Journal of Global Optimization (2005)32: 349–365
DOI 10.1007/s10898-004-1947-9

On Solving Nonconvex Optimization Problems by Reducing The Duality Gap

HOANG TUY
Institute of Mathematics, 18 Hoang Quoc Viet, 10307, Hanoi, Vietnam
(e-mail: htuy@math.ac.vn)

(Received 5 November 2002; accepted in revised form 26 July 2004)

Abstract. Lagrangian bounds, i.e. bounds computed by Lagrangian relaxation, have been used successfully in branch and bound bound methods for solving certain classes of non-convex optimization problems by reducing the duality gap. We discuss this method for the class of partly linear and partly convex optimization problems and, incidentally, point out incorrect results in the recent literature on this subject.

Key words: Branch and bound algorithm, Convergence conditions, Dual bound, Lagrangian bound, Partly convex programming, Partly linear

1. Introduction

Branch and bound methods are of frequent use in nonconvex global optimization, especially for solving large scale problems. A class of problems for which these methods have proved to be particularly useful includes the following general *partly convex optimization problem*

$$\min\{F(x,y)|\ G_i(x,y)\leqslant 0\ i=1,\dots,m,\ x\in C,\ y\in D\}, \tag{GPC}$$

where C is a nonempty compact convex subset of a (hyper)rectangle $X\subset\mathbb{R}^n_+, D$ is a nonempty closed convex subset of a (hyper)rectangle $Y\subset\mathbb{R}^p_+$, while $F(x,y): X\times Y\to\mathbb{R}$, and $G_i(x,y): X\times Y\to\mathbb{R}$, $i=1,\dots,m$, are lower semi-continuous functions, convex in y for every fixed x. As is well known, problems of this class are encountered in various forms in a multitude of applications: pooling and blending in oil refinery, optimal design of water distribution networks, structural design, signal processing, robust stability analysis, design of chips, etc. (see e.g. [7]).

Two basic operations involved in a branch and bound procedure for global minimization are branching (successive partition) and bounding (estimating a lower bound for the objective function value over the feasible portion contained in each partition set). The partial convex structure allows the problem (GPC) to be decomposed into a sequence of problems of smaller dimension by branching upon the nonconvex variables x, rather than upon the total set of variables x, y. The successive partition of the

x-space is usually required to be exhaustive, so that any filter (infinite sequence of nested partition sets) shrinks to a single point [15]. As for bounding, one possible method consists in solving for each partition set the Lagrangian dual of the problem restricted to this partition set. The lower bound provided by solving a Lagrangian relaxation is often referred to as a *Lagrangian bound* or sometimes, a *dual bound*. It has been observed that in most cases the difference between the exact optimal value and the Lagrangian bound, i.e. the duality gap, decreases when the partition set becomes smaller. One may hope that a suitably organized branch and bound process will generate a filter of partition sets such that the duality gap for the subproblems associated with the partition sets in this filter tends to zero, yielding at the limit a convex subproblem with zero duality gap. Any optimal solution of the latter subproblem will then provide an optimal solution of the original problem.

Three issues arise in this methodology: (1) for which problems Lagrangian bounds can be practically computed? (2) when Lagrangian bounds can be computed, are they the best among bounds obtained by convex relaxation? and (3) under which conditions a branch and bound of the above described type is guaranteed to converge to a global optimal solution?

Due to their interest from both a theoretical and practical point of view, these issues have, in the past two decades, attracted the attention of an increasing number of researchers. Partial answers to these issues were obtained in [4] and [5], where the idea of solving nonconvex optimization problems by reducing the duality gap was for the first time put forward and implemented on concrete problems. However, to our knowledge, the first comprehensive account on the use of Lagrangian bounds in global continuous and discrete optimization was given by Shor and Stesenko [10]. This booklet contains, among other things, several useful remarks that could have helped to avoid regrettable errors in subsequent papers, [11, 12], by other authors. In a more recent paper [2], the questions of when Lagrangian bounds can be easily computed and when the associated branch and bound algorithm is guaranteed to converge were studied in the more restricted context of partly linear programs, which constitute a subclass of the class (GPC). Lagrangian duality results, together with their applications to biconvex optimization problems (another subclass of (GPC)), were discussed at length in [7]. On the other hand, due to quite a few incorrect and misleading results, e.g. [11, 12], a rather confuse situation has resulted in this important area of research.

The aim of the present paper is to clarify this situation by investigating the above issues concerning the Lagrangian bound approach for partly convex optimization problems. After the Introduction, in Section 2, addressing the first issue, we show that the Lagrangian dual to a partly linear optimization problem is a convex problem, and under some additional assumptions, even a linear program. Since linear programs can be solved very efficiently, this seems to be an advantage of Lagrangian bounds for this class of problems. However, in the next Section 3 we show how these bounds may be very poor if the Lagrangian dual is formed from a bad reformulation of the original problem. Section 4 discusses the last issue of convergence conditions for Lagrangian bound methods for partly convex global optimization. Incidentally, several incorrect results that have been published recently on this subject will be pointed out and corrected. Section 5 concludes the paper with some applications illustrating the usefulness of the method when it is well founded.

2. When Lagrangian Bounds can be Practically Computed?

Lagrangian bounds are generally difficult to compute. In fact, more often than not, the Lagrangian dual to a nonconvex problem is itself a nonconvex problem. Therefore, the first important issue to be dealt with is when Lagrangian bounds can be computed practically. In this section we examine a class of problems for which the Lagrangian dual is a convex, or even a linear program. This class includes partly linear optimization problems that have the following general formulation:

$$f^*_{[r,s]} = \min\{\langle c(x), y\rangle + \langle c^0, x\rangle \mid A(x)y + Bx \leqslant b, r \leqslant x \leqslant s, y \geqslant 0\}, \qquad \text{(GPL)}$$

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^p$, $c : \mathbb{R}^n \to \mathbb{R}^p$, $c^0 \in \mathbb{R}^n$, $A := \mathbb{R}^n \to \mathbb{R}^{m\times p}$, $B \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $r, s \in \mathbb{R}^n_+$. Setting $A(x) = [a_{ij}(x)]$, and denoting the i-th row of B by B_i, this problem can also be written in the expanded form as

$$\begin{aligned} \min \quad & \sum_{j=1}^{p} y_j c_j(x) + \langle c^0, x\rangle, \\ \text{s.t.} \quad & \sum_{j=1}^{p} y_j a_{ij}(x) + \langle B_i, x\rangle \leqslant b_i \quad i = 1, \ldots, m, \\ & y \geqslant 0, \quad r \leqslant x \leqslant s. \end{aligned} \qquad \text{(GPL)}$$

THEOREM 1. *The Lagrangian dual of* (*GPL*) *with respect to the nonlinear constraints*, *i.e.*,

$$\varphi^*_{[r,s]} = \sup_{\lambda \geqslant 0} \inf\{\langle y, c(x)\rangle + \langle c^0, x\rangle + \langle \lambda, A(x)y + Bx - b\rangle \mid x \in [r,s],\ y \geqslant 0\}, \tag{1}$$

is the convex program

$$\varphi^*_{[r,s]} = \langle c^0, r\rangle + \max_{\lambda,t}[\langle r-s, t\rangle + \langle Br-b, \lambda\rangle], \tag{2}$$

$$\text{s.t.} \quad g(\lambda)\geqslant 0,\ \ \lambda\geqslant 0, t\geqslant 0, t + B^T\lambda + c^0 \geqslant 0, \tag{3}$$

where $g(\lambda) = \min_{j=1,\ldots,p}\min_{r\leqslant x\leqslant s}[\langle A_j(x), \lambda\rangle + c_j(x)]$.

Proof. For fixed $\lambda\geqslant 0$ we have

$$\begin{aligned}&\inf\{\langle y, c(x)\rangle + \langle c^0, x\rangle + \langle \lambda, A(x)y + Bx - b\rangle \mid x\in[r,s],\ y\geqslant 0\}\\ &\quad = -\langle b,\lambda\rangle + \inf_{x\in[r,s]}\inf_{y\geqslant 0}\{\langle Bx,\lambda\rangle + \langle c^0,x\rangle + \langle c(x) + (A(x))^T\lambda, y\rangle\}\\ &\quad = -\langle b,\lambda\rangle + h(\lambda),\end{aligned}$$

where

$$h(\lambda) = \begin{cases}\inf_{x\in[r,s]}[\langle Bx,\lambda\rangle + \langle c^0,x\rangle] & \text{if} \quad c(x) + (A(x))^T\lambda\geqslant 0 \ \ \forall x\in[r,s],\\ -\infty & \text{otherwise.}\end{cases} \tag{4}$$

Now observe that for any $q\in\mathbb{R}^n$:

$$\min_{r\leqslant x\leqslant s}\langle q,x\rangle = \langle q,r\rangle + \max\{\langle r-s,t\rangle \mid t\geqslant 0, t\geqslant -q\},$$

since $\min\{\langle q,x\rangle \mid r\leqslant x\leqslant s\} = \min\{\langle q,r\rangle + \langle q, x-r\rangle \mid 0\leqslant x-r\leqslant s-r\} = \langle q,r\rangle + \sum_{q_i<0} q_i(s_i - r_i) = \langle q,r\rangle + \max\{\langle r-s,t\rangle \mid \ t\geqslant 0, t\geqslant -q\}$. Therefore,

$$\begin{aligned}&\inf_{r\leqslant x\leqslant s}\{\langle Bx,\lambda\rangle + \langle c^0,x\rangle\}\\ &\qquad = \langle B^T\lambda + c^0, r\rangle + \max\{\langle r-s,t\rangle \mid \ \ t\geqslant 0, t\geqslant -B^T\lambda - c^0\}.\end{aligned} \tag{5}$$

Denote the j-th column of A by A_j, so that the condition $\langle A(x), \lambda\rangle + c(x)\geqslant 0 \ \ \forall x\in[r,s]$ means $g(\lambda)\geqslant 0$, where

$$g(\lambda) := \min_{j=1,\ldots,p}\min_{r\leqslant x\leqslant s}[\langle A_j(x),\lambda\rangle + c_j(x)].$$

Since for fixed x the function $\lambda\mapsto\langle A_j(x),\lambda\rangle + c_j(x)$ is affine, $g(\lambda)$ is a concave function and the problem (1) reduces to (2)–(3), which is a convex program. □

COROLLARY 1. Assume that

(*) *For every j the functions $c_j(x), a_{ij}(x), i = 1,\ldots,m$, are either all increasing on $[r,s]$ or all decreasing on $[r,s]$.*

Then the Lagrangian dual of (GPL) is the dual of its LP relaxation:

$$\min \quad \left\{ \sum_{j\in J_+} c_j(r)y_j + \sum_{j\in J_-} c_j(s)y_j + \langle c^0, x\rangle \right\}, \tag{6}$$

$$\text{s.t.} \quad \sum_{j\in J_+} a_{ij}(r)y_j + \sum_{j\in J_-} a_{ij}(s)y_j + \langle B_i, x\rangle \leqslant b_i \quad i = 1, \ldots, m, \tag{7}$$

$$y \geqslant 0, \;\; r \leqslant x \leqslant s, \tag{8}$$

where J_+ is the set of all j such that all $c_j(x), a_{ij}(x), i = 1, \ldots, m$, are increasing, and J_- is the set of all j such that all $c_j(x), a_{ij}(x), i = 1, \ldots, m$, are decreasing.

Proof. By hypothesis $J_+ \cup J_- = \{1, \ldots, p\}$, so for every $j = 1, \ldots, p$, we have

$$\min_{r\leqslant x\leqslant s} \left[\langle A_j(x), \lambda\rangle + c_j(x) \right] \geqslant 0$$

$$\Leftrightarrow \min_{r\leqslant x\leqslant s} \left[\sum_{i=1}^{m} \lambda_i a_{ij}(x) + c_j(x) \right] \geqslant 0$$

$$\Leftrightarrow \begin{cases} \sum_{i=1}^{m} \lambda_i a_{ij}(r) + c_j(r) \geqslant 0 & \text{if } j \in J_+ \ , \\ \sum_{i=1}^{m} \lambda_i a_{ij}(s) + c_j(s) \geqslant 0 & \text{if } j \in J_- \ . \end{cases} \tag{9}$$

In view of (2), (3), (5) the problem (1) thus reduces to the linear program

$$\langle c^0, r\rangle + \max \left\{ \langle r - s, t\rangle + \sum_{i=1}^{m} \lambda_i [\langle B_i, r\rangle - b_i] \right\},$$

$$\text{s.t.} \quad \left| \begin{array}{l} -\sum_{i=1}^{m} \lambda_i B_i - t \leqslant c^0, \\ -\sum_{i=1}^{m} \lambda_i a_{ij}(r) \leqslant c_j(r) \quad j \in J_+ \ , \\ -\sum_{i=1}^{m} \lambda_i a_{ij}(s) \leqslant c_j(s) \quad j \in J_- \ , \\ \lambda \geqslant 0, \;\; t \geqslant 0. \end{array} \right.$$

whose dual is

$$\langle c^0, r\rangle + \min \left\{ \sum_{j\in J_+} c_j(r)y_j + \sum_{j\in J_-} c_j(s)y_j + \langle c^0, z\rangle \right\},$$

$$\text{s.t.} \quad \sum_{j\in J_+} a_{ij}(r)y_j + \sum_{j\in J_-} a_{ij}(s)y_j + \langle B_i, r + z\rangle \leqslant b_i \quad i = 1, \ldots, m,$$

$$y \geqslant 0, \quad 0 \leqslant z \leqslant s - r.$$

By setting $x = r + z$, the latter problem becomes (6–7–8). □

COROLLARY 2. *Assume that:*

(#) For fixed λ each function $\langle A_j(x)\lambda + c_j(x)$ is quasiconcave.

Then the Lagrangian dual of (GPL) is the linear program

$$\langle c^0, r\rangle + \max\left\{\langle r-s, t\rangle + \sum_{i=1}^m \lambda_i[\langle B_i, r\rangle - b_i]\right\},$$

$$\text{s.t.}\quad \left|\begin{array}{l} -\sum_{i=1}^m \lambda_i B_i - t \leqslant c^0, \\ -\sum_{i=1}^m \lambda_i a_{ij}(v) \leqslant c_j(v) \quad v \in V,\ j=1,\dots,p, \\ \lambda \geqslant 0,\ t \geqslant 0, \end{array}\right.$$

where V denotes the vertex set of the hyperrectangle $[r,s]$.

Proof. Indeed, the condition $\min_{r\leqslant x\leqslant s}[\langle A_j(x), \lambda\rangle + c_j(x)] \geqslant 0$ is then equivalent to saying that $\min_{v\in V}[\langle A_j(v), \lambda\rangle + c_j(v)] \geqslant 0$. □

Thus, a Lagrangian bound for (GPL) is obtained by solving a convex program which reduces to a linear program if assumption (*) or (#) is satisfied. In the case when (*) is satisfied, this linear program is nothing but a LP relaxation of (GPL) and since a LP relaxation may not be the tightest convex relaxation, it is unlikely that this dual bound should be the best among all bounds obtained by convex relaxation. In the next section we will show a counter-example in the class of quadratic programming problems.

3. An Error About Lagrangian Bounds for Quadratic Programs

The best results on Lagrangian bounds for quadratic optimization under quadratic constraints have been obtained in [8–10]. An important feature that has been exploited very efficiently in these papers is that, by adding certain superfluous constraints (i.e. constraints implied by already existing constraints), one may drastically improve the Lagrangian bound and in some favorable cases even obtain the exact minimum. That is, the quality of a Lagrangian bound very much depends on the specific equivalent formulation used for the given problem when forming the Lagrangian dual. While a good reformulation may give tight dual bounds, a bad reformulation may lead to rather poor dual bounds. This can be illustrated by considering the dual bounds, developed in [11], which have been misleadingly claimed to be in a sense the best among all bounds obtained by convex relaxation.

By restricting our attention to the typical case of just one nonlinear constraint, consider as in [11] the quadratic programming problem:

$$\min\{\langle x, Qx\rangle + \langle q, x\rangle |\ Ax + d \leqslant 0,\ \langle x, Cx\rangle + \langle c, x\rangle + h \leqslant 0,\ a \leqslant x \leqslant b\}, \tag{Q}$$

where Q, C are $n\times n$ real matrices, $A \in \mathbb{R}^{m\times n}, q, c \in \mathbb{R}^n, d \in \mathbb{R}^m, h \in \mathbb{R}$, $a, b \in \mathbb{R}^n_+$ and $a \leqslant b$. Let Q_i, C_i be the i-th row of Q, C resp., $R = \{x \in \mathbb{R}^n | a \leqslant x \leqslant b\}$,

$$\alpha_i = \min\{\langle Q_i, x\rangle |\ x \in R\}, \quad \bar{\alpha}_i = \max\{\langle Q_i, x\rangle |\ x \in R\} - \alpha_i,$$
$$\beta_i = \min\{\langle C_i, x\rangle |\ x \in R\}, \quad \bar{\beta}_i = \max\{\langle C_i, x\rangle |\ x \in R\} - \beta_i.$$

With the help of two additional variables $y, z \in \mathbb{R}^n$ it can easily be shown that the problem (Q) can be rewritten equivalently as

$$\begin{aligned} \min \quad & \langle x, y\rangle + \langle q + \alpha, x\rangle, \\ \text{s.t.} \quad & Qx - y - \alpha \leqslant 0, \\ & Ax + d \leqslant 0, \\ & Cx - z - \beta \leqslant 0, \\ & \langle x, z\rangle + \langle \beta + c, x\rangle + h \leqslant 0, \\ & y \leqslant \bar{\alpha},\ z \leqslant \bar{\beta}, \\ & y \geqslant 0,\ z \geqslant 0,\ a \leqslant x \leqslant b. \end{aligned} \tag{P}$$

The following claim was made in [11]:

A1 (Proposition 2.2 in ([11]). *The Lagrangian dual of (P) is the dual of the linear program*

$$\begin{aligned} \min \quad & \langle a, y\rangle + \langle q + \alpha, x\rangle, \\ \text{s.t.} \quad & Qx - y - \alpha \leqslant 0, \\ & Ax + d \leqslant 0, \\ & Cx - z - \beta \leqslant 0, \\ & \langle a, z\rangle + \langle \beta + c, x\rangle + h \leqslant 0, \\ & y \leqslant \bar{\alpha}, \\ & z \leqslant \bar{\beta}, \\ & a \leqslant x \leqslant b, \quad y \geqslant 0, \quad z \geqslant 0. \end{aligned} \tag{LP}$$

A2 (Propositions 2.3 in [11]). *The Lagrangian bound for* (*P*) *furnished by the optimal value* $\beta(P)$ *of* (*LP*) *is at least as good as any bound obtained by convex relaxation of* (*P*).

Since it can easily be checked that (P) is a problem of the class (GPL), satisfying condition (*), assertion (A1) is a straightforward consequence of Corollary 1. As we saw, the proof of this fact, even for a more general problem, is very simple and does not need cumbersome computations as in [11]. On the other hand, assertion (A2) is wrong (so the proof of it in [11] is invalid), as should be expected from the discussion in the previous section and is clear from the following simple counter-example:[1]

Consider the quadratic program with one variable $x \in \mathbb{R}$:

[1]Example given by Nguyen thi Hoai Phuong.

$$\min\{x^2 - 3x| \ -x^2 + 4x \leqslant 0, \ 0 \leqslant x \leqslant 4\}. \tag{10}$$

whose optimal value is obviously 0. As can easily be verified, here $\alpha = 0$, $\bar{\alpha} = 4, \beta = -4, \bar{\beta} = 4$, so $\langle q + \alpha, a \rangle = 0$, and the associated problem (P) is

$$\begin{aligned} \min \quad & xy - 3x, \\ \text{s.t.} \quad & x - y \leqslant 0, \\ & -x - z + 4 \leqslant 0, \\ & xz \leqslant 0, \\ & 0 \leqslant x \leqslant 4, \ 0 \leqslant y \leqslant 4, \ 0 \leqslant z \leqslant 4. \end{aligned}$$

The Lagrangian dual of this problem is the dual of its LP relaxation and since the objective function of this LP relaxation is $0y - 12$, its optimal value is $\beta(P) = -12$, much inferior to the optimal value 0 of (Q). Furthermore, by replacing xy and xz with their convex envelopes, as in the method earlier proposed by F.A. Al-Khayyal et al. [1] (see also [15], p. 299), i.e.,

$$\max\{0, 4x + 4y - 16\}, \quad \max\{0, 4x + 4z - 16\},$$

we obtain a convex relaxation of (P), which reduces to the linear program

$$\begin{aligned} \min \quad & t - 3x, \\ \text{s.t.} \quad & 4x + 4y - 16 \leqslant t, \ x - y \leqslant 0, \\ & -x - z + 4 \leqslant 0, \ 4x + 4z - 16 \leqslant 0, \\ & 0 \leqslant x \leqslant 4, \ 0 \leqslant y \leqslant 4, \ 0 \leqslant z \leqslant 4, \ 0 \leqslant t, \end{aligned}$$

with optimal value -6, yielding a much better bound than $\beta(P) = -12$.

Furthermore, the Lagrangian dual of problem (10) itself is

$$\sup_{u \geqslant 0} \inf\{x^2 - 3x + (-x^2 + 4x)u| \ 0 \leqslant x \leqslant 4\}.$$

Since for $u = 1 : \inf\{x^2 - 3x + (-x^2 + 4x)u| \ 0 \leqslant x \leqslant 4\} = \inf\{x| \ 0 \leqslant x \leq 4\} = 0$, while the optimal value of (10) is 0, we have, by the weak duality theorem

$$\sup_{u \geqslant 0} \inf\{x^2 - 3x + (-x^2 + 4x)u| \ 0 \leqslant x \leqslant 4\} = 0.$$

Thus, in this example the duality gap is zero for the original quadratic problem (Q), while it is positive (and rather large) for the equivalent formulation (P). Also by adding the superfluous constraint $x(x-4) \leqslant 0$ (implied by $0 \leqslant x \leqslant 4$), we get another equivalent formulation of (10) with no duality gap:

$$\begin{aligned}&\sup_{u\in\mathbb{R}}\inf\{x^2-3x+(-x^2+4x)u|\ 0\leqslant x\leqslant 4\}\\&\quad=\min\{x^2-3x|\ -x^2+4x=0,\ 0\leqslant x\leqslant 4\}\\&\quad=\min\{x|\ 0\leqslant x\leqslant 4\}=0.\end{aligned}$$

This demonstrates that the bounds computed by the method in [11] are in general very poor actually, and even much inferior to the well known bounds earlier proposed in the literature for the same problem.

4. When are Lagrangian Bound Methods Guaranteed to Converge?

The last issue mentioned in the Introduction is when Lagrangian bounds can be incorporated into branch and bound procedures to produce successful convergent solution methods. In this section we discuss this issue for the general problem (GPC):

$$\inf\{F(x,y)|\ G_i(x,y)\leqslant 0\quad i=1,\dots,m,\quad x\in C,\ y\in D\},\tag{GPC}$$

where, as already stated in the Introduction, C is a nonempty compact convex subset of a (hyper) rectangle $X\subset\mathbb{R}^n_+, D$ is a nonempty closed convex subset of a (hyper) rectangle $Y\subset\mathbb{R}^p_+$, while $F(x,y):X\times Y\to\mathbb{R}$, and $G_i(x,y):X\times Y\to\mathbb{R},\ i=1,\dots,m,$ are lower semi-continuous functions, convex in y for every fixed x.

Define $G(x,y)=(G_1(x,y),\dots,G_m(x,y))$, and write $G(x,y)\leqslant 0$ to mean $G_i(x,y)\leqslant 0,\ i=1,\dots,m.$ Suppose a rectangular branch and bound algorithm is applied to solve (GPC) in which, following a general principle (see [15], Section 7.2.2), branching is performed upon the *nonconvex* variable x. For any rectangle M in the x-space let $\beta(M)$ be a lower bound computed for $\inf\{F(x,y)|\ G(x,y)\leqslant 0,\ x\in M\cap C, y\in D\}$, according to some chosen bounding rule. We will assume that the subdivision rule is exhaustive (see [15]), while the bounding rule satisfies the natural conditions: $\inf\{F(x,y)|\ x\in M\cap C, y\in D\}\leqslant\beta(M),\quad \beta(M')\geqslant\beta(M)$ whenever $M'\subset M$ (these conditions are obvious for Lagrangian bounds). At iteration k, if α_k is the objective function value of the best feasible solution known so far ($\alpha_k=\alpha>\sup\{F(x,y)|x\in C, y\in D\}$, if no feasible solution has been known), then all partition sets M with $\beta(M)\geqslant\min\{\alpha,\alpha_k\}$ are removed, and a partition set with smallest $\beta(M)$ among all remaining partition sets is selected for further subdivision. The algorithm terminates when no partition set remains for consideration: then an optimal solution is the current best feasible solution if $\alpha_k<\alpha$, or the problem is infeasible if $\alpha_k=\alpha$.

We say that the bounds are *eventually exact* if for any filter (infinite nested sequence of partition sets) $\{M_{k_\nu}\}$ collapsing to a single point x^* we have

$$\lim_{v\to+\infty} \beta(M_{k_v}) = \inf\{F(x^*,y)|\ G(x^*,y)\leqslant 0, y\in D\}. \tag{11}$$

THEOREM 2. *If the bounds are eventually exact then whenever the algorithm is infinite it generates a filter* $\{M_{k_v}\}$ *collapsing to a point* $x^* \in C$ *such that*

$$\inf\{F(x^*,y)|\ G(x^*,y)\leqslant 0,\ y\in D\} = \inf(GPC). \tag{12}$$

Any optimal solution y^* of the latter convex program then yields an optimal solution (x^*, y^*) of (GPC).

Proof. The existence of a filter $\{M_{k_v}\} \subset \{M_k\}$ follows from the general theory of branch and bound algorithms (see e.g. [15]). By exhaustiveness, such a filter collapses to a point x^*. Since every partition set M with $\beta(M) = +\infty$ is pruned, one must have $\beta(M_{k_v}) < +\infty$ and hence, $M_{k_v} \cap C \neq \emptyset\ \forall v$. The sets $M_{k_v} \cap C$ then form a nested sequence of nonempty compact sets, so by Cantor theorem, $\cap_{v=1}^{+\infty}(M_{k_v} \cap C) = (\cap_{v=1}^{+\infty}(M_{k_v})\cap\ C \neq \emptyset$. Therefore, $x^* \in C$. Since the sequence $\beta(M_{k_v})$ is nondecreasing, the limit $\beta^* = \lim_{v\to+\infty}\beta(M_{k_v})$ exists and $\beta^* \leqslant \alpha$. Let

$$\gamma := \inf\{F(x,y)|\ G(x,y)\leqslant 0,\ x\in C,\ y\in D\}. \tag{13}$$

Then, obviously, $\beta^* \leqslant \gamma$. But in view of (11),

$$\begin{aligned}\beta^* &= \inf\{F(x^*,y)|\ G(x^*,y)\leqslant 0,\ y\in D\}\\ &\geqslant \inf\{F(x,y)|\ G(x,y)\leqslant 0,\ x\in C,\ y\in D\}\\ &= \gamma,\end{aligned}$$

hence $\beta^* = \gamma$. □

It follows from this theorem that, if the bounds are eventually exact and the problem is infeasible then, with $\alpha < +\infty$, the branch and bound cannot be infinite, and infeasibility will be detected after finitely many steps. In other words, *an algorithm using eventually exact bounds is guaranteed to converge (in finitely or infinitely many steps) to a global optimal solution of (GPC), or else to detect infeasibility in finitely many steps*. The important convergence issue thus reduces to investigating conditions under which the Lagrangian bounds

$$\beta(M) = \sup_{\lambda\in\mathbb{R}^m_+} \inf\{F(x,y) + \langle\lambda, G(x,y)\rangle|\ x\in M\cap C, y\in D\} \tag{14}$$

are eventually exact.

THEOREM 3. *In* (*GPC*), *assume*, *in addition*, *that* D *is compact*. *Then the Lagrangian bounds are eventually exact*.

Proof. For $x \in C$ fixed the function $(y, \lambda) \mapsto F(x, y) + \langle \lambda, G(x, y) \rangle$ is convex, lower semi-continuous in y and linear in λ, so if D is compact then we have by the classical minimax equality (see e.g. [6] or [15]):

$$\min_{y \in D} \sup_{\lambda \in \mathbb{R}^m_+} \{F(x, y) + \langle \lambda, G(x, y) \rangle\} = \sup_{\lambda \in \mathbb{R}^m_+} \min_{y \in D} \{F(x, y) + \langle \lambda, G(x, y) \rangle\}. \tag{15}$$

Since clearly

$$\sup_{\lambda \in \mathbb{R}^m_+} \{F(x, y) + \langle \lambda, G(x, y) \rangle\} = \begin{cases} F(x, y) & \text{if } G(x, y) \leqslant 0, \\ +\infty & \text{otherwise}, \end{cases}$$

we have from (15), for every $x \in C$,

$$\min\{F(x, y) |\ G(x, y) \leqslant 0,\ y \in D\} = \sup_{\lambda \in \mathbb{R}^m_+} \min_{y \in D} \{F(x, y) + \langle \lambda, G(x, y) \rangle\}. \tag{16}$$

Consider now a filter $\{M_k\}$ collapsing to a point x^* (to simplify the notation we write M_k instead of M_{k_v}). As we saw in the proof of Theorem 2, $x^* \in C$, while by virtue of (16):

$$\min\{F(x^*, y) |\ G(x^*, y) \leqslant 0,\ y \in D\} = \sup_{\lambda \in \mathbb{R}^m_+} \min_{y \in D} \{F(x^*, y) + \langle \lambda, G(x^*, y) \rangle\}. \tag{17}$$

Let us show that this implies (11), i.e.,

$$\lim_{k \to \infty} \beta(M_k) = \min\{F(x^*, y) |\ G(x^*, y) \leqslant 0,\ y \in D\}. \tag{18}$$

From the obvious inequalities

$$\begin{aligned} \beta(M_k) \leqslant\ & \min\{F(x, y) |\ G(x, y) \leqslant 0,\ x \in M_k \cap C, y \in D\} \\ \leqslant\ & \min\{F(x^*, y) |\ G(x^*, y) \leqslant 0,\ y \in D\} \end{aligned}$$

it follows that

$$\beta(M_k) \nearrow \beta^* \leqslant \min\{F(x^*, y) |\ G(x^*, y) \leqslant 0,\ y \in D\}.$$

Arguing by contradiction, suppose that (18) does not hold, i.e.,

$$\min\{F(x^*, y) |\ G(x^*, y) \leqslant 0,\ y \in D\} > \beta^*. \tag{19}$$

Then by virtue of (16),

$$\sup_{\lambda \in \mathbb{R}^m_+} \min_{y \in D} \{F(x^*, y) + \langle \lambda, G(x^*, y) \rangle\} > \beta^*, \tag{20}$$

so there exists $\tilde{\lambda}$ satisfying

$$\min_{y \in D} \{F(x^*, y) + \langle \tilde{\lambda}, G(x^*, y) \rangle > \beta^*.$$

Using the lower semi-continuity of the function $(x, y) \mapsto \{F(x, y) + \langle\tilde{\lambda}, G(x, y)\rangle\}$ we can then find, for every fixed $y \in D$, an open ball U_y in $\mathbb{R}^n$ around x^* and an open ball V_y in $\mathbb{R}^p$ around y such that

$$F(x', y') + \langle\tilde{\lambda}, G(x', y')\rangle > \beta^* \quad \forall x' \in U_y \cap C, \forall y' \in V_y.$$

Since the balls $V_y, y \in D$, form a covering of the compact set D, there is a finite set $E \subset D$ such that the balls V_y, $y \in E$, still form a covering of D. If $U = \cap_{y \in E} U_y$ then for every $y \in D$ we have $y \in V_{y'}$ for some $y' \in E$, hence

$$F(x, y) + \langle\tilde{\lambda}, G(x, y)\rangle > \beta^* \quad \forall x \in U \cap C,\ \forall y \in D.$$

But $M_k \subset U$ for all sufficiently large k, because $\cap_k M_k = \{x^*\}$. Then the just established inequality implies that

$$\sup_{\lambda \in \mathbb{R}^m_+} \min\{F(x, y) + \langle\lambda, G(x, y)\rangle |\ x \in M_k \cap C, y \in D\} > \beta^*.$$

Hence, $\beta(M_k) > \beta^*$, a contradiction. This completes the proof. □

REMARK 1. Theorem 3 includes the results in [3] and [4] as special cases when D is a singleton and the constraints are linear.

REMARK 2. For the validity of Theorem 3 the lower semi-continuity of $F(x, y)$ is essential while, as was shown above, the condition (15) is implied by other assumptions (in particular the compactness of D), hence is superfluous. In the paper [12] this superfluous condition on zero duality gap at x^* is required, but the lower semi-continuity of $F(x, y)$ is replaced by the following weaker one:

(B) $F(x, y) > -\infty$ for $x \in C, y \in D$ and there exists an optimal solution whenever the feasible set of (GPC) is nonempty.

However, easily constructed counter examples show that the resulting theorems (Theorems 1 and 2 in [12]) are wrong (for details see [17]). In fact, the proofs of these theorems, which use only condition **(B)**, contain such obviously incorrect arguments as claiming that two disjoint closed convex sets T_1, T_2 in $\mathbb{R}^m$ can be separated by an open convex set $\Omega \supset T_2$ such that $\Omega \cap T_1 = \emptyset$.

The assumption on the compactness of D is too restrictive for many applications. For instance, this assumption is not satisfied for the problem investigated in [2] (so even if the results in [12] were correct, they could not be legitimately applied to the problem investigated in [2], as was done in [12]). In the next theorem the compactness of D is replaced by a weaker condition, at the expense, however, of requiring the continuity, and not merely the lower semi-continuity, of the functions $F(x, y)$ and $G(x, y)$.

THEOREM 4. *In* (GPC) *assume that the functions* $F(x,y), G(x,y)$ *are continuous on* $X \times Y$ *and in addition, that*

(S) *For every* $x \in C$ *there is* $\lambda \in \mathbb{R}^m_+$ *such that* $F(x,y) + \langle \lambda, G(x,y) \rangle \to +\infty$ *as* $y \in D$, $\|y\| \to +\infty$.

Then the Lagrangian bounds are eventually exact.

Proof. In view of Theorem 3 we may assume D unbounded. Consider a filter $\{M_k\}$ collapsing to a point x^*. As we saw previously, $x^* \in C$, and, as $k \to +\infty$,

$$\beta(M_k) \nearrow \beta^* \leq \overline{\beta} := \inf\{F(x^*, y) | G(x^*, y) \leq 0, y \in D\}.$$

According to (S) there exists $\tilde{\lambda} \in \mathbb{R}^m_+$ satisfying

$$F(x^*, y) + \langle \tilde{\lambda}, G(x^*, y) \rangle \to +\infty \text{ as } y \in D, \|y\| \to +\infty. \tag{21}$$

By virtue of a known minimax theorem ([6], Chapter VI, Proposition 2.3; see also [14]), this ensures that

$$\tilde{\beta} = \min_{y \in D} \sup_{\lambda \in \mathbb{R}^m_+} [F(x^*, y) + \langle \lambda, G(x^*, y) \rangle] = \sup_{\lambda \in \mathbb{R}^m_+} \inf_{y \in D} [F(x^*, y) + \langle \lambda, G(x^*, y) \rangle]. \tag{22}$$

We show that $\beta^* = \overline{\beta}$. Suppose the contrary, i.e.,

$$\beta^* < \overline{\beta}. \tag{23}$$

Obviously $\theta\lambda + (1-\theta)\tilde{\lambda} \in \mathbb{R}^m_+$ for every (λ, θ) with $\lambda \in \mathbb{R}^m_+, 0 \leq \theta < 1$. Let

$$\Lambda := \{\lambda \in \mathbb{R}^m_+ | \inf_{y \in D} [F(x^*, y) + \langle \lambda, G(x^*, y) \rangle] > -\infty\}. \tag{24}$$

For every k, since $\beta(M_k) \leq \beta^*$, we have

$$\forall (\lambda, \theta) \in \Lambda \times [0,1) \inf_{x \in C \cap M_k, y \in D} [F(x,y) + \langle \theta\lambda + (1-\theta)\tilde{\lambda}, G(x,y) \rangle] \leq \beta^*.$$

Hence, if ϵ denotes an arbitrary number satisfying $0 < \epsilon < \overline{\beta} - \beta^*$, then for every $(\lambda, \theta) \in \Lambda \times [0,1)$ and every k there exists $x^{(k,\lambda,\theta)} \in C \cap M_k$, $y^{(k,\lambda,\theta)} \in D$ satisfying

$$F(x^{(k,\lambda,\theta)}, y^{(k,\lambda,\theta)}) + \langle \theta\lambda + (1-\theta)\tilde{\lambda}, G(x^{(k,\lambda,\theta)}, y^{(k,\lambda,\theta)}) \rangle \leq \beta^* + \epsilon. \tag{25}$$

We contend that not for every $(\lambda, \theta) \in \Lambda \times [0,1)$ the sequence $\{y^{(k,\lambda,\theta)}, k = 1, 2, \ldots\}$ is bounded. Indeed, otherwise we could assume that, for every $(\lambda, \theta) \in \Lambda \times [0,1)$: $y^{(k,\lambda,\theta)} \to \overline{y}^{(\lambda,\theta)} \in D$, as $k \to +\infty$. Since $x^{(k,\lambda,\theta)} \in M_k$ while $\cap_{k=1}^{+\infty} M_k = \{x^*\}$, we would have $x^{(k,\lambda,\theta)} \to x^*$. Then letting $k \to +\infty$ in (25) would yield

$$F(x^*, \overline{y}^{(\lambda,\theta)}) + \langle \theta\lambda + (1-\theta)\tilde{\lambda}, G(x^*, \overline{y}^{(\lambda,\theta)}) \rangle \leq \beta^* + \epsilon,$$

for every $(\lambda, \theta) \in \Lambda \times [0,1)$, hence

$$\sup_{\lambda\in\Lambda,0\le\theta<1}\inf_{y\in D}[F(x^*,y)+\langle\theta\lambda+(1-\theta)\tilde{\lambda},G(x^*,y)\rangle]\le\beta^*+\epsilon<\overline{\beta}.$$

But for every $\lambda\in\Lambda$ the concave function $\varphi(\theta)=\inf_{y\in D}[F(x^*,y)+\langle\theta\lambda+(1-\theta)\tilde{\lambda},G(x^*,y)\rangle]$ satisfies $\sup_{0\le\theta<1}\varphi(\theta)\ge\varphi(1)$, i.e.,

$$\sup_{0\le\theta<1}\inf_{y\in D}[F(x^*,y)+\langle\theta\lambda+(1-\theta)\tilde{\lambda},G(x^*,y)\rangle]\ge\inf_{y\in D}[F(x^*,y)+\langle\lambda,G(x^*,y)\rangle].$$

Therefore, we would have

$$\begin{aligned}&sup_{\lambda\in\mathbb{R}^+_m}\inf_{y\in D}[F(x^*,y)+\langle\lambda,G(x^*,y)\rangle]\\&\quad=\sup_{\lambda\in\Lambda}\inf_{y\in D}[F(x^*,y)+\langle\lambda,G(x^*,y)\rangle]\quad\text{(because of (24))}\\&\quad=\sup_{\lambda\in\Lambda,0\le\theta<1}\inf_{y\in D}[F(x^*,y)+\langle\theta\lambda+(1-\theta)\tilde{\lambda},G(x^*,y)\rangle]<\overline{\beta},\end{aligned}$$

contradicting (22). Thus, there exists $(\bar{\lambda},\bar{\theta})\in\Lambda\times[0,1)$ such that, by passing to a subsequence if necessary, $\|y^{(k,\bar{\lambda},\bar{\theta})}\|\to+\infty(k\to+\infty)$. Now, for an arbitrary point $y^*\in D$, let $\rho:=F(x^*,y^*)+\langle\bar{\theta}\bar{\lambda}+(1-\bar{\theta})\tilde{\lambda},G(x^*,Y^*)\rangle$. Then there is k_0 such that for all $k\ge k_0$,

$$F(x^{(k,\bar{\lambda},\bar{\theta})},y^*)+\langle\bar{\theta}\bar{\lambda}+(1-\bar{\theta})\tilde{\lambda},G(x^{(k,\bar{\lambda},\bar{\theta})},y^*)\rangle<\rho+\epsilon.\tag{26}$$

For simplicity of notation, let us write $\bar{x}^k,\bar{y}^k$ for $x^{(k,\bar{\lambda},\bar{\theta})},y^{(k,\bar{\lambda},\bar{\theta})}$. For any $l>\|y^*\|$ define

$$\bar{y}^{kl}=\frac{l}{\|\bar{y}^k\|}\bar{y}^k+(1-\frac{l}{\|\bar{y}^k\|}y^*).$$

Clearly, there is $k_1\ge k_0$ such that for all $k\ge k_1$,

$$0<l/\|\bar{y}^k\|<1,\quad l-\|y^*\|\le\|\bar{y}^{kl}\|\le l+\|y^*\|.$$

From the inequalities (25) and (26) and the convexity of the function $y\to F(\bar{x}^k,y)+\langle\bar{\theta}\bar{\lambda}+(1-\bar{\theta})\tilde{\lambda},G(\bar{x}^k,y)\rangle$, we then deduce

$$F(\bar{x}^k,\bar{y}^{kl})+\langle\bar{\theta}\bar{\lambda}+(1-\bar{\theta})\tilde{\lambda},G(\bar{x}^k,\bar{y}^{kl})\rangle\le\max\{\beta^*,\rho\}+\varepsilon<+\infty.$$

Since $\{\bar{y}^{kl}\}$ is bounded, we can assume $\bar{y}^{kl}\to u^l\in D$ as $k\to+\infty$. This yields

$$\begin{aligned}&\bar{\theta}[F(x^*,u^l)+\langle\bar{\lambda},G(x^*,u^l)\rangle]+(1-\bar{\theta})[F(x^*,u^l)+\langle\tilde{\lambda},G(x^*,u^l)\rangle]\\&\qquad\le\max\{\beta^*,\rho\}+\varepsilon,\end{aligned}$$

where clearly $\|u^l\|\ge l-\|y^*\|\to+\infty$ as $l\to+\infty$. Since $\bar{\lambda}\in\Lambda$, by letting $l\to+\infty$ in the above inequality and taking account of (21) and (24), we get $+\infty\le\max\{\beta^*,\rho\}+\varepsilon$, which is a contradiction. Therefore $\beta^*=\bar{\beta}$, as was to be proved. □

REMARK 3. Actually it suffices to require conditions (S) for the point $x^* \in C$ such that $\{x^*\} = \cap_k M_k$, where $\{M_k\}$ is a filter of partition sets generated when solving (GPC) by the above described algorithm.

5. Applications

To illustrate the usefulness of the above results we just consider two examples of applications.

I. *The Pooling and Blending Problem*. It was shown in [2] that this problem from petrochemical industry can be given the form

$$\min\{c^T y|\ A(x)y \leqslant b,\ y \geqslant 0,\ x \in X\},$$

where X is a hyperrectangle in $\mathbb{R}^n$ and $A(x)$ is an $m \times p$ matrix whose elements $a_{ij}(x)$ are continuous functions of x. Since this is a special case of (GPL) satisfying condition (#) of Corollary 2, the Lagrangian bound can be computed by solving a linear program. Furthermore, for this (GPL) condition (S) in Theorem 4 now reads

$$(\forall x \in X) \quad (\exists \lambda \in \mathbb{R}^m_+) \quad c^T y + \langle \lambda, A(x)y - b\rangle \to +\infty \ as\ y \to +\infty,$$

where the last condition holds if and only if $\langle A(x), \lambda\rangle + c > 0$. Therefore, if $(\forall x \in X)$ $(\exists \lambda \in \mathbb{R}^m_+)$ $\langle Ax, \lambda\rangle + c > 0$, then the branch and bound algorithm using Lagrangian bounds converges. Thus, the above results, much more general than those in [2], are obtained in a much shorter way. Another short proof of these results can also be found in [15].

II. *The Bilinear Matrix Inequalities Problem*. The general bilinear matrix inequalities (BMI) problem in control theory can be formulated as (see e.g. [13]):

$$\min \langle c, x\rangle + \langle d, y\rangle \tag{27}$$

$$s.t. \quad G_0 + \sum_{j=1}^m y_j G_j \preceq 0 \tag{28}$$

$$L_0 + \sum_{i=1}^n x_i L_{i0} + \sum_{j=1}^m y_j L_{0j} + \sum_{i=1}^n \sum_{j=1}^m x_j y_j L_{ij} \prec 0 \tag{29}$$

$$x \in X = [p, q] \subset \mathbb{R}^n,\ y \in \mathbb{R}^m_+ \tag{30}$$

where x, y are the decision variables, $G_0, G_j, L_0, L_{0i}, L_{j0}, L_{ij}$ are symmetric matrices of appropriate sizes, and the notation $G \preceq 0, L \prec 0$ means that G is a semidefinite negative matrix, L is a definite negative matrix.

For ease of notation we write

$$\begin{bmatrix} A \\ B \end{bmatrix}_d \quad \text{for} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

and define

$$A_{00}(x) = \begin{bmatrix} G_0 \\ L_0 + \sum_{i=1}^n x_i L_{i0} \\ \langle x, c \rangle \end{bmatrix}_d, \; A_{j0}(x) = \begin{bmatrix} G_j \\ L_{0j} + \sum_{i=1}^n x_i L_{ij} \\ d_j \end{bmatrix}_d, \; Q_{00} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_d.$$

Then, as was shown in [13], this problem can be converted to the form

$$\min\{t|\ A_0(x,p,q) + \sum_{j=1}^m y_j A_j(x,p,q) \preceq tQ,\ y \geqslant 0,\ x \in X\},$$

where

$$A_j(x,p,q) = \begin{bmatrix} A_{j0}(x) \\ A_{j1}(x,p,q) \end{bmatrix}_d,\ Q = \begin{bmatrix} Q_{00} \\ Q_{01} \end{bmatrix}_d,\ Q_{01} = 0,$$

$$A_{j1}(x,p,q) = \begin{bmatrix} (x_1 - p_1)G_j \\ (q_1 - x_1)G_j \\ \cdots \\ (x_n - p_n)G_j \\ (q_n - x_n)G_j \end{bmatrix}_d \quad j = 0, 1, \ldots, n.$$

In this form the (BMI) problem appears to be a problem (GPL). Condition (S) in Theorem 4 can be formulated as

$$(\forall x \in X)(\exists Z_1 \succeq 0) \qquad \mathrm{Tr}(Z_1 Q_{00}) = 1, \quad \mathrm{Tr}(Z_1 A_{j0}(x) > 0 \quad j = 1, \ldots, m.$$

Therefore, by Theorem 4, under this assumption the BMI problem can be solved by a convergent branch and bound algorithm using Lagrangian bounds, as proposed in [13]. Note that the Lagrangian dual of the problem

$$\max_{Z \succeq 0} \min_{t \in \mathbb{R}, y \geqslant 0, x \in M} \left\{ t + \mathrm{Tr}\left[Z(A_0(x,p,q) + \sum_{j=1}^m y_j A_j(x,p,q) - tQ) \right] \right\}$$

has been shown in [13] to be equivalent to the LMI program

$$\max\{t|\ \mathrm{Tr}(ZA_0(x,p,q) \geqslant t,\ \mathrm{Tr}(ZA_j(x,p,q)) \geqslant 0\ \forall x \in \mathrm{vert}X,\ j = 1, \ldots, m,$$
$$\mathrm{Tr}(ZQ) = 1,\ Z \succeq 0\}$$

where vert$X$ denotes the vertex set of X.

References

1. Al-Khayyal, F.A., Larsen, C. and Van Voorhis, T. (1995), A relaxation method for nonconvex quadratically constrained quadratic programs. *Journal of Global Optimization* 6, 215–230.
2. Ben-Tal, A. et al. (1994), Global minimization by reducing the duality gap. *Mathematical Programming* 63, 193–212.
3. Dür, M. (2001), Dual bounding procedures lead to convergent branch-and-bound algorithms. *Mathematical Programming* Series A 91, 117–2001.
4. Falk, J.E. (1969), Lagrange multipliers and nonconvex programs. *SIAM Journal Control* 7, 534–545.
5. Falk, J.E. and Soland, R.M. (1969), An algorithm for separable nonconvex programming problems. *Management Science* 15, 550–569.
6. Ekeland, I. and Temam, R. (1976), *Convex analysis and variational problems*, North-Holland, American Elsevier, Amsterdam, New York.
7. Floudas, C. (1999), *Deterministic Global Optimization*, Kluwer.
8. Shor, N.Z. (1998), *Nondifferentiable Optimization and Polynomial Problems*, Kluwer.
9. Shor, N.Z. and Stetsenko, P.I. (1989), *Quadratic extremal problems and nondifferentiable optimization*, Naukova Dumka, Kiev, (in Russian).
10. Shor, N.Z. and Stetsyuk, P.I. (2002), Lagrangian bounds in multiextremal polynomial and discrete optimization problems. *Journal of Global Optimization* 23, 1–41.
11. Thoai, N.V. (2000), Duality bound method for the general quadratic programming problem with quadratic constraints. *Journal of Optimization Theory and Applications* 107, 331–354
12. Thoai, N.V. (2002), Convergence of duality bound method in partly convex programming. *Journal of Global Optimization* 22, 263–270.
13. Tuan, H.D., Apkarian, P. and Nakashima, Y. (2000), A new Lagrangian dual global optimization algorithm for solving bilinear matrix inequalities. *International Journal Robust Nonlinear Control*, 10, 561–578.
14. Tuy, H. (2004), Minimax theorems revisited, preprint, Institute of Mathematics, Hanoi.
15. Tuy, H. (1998), *Convex Analysis and Global Optimization*, Kluwer.
16. Tuy, H. (2000), On Some recent advances and applications of D.C. optimization. In Nguyen, V.H., Strodiot, J.J. and Tossings, P. (eds.) *Optimization, Lecture Notes in Economics and Mathematical Systems*. Vol. 481, Springer, pp. 473–497.
17. Tuy, H. (2002), *Wrong results in D.C. Optimization*, Preprint, Institute of Mathematics, Hanoi.