Nonlinear Dynamics and Systems Theory, 6(3) (2006) 225–244

Lagrangian Duality Algorithms for Finding a Global Optimal Solution to Mathematical Programs with Affine Equilibrium Constraints[†]

Pham Ngoc Anh[1] and Le Dung Muu[2*]

[1] *Posts and Telecommunications Institute of Technology, Hanoi, Vietnam*
[2] *Hanoi Institute of Mathematics, Vietnam*

Received: July 18, 2005; Revised: August 8, 2006

Abstract: Mathematical programs with equilibrium constraints, shortly MPEC, are optimization problems with parametric variational inequality constraints.MPEC include bilevel convex programming problems, mathematical programs with complementarity constraints, Nash-Cournot oligopolistic market models, as well as optimization over the efficient set of an affine fractional multicriteria program as special cases. MPEC are difficult global optimization ones, since their feasible domains, in general, are not convex even not connected. In this paper we consider linear programs with affine equilibrium constraints. We use the Lagrangian duality to compute lower bounds for a decomposition branch-and-bound procedure that allows approximating a global optimal solution of problems in this class of MPEC. Application to optimization over the efficient set of a multicriteria affine fractional program is discussed.

Keywords: *Equilibrium constraints; bilevel convex program; optimization over the Pareto set; Nash-Cournot model; branch-and-bound; global optimum.*

Mathematics Subject Classification (2000): 90C29.

1 Introduction

Mathematical programs with equilibrium (or variational inequality) constraints, shortly MPEC, are optimization problems whose constraints include parametric variational inequalities. For these problems we refer the readers to the comprehensive monograph [16] and the interesting bibliography paper [8]. MPEC play an important role, for example,

[†] This work was done during the visit of the second author at the University of Karlsruhe, Germany and at ICTP, Italy.

[*] Corresponding author: ldmuu@math.ac.vn

in the design of transportation networks, in economic models (see e.g. [13, 16]). These problems also include, as special cases, mathematical program with complementarity constraints, the bilevel convex programming problem, where some variables are restricted to be in the solution set of a parametric convex optimization problem, and the optimization over the efficient set of an affine fractional multicriteria program. Mathematically, for finite dimensional case, a mathematical program with equilibrium constraints can be written as

$$\begin{cases} \text{minimize } \ f(x,y) \\ \text{subject to } x \in X, y \in Y, (x,y) \in Z \\ \text{and } y \text{ solves the parametric variational inequality,} \\ \text{find } y \in C(x) \text{ such that } \langle F(x,y), v-y \rangle \geq 0 \ \forall v \in C(x), \end{cases}$$

where $X \subseteq R^n$, $Y \subseteq R^m$, $Z \subseteq R^n \times R^m$, are nonempty closed convex sets; $f : X \times Y \to R$, $F : X \times Y \to R^m$ and $C : R^n \to 2^{R^m}$ is a multivalued mapping. The MPEC which are known to be very difficult ones, being nonsmooth and nonconvex also under very favorable assumptions. Further, the computation of the (generalized) gradients of the constraints can be difficult, except special cases.

Several heuristic and deterministic methods were developed for finding local optimal solution to the MPEC. In [13] heuristic algorithms were suggested for solving some classes of MPEC. In [24] Outrata and Zowe converted a mathematical program with equilibrium constraints into an unconstrained nonsmooth Lipschitz optimization problem. Then one may use well developed nonsmooth optimization numerical methods for locally solving the converted problem. In [9] Facchinei et al applied known methods of nonlinear optimization to a regularized reformulation of the MPEC. Based on the study of subanalytic optimization problems and with the help of the theory of error bounds, some exact penalty results for the MPEC were proved by Lin and Fukushima in [15]. Recently, in [19] Mordukhovich discussed optimality conditions for the MPEC and EPEC (equilibrium problems with equilibrium constraints) by using tools of variational analysis.

Due to its nested structure, the feasible domain of a mathematical program with equilibrium constraints, even in the linear case, in general, is disconnected and may be neither open nor closed. Thus the MPEC are difficult global optimization ones and therefore it is less hope to develop an algorithm for finding global optimal solutions to general MPEC. In [21] a branch and bound algorithm based on a binary search is proposed for globally solving a class of mathematical programs with affine equilibrium constraints. The binary search method proposed in that paper works well for the case when the number of constraints defining the variational inequality-constraint is somewhat small, but it quickly becomes expensive when this number gets larger.

In this paper, we continue our work in [21] by using the Lagrangian duality bound to develop another branch-and-bound algorithm for globally solving a class of mathematical programs with affine equilibrium constraints. By contrast to the method in [21] the global optimization operation in this algorithm takes place in the y- space rather than the space of the Lagrangian variables. Thus it is expected that the proposed method works well when the number of the y-variables is somewhat small while the number of the constraints as well as the number of the x-variables may be much larger.

The rest of this paper will be organized as follows. In the next section we state the problem to be solved and list some of its special cases such as bilevel convex programming, optimization over the efficient set and Cournot-Nash oligopolistic market models. In the third section, first we show how to use the Lagrangian duality to compute lower bounds. Then we describe in detail a branch-and-bound algorithm when the vertices of the constrained set Y are known in advance. The last section is devoted to description of a relaxation algorithm that does not require prior knowledge of these vertices. Applications of the proposed algorithms to the optimization problem over the Pareto-efficient set of a multicriteria affine fractional program are also discussed in this section.

2 The Problem Statement and Examples

In what follows we restricted our attention to a special class of MPEC; namely, we consider the following affine MPEC that we call shortly AMPEC:

$$(\mathrm{P}) \quad \begin{cases} \text{minimize } \{f(x,y) := a^T x + b^T y\} \\ \text{subject to } \ x \in X, y \in Y, (x,y) \in Z := \{(x,y) \ : \ Mx + Ny + p \le 0, \} \\ \text{and } y \text{ solves the parametric variational inequality} \\ \text{find } y \in Y \text{ such that } \langle P(y)x + Qy + q, v - y\rangle \ge 0 \ \forall v \in Y \qquad \mathrm{VIP}(x) \end{cases}$$

where $X \subseteq R^n$, $Y \subseteq R^m$ are nonempty closed convex sets, $p \in R^l$, $a \in R^n, q, b \in R^m$ and for each $y \in Y$, $P(y)$, Q, M and N are given appropriate matrices.

Let $S(x)$ denote the solution-set of VIP(x). As usual we call a couple (x, y) such that $(x, y) \in Z$, $x \in X, y \in Y$, $y \in S(x)$ a feasible solution to Problem (P).

First we mention some important special cases of this problem.

Example 2.1 (Convex quadratic bilevel program). We consider the parametric variational inequality VIP(x), where $P(y) \equiv P$, and Q is a symmetric positive semidefinite matrix. In this case, since Y is convex, it is well-known that VIP(x) is equivalent to the convex programming problem

$$\min\left\{\frac{1}{2}y^T Qy + (Px)^T y + q^T y \ : y \in Y\right\}.$$

Thus AMPEC problem (P) can be equivalently rewritten as a convex bilevel problem of the form

$$\min\left\{f(x,y) := a^T x + b^T y\right\}$$

subject to

$$(x,y) \in Z_1 := \{(x,y) \ : \ Mx + Ny + p \le 0, x \in X, y \in Y\},$$

where y solves the convex quadratic program

$$\min\left\{\frac{1}{2}y^T Qy + (Px)^T y + q^T y \ : y \in Y\right\}. \tag{C_x}$$

Example 2.2 (Optimization over the weakly efficient set). Other examples for the AMPEC are optimization problems over the efficient (Pareto) and weakly efficient sets of a multicriteria (vector) affine fractional program. These problems have been recently considered by some authors (see e.g. [17, 20, 22, 29]). The problems can be formulated in forms of AMPEC. To this end, consider the affine fractional vector optimization problem

$$\mathrm{vmin}\{F(v) := \Big(\frac{A_1^T v + s_1}{B_1^T v + t_1}, ..., \frac{A_\rho^T v + s_\rho}{B_\rho^T v + t_\rho}\Big) : \quad v \in V\}, \tag{VP}$$

where $V \subset R^m$ is a bounded polyhedral convex set, A_i, B_i are m-dimensional vectors, s_i, t_i $(i = 1, ..., \rho)$ are real numbers. As usual we assume that $B_i^T v + t_i > 0$ for all $v \in V$ and all $i = 1, ..., \rho$. Thus F is continuous on V. We recall that a point $v \in V$ is said to be an (Pareto) efficient (resp. weakly efficient) solution of (VP) if there does not exist $w \in Y$ such that $F(w) \le F(v)$, $F(w) \ne F(v)$ (resp. $F(w) < F(v)$), By $E(F, V)$ (resp. $WE(F, V)$) we will denote the set of all efficient (resp. weakly efficient) solutions of (VP). It is well-known (see e.g. [27]) that if V is compact, then the efficient set is nonempty. Hence so is the weakly efficient set, since $E(F, V) \subseteq WE(F, V)$. An optimization problem over the efficient set (resp. weakly efficient set) is the problem of optimizing (minimizing or maximizing) a real-valued function f over the efficient (resp. weakly efficient) set of (VP). These minimization problems can be written respectively as

$$\min\{f(v) \ : v \in E(F, V)\}, \tag{2.1}$$

$$\min\{f(v) \ : v \in WE(F, V)\}. \tag{2.2}$$

Note that, in general, both the efficient and weakly efficient sets are not convex. The weakly efficient set is closed but the efficient set may be neither closed nor open [7]. Thus these problems are difficult global optimization ones.In order to formulate these problems in the form of AMPEC we use the following theorem due to Malivert [17].

Theorem 2.1 ([17]) *A vector $v \in V$ is efficient (resp. weakly efficient) if and only if there exist real numbers $u_i > 0$ (resp. $u_i \ge 0$ not all zero) for all $i = 1, ..., \rho$ such that*

$$\sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v - w \Big\rangle \le 0 \quad \forall w \in V.$$

In virtue of this theorem the problems (2.1) and (2.2) can be written as

$$\begin{cases} \min f(v) \ subjectto \ \ v \in V, \ u_i > 0 \ \text{n}otallzero \ \forall i = 1, ..., \rho, \\ \sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v - w \Big\rangle \le 0 \quad \forall w \in V, \end{cases}$$

and

$$\begin{cases} \min f(v) \text{ subject to } \ v \in V, \ u_i \ge 0 \text{ not all zero } \forall i = 1, ..., \rho, \\ \sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v - w \Big\rangle \le 0 \quad \forall w \in V \end{cases}$$

respectively.

Define the $(m \times \rho)$-matrix $P(v)$ by setting

$$P(v) := \Big\{(B_1^T v + t_1)A_1 - (A_1^T v + s_1)B_1, ..., (B_\rho^T v + t_\rho)A_\rho - (A_\rho^T v + s_\rho)B_\rho\Big\}.$$

Then we can rewrite these problems in the forms

$$\min\{f(v) : \quad v \in V, u > 0, \ \langle P(v)u, w - v\rangle \ge 0 \ \forall w \in V\} \tag{2.1a}$$

and

$$\min\{f(v) : \quad v \in V, u \ge 0, \ \langle P(v)u, w - v\rangle \ge 0 \ \forall w \in V\} \tag{2.2a}$$

respectively.

Clearly, the latter problem is of the form of AMPEC where v and u play the roles of y and x respectively, and $M \equiv 0$, $N \equiv 0$, $Q \equiv 0$, $p \equiv 0$, $q \equiv 0$. If in problem (2.1a) we replace the constraint $u > 0$ by $u \geq \delta$ with $\delta > 0$ sufficiently small as desired, we obtain an approximation problem to (2.1a) that is of the form of AMPEC.

In an important special case where $B_i = 0$ and $s_i = 1$ for all i, Problem (VP) becomes a linear vector program. In this case, it is well known [17, 27], both the efficient and weakly efficient sets are closed, but, in general, not convex. Thus Problems (2.1) and (2.2) remain global optimization ones, since there are local optimal solutions that are not global optimal ones. In this linear case due to a theorem of Philip [25] Problem (2.1a) can take the form of AMPEC as

$$\min\{f(v) : \; v \in V, u \geq \delta, \; \langle P(v)u, w - v\rangle \geq 0 \; \forall w \in V\},$$

where $\delta > 0$ is sufficiently small. Note that in this linear case

$$P(v) \equiv P := \Big(t_1A_1, t_2A_2, . \;\; . \,.t_\rho A_\rho\Big)$$

is independent of v.

Example 2.3 (Nash-Cournot market model). The third section of AMPEC is a Nash-Cournot oligopolistic market model (see e.g. [10, 12]). The model can be described as follows.

Suppose that there are m-firms (sectors) that supply a homogeneous product whose price p_j at sector j $(j = 1, ..., m)$ depends on total producing quantity and is given by

$$p_j(\sum_{i=1}^{m} y_i) = \alpha_j - \beta_j \sum_{i=1}^{m} y_i,$$

where $\alpha_j \geq 0$, $\beta_j > 0$ are given constants whereas y_i is the quantity of goods supplied by firm i that we have to determine. Suppose further that to produce the goods the firms need n-different materials represented by a nonnegative vector $x \in R^n$. Let x_i $(i = 1, ..., n)$ be the quantity of material i needed to produce a unique of goods. Let $c_{ij} > 0$ denote the price of a unit material i for firm j $(i = 1, ..., n,\ j = 1, ..., m)$. Assume that the cost of firm j is given by

$$h_j(x, y_j) := y_j \sum_{i=1}^{n} c_{ij}x_i + \delta_j, \quad j = 1, ..., m,$$

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

$$u_j(x, y) := p_j(\sum_{i=1}^{m} y_i)y_j - h_j(x, y_j).$$

Let

$$X_i := \{t \; : \; 0 \leq t \leq \xi_i\} \; (i = 1, ..., n),$$

$$Y_j := \{\tau \; : \; 0 \leq \tau \leq \eta_j\} \; (j = 1, ..., m),$$

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

Let

$$X := X_1 \times X_2 ... \times X_n, \quad Y = Y_1 \times ... \times Y_m$$

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

Given $x \in X$ each firm j seeks to find its producing quantity y_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.

By definition, a vector $(y_1^*, ..., y_m^*) \in Y_1 \times Y_2 ... \times Y_m$ is said to be a (Nash) equilibrium point with respect to $x^* \in X$ if

$$\begin{cases} u_j(x^*, y_1^*, ..., y_{j-1}^*, y_j, y_{j+1}^*, ..., y_m^*) \\ \leq u_j(x^*, y_1^*, ..., y_{j-1}^*, y_j^*, y_{j+1}^*, ..., y_n^*) \ \forall y_j \in Y_j, \forall j. \end{cases} \tag{2.3}$$

We will refer to a pair (x^*, y^*) satisfying (2.3) as an *equilibrium pair* of the model.

Besides the utility function of each firm, there is another cost function (leader's objective function) $f(x, y)$ depending on x and the quantity y of the goods. The problem to be solved is of finding an equilibrium pair that minimizes leader's objective function over the set of all equilibrium pairs. This problem can be formulated as a mathematical program with affine equilibrium. To this end let

$$\begin{cases} H_j(x, y) := \nabla_{y_j} h_j(x, y_j) \ (j = 1, ..., m), \\ e := (1, ..., 1)^T, \ \sigma_y := \sum_{j=1}^m y_j. \end{cases} \tag{2.4}$$

Applying Proposition 3.2.6 in [12] we see that a point $(y_1, ..., y_m)$ is equilibrium with respect to x if and only if it is a solution to the variational inequality problem

$$\text{Find } y \in Y : \ \langle F(x, y), z - y \rangle \geq 0 \ \ \forall z \in Y,$$

where $F(x, y)$ is m-dimensional vector whose jth component is

$$F_j(x, y) := H_j(x, y) - p_j(\sigma_y)e - \nabla p_j(\sigma_y)y. \tag{2.5}$$

Using (2.4) and (2.5) we have

$$F(x, y) = \begin{pmatrix} \sum_{i=1}^n c_{i1}x_i - \alpha_1 + \beta_1 \sum_{j=1}^m y_j + \beta_1 y_1 \\ ...... \\ \sum_{i=1}^n c_{im}x_i - \alpha_m + \beta_m \sum_{j=1}^m y_j + \beta_m y_m \end{pmatrix}$$

$$= P(y)x + Qy + q,$$

where

$$Q = \begin{pmatrix} 2\beta_1 & \beta_1 & ... & \beta_1 \\ \beta_2 & 2\beta_2 & ... & \beta_2 \\ ... & ... & ... & ... \\ \beta_m & \beta_m & ... & 2\beta_m \end{pmatrix} \tag{2.6}$$

and $P(y)$ is the $n \times m$ matrix independent of y whose P_{ij} entry is

$$P_{ij} = c_{ij}, \ j = 1, ..., m, \ i = 1, ..., n, \tag{2.7}$$

and

$$q = (\delta_1 - \alpha_1, ..., \delta_m - \alpha_m)^T. \tag{2.8}$$

Thus problem to be solved can take the form

$$\begin{cases} \min f(x,y) \text{ subject to} \\ x \in X = X_1 \times \dots \times X_n,\ y \in Y = Y_1 \times \dots \times Y_n, \\ \text{where } y \text{ solves the parametric variational inequality} \\ \langle Px + Qy + q, z - y\rangle \geq 0 \quad \forall z \in Y, \end{cases}$$

with Q, P and q being given by (2.6), (2.7) and (2.8) respectively. Clearly, this problem is of form (P) with $M = 0, N = 0, p = 0$.

3 A Lagrangian Bounding Algorithm

The algorithm to be described in this section relies on the branch-and-bound strategy. Two main operations in a branch-and-bound algorithm are bounding and branching ones. The Lagrangian bound is widely used in global optimization as well as in discrete programming [6, 11, 14, 26]. To the algorithm we are going to describe for AMPEC problem (P) we also use the Lagrangian bounding operation.

3.1 The Lagrangian Bound

First, we consider the case when Y is a polytope and all of its vertices are known in advance. This case occurs frequently, for instant, in economics equilibrium models where Y is a simplex or a box. Let $y^1, y^2, ..., y^s$ be the vertices of polytope Y. It is easy to verify that

$$\langle P(y)x + Qy + q, z - y\rangle \geq 0 \ \forall z \in Y$$

if and only if

$$\langle P(y)x + Qy + q, y^k - y\rangle \geq 0 \ \forall k = 1, ..., s.$$

Thus, AMPEC problem (P) can be rewritten equivalently as

$$(P) \quad \begin{cases} \text{minimize } \{f(x,y) := a^T x + b^T y\} \\ \text{subject to } x \in X, y \in Y, (x,y) \in Z := \{(x,y) \ : \ Mx + Ny + p \leq 0\}, \\ \text{and } y \in Y \text{ satisfying inequalitis} \\ \ \langle P(y)x + Qy + q, y^k - y\rangle \geq 0 \ \forall k = 1, ..., s. \end{cases}$$

Let

$$\hat{Y} := \{y \in Y : \exists x \in X \text{ such that } Mx + Ny + p \leq 0,$$
$$\langle P(y)x + Qy + q, y^k - y \geq 0 \ \forall y = 1, ..., s\}.$$

Note that if $\langle P(y)x, y\rangle$ is convex with respect to y, in particular when $X \subset R^n_+$ and $P(y) = P$ (see examples 2.1, 2.2 for linear case and 2.3), or when $P(y) = \mathrm{Diag}(y)$, then $\hat{Y}$ is convex.

Define the function $\varphi : \hat{Y} \to I\!R$ by setting, for each $y \in \hat{Y}$,

$$\begin{cases} \varphi(y) := \min_x\{f(x,y) := a^T x + b^T y\} \\ \text{s. t.} \quad x \in X, (x,y) \in Z := \{(x,y) \ : \ Mx + Ny + p \leq 0\}, \\ \langle P(y)x + Qy + q, y^k - y\rangle \geq 0 \ \forall k = 1, ..., s. \end{cases} \qquad (P_y)$$

Then the master problem

$$\min\{\varphi(y) \ : \ y \in \hat{Y}\} \qquad (MP)$$

is equivalent to Problem (P) in the sense of the following proposition whose proof is obvious directly from the definitions.

Proposition 3.1 *A point (x^*, y^*) is optimal to Problem (P) if and only if y^* is optimizer to (MP) and $f(x^*, y^*) = \varphi(y^*)$.*

Note that, unlike global optimization problems having nonconvex feasible domains, feasible points of a MPEC problem can be computed by available methods of variational inequalities (see e. g. [1, 2, 10, 12] and the references therein). However for Problem (P), a feasible point can be obtained by solving a suitable linear program. In fact, if $y \in \hat{Y}$ is fixed and x^y is an optimal solution of the linear problem (P_y) then (x^y, y) is feasible for (P). So upper bounds for the optimal value w_* of (P) can be computed by solving a linear program. As the algorithm executes more feasible points can be found, and thereby upper bounds for w_* can be iteratively improved.

We now compute a tight lower bound for w_* by using Lagrangian duality. To be specific suppose that X is given explicitly as

$$X := \{x \in R^n : \ x \geq 0, Ax + d \leq 0\},$$

where $d \in R^l$ and A is $l \times n$-matrix. Let S be a fully dimensional simplex or a rectangle in y-space such that $S \cap \hat{Y} \neq \emptyset$. Consider Problem (P) restricted on $S \cap \hat{Y}$, i.e.,

$$w_*(S) = \min_{x,y} f(x, y) := a^T x + b^T y$$

subject to

$$\begin{cases} Mx + Ny + p \leq 0 \\ Ax + d \leq 0, \ x \geq 0, \ y \in S \cap \hat{Y} \\ \langle P(y)x + Qy + q, y - y^k\rangle \leq 0 \ \ \forall k = 1, ..., s. \end{cases} \tag{P_S}$$

Let $L(x, y, \lambda, \mu, \xi)$ be the Lagrangian function of this problem associated with all constraints except the constraints $x \geq 0$, $y \in S \cap \hat{Y}$. That is

$$L(x, y, \lambda, \mu, \xi) := a^T x + b^T y + \sum_{k=1}^{s} \lambda_k \langle P(y)x + Qy + q, y - y^k\rangle + \mu^T(Mx + Ny + p) + \xi^T(Ax + d).$$

Define the function $m(y, \lambda, \mu, \xi)$ as

$$m(y, \lambda, \mu, \xi) := \inf_{x \geq 0} L(x, y, \lambda, \mu, \xi).$$

From the Lagrangian duality theorem we have

$$m(y, \lambda, \mu, \xi) \leq \varphi(y) \ \ \forall \lambda \geq 0, \mu \geq 0, \xi \geq 0, y \in S \cap \hat{Y}. \tag{3.1}$$

Since for each fixed y the functions $Mx + Ny + p$, $Ax + d$, and $\langle P(y)x + Qy + q, y - y^k\rangle$ $\forall k = 1, ..., s$ are affine with respect to x, by the Lagrangian duality theorem, we have

$$\sup_{\lambda, \mu, \xi \geq 0,} m(y, \lambda, \mu, \xi) = \varphi(y) \ \ \forall y \in S \cap \hat{Y}.$$

Let

$$\gamma_S(\lambda, \mu, \xi) = \min_{y \in S \cap \hat{Y}} m(y, \lambda, \mu, \xi).$$

Then from (3.1) it follows that

$$\gamma_S(\lambda,\mu,\xi) \le \min_{y\in S\cap\hat{Y}} \varphi(y) = w_*(S) \quad \forall\lambda\ge 0, \mu\ge 0, \xi\ge 0.$$

Thus

$$\sup_{\lambda,\mu,\xi\ge 0} \gamma_S(\lambda,\mu,\xi) \le w_*(S).$$

Hence

$$\beta(S) := \sup_{\lambda,\mu,\xi\ge 0} \gamma_S(\lambda,\mu,\xi) \tag{3.2}$$

is a lower bound $\beta(S)$ for $w_*(S)$. The following lemma states that this lower bound can be computed by minimizing a certain convex function on $S\cap\hat{Y}$.

Lemma 3.1 *Suppose that Q is positive semidefinite matrix. Then*

$$\beta(S) = \min_{y\in S\cap\hat{Y}} g_S(y),$$

where

$$g_S(y) := \sup_{(\lambda,\mu,\xi)\in\Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy+q, y-y^k\rangle + \mu^T(Ny+p) + \xi^T d\}$$

is convex on S and

$$\Omega(S) := \{(\lambda,\mu,\xi):\ \lambda\ge 0, \mu\ge 0, \xi\ge 0, G_v(\lambda,\mu,\xi)\ge 0\ \forall v\in S\cap\hat{Y}\}$$

with

$$G_v(\lambda,\mu,\xi) := a + \mu M + \xi A + \sum_{k=1}^{s} \lambda_k (P(v))^T (v-y^k).$$

Proof From (3.1) and the definition of $\nu_S(\lambda,\mu,\xi)$, it follows that

$$\beta(S) = \sup_{\lambda,\mu,\xi\ge 0} \gamma_S(\lambda,\mu,\xi) = \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} m(y,\lambda,\mu,\xi).$$

Hence

$$\begin{cases} \beta(S) = \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} \min_{x\ge 0} L(x,y,\lambda,\mu,\xi) = \\ \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} \Big[\min_{x\ge 0}\{a^T x + b^T y + \sum_{k=1}^{s} \lambda_k \langle P(y)x + Qy + q, y-y^k\rangle \\ +\mu^T(Mx+Ny+p) + \xi^T(Ax+d)\}\Big] = \\ \sup_{\lambda,\mu,\xi\ge 0} \min_{y\in S\cap\hat{Y}} \Big[\{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy+q, y-y^k\rangle + \mu^T(Ny+p) + \xi^T d\} \\ +\min_{x\ge 0}\{a^T x + \mu^T Mx + \xi^T Ax + \sum_{k=1}^{s} \lambda_k \langle P(y)x, y-y^k\rangle\}\Big]. \end{cases} \tag{3.3}$$

We now consider the last term of (3.3), that is

$$\min_{x\ge 0}\{a^T x + \mu^T Mx + \xi^T Ax + \sum_{k=1}^{s} \lambda_k \langle P(y)x, y-y^k\rangle\} =$$

$$\min_{x\geq 0}\{a^Tx+\mu^TMx+\xi^TAx+\sum_{k=1}^{s}\lambda_k\langle x,P^T(y)(y-y^k)\rangle\}=$$

$$\min_{x\geq 0}\{a^Tx+\mu^TMx+\xi^TAx+\langle x,\sum_{k=1}^{s}\lambda_kP^T(y)(y-y^k)\rangle\}=$$

$$\min_{x\geq 0}\langle x,a+M^T\mu+A^T\xi+\sum_{k=1}^{s}\lambda_kP^T(y)(y-y^k)\rangle,$$

here and afterward, $P^T(y)$ denotes the transportation of the matrix $P(y)$.

If there exists $v\in S\cap\hat{Y}$ such that

$$a+M^T\mu+A^T\xi+\sum_{k=1}^{s}\lambda_kP^T(v)(v-y^k)\ \not\geq 0\ \ \forall\lambda\geq 0,\mu\geq 0,\xi\geq 0,$$

then

$$\min_{x\geq 0}\langle x,a+\mu M+\xi A+\sum_{k=1}^{s}\lambda_kP^T(v)(v-y^k)\rangle=-\infty.$$

So, the supremum in (3.3) can be taken over, all $\lambda\geq 0$, $\mu\geq 0$ and $\xi\geq 0$ satisfying

$$a+M^T\mu+A^T\xi+\sum_{k=1}^{s}\lambda_kP^T(v)(v-y^k)\geq 0\ \ \forall v\in S\cap\hat{Y}.$$

Clearly, under the condition

$$a+\mu M+\xi A+\sum_{k=1}^{s}\lambda_kP^T(y)(y-y^k)\geq 0\ \ \forall y\in S\cap\hat{Y},$$

one has

$$\min_{x\geq 0}\langle x,a+M^T\mu+A^T\xi+\sum_{k=1}^{s}\lambda_kP^T(y)(y-y^k)\rangle=0.$$

Thus we deduce from (3.3) that

$$\begin{cases}\beta(S)=\sup_{\lambda,\mu,\xi\geq 0}\min_{y\in S\cap\hat{Y}}\left[b^Ty+\sum_{k=1}^{r}\lambda_k\langle Qy+q,y-y^k\rangle+\mu^T(Ny+p)+\xi^Td\right]\\ \text{subject to}\\ a+M^T\mu+A^T\xi+\sum_{k=1}^{s}\lambda_kP^T(v)(v-y^k)\geq 0\ \ \forall v\in S\cap\hat{Y}.\end{cases}\tag{3.4}$$

Let

$$\Omega(S):=\{(\lambda,\mu,\xi):\ \lambda\geq 0,\mu\geq 0,\xi\geq 0,G_v(\lambda,\mu,\xi)\geq 0\ \forall v\in S\cap\hat{Y}\}.$$

Then $\Omega(S)$ is a closed convex set and

$$\begin{cases}\beta(S)=\sup_{(\lambda,\mu,\xi)\in\Omega(S)}\min_{y\in S\cap\hat{Y}}\{b^Ty+\sum_{k=1}^{r}\lambda_k\langle Qy+q,y-y^k\rangle+\\ \mu^T(Ny+p)+\xi^Td\}.\end{cases}\tag{3.5}$$

Since, by the assumption, Q is positive semidefinite, the function

$$b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d$$

is convex-linear on $(S \cap \hat{Y}) \times \Omega(S)$. Then, by the well known minimax theorem, we can interchange the supremum and infimum in (3.5) to obtain

$$\beta(S) = \min_{y \in S \cap \hat{Y}} \sup_{(\lambda,\mu,\xi) \in \Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d\}.$$

Note that, since

$$b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d$$

is convex with respect to y, the function

$$g_S(y) := \sup_{(\lambda,\mu,\xi) \in \Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d\}$$

is convex on $S \cap \hat{Y}$. Hence

$$\beta(S) = \min_{y \in S \cap \hat{Y}} g_S(y)$$

is a convex program. □

Remark 3.1 (a) If $P^T(y)y - P^T(y)y^k$ is concave on Y (in the sense that its every component is concave), then $\Omega(S)$ is a polyhedral convex set, since it can be represented by a finite affine inequalities. Indeed, since $P^T(y)y - P^T(y)s^k$ is concave, it is easy to verify that

$$a + M^T \mu + A^T \xi + \sum_{k=1}^{s} \lambda_k P^T(y)(y - y^k) \geq 0 \quad \forall y \in Y \cap S$$

if and only if

$$a + M^T \mu + A^T \xi + \sum_{k=1}^{s} \lambda_k P^T(y)(y - y^k) \geq 0 \quad \forall y \in V(Y \cap S),$$

where $V(Y \cap S)$ denotes the vertex-set of $Y \cap S$.

In the case $S \subseteq Y$ the last inequalities can be written as

$$a + M^T \mu + A^T \xi + \sum_{k=1}^{s} \lambda_k P^T(v^j)(v^j - y^k) \geq 0 \quad \forall j = 1, 2, ..., r,$$

where $v^1, ..., v^r$ are the vertices of S.

(b) From the presentation of Section 2 we can see that

• For bilevel quadratic convex problem, $P(y) \equiv P$ (constant matrix).

• For optimization problem over the efficient set of an affine vector optimization program $P^T(y)y^k$ is affine and $P^T(y)y \equiv 0$.

3.2 Simplicial and Rectangular Bisections

At each iteration k of algorithm to be described below, a partition simplex (or rectangle) will be bisected into subsimplices (or subrectangles) such a way so that as the algorithm executes the obtained lower and upper bounds tend to the same limit. This can be achieved by using the following exhaustive simplicial (or rectangular) bisection that is commonly known in global optimization (see e. g. [14]).

Simplicial Bisection. We will use the following simplicial bisection [14].

Let S_k be a subsimplex of full dimension that we want to bisect at iteration k. Let v^k, w^k be two vertices of S_k such that the edge joining these vertices is longest. Let $u^k = t_k v^k + (1 - t_k) w^k$ with $0 < t_k < 1$. Let S_{k_1}, S_{k_2} be the subsimplices obtained from S_k by replacing v^k and w^k respectively by u^k. It is well known from [14] that $S_k = S_{k_1} \cup S_{k_2}$, and that if $\{S_k\}$ is an infinite sequence of nested simplices generated by this simplicial bisection process such that $0 < \delta_0 < t_k < \delta_1 < 1$ for every k, then the sequence $\{S_k\}$ shrinks to a singleton.

Rectangular Bisection. Suppose that the partition set is a rectangle given by

$$S_k := \{y = (y_1, ..., y_m) \in R^m : \ a_i \le y_i \le b_i \ i = 1, ..., m\}.$$

Let $[a_{i_k}, b_{i_k}]$ be a longest edge of S_k and $u_{i_k} = t_{i_k} a_{i_k} + (1 - i_k b_{i_k}$ with $0 < t_{i_k} < 1$. Then we bisect S_k into two rectangles S_{k_1} and S_{k_2} where

$$S_{k_1} = \{y \in R^m : \ a_i \le y_i \le b_i \ \forall i \ne i_k, \ a_{i_k} \le y_{i_k} \le u_{i_k}\},$$

and

$$S_{k_2} = \{y \in R^m : \ a_i \le y_i \le b_i \ \forall i \ne i_k, \ u_{i_k} \le y_{i_k} \le b_{i_k}\}.$$

As before, we have $S_k = S_{k_1} \cup S_{k_2}$, and that if $\{S_k\}$ is an infinite sequence of nested rectangles generated by this bisection process such that $0 < \delta_0 < t_k < \delta_1 < 1$ for every k, then the sequence $\{S_k\}$ shrinks to a singleton.

Now we are in a position to describe the algorithm for solving AMPEC problem (P), where Y is a polytope. We suppose that $\langle P(y)x, y\rangle$ is convex in Y with respect to y. In the sequel, as usual, we call (x, y) an ϵ-global optimal solution to (P) if it is feasible and $f(x, y) - w_* \le \epsilon(|f(x, y)| + 1)$ where w_* stands for its optimal value. Having the vertices $y^1, ..., y^s$ of Y we can describe the algorithm as follows.

Algorithm 1.

Initialization. Choose a tolerance $\epsilon > 0$ and a simplex or a rectangle S_0 containing Y. Compute the lower bound $\beta(S_0)$ by solving the convex program

$$\beta(S_0) := \min_{y \in \hat{Y}} \Big\{ g_0(y) := \sup_{(\lambda,\mu,\xi)\in\Omega(S_0)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k\rangle + \mu^T(Ny + p) + \xi^T d\} \Big\}.$$

Let $y^0 \in S_0$ be the obtained solution.

Solve the linear program

$$\min_x \{f(x, y^0) := a^T x + b^T y^0\}$$

subject to

$$\begin{cases} Mx + Ny^0 + p \le 0 \\ Ax + d \le 0, x \ge 0 \\ \langle P(y^0)x + Qy^0 + q, y^0 - y^k\rangle \le 0 \ \ \forall k = 1, ..., s \end{cases}$$

to obtain x^0 (hence (x^0, y^0) is feasible). Let $\alpha_0 := f(x^0, y^0)$ (an upper bound for the optimal value w_*) and $\beta_0 := \beta(S_0)$ (a lower bound for w_*). Take

$$\Gamma_0 := \begin{cases} \{S_0\} \text{ if } \alpha_0 - \beta_0 > \epsilon(|\alpha_0| + 1), \\ \emptyset \quad\quad \text{otherwise} \end{cases}$$

and go to iteration k with $k := 0$.

Iteration ($k = 0, 1...$). At the beginning of each iteration k we have family Γ_k of partition sets to each element $S \in \Gamma_k$ a real number $\beta(S)$ has been computed that serves as a lower bound for Problem (P) restricted in S. Moreover we have a lower bound β_k for w_*, a currently best feasible point (x^k, y^k) and an upper bound $\alpha_k = f(x^k, y^k)$ for w_*.

Step 1 (selection):
(i) If $\Gamma_k = \emptyset$ then terminate, (x^k, y^k) is an ϵ-global optimal solution and α_k is the ϵ-optimal value to Problem (P).
(ii) If $\Gamma_k \neq \emptyset$, then select $S_k \in \Gamma_k$ such that

$$\beta_k = \beta(S_k) = \min\{\beta(S) \ : \ S \in \Gamma_k\}.$$

Step 2 (bisection): Use the simplicial bisection (if S_k is simplicial) or use the rectangular bisection (if S_k is rectangular) to bisect S_k into two sets S_{k_1} and S_{k_2}.

Step 3 (bounding): For each newly generated sets S_{k_j} $(j = 1, 2)$ satisfying $S_{k_j} \cap \hat{Y} \neq \emptyset$, compute

$$\beta(S_{k_j}) := \min_{y \in S_{k_j} \cap \hat{Y}} g_{S_{k_j}}(y),$$

where

$$g_{S_{k_j}}(y) := \sup_{(\lambda, \mu, \xi) \in \Omega(S_{k_j})} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k \rangle + \mu^T (Ny + p) + \xi^T d\}.$$

Let y^{k_j} be the obtained solution.

Step 4 (updating upper bound): Solve the linear programs, one for each y^{k_j}, $j = 1, 2$

$$\min_x \{f(x, y^{k_j}) := a^T x + b^T y^{k_j}\}$$

subject to

$$\begin{cases} Mx + Ny^{k_j} + p \leq 0 \\ Ax + d \leq 0, x \geq 0 \\ \langle P(y^{k_j})x + Qy^{k_j} + q, y^{k_j} - y^k \rangle \leq 0 \quad \forall k = 1, ..., s \end{cases}$$

to obtain new feasible points. Use these feasible points to update the upper bound. Let (x^{k+1}, y^{k+1}) be the currently best feasible point among (x^k, y^k) and the newly generated feasible points. Set $\alpha_{k+1} := f(x^{k+1}, y^{k+1})$ and

$$\Gamma_{k+1} := \Big\{ S \in (\Gamma_k \backslash \{S_k\}) \cup \{S_{k_1}, S_{k_2}\} \ : \ \alpha_{k+1} - \beta(S) > \epsilon(|\alpha_{k+1}| + 1) \Big\}.$$

Increase k by 1 and go to Step 1 of iteration k.

Theorem 3.1 *a) If Algorithm terminates at iteration k, then (x^k, y^k) is an ϵ-global optimal solution to Problem (P).*
b) If the algorithm does not terminate, then $\beta_k \nearrow w_*, \alpha_k \searrow w_*$ as $k \to +\infty$, and any cluster point of the sequence $\{(x^k, y^k)\}$ is a global optimal solution to (P).

Proof a) If the algorithm terminates at iteration k then $\Gamma_k = \emptyset$. This implies that $\alpha_k - \beta_k \leq \epsilon(|\alpha_k| + 1)$. Since $\beta_k \leq w_*$ and $\alpha_k = f(x^k, y^k) \geq w_*$, it follows that $f(x^k, y^k) - w_* \leq \epsilon(|f(x^k, y^k)| + 1)$. Hence (x^k, y^k) is an ϵ-global optimal solution.

b) Suppose that the algorithm does not terminate. First, note that since $S_k = S_{k_1} \cup S_{k_2}$, by the rule for computing lower bound $\beta(S_k)$ we have

$$\beta_k = \beta(S_k) \leq \beta(S_{k+1}) = \beta_{k+1} \quad \forall k.$$

Also, by definition of α_k, we have $\alpha_{k+1} \leq \alpha_k \;\; \forall k$. Thus, both $\beta_* = \lim \beta_k$ and $\alpha_* = \lim \alpha_k$ exist and satisfying

$$\beta_* \leq w_* \leq \alpha_*. \tag{3.6}$$

Since the algorithm does not terminate, it generates an infinite sequence of nested partition sets that, for simplicity of notation, we also denote by $\{S_k\}$. Since the subdivision is exhaustive, this sequence strinks to a singleton, say $y^* \in \hat{Y}$. By the rule for computing lower bound β_k we have

$$\beta_k = \sup_{\tau \geq 0} \min_{y \in S_k \cap \hat{Y}} m(y, \tau) \geq \min_{y \in S_k} m(y, \tau) \quad \forall \tau \equiv (\lambda, \mu, \xi) \geq 0.$$

Since the sequence $\{S_k\}$ shrinks to y^* as $k \to +\infty$, we obtain

$$\beta_* = \lim \beta_k \geq m(y^*, \tau) \quad \forall \tau \geq 0.$$

By definition, since $\varphi(y^k)$ is an upper bound determined at iteration k and α_{k+1} is the currently smallest upper bound obtained at this iteration, we can write

$$\alpha_{k+1} \leq \varphi(y^k) \quad \forall k.$$

From $y^k \to y^*$, it follows, by the continuity of φ (see e. g. [3, 5]), that

$$\alpha_* = \lim \alpha_k = \lim \alpha_{k+1} \leq \lim \varphi(y^k) = \varphi(y^*).$$

On the other hand, by Lagrangian duality theorem for the convex program determining $\varphi(y^*)$, we have

$$\sup_{\tau \geq 0} m(y^*, \tau) = \varphi(y^*).$$

Hence

$$\alpha_* \leq \varphi(y^*) \leq \beta_*$$

which together with (3.6) implies

$$\beta_* = w_* = \alpha_* = \varphi(y^*).$$

Let (x^*, y^*) be any cluster point of the sequence $\{(x^k, y^k)\}$. By the definition we have $\alpha_k = f(x^k, y^k)$. Since $\alpha_k \searrow w_*$, it follows from continuity of f that $w_* = f(x^*, y^*)$. Hence (x^*, y^*) is a globally optimal solution to Problem (P). □

Remark 3.2 Note that when Q is positive semidefinite, the function

$$g_S(y) := \max_{(\lambda,\mu,\xi)\in\Omega(S)} \{b^T y + \sum_{k=1}^{s} \lambda_k \langle Qy + q, y - y^k\rangle + \mu^T(Ny + p) + \xi^T d\}$$

is convex and subdifferentiable, since it is the maximum of a family of convex functions. The subgradient of g_S at a point y can be obtained by taking the convex envelope of the gradients of those quadratic functions in the family at which $g_S(y)$ is attained [4].

3.3 Optimization over the Weakly Efficient Set

Now we return to the optimization over the weakly efficient set mentioned in the previous section. By using again the necessary and sufficient condition due to Malivert (see Theorem 2.1) the minimization problem over the weakly efficient set of a multicriteria affine fractional program can be written as

$$\min f(v)$$

subject to

$$v \in V,\ \ u \in U,$$

$$\sum_{i=1}^{\rho} \Big\langle u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], w - v\Big\rangle \ge 0 \ \ \forall w \in V,$$

where

$$U := \{u \in R^\rho :\ u \ge 0, \sum_{i=1}^{p} u_i = 1\}$$

is a simplex in the criteria space.

Since

$$\Big\langle \sum_{i=1}^{\rho} u_i B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], w - v\Big\rangle$$

is affine with respect to w, we can easily check that

$$\Big\langle \sum_{i=1}^{\rho} u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], w - v\Big\rangle \ge 0 \ \ \forall w \in V$$

if and only if

$$\Big\langle \sum_{i=1}^{\rho} u_i[(B_i^T v + t_i)A_i - (A_i^T v + s_i)B_i], v^j - v\Big\rangle \ge 0 \ \ \forall j = 1, ..., r,$$

where v^j $(j = 1, ..., r)$ are vertices of V. Thus we can write this problem as

$$\begin{cases} \min f(v):\ \ \text{s. t. } v \in V,\ u \in U, \\ \Big\langle \sum_{i=1}^{\rho} u_i[(B_i^T v + t_i)A_i - (A_i^T + s_i)B_i], v^j - v\Big\rangle \ge 0 \ \ \forall j = 1, ..., r. \end{cases} \tag{3.7}$$

It is worth pointing out that, in contrast to the linear case, this problem does not necessarily attain its optimal solution among the vertices of V.

In this formulation we require that all vertices of V are known. This case appeared already in some applications in economics (see e. g. [18, 27, 28]) where each component v_j of the decision variable v represents the ratio of ith quantity to be determined. In such practical models, V is a simplex given by

$$V := \Big\{ v^T = (v_1, ..., v_m) \ : \sum_{i=1}^{m} v_i = 1, v_j \geq 0 \ \forall j = 1, ..., m \ \Big\},$$

whose vertices are easy to compute. Generally, let us assume that V is a polytope given explicitly as

$$V := \{ v \in R^m : \ v \geq 0, \ Gv - g \leq 0 \}.$$

For simplicity of notation, for each vertex v^j, we take

$$M_j(u, v) := \Big\langle \sum_{k=1}^{\rho} u_k [(B_k^T v + t_k) A_k - (A_k^T v + s_k) B_k], v - v^j \Big\rangle,$$

$$G_j(u) := \sum_{k=1}^{\rho} u_k [(B_k^T v^j + t_k) A_k^T - (A_k^T v^j + s_k) B_k^T],$$

$$g_j(u) := \sum_{k=1}^{\rho} u_k [t_k A_k^T - s_k B_k^T] v^j.$$

Let $G(u)$ denote the $(r \times m)$-matrix whose ith row is $G_j(u)$ $(j = 1, ..., r)$, and $g(u)$ denote the r-dimensional vector whose jth entry is $g_j(u)$. Now let

$$H(u) := \begin{pmatrix} G \\ G(u) \end{pmatrix}, \quad h(u) := \begin{pmatrix} g \\ g(u) \end{pmatrix}.$$

Under these notations we can write the problem (3.7) in the form

$$\min\{ f(v) : \ H(u)v - h(u) \leq 0, v \geq 0, u \in U \}. \tag{3.8}$$

To apply the Lagrangian duality we take the Lagrangian function for this problem with respect to the constraint $H(u)v - h(u) \leq 0$, that is

$$L(\lambda, u, v) := f(v) + \lambda^T \Big(H(u)v - h(u) \Big).$$

Using the fact that both $H(u)$ and $h(u)$ are affine, by a similar argument as in the proof of Lemma 2.1, we can compute lower bounds by solving linear programs as stated by the following lemma.

Lemma 3.2 *Suppose $f(v) = b^T v$. Let S be the subsimplex of the simplex U, and s^j $(j = 1, ..., \rho)$ be the vertices of S. Then $\beta(S) = \min\{\beta(s^j) \ : j = 1, ..., \rho\}$ where, for each fixed s^j, $\beta(s^j)$ is the optimal value of the linear program*

$$\beta(s^j) := \max\{ -h^T(s^j) u \ : H^T(s^j) u + b \geq 0 \}.$$

Having this lower bounding operation we can use Algorithm 1 with the exhaustive simplicial bisection taking over subsimplices of the simplex U to solve problem (3.8). Note that in this case, if (u^S, v^S) is a solution obtained by computing lower bound $\beta(S)$ according to Lemma 3.2, then v^S is weakly efficient. Hence (u^S, v^S) can serve for updating upper bound in the algorithm.

4 A Relaxation Algorithm

In the algorithm presented in the preceding section we required that all vertices of the polytope Y are known in advance. In the case computing all of these vertices is expensive, we recommend to use another algorithm that allows that these vertices can be computed iteratively. It is expected that the algorithm finds an approximate solution without computing all of these vertices. In order to present the algorithm, suppose that we know already some vertices $v^1, ..., v^r$ of Y. Having these vertices we form the relaxation problem

$$\begin{cases} r_* = \min f(x,y) := a^T x + b^T y \\ \text{s.t. } Mx + Ny + p \le 0 \\ Ax + d \le 0,\ x \ge 0,\ y \in Y \\ \langle P(y)x + Qy + q, v^k - y\rangle \ge 0 \ \ \forall k = 1, ..., r. \end{cases} \tag{RP}$$

Clearly, the feasible domain of this problem contains that of Problem (P) presented in Section 2 with $X = \{x \ge 0 \ : \ Ax + d \le 0\}$. Applying Algorithm 1 to this problem we obtain an ϵ-global optimal solution of (RP). If this solution satisfies the variational inequality constraint, then it is also an ϵ-global solution to the original problem (P). Otherwise, it violates at least one constraint. Then we add one or more violated constraints to obtain new relaxation problem and repeat the process. The algorithm can be described in detail as follows.

Algorithm 2.

Step 1. Choose distinct vertices $v^1, ..., v^r$ of Y and a tolerance $\epsilon > 0$.

Step 2. Use Algorithm 1 to solve (RP) to obtain an ϵ-global optimal solution (x^r, y^r) to (RP).

Step 3. Solve the linear program

$$\min\{\langle P(y^r)x^r + Qy^r + q, y\rangle \ : y \in Y\} \tag{L_r}$$

to obtain an basis (vertex) solution y^{r+1}.

(a) If

$$\langle P(y^r)x^r + Qy^r + q, y^r\rangle \le \langle P(y^r)x^r + Qy^r + q, y^{r+1}\rangle$$

(hence y^r also solves (L_x)), then terminate: (x^r, y^r) is an ϵ-global optimal solution to the original problem.

(b) Otherwise, take $v^{r+1} := y^{r+1}$. Add v^{r+1} to the list of known vertices of Y to form the new relaxation problem (RP) and go back to Step 2.

The following theorem shows validity and finiteness of this algorithm.

Theorem 4.1 *(i) If the algorithm terminates at case (a) of Step 3, then (x^r, y^r) is an ϵ-global optimal solution.*

(ii) The algorithm terminates after a finite number of Step 2 yielding an ϵ-global optimal solution to the original problem (P).

Proof (i) Note that Problem (P) can be written as

$$w_* := \min\{f(x,y) := a^T x + b^T y\}$$

subject to

$$\begin{cases} Mx + Ny + p \le 0 \\ Ax + d \le 0, x \ge 0, y \in Y \\ \langle P(y)x + Qy + q, v - y\rangle \ge 0 \ \ \forall v \in V(Y) \end{cases}$$

while the relaxation problem is

$$r_* := \min\{f(x,y) := a^T x + b^T y\}$$

subject to

$$\begin{cases} Mx + Ny + p \leq 0 \\ Ax + d \leq 0, x \geq 0, \ y \in Y \\ \langle P(y)x + Qy + q, v^k - y\rangle \geq 0 \ \ \forall k = 1, ..., r \end{cases}$$

with v^k $(k = 1, ..., r)$ being some vertices of Y.

In the case (a) we have

$$\langle P(y^r)x^r + Qy^r + q, y^r\rangle \leq \langle P(y^r)x^r + Qy^r + q, y^{r+1}\rangle.$$

Since y^{r+1} is an optimal solution of (L_r), we have

$$\langle P(y^r)x^r + Qy^r + q, y^{r+1}\rangle \leq \langle P(y^r)x^r + Qy^r + q, y\rangle \ \forall y \in Y.$$

Thus

$$\langle P(y^r)x^r + Qy^r + q, y - y^r\rangle \geq 0 \ \forall y \in Y.$$

Hence (x^r, y^r) is feasible for (P). But, since (x^r, y^r) is an ϵ-global optimal solution to the relaxed problem (RP), it must be an ϵ-global solution to (P).

(ii) Note that if $y^{r+1} = v^j$ for some $j \leq r$, then we have the case (a), and therefore, the algorithm terminates. Thus, if situation (a) is not the case, we have $v^{r+1} \neq v^j$ for all r and all $j \leq r$, since the number of the vertices of Y is finite, the algorithm must terminate with case (a). $\square$

5 Conclusion

We have considered a class of mathematical programs with affine variational inequality constraints and presented some of its important special cases such as bilevel convex programming, optimization over the efficient set and Cournot-Nash oligopolistic market model. We have developed two decomposition branch-and-bound algorithms for globally solving this class of mathematical programs with affine equilibrium constraints. The proposed algorithms use the Lagrangian bound and exhaustive simplicial and rectangular subdivisions widely used in global optimization. The main subproblems needed to be solved in the algorithms are convex programs that can be solved by well developed methods of nonsmooth convex programming.

Acknowledgement

The second author wishes to thank the financial supports from the Alexander-von-Humboldt Foundation, Germany and from the Associate and Federation Schemes, ICTP, Italy.

References

[1] Anh, P.N., Muu, L.D., Nguyen, V.H. and Strodiot, J.J. On the contraction and non-expensiveness properties of the marginal mapping in generalized variational inequalities involving cocoercive operators. In: *Generalized Convexity, Generalized Monotonicity and Applications*. (Eberhard, A., Hadjisavvas, H. and Luc. D.T., eds.). Springer, Berlin, 2005, 89–111.

[2] Anh, P.N., Muu, L.D., Nguyen, V.H. and Strodiot, J.J. Using the Banach contraction principle to implement the proximal point method for solving multivalued monotone variational inequalities. *J. of Optimization Theory and Applications* **124** (2005) 285–306.

[3] Aubin, J.P. and Ekeland, I. *Applied Nonlinear Analysis*. Marcel Dekker, New York, 1984.

[4] Auslender, A. *Optimisation: Methodes Numeŕiques*. Masson, Paris, 1976.

[5] Berge, C. *Topological Spaces*. Macmillan, New York, 1968.

[6] Bental, A., Eiger, G. and Gershovitg, V. Global minimization by reducing the duality gap. *Mathematical Programming* **63** (1994) 193–212.

[7] Choo, E.U. and Atkins, D.H. Connectedness in multiple linear fractional programming. *Management Science* **29** (1983) 250–255.

[8] Dempe, S. Annotated Bibliography on Bilevel Programming and Mathematical Programs with Equilibrium Constraints. *Optimization* **52** (2003) 333–359.

[9] Facchinei, F., Jiang, H. and Qi, Q. A smoothing method for mathematical programs with equilibrium constraints. *Mathematical Programming* **85** (1999) 107–134.

[10] Facchinei, F. and Pang, J. S. *Finite Dimesional Variational Inequality and Complementarity Problems*. Springer, Berlin, 2002.

[11] Falk, J.E. Lagrange multipliers and nonconvex programs. *SIAM J. Control* **7** (1969) 534–545.

[12] Konnov, I.V. *Combined relaxation methods for variational inequalities*. Springer, Berlin, 2001.

[13] Friesz, T.L., Tobin, R.L., Cho, H.J. and Mehta, N.J. Sensitivity analysis based heuristic algorithm for mathematical programs with variational inequality constraints. *Mathematical Programming* **48** (1990) 265–284.

[14] Horst, R. and Tuy, H. *Global Optimization: Deterministic Approach*. Springer, Berlin, 2003.

[15] Lin, G.H. and Fukushima, M. Some exact penalty results for nonlinear programs and mathematical programs with equilibrium constraints. *J. Optimization Theory and Applications* **118** (2003) 67–80.

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

[17] Malivert, C. Multicriteria fractional optimization. In: *The 2-nd Catalans Days on Appl.* Presses Universitaires de Perpignan, 1995, 189-198.

[18] Marcotte, P. Network design problem with congestion effects: A case of bilevel programming. *Mathematical Programming* **34** (1986) 142–162.

[19] Mordukhovich, B.S. Optimization and equilibrium problems with equilibrium constraints. *Omega* **33** (2005) 379–384.

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

[21] Muu, L.D. and Quy, N.V. Method for finding global optimal solution to linear programs with equilibrium constraints. *Acta Mathematica Vietnamica* **26** (2001) 314–333.

[22] Muu,. L.D. and Tuyen, H.Q. Bilinear programming approach to optimization over the efficient set of a vector affine fractional problem. *ACTA Mathematica Vietnamica* **27** (2002) 119–139.

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

[24] Outrata, J.V. and Zowe, J. A numberical approach to optimization problems with variational inequality constraints. *Mathematical Programming* **68** (1995) 105–130.

[25] Philip, J. Algorithms for the vector maximization problem. *Mathematical Programming* **2** (1972) 207–228.

[26] Shor, N.Z. and Stetsyuk, P.I. Lagrangian bounds in multiextremal polynomial and discrete optimization problems. *J. of Global Optimization* **23** (2002) 1–41.

[27] R. E. Steuer, *Multiple Criteria Optimization: Theory, Computation, and Applications*, John Willey and Sons, New York, 1986.

[28] Tobin, R.L. Uniqueness results and algorithms for Stackelberg-Cournot-Nash equilibria. In: *Hierarchical Optimization*. (Anandalingam, G. and Friesz, T., eds.), Ann. Oper. Res., 1992, 21–36.

[29] Tuyen, H.Q. and Muu, L.D. Biconvex programming approach to optimization over the weakly efficient set of a multiple objective affine fractional problem. *Operations Research Letters* **28** (2001) 81–92.