Volume 37, Number 2, 2012, pp. 201–217

A NEW DECOMPOSITION ALGORITHM FOR GLOBALLY SOLVING MATHEMATICAL PROGRAMS WITH AFFINE EQUILIBRIUM CONSTRAINTS

L. D. MUU, Q. TRAN DINH, L. T. H. AN AND P. D. TAO

Abstract. This paper proposes a new decomposition method for globally solving mathematical programming problems with affine equilibrium constraints (AMPEC). First, we view AMPEC as a bilevel programming problem where the lower level problem is a parametric affine variational inequality. Then we use a regularization technique to formulate the resulting problem as a mathematical program with an additional constraint defined by the difference of two convex functions (DC function). A main feature of this DC decomposition is that the second component depends upon only the parameter in the lower level problem. This property allows us to develop branch-and-bound algorithms for globally solving AMPEC where the adaptive rectangular bisection takes place only in the space of the parameters. As an example, we use the proposed algorithm to solve a bilevel Nash-Cournot equilibrium market model. Computational results show the efficiency of the proposed algorithm.

1. Introduction

In this paper, we consider the following mathematical programming problem with affine (not necessarily monotone) variational inequality constraints (AMPEC):

$$\min_{x\in\mathbb{R}^n, y\in\mathbb{R}^m} \quad f(x,y) \tag{1.1}$$

$$\text{s.t.} \quad (x,y)\in S, \tag{1.2}$$

$$x\in C,\ (Ax+By+a)^T(v-x)\geq 0,\ \forall v\in C, \tag{1.3}$$

where $\emptyset \neq S \subseteq \mathbb{R}^{n+m}$, $\emptyset \neq C \subseteq \mathbb{R}^n$ are two closed convex sets, $f:\mathbb{R}^{m+n}\to\mathbb{R}$ is a convex function, A, B are given appropriate real matrices and $a\in\mathbb{R}^n$. This class of optimization problems is known to be very difficult to solve due to its nonconvexity, nondifferentiability and loss of the constraint qualification. However such problems arise frequently in many applications, for example, in

Received January 27, 2011.

2010 *Mathematics Subject Classification.* 49M37, 90C26, 65K15.

Key words and phrases. Mathematical programs with affine equilibrium constraints, regularization, bilevel convex quadratic programming, DC formulation, global optimization, Nash-Cournot model.

This work is supported by the Vietnam National Foundation for Science Technology Development (NAFOSTED) under grant number 10102-2011.19.

shape optimization, design transportation network, economic modeling and data mining. A natural way to handle a nested problem such as (1.1)-(1.3) is to reduce it into an one-level optimization problem by using the Karush-Kuhn-Tucker theorem for the lower level variational inequality. Several algorithms for globally solving the reduced mathematical programs with complementarity constraints were proposed, see, e.g. [3, 5, 19, 22]. Since the number of the complementarity constraints is just equal to the number of constraints defining the set C in the lower variational inequality problem, these global optimization algorithms become expensive when the number of constraints is high, for example, when $C := \{x \in \mathbb{R}^n \mid x \geq 0,\ c_j(x) \leq 0,\ j = 1, \ldots, p\}$ with $c_j,\ j = 1, ..., p$ being continuous functions (even linear) and either n or p are somewhat large (often appears in practice).

In this paper, we propose another solution-approach to AMPEC without using the Karush-Kuhn-Tucker theorem for the lower level variational inequality. We use a regularization technique to formulate AMPEC as a mathematical program with an additional constraint defined by $g_1(x, y) - h_1(x, y) \leq 0$, where g_1 and h_1 are differentiable convex functions. The main feature of this constraint is that the second component h_1 can be chosen such that it only depends upon the parameter y. Moreover, in some special important cases such as bilevel convex quadratic problems, h_1 is separable. This formulation allows us to develop a decomposition branch-and-bound algorithm for globally solving AMPEC where the branching operation involving only the parameters in the lower level variational inequality. Unlike the existing global optimization algorithms mentioned above, the proposed algorithm can solve AMPEC where the constraint set C is given as $C := \{x \in \mathbb{R}^n \mid x \geq 0,\ c_j(x) \leq 0,\ j = 1, \ldots, p\}$ with n and p relatively large. As an example, we use the proposed algorithm to find a global optimal equilibrium pair to a bilevel Nash-Cournot equilibrium market model. We test the proposed algorithm by some randomly generated data. The numerical results show that our algorithm can solve this bilevel model for high dimensional problems.

The rest of the paper is organized as follows. In the next section we give a DC formulation to AMPEC by using suitable regularization matrices. Some important special cases of AMPEC are presented at the end of this section. The third section is devoted to description of a branch-and-bound algorithm for globally solving a bilevel Nash-Cournot equilibrium market model by using a DC decomposition, where the second component is separable and depends upon only the parameter y. Section 4 presents a numerical example. The paper is ended by some conclusions.

2. DC formulations and examples

Let us recall the AMPEC problems (1.1)-(1.3). As usual, we will refer to x as a primary variable or decision variable and y as a parameter. We call (x, y) a feasible point to (1.1)-(1.3) if $(x, y) \in S$ and x solves the lower level variational inequality (1.3). Note that when A is symmetric positive semidefinite, the variational inequality (1.3) is equivalent to the following parametric convex quadratic problem

$$\min\left\{\varphi(x,y) := \frac{1}{2}x^T Ax + (By+a)^T x \mid x \in C\right\}. \tag{2.1}$$

In this particular case, (1.1)-(1.3) becomes a bilevel convex program

$$\min\{f(x,y) \mid (x,y) \in S\}, \tag{BP}$$

where x solves the convex quadratic program

$$\min\left\{\varphi(x,y) := \frac{1}{2}x^T Ax + (By+a)^T x \mid x \in C\right\}. \tag{2.2}$$

In the general case, when A is indefinite, the variational inequality (1.3) is not necessarily equivalent to problem (2.2). Therefore, Problem (1.1)-(1.3), in general, can not be reformulated as a bilevel problem of the form (BP), see, e.g. [7, 14].

2.1. **DC Formulations.** The main difficulty of solving problem (1.1)-(1.3) is that the constraint defined by the variational inequality (1.3) is neither convex nor given explicitly as a constraint set of a standard mathematical programming problem. A natural way for solving (1.1)-(1.3) is to reformulate it into a standard mathematical programming problem. In this paper, we shall reformulate problem (1.1)-(1.3) as a smoothly DC (difference of two convex functions) program. We recall that a function f is said to be DC on a convex set D if it can be expressed as the difference of two convex functions on D, i.e. $f = g - h$, where g and h are convex on D. In order to reformulate (1.3) as a DC constraint, we use a gap function proposed in [21] to formulate the variational inequality (1.3) into a nonlinear equation defined by a smoothly DC function. More precisely, for each (x,y), we define the function $g(x,y)$ by letting

$$g(x,y) := \max_{v\in C}\left\{(x-v)^T(Ax+By+a) - \frac{1}{2}(v-x)^T G(v-x)\right\}, \tag{2.3}$$

where G is an arbitrary $n \times n$-symmetric positive definite matrix. We refer to G as a *regularization matrix*. Since G is positive definite, the optimization problem in (2.3) is uniquely solvable for every (x,y), i.e. g is well-defined.

The following lemma provides the properties of the gap function g. The proof of this lemma can be found in [21].

Lemma 2.1. *Let g be defined by* (2.3). *Then*

(i) $g(x,y) \geq 0$ *for every* $(x,y) \in C \times \mathbb{R}^m$,
(ii) $(x,y) \in S$, $x \in C$, $g(x,y) = 0$ *if and only if* (x,y) *is a feasible solution of* (1.1)-(1.3).

For any symmetric matrix A, it can be expressed as $A = A_1 - A_2$, where A_1 is symmetric positive definite and A_2 is symmetric. In what follows by $\mathrm{diag}(\alpha)$ we denote the diagonal matrix whose every diagonal entry is α. The following proposition shows that, with a suitable choice of the regularization matrix G, the function g can be represented as a DC representation.

Proposition 2.2. *Suppose that A is symmetric and $A = A_1 - A_2$, where A_1 is a symmetric positive definite matrix and A_2 is a symmetric matrix such that $A_2 + \frac{1}{2}U^TU$ is positive (semi) definite, and U, V are two appropriate matrices satisfying $U^TV = B$. Let $G = 2A_1$. Then*

$$g(x,y) = g_1(x,y) - h_1(x,y), \tag{2.4}$$

where g_1 and h_1 are two differentiable convex functions given by

$$\begin{aligned} g_1(x,y) \quad &:= \frac{1}{2}\|Ux+Vy\|^2 + a^Tx \\ &+ \max_{v\in C}\left\{[(A_1+A_2)x - By - a]^Tv - v^TA_1v\right\}, \end{aligned} \tag{2.5}$$

and

$$h_1(x,y) := \frac{1}{2}x^T(2A_2 + U^TU)x + \frac{1}{2}\|Vy\|^2. \tag{2.6}$$

Proof. With a simple arrangement from (2.3), it shows that

$$\begin{aligned} g(x,y) \quad &= x^TAx - \frac{1}{2}x^TGx + x^TBy + a^Tx \\ &+ \max_{v\in C}\{-v^TAx - v^TBy - a^Tv - \frac{1}{2}v^TGv + x^TGv\}. \end{aligned} \tag{2.7}$$

Since $A = A_1 - A_2$ and $G = 2A_1$, the last expression implies

$$\begin{aligned} g(x,y) \quad &= -x^TA_2x + x^TBy + a^Tx \\ &+ \max_{v\in C}\{-v^TA_1v + [(A_1+A_2)x - By - a]^Tv\}. \end{aligned} \tag{2.8}$$

On the other hand, since $B = U^TV$ we can express

$$2x^TBy = 2x^TU^TVy = \|Ux+Vy\|^2 - \|Ux\|^2 - \|Vy\|^2.$$

Substituting this expression into (2.8) we get

$$\begin{aligned} g(x,y) = -x^TA_2x + \frac{1}{2}\|Ux+Vy\|^2 - \frac{1}{2}\|Ux\|^2 - \frac{1}{2}\|Vy\|^2 + a^Tx \\ + \max_{v\in C}\{-v^TA_1v + [(A_1+A_2)x - By - a]^Tv\}. \end{aligned}$$

Hence

$$g(x,y) = g_1(x,y) - h_1(x,y),$$

where g_1 and h_1 are two functions given by (2.5) and (2.6), respectively. Since $A_2 + \frac{1}{2}U^TU$ is positive semidefinite, h_1 is convex. Clearly, h_1 is differentiable everywhere, while g_1 is differentiable everywhere because the convex program (strongly quadratic concave maximization):

$$\max_{v\in C}\left\{ - v^TA_1v + [(A_1+A_2)x - By - a]^Tv\right\}$$

is uniquely solvable for any (x,y). □

Remark 2.3. From (2.5), by a simple computation, we have

$$\nabla_x g_1(x, y) = U^T(Ux + Vy) + a + (A_1 + A_2)^T z(x, y), \tag{2.9}$$

$$\nabla_y g_1(x, y) = V^T(Ux + Vy) - B^T z(x, y), \tag{2.10}$$

where $z(x, y)$ is a unique solution of the strongly convex quadratic program

$$\max_{v \in C} \{ - v^T A_1 v + [(A_1 + A_2)x - By - a]^T v \}.$$

Remark 2.4. Since matrices U and V in Proposition 2.2 can be arbitrary, we can choose U and V such that V has a simple form. For example, if we choose $U = [(\Sigma B^+)^T]^+$, where B^+ is the (Moore-Penrose) pseudo-inverse of B and Σ is a diagonal matrix, then V is a diagonal matrix, precisely, $V = \Sigma$.

We call the DC decomposition $g(x, y) = g_1(x, y) - h_1(x, y)$, where g_1 and h_1 are given by (2.5) and (2.6), respectively, a *spectral decomposition.* In this decomposition, the function h_1 is a quadratic form, even separable quadratic if $2A_2 + U^T U$ is diagonal. The separable quadratic property of h_1 is useful when applying to global algorithms that use the convex envelope of $-h_1$ (see Section 3 below).

By using Proposition 2.2, problem (1.1)-(1.3) can be reformulated equivalently to a convex optimization problem with an additional DC constraint of the form:

$$(\mathrm{P}_1) \qquad \min_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} f(x, y)$$

$$\text{s.t. } (x, y) \in S, \; x \in C \tag{2.11}$$

$$g(x, y) = g_1(x, y) - h_1(x, y) \leq 0, \tag{2.12}$$

where g_1 and h_1 are given by (2.5) and (2.6), respectively.

Formulation (P_1) allows us to apply theory and methods in smooth and DC optimization both local and global to mathematical programs with affine equilibrium constraints.

2.2. **Special Cases.** In this subsection, we consider some special, but important, cases of problem (1.1)-(1.3) and their reformulation in the form of (P_1).

2.2.1. *Linear program with linear complementarity constraints.* Note that when $C = \mathbb{R}^n_+$, S is a polyhedron defined by

$$S := \{(x, y) \; : \; Ax + By + a \geq 0\},$$

and $f(x, y) = c^T x + c^T y$, Problem (1.1)-(1.3) becomes a linear program with an additional linear complementarity constraint of the form

$$(\mathrm{CP}) \qquad \min_{(x,y)} f(x, y),$$

$$\text{s.t. } x \geq 0, \; Ax + By + a \geq 0, \; x^T(Ax + By + a) = 0. \tag{2.13}$$

For this program, the following gap function has been used [4, 16, 3]:

$$p(x, y) = \sum_{j=1}^{n} \min\{x_j, (Ax + By + a)_j\}.$$

It has been shown that if f is bounded from below, then there exists $t_* > 0$ such that for every $t \geq t_*$, Problem (CP) is equivalent to the following concave minimization problem

$$\min_{(x,y)} \{f(x,y) + tp(x,y)\}$$
$$\text{s.t. } x \geq 0,\ Ax + By + a \geq 0,$$

in the sense that their solution sets coincide. In [16], Mangasarian and Pang replaced p by the differentiable function

$$\min\{\sum_{j=1}^{n} r_j x_j + s_j(Ax + By + a)_j \mid r_j, s_j \geq 0, r_j + s_j = 1, j = 1, \dots, n\}.$$

Note that the DC function $g(x,y) = g_1(x,y) - h_1(x,y)$ with g_1 and h_1 given as in Proposition 2.2 is a differentiable merit DC function for (CP) without introducing $2n$-extra variables r and s.

2.2.2. *Linear optimization over the Pareto-efficient set.* Let $X \subset \mathbb{R}^n$ be a nonempty bounded polyhedron and W be a $(p \times n)$-real matrix. Consider the vector optimization problem of the form

$$\min\{Wx \mid x \in X\}. \tag{2.14}$$

We recall that a point $x^* \in X$ is said to be an *efficient solution* or a *Pareto solution* to (2.14), if whenever $x \in X, Wx \leq Wx^*$, then $Wx = Wx^*$. Let $E(W,X)$ denote the set of all efficient solutions to (2.14). Consider the optimization over the efficient set

$$\min\{f(x) \mid x \in E(W,X)\}, \tag{PP}$$

where f is a real valued convex function on $\mathbb{R}^n$. This problem has some applications in decision making and recently has been studied in many research articles, see, e.g.[1, 2, 6, 15, 17, 20] and the references therein. Since the efficient set is rarely convex, this problem is a nonconvex optimization problem.

It has been shown in [20] that one can find a simplex Y in $\mathbb{R}^p$ such that a point x^* is efficient for (2.14) if and only if there exists $y^* \in Y$ such that

$$(W^T y^*)^T (x - x^*) \geq 0,\ \forall x \in X.$$

Thus the above optimization problem over the efficient set can be formulated as a mathematical program with affine equilibrium constraints of the form:

$$\min\{f(x) \mid (x,y) \in X \times Y,\ (W^T y)^T (v - x) \geq 0,\ \forall v \in X\}. \tag{EP}$$

Consequently, a point x^* is a global optimization to (PP) if and only if there exists $y^* \in Y$ such that (x^*, y^*) is a global optimal solution to (EP). The latter problem is of the form (1.1)-(1.3) with $S = X \times Y$, $C = X$ and $A = 0$, $B = W^T$, $a = 0$. Since $A = 0$, we can apply Proposition 2.2, for example, with $A_1 = A_2 = I$, where I is the identity matrix. Since $B = W^T$, $a = 0$, from Proposition 2.2 we have $g(x) = g_1(x) - h_1(x)$ with

$$g_1(x,y) = \frac{1}{2}\|Ux + Vy\|^2 + a^T x + \max_{v\in C}\left\{(2Ix - W^T y)v - v^T A_1 v\right\},$$

$$h_1(x,y) = \frac{1}{2}x^T(2A_2 + U^T U)x + \frac{1}{2}\|Vy\|^2,$$

where $U^T V = W^T$. Thus, by Lemma 2.1, we can formulate (PP) as the following optimization problem with a DC constraint

$$\min\{f(x) \mid (x,y) \in X \times Y,\ g_1(x,y) - h_1(x,y) \le 0\}.$$

2.2.3. *A bilevel Nash-Cournot oligopolistic equilibrium market model.* Suppose that there are n-firms (sectors) that supply a homogeneous product whose price p at each sector j ($j = 1,\dots,n$) depends on total producing quantity and is given by

$$p(\sum_{j=1}^{n} x_j) = \alpha - \beta \sum_{j=1}^{n} x_j,$$

where $\alpha > 0$, $\beta > 0$ are given constants, x_j is the quantity of goods supplied by firm j that we have to determine. Suppose further that, to produce the goods, the firms need m-different materials represented by a vector $y \in \mathbb{R}^m$. Let y_i be the quantity of material i needed to produce a unique of goods ($i = 1,\dots,m$). Let c_{ji} denote the price of a unit material i for firm j ($i = 1,\dots,m$, $j = 1,\dots,n$). When $c_{ji} \le 0$, it means that firm j is encouraged to use material i; for example, it is a waste material. Assume that the cost of firm j is given by

$$h_j(x_j, y) := x_j \sum_{i=1}^{m} c_{ji} y_i + \delta_j, \quad j = 1,\dots,n,$$

where $\delta_j \ge 0$ is the fixed charge cost at firm j. Then the utility function of firm j can be given by

$$u_j(x,y) := p(\sum_{i=1}^{n} x_i) x_j - h_j(x_j, y).$$

Let

$$Y_i := \{y_i \ :\ 0 \le y_i \le \xi_i\}\ (i = 1,\dots,m),$$
$$X_j := \{\tau \ :\ 0 \le \tau \le \eta_j\}\ (j = 1,\dots,n),$$

where ξ_i is the upper bound of the material i, and η_j is the upper bound of the quantity of the goods produced by firm j.

Let

$$Y := Y_1 \cdots \times Y_m, \quad X = X_1 \times \cdots \times X_n$$

be the feasible (strategy)-sets of the model.

Given $y \in Y$, each firm j seeks to find its producing quantity x_j such that its benefit $u_j(x,y)$ is maximal. However, a maximal policy for all firms altogether, in general, does not exist. So they agree with an equilibrium point in the sense of Nash [8].

By the definition, a vector $(x_1^*, \dots, x_n^*) \in X_1 \times \cdots \times X_n$ is said to be a Nash-equilibrium point with respect to $y^* \in Y$ if, for all $x_j \in X_j$ and j,

$$u_j(x_1^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*, y^*) \le u_j(x_1^*, \dots, x_{j-1}^*, x_j^*, x_{j+1}^*, \dots, x_n^*, y^*).$$

We will refer to such a pair (x^*, y^*) as an *equilibrium pair* of the model.

Beside the utility function of each firm, there is another cost function (leader's objective function) $f(x, y)$ depending on y and the quantity x of the goods. The problem needs to be solved is to find an equilibrium pair that minimizes leader's objective function over the set of all equilibrium pairs. We call such a pair (x^*, y^*) a *global optimal equilibrium pair* of the model. This problem can be reformulated as a mathematical program with affine equilibrium constraints. To this end, let

$$H_j(x_j, y) := \nabla_{x_j} h_j(x_j, y) \ (j = 1, \dots, n),$$

$$e := (1, \dots, 1)^T,$$

and

$$\sigma_x := \sum_{j=1}^{n} x_j.$$

Applying Proposition 3.2.6 in [12] we see that a point $(x_1, \dots, x_n)$ is equilibrium with respect to y if and only if it is a solution of the following variational inequality problem

$$\text{Find } x \in X \text{ such that: } F(x, y)^T (z - x) \ge 0, \text{ for all } z \in X,$$

where $F(x, y)$ is an n-dimensional vector function whose j-th component is defined by

$$F_j(x, y) := H_j(x, y) - p(\sigma_x) - \nabla p(\sigma_x) x_j. \tag{2.15}$$

Using (2.15) and the definition of $H_j(x, y)$ we have

$$F_j(x, y) = \sum_{i=1}^{m} c_{ji} y_i - \alpha + \beta \sum_{k=1}^{n} x_k + \beta x_j \ (j = 1, \dots, n).$$

Thus

$$F(x, y) = Ax + By + a,$$

where

$$A = \begin{bmatrix} 2\beta & \beta & \beta & \cdots & \beta \\ \beta & 2\beta & \beta & \cdots & \beta \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta & \beta & \beta & \cdots & 2\beta \end{bmatrix} \tag{2.16}$$

and B is an $(n \times m)$ matrix (independent of x) whose B_{ij} entry is

$$B_{ji} = c_{ji}, \ j = 1, \dots, n, \ i = 1, \dots, m, \tag{2.17}$$

and

$$a = (-\alpha, \dots, -\alpha)^T \in \mathbb{R}^n. \tag{2.18}$$

Finally, the problem needs to be solved takes the form:

$$\begin{array}{ll} \min\limits_{x,y} & f(x,y) \\ \text{s.t.} & y \in Y := Y_1 \times \cdots \times Y_m,\ x \in X := X_1 \times \cdots \times X_n \\ & \text{where } x \text{ solves the parametric variational inequality:} \\ & (Ax + By + a)^T(v - x) \geq 0,\ \forall v \in X, \end{array}$$

with A, B and a being given by (2.16), (2.17) and (2.18), respectively. This problem is indeed in the form of (1.1)-(1.3), and therefore, we can use Proposition 2.2 to obtain its DC formulation.

2.2.4. *Optimization over the solution-set of a variational inequality.* Let us consider a particular case of problem (1.1)-(1.3) when the variable y is absent. In this case, with $S = \mathbb{R}^{m+n}$, it takes the form

$$(\mathrm{P}_2) \qquad \begin{array}{l} \min f(x) \\ \text{s.t. } x \in C,\ (Ax + a)^T(v - x) \geq 0,\ \forall v \in C, \end{array}$$

where, as before, f is a real valued convex function on C and $\emptyset \neq C \subseteq \mathbb{R}^n$ is a closed convex set. Problems over the solution-set of a pseudomonotone variational inequality were studied in [11] (notions of pseudomonotonicity and monotonicity can be found in [12, 13]). Here, we do not require any assumption on monotonicity. Note that without monotonicity of A, the solution-set of the variational inequality constraint in (P_2) is not necessarily convex. Therefore, this problem remains a nonconvex optimization one. By Lemma 2.1 we can rewrite (P_2) as

$$\min\big\{f(x) \mid x \in C,\ g(x) \leq 0\big\},$$

where, by (2.3),

$$g(x) = x^T Ax - \frac{1}{2}x^T Gx + a^T x + \max_{v \in C}\big\{ - v^T Ax - a^T v - \frac{1}{2}v^T Gv + x^T Gv\big\}.$$

If A is symmetric, we express A as $A = A_1 - A_2$ with A_1 being symmetric positive definite and A_2 symmetric positive (semi) definite. From Proposition 2.2 we have

$$g_1(x) = a^T x + \max_{v \in C}\big\{[(A_1 + A_2)x - a]^T v - v^T A_1 v\big\}, \tag{2.19}$$

and

$$h_1(x,y) = x^T A_2 x. \tag{2.20}$$

Note that when f is constant, Problem (P_2) becomes an affine variational inequality of the form [8, 14]:

$$\text{Find } x \in C \text{ such that: } (Ax + a)^T(v - x) \geq 0, \text{ for all } v \in C.$$

By Lemma 2.1, x is a solution to this problem if and only if it is a global optimal solution to the differentiable DC program:

$$0 = \min\{g(x) := g_1(x) - h_1(x) :\ x \in C\},$$

where g_1 and h_1 are given as in Propositions 2.2.

3. On global optimization methods for AMPEC

Theoretically, the global optimization methods such as branch-and-bound, outer and inner approximations, e.g., [10], can be applied to AMPEC by using the DC formulations obtained in the preceding sections. Note that AMPEC can be equivalently converted into an one-level mathematical program with an additionally complementarity constraint by applying the Karush-Kuhn-Tucker theorem to the lower level variational inequality. Branch-and-Bound-type algorithms developed in [3, 5, 19, 22] can be applied to globally solve the latter problem. These existing algorithms use different subdivisions, but all of them take place in a space whose dimension is equal to the number of the Lagrangian multipliers. The latter number is large when the feasible set of the lower level affine variational inequality is given, as usual, as $C := \{x \in \mathbb{R}^n \mid x \geq 0,\ Px = q\}$ with n large (often in practical problems). However, it is well recognized that global optimization algorithms are only recommended to the case when the dimension of the space, where the global optimization operations such as subdivision take place, is relatively small.

It can be observed that in AMPEC problem (1.1)-(1.3), where A is monotone on C, only the variable y makes the problem nonconvex. In fact, when A is monotone and y is absent, the solution-set of the lower variational inequality is convex. This observation suggests us to look for DC decompositions of g where the second component h_1 that makes g nonconvex depends upon only y. From (2.6) in Proposition 2.2 and Remark 2.4 we see that if we choose $U = [(\Sigma B^+)^T]^+$ and A_2 such that $2A_2 + U^TU = 0$, then h_1 is independent of x and separable. In some models such as bilevel strongly convex quadratic problem [18] and Nash-Cournot equilibrium model (Example 2.2.c), since A is positive definite, one can choose $A_2 = -(1/2)U^TU$. Then, by virtue of Proposition 2.2, we have $h_1(x, y) = \frac{1}{2}\|\Sigma y\|^2$ is independent of x and separable.

As an example, we now describe a branch-and-bound algorithm for minimizing a convex function over the equilibrium set of the Nash-Cournot equilibrium market model that we have studied in Subsection 2.2. In practical Nash-Cournot models, the number m of the materials that the producers need to produce the goods is much less than the number n of the firms, for example, in electricity production, it takes only oil and coal as two main materials into account.

This fact suggests that we should choose a DC decomposition such that the function h_1, which makes the problem nonconvex, only depends upon y variable. For this purpose we choose the DC decomposition given in Proposition 2.2 with

$$A_1 = \begin{bmatrix} 2\beta & \beta & \beta & \cdots & \beta \\ \beta & 2\beta & \beta & \cdots & \beta \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta & \beta & \beta & \cdots & 2\beta \end{bmatrix} - \frac{1}{2}U^TU. \tag{3.1}$$

and

$$A_2 = -\frac{1}{2}U^TU. \tag{3.2}$$

Note that since $\lambda_{\min}(A) = \beta > 0$, where $\lambda_{\min}(A)$ is the smallest eigenvalue of A, the matrix A is positive definite. If we choose Σ such that $\lambda_{\max}(\Sigma) < \beta$, where $\lambda_{\max}(\Sigma)$ is the largest eigenvalue of Σ, then A_1 is still positive definite. By Proposition 2.2, one has

(3.3)

$$g_1(x,y) = \frac{1}{2}\|Ux + Vy\|^2 + a^T x + \max_{v\in C}\big\{-v^T A_1 v + [(A - U^T U)x - By - a]^T v\big\},$$

and

$$h_1(x,y) = h_1(y) := \frac{1}{2}\|\Sigma y\|^2 \text{ (separable and depending on } y \text{ only).} \tag{3.4}$$

Thus computing global optimal Nash equilibrium pairs of the bilevel Nash-Cournot equilibrium market model presented in Subsection 2.2 leads to the problem:

$$\text{(NC)} \qquad \alpha_* := \min\big\{f(x,y) \mid x \in X,\ y \in Y,\ g_1(x,y) - h_1(y) \le 0\big\},$$

where g_1 and h_1 are given by (3.3) and (3.4), respectively.

The separability property of h_1 suggests us to use the convex envelope of h_1 on the box (rectangle) Y to compute the lower bounds in the branch-and-bound algorithm to be described below. Moreover, since h_1 only depends upon the variable $y \in Y$, one can use an adaptive rectangular bisection that only takes place in the y-space.

Now we describe in detail these bounding and branching operations.

3.1. **Bounding by the convex envelope.** We recall [9, 10] that a function $l(y)$ is said to be the convex envelope of a function $q(y)$ on a convex set Y if l is convex on Y, $l(y) \le q(y)$ for every $y \in Y$ and if $p(y)$ is a convex function on Y such that $p(y) \le q(y)$ for every $y \in Y$ then $p(y) \le l(y)$ for every $y \in Y$. In general, computing the convex envelope of a function on an arbitrary convex set, even polyhedron, is expensive. Fortunately, in our case, since h_1 given in (3.4) is separable, concave, and Y is a box, its convex envelope is an affine function that can be given explicitly (see, e.g. [9]). Namely, suppose that $h_1(y) = \sum_{j=1}^m \xi_j y_j^2$, $(\xi \ge 0)$. Let l^R denote the convex envelope of $-h_1$ on the box

$$R := \{y = (y_1, \dots, y_m)^T \mid a_j \le y_j \le b_j,\ j = 1, \dots, m\} \subseteq Y.$$

Then $l^R(y) = \sum_{j=1}^m l_j^R(y_j)$, where l_j^R is the convex envelope of the function $-\xi_j y_j^2$ on the interval $[a_j, b_j]$ $(j = 1, \dots, m)$. The latter in turn is the affine function joining a_j and b_j.

Let $\alpha(R)$ and $\beta(R)$ denote the optimal value of problem (NC) restricted on R and the optimal value of its relaxed problem, respectively, that is

$$(\text{NC}_R) \qquad \alpha(R) := \min\big\{f(x,y) \mid x \in X,\ y \in R,\ g_1(x,y) - h_1(y) \le 0\big\},$$

$$(\text{RNC}_R) \qquad \beta(R) := \min\big\{f(x,y) \mid x \in X,\ y \in R,\ g_1(x,y) + l^R(y) \le 0\big\}.$$

Since $l^R(y) \le -h_1(y)$ for every $y \in R$, we have $\beta(R) \le \alpha(R)$.

3.2. **An adaptive rectangular bisection.** It is clear that if $\beta(R) = \alpha(R)$ then the minimum of f over the set $x \in X, y \in R, g_1(x,y) - h_1(y) \leq 0$ has been found. Otherwise, if $\beta(R) < \alpha(R)$ then there must exist at least one index j such that $l_j^R(y_j^*) < -\xi_j y_j^{*2}$, where y_j^* denotes the j^{th} entry of an optimal solution to the relaxed problem defining $\beta(R)$. Let j_R be an index such that

$$\delta(R) := -\xi_{j_R} y_{j_R}^{*2} - l_j^R(y_{j_R}^*) = \max_{1 \leq j \leq m} \big\{ -\xi_j y_j^{*2} - l_j^R(y_j^*) \big\}.$$

Note that at the ends of each edge of the box R, the value of the function $-\xi_j y_j^2$ and of its convex envelope coincide. Thus $\delta(R) \neq 0$ implies that $y_{j_R}^*$ is not an endpoint of the edge j_R of R.

Using j_R and $y_{j_R}^*$ we bisect R into two sub-boxes R^+ and R^- by setting

$$R^+ := \{y = (y_1, \ldots, y_m)^T \in R \mid y_{j_R} \geq y_{j_R}^*\}, \tag{3.5}$$

$$R^- := \{y = (y_1, \ldots, y_m)^T \in R \mid y_{j_R} \leq y_{j_R}^*\}. \tag{3.6}$$

Clearly, both R^+ and R^- are not empty. For this bisection we have the following lemma whose proof can be found, e.g., in [17]:

Lemma 3.1. *Let $\{R_k\}$ be an infinite sequence of boxes generated by the bisection process defined by* (3.5) *and* (3.6). *Suppose that $R_{k+1} \subset R_k$ for every k. Then*

$$\lim_{k \to \infty} \big\{\alpha(R_k) - \beta(R_k)\big\} = 0.$$

3.3. **Computing an upper bound.** Note that a feasible point of the AMPEC problem (1.1)-(1.3) can be computed whenever the lower level problem is solved. In the Nash-Cournot equilibrium market model described in Subsection 2.2, the lower level problem can be solved efficiently with available codes, since it is a strongly convex quadratic program over the polyhedron X. In fact, with a fixed $y \in Y$, the lower level problem is the strongly monotone variational inequality

$$\text{Find } x \in X \text{ such that: } (Ax + By + a)^T(v - x) \geq 0, \text{ for all } v \in X, \tag{VI$_y$}$$

where A is given by (2.16) and $a = (-\alpha, \ldots, -\alpha)^T$. This variational inequality is reduced to the strongly convex quadratic program (see, e.g. [12]):

$$\min \big\{ \frac{1}{2} x^T A x + \sum_{k=1}^{n} (\mu_k + \alpha) x_k \mid x \in X \big\},$$

where $\mu_k = \sum_{i=1}^m c_{ki} y_i$. Hence, if x is the optimal solution to this problem then (x, y) is a feasible point to the model, and therefore, $f(x, y)$ is an upper bound for the optimal value α_*. Now we are able to describe in detail an algorithm for globally solving Problem (NC) thereby obtaining a global optimal equilibrium pair to the bilevel Nash-Cournot equilibrium market model presented in Subsection 2.2.

The B&B algorithm is described as follows:

B&B Algorithm.

Initialization. Choose a tolerance $\varepsilon \geq 0$, take $R_0 = Y$ and solve the relaxed problem (RNC$_R$) with $R = R_0$ to obtain the optimal value $\beta_0 := \beta(R_0)$ and an optimal solution (x^{R_0}, y^{R_0}). If $l_{R_0}(y^{R_0}) = h_1(y^{R_0})$ then (x^{R_0}, y^{R_0}) is a global optimal solution to Problem (NC) and terminate the algorithm. Otherwise, solve the lower level problem (VI_y) with $y = y^{R_0}$ to obtain a feasible point. Let (x^0, y^0) be the currently best feasible point and $\alpha_0 = f(x^0, y^0)$ be the currently best upper bound (we also call it the score). Set

$$\Gamma_0 := \begin{cases} \{R_0\} & \text{if } \alpha_0 - \beta_0 > \varepsilon(|\alpha_0| + 1), \\ \emptyset & \text{otherwise.} \end{cases}$$

Iteration k ($k = 0, 1, \dots$). At the beginning of each iteration k we have a family Γ_k of subboxes of Y, a lower bound β_k, an upper bound α_k for the optimal value α_* and a feasible point (x^k, y^k) such that $\alpha_k = f(x^k, y^k)$.

a) If $\Gamma_k = \emptyset$, then terminate: α_k is an ε-optimal value and (x^k, y^k) is an ε-global optimal solution.

b) If $\Gamma_k \neq \emptyset$, choose $\mathbb{R}_k \in \Gamma_k$ such that

$$\beta(R_k) = \min\{\beta(R) \mid R \in \Gamma_k\}.$$

Bisect R_k into two rectangles R_{k1} and R_{k2} according to the bisection (3.5) and (3.6). For each $(j = 1, 2)$, compute

$$(\mathrm{RNC}_{R_{kj}}) \quad \beta(R_{kj}) := \min\big\{f(x, y) \mid x \in X,\ y \in R_{kj},\ g_1(x, y) + l^{R_{kj}}(y) \leq 0\big\}.$$

Let $(x^{R_{kj}}, y^{R_{kj}})$ be the obtained optimal solution to this subproblem. Use $y^{R_{kj}}$ $(j = 1, 2)$ to compute new feasible points by solving the strongly monotone variational inequalities (VI_y) with $y = y^{R_{kj}}$ $(j = 1, 2)$. Let (x^{k+1}, y^{k+1}) be the currently best feasible point and $\alpha_{k+1} = f(x^{k+1}, y^{k+1})$ be the new upper bound (new score). Delete all $R \in \Gamma_k$ such that

$$\alpha_{k+1} - \beta(R) \leq \varepsilon(|\alpha_{k+1}| + 1).$$

Let Γ_{k+1} be the remaining set of subrectangles (may be empty). Then go to iteration k with $k := k + 1$. □

The following theorem shows the convergence of the B&B algorithm.

Theorem 3.2. *Suppose that the sequence* $\{(x^k, y^k)\}_k$ *is generated by the* B&B *algorithm. Then*

(i) *If the algorithm terminates at some iteration* k *then* (x^k, y^k) *is an* ε*-global optimal equilibrium pair to the Nash-Cournot equilibrium market model.*

(ii) *If the algorithm does not terminate then* $\alpha_k \searrow \alpha_*$, $\beta_k \nearrow \alpha_*$ *as* $k \to +\infty$ *and any limit point of the sequence* $\{(x^k, y^k)\}$ *is a global optimal equilibrium pair to the model.*

Proof. We only give a sketch for the proof, because it can be done by using Lemma 3.1 and by a standard argument commonly used in global optimization.

The statement (i) is obvious, since if the algorithm is terminated at iteration k, then $\Gamma_k = \emptyset$. In this case $\alpha_k - \beta_k \leq \varepsilon(|\alpha_k| + 1)$. Hence, α_k is an ε-global optimal value and (x^k, y^k) is an ε-global optimal solution.

Now, we prove (ii). If the algorithm does not terminate, then it generates an infinite sequence of iterates (x^k, y^k). Let (x^*, y^*) be any limit point of this sequence. Suppose $(x^{k_q}, y^{k_q}) \to (x^*, y^*)$ as $q \to \infty$. Then the corresponding sequence of the rectangles has a nested subsequence, which for simplicity of notation, we denote by $\{R_{k_q}\}$. Then, by Lemma 3.1, $\lim_{q\to\infty}(\alpha_{k_q} - \beta_{k_q}) = 0$. Since both sequences $\{\alpha_k\}$, $\{\beta_k\}$ are monotone, it follows that $\lim \alpha_k = \lim \beta_k$. By the definition, $\alpha_k \geq \alpha_*$ and $\beta_k \leq \alpha_*$ for every k, we have $\alpha_k \searrow \alpha_*$, $\beta_k \nearrow \alpha_*$. Note that $\alpha_{k_q} = f(x^{k_q}, y^{k_q})$, by the continuity of f, we can deduce that (x^*, y^*) is a global optimal solution of the problem. □

4. Numerical results

We test the proposed B&B algorithm for the bilevel Nash-Cournot equilibrium market model presented in Subsection 2.2.3. All the tests are implemented in Matlab 7.8.0 (R2009a) for Linux running on a PC Desktop Intel(R) Core(TM)2 Quad CPU Q6600 with 2.4GHz and 3Gb RAM. The input data of the problem is generated randomly.

- The objective function is chosen by a convex quadratic form $f(x, y) = \frac{1}{2}x^TQ_1x + \frac{1}{2}y^TQ_2y + q_1^Tx + q_2^Ty$, where Q_1, Q_2, q_1 and q_2 are generated randomly. The parameters $\beta = 0.125$, $\alpha = 10$ whereas $B = (c_{ij})_{n\times m}$ is generated randomly in $(0, 1)$. The convex sets $X = [0, 5]^n$ and $Y = [0, 5]^m$,
- For computing the lower bound, we used the interior point method implemented in the built-in Matlab solver FMINCON with maximum of iterations being 500 to solve the convex subproblems. The convex quadratic problems are solved by QUADPROG (a built-in Matlab solver) and the CVX package with the Sedumi solver (a freely available Malab code for convex programming at http://cvxr.com/cvx/).
- For computing the upper bound, a local optimization method in DC optimization is used that proves a feasible point to the problem (1.3).

We perform the B&B algorithm for 20 random problems with different sizes. The results are reported in Table 1, where m, n are the sizes of the problem; `iter` is the number of iterations; `cbval` is the currently best upper bound (score); `lbval` is the lower bound for the optimal value; `cputime` is the CPU time in second; `status` is the status of stopping criterion (`solved` shows that an ε-global optimal solution is found, `incomp.` indicates that the solver is stopped when the lower bound is improved too slowly, `exceed` means that the running time exceeds the limit 36.000 seconds); and `node` is the maximum number of the nodes in the B&B tree that have been stored.

From the computational results we can observe the following preliminary remarks:

(1) The proposed B&B algorithm can solve globally AMPEC, in particular, bilevel convex quadratic problems, with several hundreds of decision variables while the number of the parameters is relatively small.
(2) The numbers of iterations in Table 1 indicates that the adaptive rectangular bisection used is rather effective.
(3) Almost CPU time spends on solving the general convex subproblems for computing lower and upper bounds. Note that at each iteration in the

Table 1. Computational results of the B&B algorithm for Nash-Cournot Problem

Problem Info.			Branch & Bound algorithm					
N^0	m	n	cbval	lbval	iter	time(s)	node	status
1	5	10	1338.2220	1338.2021	17	88.46	7	solved
2	10	10	1576.4746	1576.4407	154	962.99	56	solved
3	5	20	3711.1289	3711.1289	43	521.23	7	solved
4	5	30	3537.2899	3537.2899	46	693.88	7	solved
5	8	50	3994.0027	3992.7705	99	7944.81	24	solved
6	5	100	3162.2176	3160.5017	47	7004.01	8	solved
7	6	100	4073.9795	4049.4880	62	9822.34	12	incomp.
8	7	100	3825.4430	3825.2157	73	11194.71	17	solved
9	5	150	2731.9005	2692.1867	43	14730.51	8	incomp.
10	6	150	3781.3484	3711.8269	73	20531.79	14	incomp.
11	1	200	3173.2954	3173.2954	9	4662.39	2	solved
12	2	200	2738.1198	2738.1198	19	5123.47	6	solved
13	3	200	2391.6111	2391.6111	18	6089.43	4	solved
14	4	200	2869.7684	2869.7684	22	10175.95	4	solved
15	5	200	3726.2399	3726.2399	55	26477.86	9	solved
16	6	200	2759.8484	2751.9396	75	36107.07	14	exceed
17	7	200	2459.9965	2390.6909	78	36270.34	21	exceed
18	8	200	3333.2645	3102.4295	80	36456.48	34	exceed
19	2	300	3008.2311	2975.1594	14	14963.37	2	incomp.
20	3	300	3275.0818	3275.0818	29	29976.60	6	solved

interior point algorithm for solving the convex subproblem one needs to solve a strongly convex quadratic program.

5. Conclusion

We have formulated some classes of bilevel programming problems in the form of AMPEC. We have also used a regularization technique to obtain smoothly DC optimization formulations to AMPEC. A suitable regularization matrix results into a DC decomposition, where the second component depends upon only the parameter of the lower problem. We have described a decomposition branch-and-bound algorithm for globally solving AMPEC. This algorithm uses an adaptive rectangular bisection involving only the parameter which is often much less than the number of the decision variables in practical problems. Computational results on a bilevel Nash-Cournot equilibrium market model show efficiency of the proposed algorithm.

References

[1] Le Thi H. An, Dinh T. Pham and L. D. Muu, Numerical solution for optimization over the efficient set by D. C. optimization algorithm, *Oper. Res. Lett.* **19** (1996), 117–128.

[2] L. T. H. An, D. T. Pham and L. D. Muu, Simplicially constrained D. C. optimization over the efficient and weakly efficient sets, *J. Optim. Theory Appl.* **117** (2003), 503–521.

[3] L. T. H. An, D. T. Pham, N. C. Nam and N. V. Thoai, DC programming techniques for solving a class of nonlinear Bilevel programs, *J. Global Optim.* **44**(3) (2009), 313-337.

[4] L. T. H. An, P. D. Tao and L. D. Muu, Exact penalty in DC programming, *Vietnam J. Math.* **27**(2) (1999), 169–178.

[5] P. N. Anh and L. D. Muu, Lagrangian duality algorithms for finding a global optimal solution to mathematical programs with affine equilibrium constraints, *Nonlinear Dyn. Syst. Theory* **6** (2006), 225–244.

[6] H. Benson, Optimization over the efficient set, *J. Math. Anal. Appl.* **98** (1984), 562–580.

[7] S. Dempe, Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints, *Optimization* **52** (2003), 333–359.

[8] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vol. I, II. Springer-Verlag, New York, 2003.

[9] J. E. Falk and R. M. Soland, An algorithm for separable nonconvex programming problems, *Manage. Sci.* **15** (1969), 550–569.

[10] R. Horst and H. Tuy, *Global Optimization: Deterministic Approach*, Springer-Verlag, Berlin, 1992.

[11] V. V. Kalashikov and N. J. Kalashikova, Solving two-level variational inequalities, *J. Global Optim.* **8**(3) (1996), 289–294.

[12] I. V. Konnov, *Combined Relaxation Methods for Variational Inequalities*, Springer-Verlag, Berlin, 2000.

[13] I. V. Konnov and S. Kum, Descent methods for mixed variational inequalities in a Hilbert spaces, *Nonlinear Anal.* **47** (2001), 561–572.

[14] G. M. Lee, N. N. Tam and N. D. Yen, *Quadratic Programming and Affine Variational Inequalities: A Qualitative Study*, Springer-Verlag, 2005.

[15] J. Q. Luo, J. S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, 1996.

[16] O. L. Mangasarian and J. S. Pang, Exact penalty functions for linear programs with linear complementarity constraints, *Optimization* **42** (1997), 1–8.

[17] L. D. Muu and W. Oettli, Optimization over equilibrium sets, *Optimization* **49** (2000), 179–189.

[18] L. D. Muu and N. V. Quy, A global optimization method for solving convex quadratic bilevel programming problems, *J. Global Optim.* **26** (2003), 199–219.

[19] L. D. Muu and N. V. Quy, On branch-and-bound algorithms for global optimal solutions to mathematical programs with affine equilibrium constraints, *Vietnam J. Math.* **35**(4) (2007), 523–539.

[20] J. Philip, Algorithms for the vector maximization problems, *Math. Program.* **2** (1972), 207–229.

[21] K. Taij and M. Fukushima, A new merit function and a successive quadratic programming algorithm for variational inequalities problems, *SIAM J. Optim.* **6** (1996), 703–713.

[22] N. V. Thoai, Y. Yamamoto and A. Yoshise, Global optimization method for solving mathematical programs with complementarity constraints, *J. Optim. Theory Appl.* **24** (2005), 467–490.

Institute of Mathematics, VAST, Hanoi, Vietnam
E-mail address: `ldmuu@math.ac.vn`

Hanoi University of Science, Hanoi, Vietnam.
Current address: Department of Electrical Engineering (ESAT/SCD) and
Optimization in Engineering Center (OPTEC)
KU Leuven, Belgium
E-mail address: `quoc.trandinh@esat.kuleuven.be`

Metz University, France
E-mail address: `lethi.at.univ-metz.fr`

INSA, Rouen, France
E-mail address: `pham@insa-rouen.fr`