857, 0)	0.021261	0.009170	0.012091
5	(0.130662, 0.156198, 1.558087, 0.142857, 0)	0.012365	0.009170	0.003195

Table 7. The computational result of Example 6.7.

k	v^k	α^k	β^k	Gap
1	(0.154919, 0.170694, 0.006096, 0.243734, 0.424554)	-0.146438	$-\infty$	∞
2	(0.154760, 0.170508, 0.006114, 0.243741, 0.424874)	-0.146490	-0.146494	0.000004
3	(0.154757, 0.170505, 0.006114, 0.243741, 0.424879)	-0.146494	-0.146494	0.000000

$$-\frac{L^T x - 0.022}{\sqrt{x^T Q x}}, f_1(x) = -L^T x, f_2(x) = x^T Q x, X = \{x \in \mathbb{R}_+^5 | L^T x \geq 0.25\}, \quad g(x) = x^T Q x - 2.5,$$

$$Q = \begin{bmatrix} 4.415125 & 1.124907 & 2.310423 & 1.443982 & 1.393465 \\ 1.124907 & 4.074815 & 1.963056 & 1.287082 & 1.535600 \\ 2.310423 & 1.963056 & 9.139115 & 2.338314 & 1.983779 \\ 1.443982 & 1.287082 & 2.338314 & 4.431688 & 1.670681 \\ 1.393465 & 1.535600 & 1.983779 & 1.670600 & 5.314346 \end{bmatrix}$$

$$L = [0.156723, 0.158738, 0.204619, 0.216932, 0.348760]^T$$

At the initialization step, the box $[\omega, \Omega] = [-0.15, 2.45, -0.15, 2.50]$, $\hat{d} = (1, 1)^T$ and $\varepsilon = 1e - 6$. The initial bounds $\alpha_0 = 5.782681$, $\beta_0 = \varphi(\Omega) = 1.322750$. The algorithm terminates after 3 iterations and returns the optimal solutions $x^* = (0.154757, 0.170505, 0.006114, 0.243741, 0.424879)$ with objective value $h(x^*) = -0.253628$. From this, we calculate $\mathcal{E}(x)$, $\mathcal{S}(x)$, and $\mathcal{V}(x)$ as 0.253628, 0.146494, 2.500000. This result is approximately the same as that reported in [37]. The computation details are given in Table 7.

7. Conclusion

This study presents an exact algorithm for solving a class of pseudoconvex semivectorial bilevel optimization problems where the upper objective function is scalar pseudoconvex, and the lower one is a pseudoconvex vector function. The algorithm's efficiency stems from using the outcome space-cutting cone through a monotonic optimization approach and the neural dynamic approach to solving subproblems, which is computationally efficient and does not necessitate particular characteristics of the objective functions. The algorithm's convergence has been demonstrated, and its effectiveness is supported by computational experiments.

Acknowledgements

The authors thank the anonymous peer reviewers and the editor for their constructive comments which helped to improve the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by Rikkeisoft Corporation and supported by the Center for Digital Technology and Economy (BK Fintech), Hanoi University of Science and Technology.

Notes on contributors

Tran Ngoc Thang, Ph. D., is a teaching staff member at the Faculty of Mathematics and Informatics, Hanoi University of Science and Technology, Hanoi, Vietnam. His research interests are applied mathematics and informatics, including multi-objective optimization, stochastic optimization, machine learning, and data analysis. He has a publication record, having authored and co-authored over 30 research papers in esteemed journals and conferences.

Dao Minh Hoang is a Ph.D. candidate at the Dresden University of Technology in Germany. His program is funded by The School of Embedded Composite Artificial Intelligence (SECAI). Before that, he graduated from his master's program in Applied Mathematics and Informatics at Hanoi University of Science and Technology. He works on the mathematical characterization of high-dimensional fitness landscapes, investigating modern techniques from AI and computational statistics to discover and characterize pertinent landscape features and develop methods that improve optimization algorithms.

Nguyen Viet Dung received his Bachelor degree in the Faculty of Mathematics and Informatics at Hanoi University of Science and Technology, Hanoi, Vietnam. He is now pursuing a master degree in Mathematics and Informatics at Hanoi University of Science and Technology, Hanoi, Vietnam. His research interests involve multi-objective optimization, deep neural network models and application of optimization theory in deep learning.

References

- [1] L.M. Alberto Cambini, *Generalized Convexity and Optimization*, 1st ed., Springer, Berlin, 2008.
- [2] E. Angulo, E. Castillo, R. Garcia-Rodenas, and J. Sanchez-Vizcaino, *A continuous bi-level model for the expansion of highway networks*, *Comput. Oper. Res.* 41 (2014), pp. 262–276.
- [3] M. Avriel, W.E. Diewert, S. Schaible, and I. Zang, *Generalized Concavity*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2010.
- [4] H. Belkhir, M.E.A. Chergui, and F.Z. Ouail, *Optimizing a linear function over an efficient set*, *Oper. Res.* 22 (2022), pp. 3183–3201.
- [5] H.P. Benson, *An outcome space algorithm for optimization over the weakly efficient set of a multiple objective nonlinear programming problem*, *J. Global Optim.* 52 (2012), pp. 553–574.
- [6] W. Bian, L. Ma, S. Qin, and X. Xue, *Neural network for nonsmooth pseudoconvex optimization with general convex constraints*, *Neural. Netw.* 101 (2018), pp. 1–14.
- [7] J. Bracken and J.T. McGill, *Mathematical programs with optimization problems in the constraints*, *Oper. Res.* 21 (1973), pp. 37–44.
- [8] L. Carosi and L. Martein, *Some classes of pseudoconvex fractional functions via the Charnes-Cooper transformation*, in *Generalized Convexity and Related Topics*, I. V. Konnov, D. T. Luc and A. M. Rubinov, eds., Springer, Berlin, 2006, pp. 177–188.
- [9] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Society for Industrial and Applied Mathematics, Philadelphia, 1990.
- [10] A. Fischer, A.B. Zemkoho, and S. Zhou, *Semismooth newton-type method for bilevel optimization: Global convergence and extensive numerical experiments*, *Optim. Methods Software.* 37 (2022), pp. 1770–1804.
- [11] L. Franceschi, P. Frasconi, S. Salzo, R. Grazzi, and M. Pontil, *Bilevel programming for hyperparameter optimization and meta-learning*, in *Proceedings of the 35th International Conference on Machine Learning*, J. Dy and A. Krause, eds., *Proceedings of Machine Learning Research*, Vol. 80, 10–15 Jul, PMLR, Stockholm, Sweden, 2018, pp. 1568–1577.
- [12] Y. Gao, G. Zhang, J. Lu, and H.M. Wee, *Particle swarm optimization for bi-level pricing problems in supply chains*, *J. Global Optim.* 51 (2011), pp. 245–254.
- [13] J.R. Hicks, *Marktforn und gleichgewicht*, *Econ. J.* 45 (1935), pp. 334–336.
- [14] L.P. Hoang, D.D. Le, T.A. Tran, and T.T. Ngoc, *Improving pareto front learning via multi-sample hypernetworks*, preprint (2022). Available at arXiv:2212.01130.

- [15] R. Horst and N.V. Thoai, *Maximizing a concave function over the efficient or weakly-efficient set* research supported by the ‘deutsche forschungsgemeinschaft’ through the ‘graduiertenkolleg mathematische optimierung’ at the university of trier and through the project decomp.1, Eur. J. Oper. Res. 117 (1999), pp. 239–252.
- [16] R. Horst, N.V. Thoai, Y. Yamamoto, and D. Zenke, *On optimization over the efficient set in linear multicriteria programming*, J. Optim. Theory Appl. 134 (2007), pp. 433–443.
- [17] J.M. Jorge, *A bilinear algorithm for optimizing a linear function over the efficient set of a multiple objective linear programming problem*, J. Global Optim. 31 (2005), pp. 1–16.
- [18] V. Kalashnikov, S. Dempe, B. Mordukhovich, and S.V. Kavun, *Bilevel optimal control, equilibrium, and combinatorial problems with applications to engineering*, Math. Probl. Eng. 2017 (2017), pp. 7190763.
- [19] N. Kim and T. Tran Ngoc, *Optimization over the efficient set of a bicriteria convex programming problem*, Pac. J. Optim. 9 (2013), pp. 103–115.
- [20] H.A. Le Thi, T. Pham Dinh, and L.D. Muu, *Simplicially-constrained dc optimization over efficient and weakly efficient sets*, J. Optim. Theory Appl. 117 (2003), pp. 503–531.
- [21] Q. Li, Y. Liu, and L. Zhu, *Neural network for nonsmooth pseudoconvex optimization with general constraints*, Neurocomputing 131 (2014), pp. 336–347.
- [22] N. Liu, J. Wang, and S. Qin, *A one-layer recurrent neural network for nonsmooth pseudoconvex optimization with quasiconvex inequality and affine equality constraints*, Neural. Netw. 147 (2022), pp. 1–9.
- [23] J. Lu, J. Han, Y. Hu, and G. Zhang, *Multilevel decision-making: A survey*, Inf. Sci. 346–347 (2016), pp. 463–487.
- [24] Z. Lukac, K. Soric, and V.V. Rosenzweig, *Production planning problem with sequence dependent setups as a bilevel programming problem*, Eur. J. Oper. Res. 187 (2008), pp. 1504–1512.
- [25] M. Mackay, P. Vicol, J. Lorraine, D. Duvenaud, and R. Grosse, *Self-tuning networks: Bilevel optimization of hyperparameters using structured best-response functions*, in *International Conference on Learning Representations*, New Orleans, Louisiana, United States, 2019.
- [26] A. Malek, N. Hosseinipour-Mahani, and S. Ezazipour, *Efficient recurrent neural network model for the solution of general nonlinear optimization problems*, Optim. Methods Software. 25 (2010), pp. 489–506.
- [27] A. Mitsos, P. Lemonidis, and P.I. Barton, *Global solution of bilevel programs with a nonconvex inner program*, J. Global Optim. 42 (2008), pp. 475–513.
- [28] L. Muu, *On the construction of initial polyhedral convex set for optimization problems over the efficient set and bilevel linear programs*, Vietnam J. Math. 28 (2000), pp. 77–82.
- [29] A. Ren and Y. Wang, *A novel penalty function method for semivectorial bilevel programming problem*, Appl. Math. Model. 40 (2016), pp. 135–149.
- [30] Y. Shehu, P.T. Vuong, and A. Zemkoho, *An inertial extrapolation method for convex simple bilevel optimization*, Optim. Methods Software. 36 (2019), pp. 1–19.
- [31] A. Sinha, P. Malo, and K. Deb, *Towards understanding bilevel multi-objective optimization with deterministic lower level decisions*, in *Evolutionary Multi-Criterion Optimization*, Springer International Publishing, Cham, 2015, pp. 426–443.
- [32] A.Z. Stephan Dempe, *Bilevel Optimization: Advances and Next Challenges*, 1st ed., Springer, Cham, 2020.
- [33] T.N. Thang, *Outcome-based branch and bound algorithm for optimization over the efficient set and its application*, in *Some Current Advanced Researches on Information and Computer Science in Vietnam: Post-proceedings of The First NAFOSTED Conference on Information and Computer Science*, Springer, Hanoi, Vietnam, 2015, pp. 31–47.
- [34] T.N. Thang and N.T.B. Kim, *Outcome space algorithm for generalized multiplicative problems and optimization over the efficient set*, JIMO 12 (2016), pp. 1417–1433.
- [35] T.N. Thang, D.T. Luc, and N.T.B. Kim, *Solving generalized convex multiobjective programming problems by a normal direction method*, Optimization 65 (2016), pp. 2269–2292.
- [36] T.N. Thang, V.K. Solanki, T.A. Dao, N. Thi Ngoc Anh, and P. Van Hai, *A monotonic optimization approach for solving strictly quasiconvex multiobjective programming problems*, J. Intell. Fuzzy Syst. 38 (2020), pp. 6053–6063. 5.

- [37] T. Thi Thanh Tuoi, T. Tuan Khang, N. Thi Ngoc Anh, and T. Ngoc Thang, Fuzzy portfolio selection with flexible optimization via quasiconvex programming, in *Intelligent Systems and Networks*, Springer Nature Singapore, Singapore, 2022, pp. 360–368.
- [38] T.A. Tran, L.P. Hoang, D.D. Le, and T.N. Tran, *A framework for controllable pareto front learning with completed scalarization functions and its applications*, preprint (2023). Available at [arXiv:2302.12487](https://arxiv.org/abs/2302.12487).
- [39] H. Tuy, *Convex Analysis and Global Optimization*, Springer, Cham, 2016.
- [40] S. Ukkusuri, K. Doan, and H.A. Aziz, *A bi-level formulation for the combined dynamic equilibrium based traffic signal control*, *Proc. Soc. Behav. Sci.* 80 (2013), pp. 729–752. 20th International Symposium on Transportation and Traffic Theory (ISTTT 2013).
- [41] D.J. White and G. Anandalingam, *A penalty function approach for solving bi-level linear programs*, *J. Global Optim.* 3 (1993), pp. 397–419.
- [42] J. Yu, A. Vexler, and K. Jalal, *A critical issue of using the variance of the total in the linearization method—In the context of unequal probability sampling*, *Stat. Med.* 38 (2019), pp. 1475–1483.