A Global Optimization Method for the Stackelberg Problem with Convex Functions via Problem Transformation and Concave Programming

Kiyotaka Shimizu and Min Lu

***Abstract*—A global optimization method is proposed for solving the Stackelberg problem through problem transformation and concave programming. The proposed method is applicable to a broad class of the Stackelberg problem, in which each function in upper-level is convex or a difference of two convex functions, and that in lower-level is convex.**

I. Introduction

The Stackelberg solution [15], [16] is well known as the most rational game solution to answer a question: "What will be the best strategy for decision-maker 1 who knows decision-maker 2's objective function and has to choose his strategy first, while decision-maker 2 chooses his strategy after announcement of decision-maker 1's strategy?"

The Stackelberg problem is formulated as a two-level optimization problem, where decision-maker 1 and 2 correspond to the upper- and lower-level problems, respectively.

The upper-level determines the optimal value of its decision variables (parameters for the lower-level) so as to minimize its objective, while the lower-level minimizes its own objective with respect to the lower-level decision variables under the given parameters. It is noticed that the objective and constraints in the upper-level are functions of both the upper-level decision variables and the optimal solution to the lower-level problem.

Generally speaking, the Stackelberg problem is not a convex program even if all objective and constraints are convex functions. Accordingly, its solution cannot be obtained by conventional mathematical programming, and furthermore, it is very difficult to find a globally optimal solution for the Stackelberg problem.

In the past, several algorithms have been proposed for solving the Stackelberg problem. The first approach is based on solving the nonlinear program which is obtained by replacing the lower-level problem with its Kuhn–Tucker conditions [2], [4], [6], [7], [11]. In this category, Fortuny–Amat and McCarl [7] developed a computational method by transforming the original problem to a mixed-integer program, and Bard [4] and Edmunds and Bard [6] developed a branch-and-bound type algorithm based on the active constraint strategy. But either of them is only effective for a particular Stackelberg problem with a linear upper-level problem and a convex quadratic lower-level one. Lu and Shimizu [11] extended such an approach to a class of problems in which the upper-level problem consists of convex functions or differences of two convex functions, but the lower-level is still limited to be convex quadratic. Further, Al-Khayyal, Horst, and Pardalos [2] treated a class of problems in which the upper-level problem is concave but the lower-level one is linear. On the other hand, the double penalty method was proposed by Shimizu and Aiyoshi [13] and Aiyoshi and Shimizu [1] where the Stackelberg problem was solved through transforming it to an unconstrained optimization problem by use of interior and exterior penalty function methods. However, this method cannot guarantee to get a globally optimal solution. Shimizu and Ishizuka [14] proposed a nondifferentiable optimization approach for the parameter design problem which is a special class of the Stackelberg problem where the objective and/or constraints in the upper-level include optimal-value functions of the lower-level problems rather than optimal solutions.

The purpose of this paper is to develop a global optimization method for the Stackelberg problem. First, the Stackelberg problem is equivalently expressed as an optimization problem with an equality and several inequality constraints by introducing the optimal-value function of the lower-level problem. Under certain conditions, this problem is transformed into a nonlinear program whose objective and constraints are convex functions. Then, by using the concept of exterior penalty method, we obtain an auxiliary problem having only inequality constraints for the transformed problem. It is proved that a solution to the auxiliary problem converges to a globally optimal solution to the transformed problem as the penalty parameter goes to infinity. Finally, we show that the auxiliary problem can be equivalently transformed into a concave program whose globally optimal solution can be found.

The proposed method is applicable to a broad class of the Stackelberg problem whose functions in the upper-level are convex functions or differences of two convex functions, and those in the lower-level are convex ones. It transforms the Stackelberg problem into a concave program that assures obtaining a globally optimal solution. In addition, by applying the cutting plane type method with subgradients, we can make an algorithm even when related functions are not differentiable.

II. Problem Formulation

The Stackelberg problem is formulated below. Let the lower-level problem be of the form

$$\min_{\boldsymbol{y}} f(\boldsymbol{x}, \boldsymbol{y}) \tag{1a}$$

$$\text{subj. to } \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) \leqq \mathbf{0} \tag{1b}$$

where $\boldsymbol{x} \in R^N$ is the given parameter, $\boldsymbol{y} \in R^n$ is the decision variable, $f\colon R^N \times R^n \to R^1$ and $\boldsymbol{g}\colon R^N \times R^n \to R^m$. Denote by $S(\boldsymbol{x})$ and $P(\boldsymbol{x})$ the feasible solution set and the optimal solution set of problem (1), respectively, and assume $S(\boldsymbol{x}) \neq \emptyset$ and $P(\boldsymbol{x}) \neq \emptyset$. Note that this problem is parameterized by $\boldsymbol{x}$.

Next, let us formulate another optimization problem with respect to the parameter $\boldsymbol{x}$, in which the answering $\boldsymbol{y}^* \in P(\boldsymbol{x})$, i.e., parametric optimal solution, is contained

$$\min_{\boldsymbol{x}} \min_{\boldsymbol{y}^*} F(\boldsymbol{x}, \boldsymbol{y}^*) \tag{2a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}^*) \leqq \mathbf{0} \tag{2b}$$

$$\boldsymbol{y}^* \in P(\boldsymbol{x}) \tag{2c}$$

where $F\colon R^N \times R^n \to R^1$ and $\boldsymbol{G}\colon R^N \times R^n \to R^M$. Let us call this the upper-level problem.

Then, the Stackelberg problem is formulated in a two-level optimization scheme, in which problem (1) is subordinated to problem (2) as a part of the constraints:

$$\min_{\boldsymbol{x}} \min_{\boldsymbol{y}^*} F(\boldsymbol{x}, \boldsymbol{y}^*) \tag{3a}$$

$$\text{subj. to } \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}^*) \leqq \mathbf{0} \tag{3b}$$

$$\boldsymbol{y}^* \in P(\boldsymbol{x}) \tag{3c}$$

Manuscript received November 19, 1993; revised September 17, 1994.

The authors are with the Faculty of Science and Technology, Keio University, Yokohama 223, Japan.

IEEE Log Number 9414505.

$$P(\boldsymbol{x}) = \{\boldsymbol{y}^* \in S(\boldsymbol{x}) | f(\boldsymbol{x}, \boldsymbol{y}^*) = \min_{\boldsymbol{y} \in S(\boldsymbol{x})} f(\boldsymbol{x}, \boldsymbol{y})\} \tag{3d}$$

where $S(\boldsymbol{x}) = \{\boldsymbol{y} \in R^n | \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) \leqq \boldsymbol{0}\}$. We assume that there exists a feasible solution satisfying (3b)–(3d). The problem is to seek an optimal solution $\boldsymbol{x}^*$ for decision-maker 1 and its corresponding $\boldsymbol{y}^* \in P(\boldsymbol{x}^*)$ for decision-maker 2.

When the optimal solution $\boldsymbol{y}^*(\boldsymbol{x})$ is unique for given $\boldsymbol{x}$, i.e., when $P(\boldsymbol{x}) = \{\boldsymbol{y}^*(\boldsymbol{x})\}$, problem (2) or (3) becomes merely as

$$\min_{\boldsymbol{x}} F(\boldsymbol{x}, \boldsymbol{y}^*(\boldsymbol{x}))$$
$$\text{subj. to } \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}^*(\boldsymbol{x})) \leqq \boldsymbol{0}$$

where $\boldsymbol{y}^*(\boldsymbol{x}) = \arg\min_{\boldsymbol{y} \in S(\boldsymbol{x})} f(\boldsymbol{x}, \boldsymbol{y})$.

In this manner, the Stackelberg problem can be represented by a hierarchical optimization problem in which one of the constraints for the upper-level problem consists of a parameterized optimization problem in the lower-level. It is noted, however, that the Stackelberg problem is much more difficult than the usual two-level programming problem that includes only optimal-value function $w(\boldsymbol{x}) \triangleq f(\boldsymbol{x}, \boldsymbol{y}^*)$ in the objective and/or constraints. In the case of linear or quadratic objective functions, however, solutions for the Stackelberg problem have been extensively studied [4], [6], [7].

III. Problem Transformation

Let us define the optimal-value function of the lower-level problem (1) as

$$w(\boldsymbol{x}) \triangleq f(\boldsymbol{x}, \boldsymbol{y}^*) = \min_{\boldsymbol{y} \in S(\boldsymbol{x})} f(\boldsymbol{x}, \boldsymbol{y}) \tag{4}$$

where $\boldsymbol{y}^* \in P(\boldsymbol{x})$ is an arbitrarily chosen solution vector of the lower-level. Then, the Stackelberg (3) is equivalently expressed as follows:

$$\min_{\boldsymbol{x}, \boldsymbol{z}} F(\boldsymbol{x}, \boldsymbol{z}) \tag{5a}$$
$$\text{subj. to } \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{z}) \leqq \boldsymbol{0} \tag{5b}$$
$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{z}) \leqq \boldsymbol{0} \tag{5c}$$
$$f(\boldsymbol{x}, \boldsymbol{z}) - w(\boldsymbol{x}) = 0 \tag{5d}$$

where $w(\boldsymbol{x})$ is the optimal-value function which is defined by (4).

Here, $\boldsymbol{z} \in R^N$ serves a substitute for the solution of the lower-level problem $\boldsymbol{y}^* \in P(\boldsymbol{x})$ and is an artificial variable newly introduced as the upper-level decision variable. Note that the constraint (5c) is added.

It is pointed out that regarding $(\boldsymbol{x}, \boldsymbol{z})$ and $\boldsymbol{y}$ as the upper-level and the lower-level decision variables, respectively, we can consider problem (5) as a special case of the parameter design problem studied in [14]. It is also noticed that in Bard [3] the Stackelberg problem is transformed into an infinitely constrained programming problem, but no computational method is given there.

Although problem (5) appears to be an ordinary nonlinear programming problem, it becomes a nondifferentiable one since $w(\boldsymbol{x})$ is not differentiable in general. Unfortunately, effective computational methods have not been proposed yet for optimization problems with nondifferentiable equality constraints such as problem (5).

Let us impose the following assumption:

Assumption 1: The functions $f(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$ are convex in $R^N \times R^n$.

The optimal-value function $w(\boldsymbol{x})$ has the following property (see Lemma 1 of [12]):

Proposition 1: The optimal-value function $w(\boldsymbol{x})$ is convex when Assumption 1 holds.

From Proposition 1 there exist subgradients of $w(\boldsymbol{x})$.

Suppose further the following assumption:

Assumption 2: The functions $F(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y})$ are convex in $R^N \times R^n$.

Under Assumptions 1 and 2, problem (5) becomes a program whose objective and the inequality constraints are convex functions and the equality constraint is a difference of two convex functions. By introducing an auxiliary variable t_1, such problem can be equivalently transformed into the following problem in which all functions are convex:

$$\min_{\boldsymbol{x}, \boldsymbol{z}, t_1} F(\boldsymbol{x}, \boldsymbol{z}) \tag{6a}$$
$$\text{subj. to } \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{z}) \leqq \boldsymbol{0} \tag{6b}$$
$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{z}) \leqq \boldsymbol{0} \tag{6c}$$
$$f(\boldsymbol{x}, \boldsymbol{z}) - t_1 = 0 \tag{6d}$$
$$w(\boldsymbol{x}) - t_1 = 0. \tag{6e}$$

If a globally optimal solution $(\boldsymbol{x}^o, \boldsymbol{z}^o, t_1^o)$ to (6) is found by some algorithm, then $(\boldsymbol{x}^o, \boldsymbol{z}^o)$ solves the Stackelberg problem (3). In the following sections, we develop a global optimization method for solving (6) instead of solving the original problem (3).

Remark 1: Problem (6) is not a convex program since it has nonlinear equality constraints. Therefore, a globally optimal solution cannot be obtained by conventional mathematical programming.

Remark 2: By applying the approach proposed in Section III-B of [11], even when $F(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y})$ are differences of two convex functions and $f(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$ are convex functions, we can transform problem (5) to a problem such as problem (6).

IV. Application of an Exterior Penalty Method

We consider to apply the concept of exterior penalty method [5] to problem (6). By adding the equality constraints (6d), (6e) to the objective function, we obtain the following auxiliary problem:

$$\min_{\boldsymbol{x}, \boldsymbol{z}, t_1} F(\boldsymbol{x}, \boldsymbol{z}) - s[f(\boldsymbol{x}, \boldsymbol{z}) + w(\boldsymbol{x}) - 2t_1] \tag{7a}$$
$$\text{subj. to } \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{z}) \leqq \boldsymbol{0} \tag{7b}$$
$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{z}) \leqq \boldsymbol{0} \tag{7c}$$
$$f(\boldsymbol{x}, \boldsymbol{z}) - t_1 \leqq 0 \tag{7d}$$
$$w(\boldsymbol{x}) - t_1 \leqq 0. \tag{7e}$$

Next, let us consider the relationship between (7) and (6). First, we give the following proposition:

Proposition 2: The functions $F(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y}), f(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$ are continuous at any $(\boldsymbol{x}, \boldsymbol{y})$, and $w(\boldsymbol{x})$ is continuous at any $\boldsymbol{x}$.

This proposition holds obviously from Assumptions 1 and 2, Proposition 1, and the continuity of convex functions.

We impose the following assumptions:

Assumption 3: The set $\{\boldsymbol{x} | \boldsymbol{G}(\boldsymbol{x}, \check{\boldsymbol{y}}(\boldsymbol{x})) \leqq \boldsymbol{0}\}$ is not empty.

Assumption 4: The function $F(\boldsymbol{x}, \boldsymbol{y})$ has a lower bound.

Then, we have the following relations between (7) and (6):

Theorem 1: Let Assumptions 1–4 hold and assume that problem (7) has a globally optimal solution for any $s > 0$. Let $\{(\boldsymbol{x}^k, \boldsymbol{z}^k, t_1^k)\}$ be a sequence of the globally optimal solutions to problem (7) corresponding to a sequence of positive numbers $\{s^k\}$ monotonously diverging to infinity. Then, any accumulation point of the sequence $\{(\boldsymbol{x}^k, \boldsymbol{z}^k, t_1^k)\}$ is a globally optimal solution to problem (6), and the following relation holds:

$$\lim_{k \to \infty} s^k[f(\boldsymbol{x}^k, \boldsymbol{z}^k) + w(\boldsymbol{x}^k) - 2t_1^k] = 0.$$

Proof: Denote any one of the accumulation points by $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{z}}, \tilde{t}_1)$ and a newly convergent subsequence of $\{(\boldsymbol{x}^k, \boldsymbol{z}^k, t_1^k)\}$ to $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{z}}, \tilde{t}_1)$ by $\{(\boldsymbol{x}^k, \boldsymbol{z}^k, t_1^k)\}$.

Feasibility of the Accumulation Point: Obviously, the following relations hold:

$$\begin{aligned} G(\tilde{x}, \tilde{z}) &\leqq 0 \\ g(\tilde{x}, \tilde{z}) &\leqq 0 \\ f(\tilde{x}, \tilde{z}) - \tilde{t}_1 &\leqq 0 \\ w(\tilde{x}) - \tilde{t}_1 &\leqq 0. \end{aligned} \tag{8}$$

Thus, it is sufficient to show that

$$f(\tilde{x}, \tilde{z}) - \tilde{t}_1 = 0 \tag{9a}$$

$$w(\tilde{x}) - \tilde{t}_1 = 0. \tag{9b}$$

Let us prove (9) by contradiction method. Suppose that (9a) or (9b) does not hold. Then from the continuity of f and w, and $s^k \to \infty$ as $k \to \infty$, we have

$$\lim_{k\to\infty} s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] = -\infty.$$

This relation and Assumption 4 implies that

$$\lim_{k\to\infty} F(x^k, z^k) - s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] = \infty. \tag{10}$$

On the other hand, there exists a $\bar{x} \in \{x \mid G(x, \check{y}(x)) \leqq 0\}$ by Assumption 3. As letting

$$(\bar{x}, \bar{z}, \bar{t}_1) = \{\bar{x}, \check{y}(\bar{x}), f[\bar{x}, \check{y}(\bar{x})]\}$$

we have

$$f(\bar{x}, \bar{z}) - \bar{t}_1 = 0$$

$$w(\bar{x}) - \bar{t}_1 = 0$$

and hence, the following relation holds:

$$\begin{aligned} \lim_{k\to\infty} & F(\bar{x}, \bar{z}) + s^k[f(\bar{x}, \bar{z}) + w(\bar{x}) - 2\bar{t}_1] \\ &= F(\bar{x}, \bar{z}) < \infty. \end{aligned} \tag{11}$$

Accordingly, there exists a positive integer K_1 such that

$$\begin{aligned} F(x^k, z^k) &- s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] \\ &> F(\bar{x}, \bar{z}) - s^k[f(\bar{x}, \bar{z}) + w(\bar{x}) - 2\bar{t}_1] \quad \forall k > K_1 \end{aligned}$$

which contradicts the fact that (x^k, z^k, t_1^k) is a globally optimal solution to problem (7) with $s = s^k$. This with problem (8) proves that $(\tilde{x}, \tilde{z}, \tilde{t}_1)$ is a feasible solution to problem (6).

Global Optimality of the Accumulation Point: Suppose that $(\tilde{x}, \tilde{z}, \tilde{t}_1)$ is not a globally optimal solution to (6), then we have the existence of (x', z', t_1') such that

$$F(x', z') < F(\tilde{x}, \tilde{z}) \tag{12a}$$

$$G(x', z') \leqq 0, \quad g(x', z') \leqq 0 \tag{12b}$$

$$f(x', z') - t_1' = 0 \tag{12c}$$

$$w(x') - t_1' = 0. \tag{12d}$$

The continuity of $F(x, y)$ and (12a) imply that there exists a positive integer K_2 such that

$$F(x', z') < F(x^k, z^k) \quad \forall k > K_2. \tag{13}$$

Hence, we have the following relation for all $k > K_2$

$$\begin{aligned} F(x', z') &- s^k[f(x', z') + w(x') - 2t_1'] \\ &= F(x', z') \\ &< F(x^k, z^k) \\ &\leqq F(x^k, z^k) - s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] \end{aligned}$$

which contradicts the fact that (x^k, z^k, t_1^k) is a globally optimal solution to (7). Therefore $(\tilde{x}, \tilde{z}, \tilde{t}_1)$ solves (6) globally.

Proof of $\lim_{k\to\infty} s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] = 0$: Suppose that

$$\lim_{k\to\infty} s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] \neq 0.$$

Then there exists a positive number ε and a positive integer K_3 such that

$$\begin{aligned} F(x^k, z^k) &- s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] \\ &> F(\tilde{x}, \tilde{z}) + \varepsilon \quad \forall k > K_3. \end{aligned}$$

Therefore, we have the following relation for all $k > K_3$:

$$\begin{aligned} F(\tilde{x}, \tilde{z}) &- s^k[f(\tilde{x}, \tilde{z}) + w(\tilde{x}) - 2\tilde{t}_1] \\ &= F(\tilde{x}, \tilde{z}) \\ &< F(\tilde{x}, \tilde{z}) + \varepsilon \\ &< F(x^k, z^k) - s^k[f(x^k, z^k) + w(x^k) - 2t_1^k] \end{aligned}$$

which contradicts the fact that (x^k, z^k, t_1^k) is a globally optimal solution to (7). ■

Theorem 1 shows that a sequence of globally optimal solutions to the auxiliary problem (7) converges to a globally optimal one to the transformed problem (6).

V. Global Optimization of the Auxiliary Problem

The auxiliary problem (7) is an inequality constrained optimization problem where the objective is a difference of two convex functions and the constraint functions are convex. We now consider the following problem related to (7):

$$\min_{x, z, t_1, t_2} \; t_2 - s[f(x, z) + w(x) - 2t_1] \tag{14a}$$

$$\text{subj. to } F(x, z) - t_2 \leqq 0 \tag{14b}$$

$$G(x, z) \leqq 0 \tag{14c}$$

$$g(x, z) \leqq 0 \tag{14d}$$

$$f(x, z) - t_1 \leqq 0 \tag{14e}$$

$$w(x) - t_1 \leqq 0. \tag{14f}$$

This is a concave program whose objective function is concave and whose constraint set is convex.

The following theorem holds in regard to the equivalence relation between (7) and (14).

Theorem 2: Let the penalty parameter $s = s^k$ be given. Then,

1) If (x^k, z^k, t_1^k) solves problem (7), then there exists a $t_2^k \in R^1$ such that (x^k, z^k, t_1^k, t_2^k) solves problem (14).
2) If (x^k, z^k, t_1^k, t_2^k) solves problem (14), then (x^k, z^k, t_1^k) solves problem (7).

Proof: See Appendix. ■

By this theorem, we can obtain a globally optimal solution to problem (7) by solving the concave program (14).

Consequently, we have the following procedure for finding a globally optimal solution to problem (6) which is equivalent to the original Stackelberg problem (3).

Main Algorithm:

Step 1) Let $\delta > 0$ be a termination scalar. Choose a penalty parameter $s^1 > 0$ and a scalar $\beta > 1$. Set $k := 1$.

Step 2) Solve the concave program (14) with $s = s^k$, and obtain (x^k, z^k, t_1^k, t_2^k) the globally optimal solution. Go to Step 3.

Step 3) If $-s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k] < \delta$, then take $(\boldsymbol{x}^k, z^k, t_1^k)$ as the globally optimal solution to (6) and terminate. Otherwise, set $s^{k+1} := \beta s^k$, $k := k+1$, and go to Step 2.

In Step 3, we terminate the computation when the penalty term $-s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k]$ gets sufficiently small. This idea has been widely used in penalty methods, e.g., [5].

We now explain how to solve the concave program (14) with $s = s^k$ given. Generally, there are several local optima in a concave program and they are attained at extreme points of the constraint set because of the concavity of the objective and the convexity of the constraint set (see Theorem I.1 of [10]). Therefore, we only need to search the set of extreme points of the constraint set for finding a globally optimal solution to a concave program.

The global optimization for the concave program (14) in Step 2 is achieved as follows.

There are two main approaches for obtaining a globally optimal solution to a concave program: 1) the outer approximation method by cutting planes and the 2) branch-and-bound method. Horst and Tuy [10] presented a survey of these approaches. Here, we apply the outer approximation method by cutting planes to solve the concave program (14) with $s = s^k$.

Let the feasible set

$$\begin{aligned} S = \{(\boldsymbol{x}, z, t_1, t_2) |\, & F(\boldsymbol{x}, z) - t_2 \leqq 0, \\ & \boldsymbol{G}(\boldsymbol{x}, z) \leqq \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}, z) \leqq \boldsymbol{0}, \\ & f(\boldsymbol{x}, z) - t_1 \leqq 0, w(\boldsymbol{x}) - t_1 \leqq 0\} \end{aligned}$$

be enclosed in a polytope $S_1 \supset S$. Instead of solving (14), we solve the relaxed problem

$$\min_{(\boldsymbol{x}, z, t_1, t_2) \in S_1} t_2 - s^k[f(\boldsymbol{x}, z) + w(\boldsymbol{x}) - 2t_1]. \tag{15}$$

Let $(\boldsymbol{x}^{k,1}, z^{k,1}, t_1^{k,1}, t_2^{k,1})$ be a globally optimal solution of the above relaxed problem. Then, obviously, $(\boldsymbol{x}^{k,1}, z^{k,1}, t_1^{k,1}, t_2^{k,1})$ solves (14) if $(\boldsymbol{x}^{k,1}, z^{k,1}, t_1^{k,1}, t_2^{k,1}) \in S$. Otherwise, we can find a hyperplane $\ell_1(\boldsymbol{x}, z, t_1, t_2) = 0$ separating S and $(\boldsymbol{x}^{k,1}, z^{k,1}, t_1^{k,1}, t_2^{k,1})$ in the sense that

$$\ell_1(\boldsymbol{x}^{k,1}, z^{k,1}, t_1^{k,1}, t_2^{k,1}) > 0$$

and

$$\ell_1(\boldsymbol{x}, z, t_1, t_2) \leqq 0 \quad \text{for all} \quad (\boldsymbol{x}, z, t_1, t_2) \in S.$$

The linear constraint $\ell_1(\boldsymbol{x}, z, t_1, t_2) \leqq 0$ is added to the system of inequalities defining S_1. We cut off the point $(\boldsymbol{x}^{k,1}, z^{k,1}, t_1^{k,1}, t_2^{k,1})$ and determine a new polytope S_2 that approximates S tighter than S_1. Then, S_1 is replaced by S_2 and the procedure repeated.

Since the functions $F(\boldsymbol{x}, \boldsymbol{y})$, $\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{y})$, $f(\boldsymbol{x}, \boldsymbol{y})$, $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})$, and $w(\boldsymbol{x})$ are all convex, the maximal component function

$$\begin{aligned} p(\boldsymbol{x}, z, t_1, t_2) = \max\{ & F(\boldsymbol{x}, z) - t_2, G_1(\boldsymbol{x}, z), \cdots, \\ & G_M(\boldsymbol{x}, z), g_1(\boldsymbol{x}, z), \cdots, g_m(\boldsymbol{x}, z), \\ & f(\boldsymbol{x}, z) - t_1, w(\boldsymbol{x}) - t_1\} \end{aligned}$$

is convex and possesses subgradients. Further, how to compute a subgradient of an optimal-value function or a maximal component function can be found in [8], [9].

However, there might exist a vector $\overline{\boldsymbol{x}} \in \{\boldsymbol{x} | (\boldsymbol{x}, z, t_1, t_2) \in S_j\}\ (j = 1, 2, \cdots)$ such that the set $\{\boldsymbol{y} | \boldsymbol{g}(\overline{\boldsymbol{x}}, \boldsymbol{y}) \leqq \boldsymbol{0}\}$ be empty, i.e., no subgradient of $w(\overline{\boldsymbol{x}})$ will be available. To cope algorithmically with this situation, we take S_1 so that $S_1 \subset \{(\boldsymbol{x}, z, t_1, t_2) | \boldsymbol{g}(\boldsymbol{x}, z) \leqq \boldsymbol{0}\}$.

Assuming that an initial polytope S_1 is given, at each iteration j define the linear function $\ell_1(\boldsymbol{x}, z, t_1, t_2)$ as follows:

$$\ell_j(\boldsymbol{x}, z, t_1, t_2) = \boldsymbol{q}^{jT} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{x}^{k,j} \\ z - z^{k,j} \\ t_1 - t_1^{k,j} \\ t_2 - t_2^{k,j} \end{pmatrix} + p(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j})$$

where $\boldsymbol{q}^j$ is a subgradient of the convex function p at the trial point $(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j})$.

Accordingly, a procedure of the outer approximation by cutting planes for solving the concave program (14) can be stated as follows. Note that the set of extreme points of a polytope is its vertex set.

Partial Algorithm for Step 2:

Step 2–1) Find a polytope S_1 so that $S_1 \supset S$ and $S_1 \subset \{(\boldsymbol{x}, z, t_1, t_2) | \boldsymbol{g}(\boldsymbol{x}, z) \leqq \boldsymbol{0}\}$. Set $j := 1$.

Step 2–2) To obtain a new trial point $(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j})$, solve

$$\min_{(\boldsymbol{x}, z, t_1, t_2) \in V(S_j)} t_2 - s^k[f(\boldsymbol{x}, z) + w(\boldsymbol{x}) - 2t_1] \tag{16}$$

where $V(S_j)$ is the vertex set of the polytope S_j. If $(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j}) \in S$, take $(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j})$ as a globally optimal solution to (14) and terminate Step 2. Otherwise, go to Step 2-3.

Step 2–3) Add the following linear constraint:

$$\ell_j(\boldsymbol{x}, z, t_1, t_2) = \boldsymbol{q}^{jT} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{x}^{k,j} \\ z - z^{k,j} \\ t_1 - t_1^{k,j} \\ t_2 - t_2^{k,j} \end{pmatrix} + p(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j}) \leqq 0$$

to the system of inequalities defining S_j, where $\boldsymbol{q}^j$ is a subgradient of p at the trial point $(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j})$. Set $S_{j+1} := S_j \cap \{(\boldsymbol{x}, z, t_1, t_2) | \ell_j(\boldsymbol{x}, z, t_1, t_2) \leqq 0\}$, $j := j+1$ and go to Step 2-2.

In Step 2-2, we need to find the vertex set $V(S_j)$ of the polytope S_j in order to solve problem (16) being equivalent to problem (15). For doing this, some procedures are found in Section II-D of [10].

It is proved (e.g., Theorem II.1 of [10]) that every accumulation point of the sequence $(\boldsymbol{x}^{k,j}, z^{k,j}, t_1^{k,j}, t_2^{k,j})$ is a globally optimal solution to problem (14).

VI. Numerical Example

In order to illustrate the effectiveness of the proposed method, we consider a simple numerical example as follows:

$$\begin{aligned} & \min_x 16x^2 + 9[\breve{y}(x)]^2 \\ & \quad \text{subj. to } -4x + \breve{y}(x) \leqq 0 \\ & \qquad x \geqq 0 \\ & \qquad [x + \breve{y}(x) - 20]^4 = \min_y (x + y - 20)^4 \\ & \qquad\quad \text{subj. to } 4x + y - 50 \leqq 0 \\ & \qquad\quad y \geqq 0. \end{aligned} \tag{17}$$

The feasible set of this problem consists of the line segments $\overline{AB}$ and $\overline{BC}$ as shown in Fig. 1. The set $\overline{ABC}$ is not convex. Furthermore, this problem has two locally optimal solutions which are $D(7.2, 12.8)^T$ and $E(11.25, 5)^T$, and the globally optimal one is $E(11.25, 5)^T$.

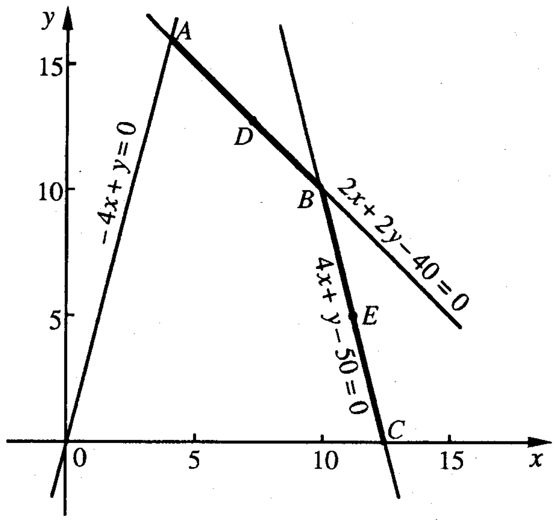

Fig. 1. Feasible solution set and optimal solutions of the example.

TABLE I
COMPUTATIONAL RESULTS OF THE EXAMPLE

k	s^k	(x^k, z^k)	(t_1^k, t_2^k)	F^k	$s^k H^k$	$F^k - s^k H^k$
1	0.1	(5.546, 10.48)	(248.6, 1481.2)	1481.2	24.86	1506.1
2	1.0	(5.899, 10.49)	(170.7, 1546.4)	1546.4	170.7	1717.2
3	10.0	(6.646, 11.58)	(9.826, 1914.3)	1914.3	98.26	2012.6
4	100.0	(11.25, 4.999)	(197.8, 2250.0)	2250.0	$7.252 \cdot 10^{-4}$	2250.0
True Values		(11.25, 5)	—	2250	—	—

Another feature of problem (17) is that objective function in the lower-level is not quadratic. Therefore, its globally optimal solution cannot be obtained by use of any other approaches. However, when the objective function in the lower-level is quadratic, such a problem can be solved by using the method proposed in [11].

Using the method proposed in this paper, we have a concave program which corresponds to problem (14) as follows:

$$\begin{aligned}
&\min_{x,z,t_1,t_2} \; t_2 - s[(x+z-20)^4 + w(x) - 2t_1] \\
&\text{subj. to } 16x^2 + 9z^2 - t_2 \leqq 0 \\
&\qquad -4x + z \leqq 0 \\
&\qquad 4x + z - 50 \leqq 0 \\
&\qquad (x+z-20)^4 - t_1 \leqq 0 \\
&\qquad w(x) - t_1 \leqq 0 \\
&\qquad x, z \geqq 0
\end{aligned} \tag{18}$$

where $w(x)$ is the optimal-value function of the lower-level problem. That is,

$$w(x) = \min_y \{(x+y-20)^4 \,|\, 4x + y - 50 \leqq 0,\; y \geqq 0\}.$$

Table I summarizes the computational results obtained by the proposed method. Each line of Table I shows the results to problem (18) with $s = s^k$. Note that k is the iteration number, and $F^k = F(x^k, z^k)$, $H^k = f(x^k, z^k) + w(x^k) - 2t_1^k$.

The data shown in Table I verified that the globally optimal solution to problem (17) was obtained.

VII. SUMMARY AND CONCLUSIONS

This paper proposed a global optimization method for the Stackelberg problem. First, by introducing the optimal-value function of the lower-level problem, the original problem was equivalently transformed into a nonlinear program with an equality and some inequality constraints where all functions are convex. Then, by using the concept of exterior penalty method, we obtained an auxiliary problem having only inequality constraints for the transformed problem. It was proved that any accumulation point of a sequence of globally optimal solutions to the auxiliary problems solves the transformed problem. After that, we showed that the auxiliary problem was equivalently transformed into a concave program for which a globally optimal solution could be found. After all, we solved the Stackelberg problem by solving a sequence of concave programs, by which global optimization is achieved.

The proposed method can be applied to a very broad class of the Stackelberg problem, and it does guarantee the global optimality. A numerical example was given to verify the proposed algorithm.

However, the conclusion of Theorem 1 is slightly weak, which does not show the existence of an exact penalty factor. This might be improved by adopting exact penalty approach. Furthermore, we applied the cutting plan method proposed in [10] to solve the concave program. In order to improve computational speed, however, we are attempting to develop more effective algorithms for solving the concave program.

APPENDIX
PROOF OF THEOREM 2

1) Let $t_2^k = F(\boldsymbol{x}^k, z^k)$. Then, $(\boldsymbol{x}^k, z^k, t_1^k, t_2^k)$ is a feasible solution to problem (14) obviously. Suppose that $(\boldsymbol{x}^k, z^k, t_1^k, t_2^k)$ is not a globally optimal solution to problem (14), then there exists a $(\boldsymbol{x}', \boldsymbol{z}', t_1', t_2')$ such that

$$\begin{aligned}
&t_2' - s^k[f(\boldsymbol{x}', \boldsymbol{z}') + w(\boldsymbol{x}') - 2t_1'] \\
&\quad < t_2^k - s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k] && \text{(A1)} \\
&\quad F(\boldsymbol{x}', \boldsymbol{z}') - t_2' \leqq 0 && \text{(A2)} \\
&\quad \boldsymbol{G}(\boldsymbol{x}', \boldsymbol{z}') \leqq \boldsymbol{0},\; \boldsymbol{g}(\boldsymbol{x}', \boldsymbol{z}') \leqq \boldsymbol{0} && \text{(A3)} \\
&\quad f(\boldsymbol{x}', \boldsymbol{z}') - t_1' \leqq 0,\; w(\boldsymbol{x}') - t_1' \leqq 0. && \text{(A4)}
\end{aligned}$$

(A3), (A4) implies that $(\boldsymbol{x}', \boldsymbol{z}', t_1')$ is a feasible solution to (7). Furthermore, we have the relation

$$\begin{aligned}
&F(\boldsymbol{x}', \boldsymbol{z}') - s^k[f(\boldsymbol{x}', \boldsymbol{z}') + w(\boldsymbol{x}') - 2t_1'] \\
&\quad \leqq t_2' - s^k[f(\boldsymbol{x}', \boldsymbol{z}') + w(\boldsymbol{x}') - 2t_1'] \\
&\quad < t_2^k - s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k] \\
&\quad = F(\boldsymbol{x}^k, z^k) - s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k]
\end{aligned}$$

which contradicts the fact that $(\boldsymbol{x}^k, z^k, t_1^k)$ is a globally optimal solution to (7). This proves that $(\boldsymbol{x}^k, z^k, t_1^k, t_2^k)$ is a globally optimal solution to (14).

2) We prove by contradiction method. Suppose that $(\boldsymbol{x}^k, z^k, t_1^k)$ is not a globally optimal solution to (7), then we have the existence of $(\boldsymbol{x}', \boldsymbol{z}', t_1')$ such that

$$\begin{aligned}
&F(\boldsymbol{x}', \boldsymbol{z}') - s^k[f(\boldsymbol{x}', \boldsymbol{z}') + w(\boldsymbol{x}') - 2t_1'] \\
&\quad < F(\boldsymbol{x}^k, z^k) - s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k] \\
&\quad \boldsymbol{G}(\boldsymbol{x}', \boldsymbol{z}') \leqq \boldsymbol{0},\; \boldsymbol{g}(\boldsymbol{x}', \boldsymbol{z}') \leqq \boldsymbol{0} \\
&\quad f(\boldsymbol{x}', \boldsymbol{z}') - t_1' \leqq 0,\; w(\boldsymbol{x}') - t_1' \leqq 0.
\end{aligned}$$

Let $t_2' = F(\boldsymbol{x}', z')$ then $(\boldsymbol{x}', z', t_1', t_2')$ is a feasible solution to (14), and the following relation holds:

$$\begin{aligned} &t_2' - s^k[f(\boldsymbol{x}', z') + w(\boldsymbol{x}') - 2t_1'] \\ &\quad = F(\boldsymbol{x}', z') - s^k[f(\boldsymbol{x}', z') + w(\boldsymbol{x}') - 2t_1'] \\ &\quad < F(\boldsymbol{x}^k, z^k) - s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k] \\ &\quad < t_2^k - s^k[f(\boldsymbol{x}^k, z^k) + w(\boldsymbol{x}^k) - 2t_1^k]. \end{aligned}$$

This contradicts the fact that $(\boldsymbol{x}^k, z^k, t_1^k, t_2^k)$ is a globally optimal solution to (14). This proves that $(\boldsymbol{x}^k, z^k, t_1^k)$ is a globally optimal solution to (7). ∎

REFERENCES

[1] E. Aiyoshi and K. Shimizu, "A solution method for the static constrained Stackelberg problem via penalty method," *IEEE Trans. Automat. Contr.*, vol. AC-29, no. 12, pp. 1111–1114, 1984.

[2] F. Al-Khayyal, R. Horst, and P. Pardalos, "Global optimization of concave functions subject to quadratic constraints: An application in nonlinear bilevel programming," *Annals of Operations Res.*, vol. 34, pp. 125–147, 1992.

[3] J. F. Bard, "Optimality conditions for the bilevel programming problem," *Naval Res. Logistics Quart.*, vol. 35, pp. 413–418, 1988.

[4] ——, "Convex two-level optimization," *Math. Programming*, vol. 40, pp. 15–27, 1988.

[5] M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms.* New York: Wiley, 1979.

[6] T. A. Edmunds and J. F. Bard, "Algorithms for the nonlinear bilevel mathematical programs," *IEEE Trans. Syst. Man Cyber.*, vol. 21, no. 1, pp. 83–89, 1991.

[7] J. Fortuny-Amat and B. McCarl, "A representation and economic interpretation of a two-level programming problem," *J. Oper. Res. Soc.*, vol. 32, pp. 783–792, 1981.

[8] J. Gauvin and F. Dubean, "Differential properties of the marginal function in mathematical programming," *Math. Programming Study*, vol. 19, pp. 101–119, 1982.

[9] W. Hogan, "Directional derivatives for extremal-value functions with applications to the completely convex case," *Oper. Res.*, vol. 21, pp. 188–209, 1973.

[10] R. Horst and H. Tuy, *Global Optimization—Deterministic Approaches.* New York: Springer-Verlag, 1990.

[11] M. Lu and K. Shimizu, "A global optimization technique for solving nonlinear programming problem with equality and inequality constraints and its application to the bilevel programming problem," *Trans. Soc. Instrument and Control Engineers*, vol. 28, no. 7, pp. 879–886, 1992, in Japanese.

[12] O. L. Mangasarian and J. B. Rosen, "Inequalities for stochastic nonlinear programming problems," *Oper. Res.*, vol. 12, pp. 143–154, 1964.

[13] K. Shimizu and E. Aiyoshi, "A new computational method for Stackelberg and min-max problems by use of a penalty method," *IEEE Trans. Automat. Contr.*, vol. AC-26, no. 2, pp. 460–466, 1981.

[14] K. Shimizu and Y. Ishizuka, "Optimality conditions and algorithms for parameter design problems with two-level structure," *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 10, pp. 986–993, 1985.

[15] M. Simaan and J. B. Cruz, Jr., "On the Stackelberg strategy in nonzero-sum games," *J. Optimization Theory and Applications*, vol. 11, no. 5, pp. 533–555, 1973.

[16] M. Simman, "Stackelberg optimization of two-level systems," *IEEE Trans. Syst. Man Cyber.*, vol. SMC-7, no. 4, pp. 554–556, 1977.

$\theta(1)$ Time Quadtree Algorithm and Its Application for Image Geometric Properties on a Mesh Connected Computer (MCC)

Jelloul Elmesbahi, Omar Bouattane, and Zouhair Benabbou

***Abstract*—Quadtree and octree decomposition of multicolored images is often used to solve problems of image processing. This paper describes a $\theta(1)$ time algorithm for the decomposition of the images into quadtree and octree on a mesh connected computer. Images will be stored in the MCC of the same size, at one pixel per processing element (PE). Certain applications of the algorithm are also described namely, the computation of such geometric properties as the area and perimeter. The complexity of these algorithms depends on the hierarchical data structure. They are carried out in $O(\log^2 m + \log m)$ times on an $n \times n$ MCC, where $m = n/k$, and k is the size of the smallest homogeneous quadrant in the image.**

I. INTRODUCTION

The quadtree representation of images has been the subject of much research, for there is vast field of applications. For example it is applied to the component labeling, perimeter computation, determination of the euler number, and to region expansion in images [1]–[3].

Quadtrees also facilitate branch structuring of sub-images in response to different criteria, thus offering the possibility of presenting the image at any given level of specificity.

Quadtrees structuring is one of the forms of representation which are based on the systematic decomposition of the image; a decomposition carried out by successive subdivisions of the image matrix into four quadrants of equal size. Each quadrant is then further subdivided into sub-quadrants if it is not homogenous: i.e., all of its constituent pixels do not have the same color. The subdivision is repeated until the sub-quadrants become homogenous or until the sub-quadrants are reduced to one pixel.

Example: Fig. 1(a) gives the image to be decomposed. In Fig. 1(b) we see the division into homogenous quadrants, and Fig. 1(c) shows the quadtree elements following the decomposition. As well as the quadtrees representation.

Quadtree problems were aborted by several authors, both in serial and parallel architectures. In our case that is related to the parallel architecture, we will present some recent works concerning quadtree representation and its applications, on different parallel architectures. For example in [4], the used architecture is a horizontally reconfigurable architecture computing system, in [5], authors use a reconfigurable cellular array of $n \times n$ processors which stores binary images of the same size. The subdivision process used therein is an $O(n)$ times algorithm, that configure the matrix of processors in order to construct the image quadtree. One can find other references in the literature concerning the parallel architectures [6], [7].

Our work which deals with an $n \times n$ reconfigurable mesh connected computer, is an algorithm that gives some information to the processing elements to indicate the position of the PE in the quadtree representation of the image; it has a complexity of $\theta(1)$ time. The information received by each PE allow us to envisage easily other

Manuscript received January 9,. 1992; revised October 7, 1994. This work was supported by Ecole Marocaine Des Sciences De l'Ingenieur (EMSI) Casablanca, Morocco.

The authors are with the Faculty of Science, Ain Chock Hassan II University, Hay Essalam Casablanca, Morocco.

IEEE Log Number 9414506.

0018–9472/95$04.00 © 1995 IEEE