

## SOLUTIONS AND OPTIMALITY CRITERIA TO BOX CONSTRAINED NONCONVEX MINIMIZATION PROBLEMS

DAVID YANG GAO

Department of Mathematics & Grado  
Department of Industrial and System Engineering  
Virginia Polytechnic Institute and State University  
Blacksburg, VA, 24061, USA

(Communicated by K.L. Teo)

**ABSTRACT.** This paper presents a canonical duality theory for solving nonconvex polynomial programming problems subjected to box constraints. It is proved that under certain conditions, the constrained nonconvex problems can be converted to the so-called canonical (perfect) dual problems, which can be solved by deterministic methods. Both global and local extrema of the primal problems can be identified by a triality theory proposed by the author. Applications to nonconvex integer programming and Boolean least squares problems are discussed. Examples are illustrated. A conjecture on NP-hard problems is proposed.

**1. Primal problem and its dual form.** The box constrained nonconvex minimization problem is proposed as a primal problem ( $\mathcal{P}$ ) given below:

$$(\mathcal{P}) : \min_{\mathbf{x} \in \mathcal{X}_a} \{P(\mathbf{x}) = Q(\mathbf{x}) + W(\mathbf{x})\} \quad (1)$$

where  $\mathcal{X}_a = \{\mathbf{x} \in \mathcal{R}^n \mid \ell^l \leq \mathbf{x} \leq \ell^u\}$  is a feasible space,  $Q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A \mathbf{x} - \mathbf{c}^T \mathbf{x}$  is a quadratic function,  $A = A^T \in \mathcal{R}^{n \times n}$  is a given symmetric matrix,  $\ell^l$ ,  $\ell^u$ , and  $\mathbf{c}$  are three given vectors in  $\mathcal{R}^n$ ,  $W(\mathbf{x})$  is a nonconvex function. In this paper, we simply assume that  $W(\mathbf{x})$  is a so-called double-well fourth order polynomial function defined by

$$W(\mathbf{x}) = \frac{1}{2} \left( \frac{1}{2} |B\mathbf{x}|^2 - \alpha \right)^2, \quad (2)$$

where  $B \in \mathcal{R}^{m \times n}$  is a given matrix and  $\alpha > 0$  is a given parameter. The notation  $|\mathbf{x}|$  used in this paper denotes the Euclidean norm of  $\mathbf{x}$ .

Problems of the form (1) appear frequently in many applications, such as semi-linear nonconvex partial differential equations [15], structural limit analysis, discretized optimal control problems with distributed parameters, information theory, and network communication. Particularly, if  $W(\mathbf{x}) = 0$ , the problem ( $\mathcal{P}$ ) is directly related to certain successive quadratic programming methods ([9, 10, 18]).

---

2000 *Mathematics Subject Classification.* 49N15, 49M37, 90C26, 90C20.

*Key words and phrases.* Global optimization; Duality; Nonconvex minimization; Box constraints; Integer programming; Boolean least squares problem, NP-hard problems.

The main result of this paper has been announced at the INFORMS Annual Meeting, Nov. 12-15, 2005, San Francisco, CA., and the International Conference on Nonlinear Programming with Applications, May 29- June 1, 2006, Shanghai, China.

Moreover, if  $\ell^l = \{0\}$  and  $\ell^u = \{1\}$ , the problem leads to one of the fundamental problems in combinatorial optimization, namely, the integer programming problem [6]. Due to the nonconvexity of the cost function and the inequality constraints, traditional KKT theory and direct methods can only be used for solving the problem ( $\mathcal{P}$ ) to local optimality (cf. [23, 25, 26]). It was shown (see Murty and Kabadi (1987) [28] and Pardalos and Schnitger (1988) [29]) that if a point  $\bar{\mathbf{x}}$  is degenerate (i.e., there is a component  $\bar{x}_i$  of  $\bar{\mathbf{x}}$  such that it lies at either the upper or lower bound and also has the gradient  $\nabla_{x_i} P(\mathbf{x}) = 0$ ), then even determining whether it is a local solution to a constrained quadratic programming problem ( $W(\mathbf{x}) = 0$ ) is an NP-complete problem. Therefore, necessary and sufficient conditions for global optimality are fundamentally important in nonconvex minimization problems. Much effort and progress have been made on solving box constrained nonconvex minimization problems during the last twenty years (see, for example, Al-Khayyal and Falk (1983), Sherali and Alameddine (1992), Akrotirianakis and Floudas (2004), Floudas (2000), and Floudas *et al* (2004, 2005) and much more).

*Canonical duality theory* is a newly developed, potentially useful methodology [12, 13]. This theory is composed mainly of a *canonical dual transformation* and a *trinality theory*. The canonical dual transformation can be used to formulate perfect dual problems with zero duality gap, while the triality theory provides sufficient conditions for identifying global and local optimizers. The canonical duality theory has been used successfully for solving many difficult global optimization problems including polynomial minimization [14, 19], nonconvex quadratic minimization with spherical, quadratical, box, and integer constraints (see [6, 16, 17, 18]). Detailed discussion on canonical duality theory and its extensive applications can be found in the recent review articles [15, 21].

In this paper, we shall demonstrate applications of the canonical duality theory by solving the general nonconvex problem given in (1) and its special cases. As indicated in [14], the key step of this canonical dual transformation is to introduce a geometrical operator such that both the nonconvex function  $W(\mathbf{x})$  and the constraints can be written in the so-called canonical form [12]. In order to do so, we assume without loss of generality that  $\ell^u = -\ell^l = \ell^{\frac{1}{2}} = \{\sqrt{\ell_i}\}$  (if necessary, a simple linear transformation can be used to convert the problem to this form). Thus, the geometrical operator  $\Lambda : \mathcal{R}^n \rightarrow \mathcal{R}^{1+n}$  and the associated canonical function  $U$  can be introduced as the following:

$$\mathbf{y} = \Lambda(\mathbf{x}) = \begin{pmatrix} \xi(\mathbf{x}) \\ \epsilon(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}|B\mathbf{x}|^2 - \alpha \\ \{x_i^2 - \ell_i\} \end{pmatrix} \in \mathcal{R}^{1+n},$$

$$U(\mathbf{y}) = \frac{1}{2}\xi^2 + \Psi(\epsilon) \quad (3)$$

where

$$\Psi(\epsilon) = \begin{cases} 0 & \text{if } \epsilon \leq 0, \\ +\infty & \text{otherwise.} \end{cases} \quad (4)$$

Clearly, the canonical function  $U(\mathbf{y})$  is convex and its effective domain is

$$\mathcal{Y} = \left\{ \mathbf{y} = \begin{pmatrix} \xi \\ \epsilon \end{pmatrix} \in \mathcal{R}^{1+n} \mid \xi \geq -\alpha, \epsilon \leq 0 \in \mathcal{R}^n \right\}.$$

Thus, the primal problem (1) can be written in the following unconstrained canonical form

$$(\mathcal{P}) : \min\{P(\mathbf{x}) = U(\Lambda(\mathbf{x})) + Q(\mathbf{x}) : \mathbf{x} \in \mathcal{R}^n\}. \quad (5)$$

Let  $\mathbf{y}^* = \begin{pmatrix} \varsigma \\ \boldsymbol{\sigma} \end{pmatrix} \in \mathcal{R}^{1+n}$  be a dual variable of  $\mathbf{y}$ . The sup-Fenchel conjugate of  $U(\mathbf{y})$  can be defined by

$$\begin{aligned} U^\sharp(\mathbf{y}^*) &= \sup_{\mathbf{y} \in \mathcal{R}^{1+n}} \{\langle \mathbf{y}^T \mathbf{y}^* - U(\mathbf{y}) \rangle\} \\ &= \sup_{\xi \geq -\alpha} \{\xi \varsigma - \frac{1}{2} \xi^2\} + \sup_{\boldsymbol{\epsilon} \in \mathcal{R}^n} \{\boldsymbol{\epsilon}^T \boldsymbol{\sigma} - \Psi(\boldsymbol{\epsilon})\} \\ &= \frac{1}{2} \varsigma^2 + \Psi^\sharp(\boldsymbol{\sigma}) \quad \text{s.t. } \varsigma \geq -\alpha, \end{aligned} \quad (6)$$

where  $\Psi^\sharp(\boldsymbol{\sigma}) = \{0 \text{ (if } \boldsymbol{\sigma} \geq 0), +\infty \text{ (if } \boldsymbol{\sigma} < 0)\}$  is the sup-Fenchel conjugate of  $\Psi(\boldsymbol{\epsilon})$ . Dually, the effective domain of  $U^\sharp$  is

$$\mathcal{Y}^* = \left\{ \mathbf{y}^* = \begin{pmatrix} \varsigma \\ \boldsymbol{\sigma} \end{pmatrix} \in \mathcal{R}^{1+n} \mid \varsigma \geq -\alpha, \boldsymbol{\sigma} \geq 0 \in \mathcal{R}^n \right\}.$$

From the theory of convex analysis, the following *extended canonical duality relations* (see [12]) hold:

$$\mathbf{y}^* \in \partial U(\mathbf{y}) \Leftrightarrow \mathbf{y} \in \partial U^\sharp(\mathbf{y}^*) \Leftrightarrow \mathbf{y}^T \mathbf{y}^* = U(\mathbf{y}) + U^\sharp(\mathbf{y}^*). \quad (7)$$

By the canonical dual transformation developed in [13, 14], the canonical dual function of  $P(\mathbf{x})$  is defined by

$$P^d(\mathbf{y}^*) = Q^\Lambda(\mathbf{y}^*) - U^\sharp(\mathbf{y}^*),$$

where  $Q^\Lambda(\mathbf{y}^*)$  is the so-called  $\Lambda$ -canonical dual transformation (see [12]) defined by

$$\begin{aligned} Q^\Lambda(\mathbf{y}^*) &= \text{sta} \{(\Lambda(\mathbf{x}))^T \mathbf{y}^* + Q(\mathbf{x}) : \mathbf{x} \in \mathcal{R}^n\} \\ &= -\frac{1}{2} \mathbf{c}^T [G(\varsigma, \boldsymbol{\sigma})]^{-1} \mathbf{c} - \alpha \varsigma - \boldsymbol{\ell}^T \boldsymbol{\sigma}, \end{aligned} \quad (8)$$

where the notation  $\text{sta} \{ * : \mathbf{x} \in \mathcal{R}^n \}$  represents finding the stationary point of the statement  $\{ * \}$  with respect to  $\mathbf{x} \in \mathcal{R}^n$ ,  $G(\varsigma, \boldsymbol{\sigma})$  is a symmetrical matrix, defined by

$$G(\varsigma, \boldsymbol{\sigma}) = A + \varsigma B^T B + 2 \text{Diag}(\boldsymbol{\sigma}) \in \mathcal{R}^{n \times n}, \quad (9)$$

and  $\text{Diag}(\boldsymbol{\sigma}) \in \mathcal{R}^{n \times n}$  denotes a diagonal matrix with  $\{\sigma_i\}$  ( $i = 1, 2, \dots, n$ ) as its diagonal entries.

Let  $\mathcal{S}_a \subset \mathcal{Y}^*$  denote the dual feasible space:

$$\mathcal{S}_a = \left\{ \mathbf{y}^* = \begin{pmatrix} \varsigma \\ \boldsymbol{\sigma} \end{pmatrix} \in \mathcal{R}^{1+n} \mid \varsigma \geq -\alpha, \boldsymbol{\sigma} \geq 0, \det G(\varsigma, \boldsymbol{\sigma}) \neq 0 \right\}. \quad (10)$$

The canonical dual problem of  $(P)$  can be proposed as the following

$$(P^d) : \text{sta} \left\{ P^d(\varsigma, \boldsymbol{\sigma}) = -\frac{1}{2} \mathbf{c}^T [G(\varsigma, \boldsymbol{\sigma})]^{-1} \mathbf{c} - \frac{1}{2} \varsigma^2 - \alpha \varsigma - \boldsymbol{\ell}^T \boldsymbol{\sigma} : (\varsigma, \boldsymbol{\sigma})^T \in \mathcal{S}_a \right\}. \quad (11)$$

**Theorem 1** (Canonical Duality Theorem). *Problem  $(P^d)$  is canonically dual to the primal problem  $(P)$  in the sense that if  $\bar{\mathbf{y}}^* = (\bar{\varsigma}, \bar{\boldsymbol{\sigma}})^T$  is a KKT point of  $(P^d)$ , then the vector defined by*

$$\bar{\mathbf{x}} = [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c} \quad (12)$$

*is a KKT point of  $(P)$ , and*

$$P(\bar{\mathbf{x}}) = P^d(\bar{\boldsymbol{\lambda}}). \quad (13)$$

*Proof.* Suppose that  $\bar{\mathbf{y}}^* = (\bar{\varsigma}, \bar{\boldsymbol{\sigma}})^T$  is a KKT point of  $(\mathcal{P}^d)$ , then the criticality condition  $\delta_{\varsigma} P^d(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) = 0$  (where  $\delta_{\varsigma} P^d$  stands for Gâteaux derivative of  $P^d$  with respect to  $\varsigma$ ) leads to

$$\delta_{\varsigma} P^d(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) = \frac{1}{2} \mathbf{c}^T [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} B^T B [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c} - \alpha - \bar{\varsigma} = 0. \tag{14}$$

In terms of  $\bar{\mathbf{x}} = [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c}$ , the equation (14) is the canonical dual equation  $\bar{\varsigma} = \frac{1}{2} |B\bar{\mathbf{x}}|^2 - \alpha = \bar{\xi}$  and the condition  $\frac{1}{2} |B\bar{\mathbf{x}}|^2 \geq 0$  implies  $\bar{\varsigma} \geq -\alpha$ . On the other hand, the criticality condition for  $\delta_{\boldsymbol{\sigma}} P^d(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) \leq 0$  leads to the KKT condition

$$\begin{aligned} \bar{x}_i^2(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) - \ell_i &\leq 0, \quad \bar{\sigma}_i \geq 0, \\ \bar{\sigma}_i(\bar{x}_i^2(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) - \ell_i) &= 0, \quad i = 1, \dots, n \end{aligned} \tag{15}$$

where  $\bar{x}_i(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})$  is the  $i$ -th component of the vector  $\bar{\mathbf{x}}(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) = [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c}$ . This shows that  $\bar{\mathbf{x}}(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})$  is also a KKT point of the primal problem  $(\mathcal{P})$ . By the complementarity condition in (15), we have  $\bar{\sigma}_i \ell_i = \bar{x}_i^2 \bar{\sigma}_i$ ,  $i = 1, 2, \dots, n$ . Thus, in term of  $\bar{\mathbf{x}} = [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c}$  and the Fenchel-Young equality  $\frac{1}{2} \bar{\varsigma}^2 + \alpha \bar{\varsigma} = \frac{1}{2} |B\bar{\mathbf{x}}|^2 \bar{\varsigma} - \frac{1}{2} (\frac{1}{2} |B\bar{\mathbf{x}}|^2 - \alpha)^2$ , we have

$$\begin{aligned} P^d(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) &= \frac{1}{2} \bar{\mathbf{x}}^T G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}}) \bar{\mathbf{x}} - \mathbf{c}^T \bar{\mathbf{x}} - \sum_{i=1}^n \bar{x}_i^2 \bar{\sigma}_i - \frac{1}{2} |B\bar{\mathbf{x}}|^2 \bar{\varsigma} + \frac{1}{2} \left( \frac{1}{2} |B\bar{\mathbf{x}}|^2 - \alpha \right)^2 \\ &= \frac{1}{2} \bar{\mathbf{x}}^T A \bar{\mathbf{x}} - \bar{\mathbf{x}}^T \mathbf{c} + \frac{1}{2} \left( \frac{1}{2} |B\bar{\mathbf{x}}|^2 - \alpha \right)^2 = P(\bar{\mathbf{x}}). \end{aligned}$$

This proves the theorem. □

Theorem 1 shows that there is no duality gap between the primal problem  $(\mathcal{P})$  and its canonical dual  $(\mathcal{P}^d)$ . In the dual feasible space  $\mathcal{S}_a$ , although it is required that  $\det G(\varsigma, \boldsymbol{\sigma}) \neq 0$ , which is essentially not a constraint since in the case that the matrix  $G(\varsigma, \boldsymbol{\sigma})$  is singular, the vector  $\mathbf{x} = [G(\varsigma, \boldsymbol{\sigma})]^{-1} \mathbf{c} = 0 \in \mathcal{R}^n$  is a trivial KKT point of the primal problem.

It is known that the KKT conditions are only necessary for local minimizers of the nonconvex quadratic programming problem  $(\mathcal{P})$ . The next section will show that the global and local extrema of the primal problem  $(\mathcal{P})$  depend on the canonical dual solutions.

**2. Global and local optimality criteria.** In order to identify global and local extrema among the KKT points of the nonconvex problem  $(\mathcal{P})$ , we need to introduce some useful feasible spaces:

$$\mathcal{S}_a^+ = \{(\varsigma, \boldsymbol{\sigma})^T \in \mathcal{S}_a \mid G(\varsigma, \boldsymbol{\sigma}) \text{ is positive definite}\}, \tag{16}$$

$$\mathcal{S}_a^- = \{(\varsigma, \boldsymbol{\sigma})^T \in \mathcal{S}_a \mid G(\varsigma, \boldsymbol{\sigma}) \text{ is negative definite}\}. \tag{17}$$

**Theorem 2 (Triality Theorem).** *Suppose that the vector  $\bar{\mathbf{y}}^* = (\bar{\varsigma}, \bar{\boldsymbol{\sigma}})^T$  is a KKT point of the canonical dual function  $P^d(\mathbf{y}^*)$  and  $\bar{\mathbf{x}} = [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c}$ .*

*If  $\bar{\mathbf{y}}^* = (\bar{\varsigma}, \bar{\boldsymbol{\sigma}})^T \in \mathcal{S}_a^+$ , then  $\bar{\mathbf{y}}^*$  is a global maximizer of  $P^d$  on  $\mathcal{S}_a^+$ , the vector  $\bar{\mathbf{x}}$  is a global minimizer of  $P$  on  $\mathcal{X}_a$ , and*

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_a} P(\mathbf{x}) = \max_{\mathbf{y}^* \in \mathcal{S}_a^+} P^d(\mathbf{y}^*) = P^d(\bar{\mathbf{y}}^*). \tag{18}$$

*If  $\bar{\mathbf{y}}^* \in \mathcal{S}_a^-$ , on the neighborhood  $\mathcal{X}_o \times \mathcal{S}_o \subset \mathcal{X}_a \times \mathcal{S}_a$  of  $(\bar{\mathbf{x}}, \bar{\mathbf{y}}^*)$ , we have that either*

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) = \min_{\mathbf{y}^* \in \mathcal{S}_o} P^d(\mathbf{y}^*) = P^d(\bar{\mathbf{y}}^*) \tag{19}$$

holds, or

$$P(\bar{\mathbf{x}}) = \max_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) = \max_{\mathbf{y}^* \in \mathcal{S}_o} P^d(\mathbf{y}^*) = P^d(\bar{\mathbf{y}}^*). \quad (20)$$

*Proof.* In the canonical form of the primal problem (5), replacing  $U(\Lambda(\mathbf{y}))$  by the Fenchel-Young equality  $(\Lambda(\mathbf{x}))^T \mathbf{y}^* - U^\sharp(\mathbf{y}^*)$ , the Gao-Strang type *total complementary function* (see [22]) associated with  $(\mathcal{P})$  can be obtained as

$$\Xi(\mathbf{x}, \mathbf{y}^*) = \frac{1}{2} \mathbf{x}^T G(\zeta, \bar{\boldsymbol{\sigma}}) \mathbf{x} - U^\sharp(\mathbf{y}^*) - \mathbf{x}^T \mathbf{c} - \alpha \zeta - \boldsymbol{\ell}^T \boldsymbol{\sigma}. \quad (21)$$

By Theorem 1 we know that the vector  $\bar{\mathbf{y}}^* \in \mathcal{S}_a$  is a KKT point of the problem  $(\mathcal{P}^d)$  if and only if  $\bar{\mathbf{x}} = [G(\bar{\zeta}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c}$  is a KKT point of the problem  $(\mathcal{P})$  and

$$P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\mathbf{y}}^*) = P^d(\bar{\mathbf{y}}^*). \quad (22)$$

Since the total complementary function  $\Xi$  is a saddle function on  $\mathcal{R}^n \times \mathcal{S}_a^+$ , i.e., it is convex in  $\mathbf{x} \in \mathcal{R}^n$  and concave in  $\mathbf{y}^* \in \mathcal{S}_a^+$ . Thus, we have

$$\begin{aligned} P^d(\bar{\mathbf{y}}^*) &= \max_{\mathbf{y}^* \in \mathcal{S}_a^+} P^d(\mathbf{y}^*) \\ &= \max_{\mathbf{y}^* \in \mathcal{S}_a^+} \min_{\mathbf{x} \in \mathcal{R}^n} \Xi(\mathbf{x}, \mathbf{y}^*) = \min_{\mathbf{x} \in \mathcal{R}^n} \max_{\mathbf{y}^* \in \mathcal{S}_a^+} \Xi(\mathbf{x}, \mathbf{y}^*) \\ &= \min_{\mathbf{x} \in \mathcal{R}^n} \left\{ \frac{1}{2} \mathbf{x}^T A \mathbf{x} - \mathbf{x}^T \mathbf{c} + \max_{\mathbf{y}^* \in \mathcal{S}_a^+} \{ (\Lambda(\mathbf{x}))^T \mathbf{y}^* - U^\sharp(\mathbf{y}^*) \} \right\} \\ &= \min_{\mathbf{x} \in \mathcal{X}_a} P(\mathbf{x}) \end{aligned}$$

due to the fact that

$$U(\Lambda(\mathbf{x})) = \max_{\mathbf{y}^* \in \mathcal{S}_a^+} \{ (\Lambda(\mathbf{x}))^T \mathbf{y}^* - U^\sharp(\mathbf{y}^*) \} = \begin{cases} \frac{1}{2} (\frac{1}{2} |B\mathbf{x}|^2 - \alpha)^2 & \text{if } \mathbf{x} \in \mathcal{X}_a \\ \infty & \text{otherwise.} \end{cases}$$

From Theorem 1 we have (18).

On the other hand, if  $\bar{\mathbf{y}}^* \in \mathcal{S}_a^-$ , the matrix  $G(\bar{\zeta}, \bar{\boldsymbol{\sigma}})$  is negative definite. In this case, the total complementary function  $\Xi(\mathbf{x}, \mathbf{y}^*)$  defined by (21) is a so-called super-Lagrangian (see [12]), i.e., it is locally concave in both  $\mathbf{x} \in \mathcal{X}_o \subset \mathcal{X}_a$  and  $\mathbf{y}^* \in \mathcal{S}_o \subset \mathcal{S}_a^-$ . Thus, by the triality theory developed in [12], we have either

$$\begin{aligned} P(\bar{\mathbf{x}}) &= \min_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) \\ &= \min_{\mathbf{x} \in \mathcal{X}_o} \max_{\mathbf{y}^* \in \mathcal{S}_o} \Xi(\mathbf{x}, \mathbf{y}^*) = \min_{\mathbf{y}^* \in \mathcal{S}_o} \max_{\mathbf{x} \in \mathcal{X}_o} \Xi(\mathbf{x}, \mathbf{y}^*) \\ &= \min_{\mathbf{y}^* \in \mathcal{S}_o} P^d(\mathbf{y}^*), \end{aligned}$$

or

$$\begin{aligned} P(\bar{\mathbf{x}}) &= \max_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) \\ &= \max_{\mathbf{x} \in \mathcal{X}_o} \max_{\mathbf{y}^* \in \mathcal{S}_o} \Xi(\mathbf{x}, \mathbf{y}^*) = \max_{\mathbf{y}^* \in \mathcal{S}_o} \max_{\mathbf{x} \in \mathcal{X}_o} \Xi(\mathbf{x}, \mathbf{y}^*) \\ &= \max_{\mathbf{y}^* \in \mathcal{S}_o} P^d(\mathbf{y}^*). \end{aligned}$$

This proves the statements (19) and (20).  $\square$

In the case that  $W(\mathbf{x}) = 0$ , the problem  $(\mathcal{P})$  is reduced to a box constrained quadratic minimization problem:

$$(\mathcal{P}_q) : \min_{\mathbf{x} \in \mathcal{X}_a} \left\{ P_q(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T A \mathbf{x} - \mathbf{x}^T \mathbf{c} \right\}. \quad (23)$$

The canonical dual problem for this box constrained quadratic minimization problem is a special case of the problem  $(\mathcal{P}^d)$  (see [18]), i.e.,

$$(\mathcal{P}_q^d) : \text{sta} \left\{ P_q^d(\boldsymbol{\sigma}) = -\frac{1}{2} \mathbf{c}^T [G_q(\boldsymbol{\sigma})]^{-1} \mathbf{c} - \boldsymbol{\ell}^T \boldsymbol{\sigma} : \boldsymbol{\sigma} \in \mathcal{S}_q \right\}, \quad (24)$$

where  $G_q(\boldsymbol{\sigma}) = A + 2\text{Diag}(\boldsymbol{\sigma})$  and

$$\mathcal{S}_q = \{\boldsymbol{\sigma} \in \mathcal{R}^n \mid \boldsymbol{\sigma} \geq 0, \det G_q(\boldsymbol{\sigma}) \neq 0\}.$$

Furthermore, we let

$$\mathcal{S}_q^+ = \{\boldsymbol{\sigma} \in \mathcal{S}_q \mid G_q(\boldsymbol{\sigma}) \text{ is positive definite}\}.$$

Then combining Theorem 1 and 2, we have the following result:

**Corollary 1.** *If  $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q$  is a KKT point of  $(\mathcal{P}_q^d)$ , then the vector  $\bar{\mathbf{x}} = [G_q(\bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c}$  is a KKT point of  $(\mathcal{P}_q)$  and*

$$P_q(\bar{\mathbf{x}}) = P_q^d(\bar{\boldsymbol{\sigma}}). \quad (25)$$

*If  $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^+$ , then  $\bar{\boldsymbol{\sigma}}$  is a global maximizer of  $P_q^d(\boldsymbol{\sigma})$  on  $\mathcal{S}_q^+$ , the vector  $\bar{\mathbf{x}}$  is a global minimizer of the problem  $(\mathcal{P}_q)$ , and*

$$P_q(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_a} P_q(\mathbf{x}) = \max_{\boldsymbol{\sigma} \in \mathcal{S}_q^+} P_q^d(\boldsymbol{\sigma}) = P_q^d(\bar{\boldsymbol{\sigma}}). \quad (26)$$

This Corollary shows that the box constrained nonconvex quadratic minimization problem  $(\mathcal{P}_q)$  can be converted to a concave maximization canonical dual problem

$$\max\{P_q^d(\boldsymbol{\sigma}) : \boldsymbol{\sigma} \in \mathcal{S}_q^+\}. \quad (27)$$

Clearly, if for a given matrix  $A \in \mathcal{R}^{n \times n}$  and  $\boldsymbol{\sigma} \in \mathcal{S}_q^+$  such that the vector  $\mathbf{c}$  is in the column space of  $G_q(\boldsymbol{\sigma}) = A + 2\text{Diag}(\boldsymbol{\sigma})$ , the canonical dual problem (27) has a unique solution  $\bar{\boldsymbol{\sigma}}$  which leads to a global minimizer  $\bar{\mathbf{x}} = G_q(\bar{\boldsymbol{\sigma}})^{-1} \mathbf{c}$ . Based on this Corollary, certain algorithms can be developed for solving the box constrained nonconvex minimization problem  $(\mathcal{P}_q)$ .

**3. Integer programming and boolean least squares problem.** As it was indicated in [10] and in Section 1 of this paper that the integer programming is a special case of box constrained optimization. Particularly, if we let  $\mathcal{X}_{ip}$  be a subset of  $\mathcal{X}_a$ , defined by

$$\mathcal{X}_{ip} = \{\mathbf{x} \in \mathcal{R}^n \mid -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{x} \in \{-1, 1\}^n\},$$

where the notation  $\{-1, 1\}^n$  denotes integer vectors of  $\mathcal{R}^n$  with components either  $-1$  or  $1$ , the primal problem  $(\mathcal{P})$  is a nonconvex integer programming problem (denoted by  $(\mathcal{P}_{ip})$ ):

$$(\mathcal{P}_{ip}) : \min_{\mathbf{x} \in \mathcal{X}_{ip}} \{P(\mathbf{x}) = Q(\mathbf{x}) + W(\mathbf{x})\}. \quad (28)$$

Canonical duality theory for solving quadratic 0 - 1 integer programming problems was first studied in the joint work with Fang *et al* [6]. In the current nonconvex problem  $(\mathcal{P}_{ip})$ , we have  $\ell_i = 1$  ( $i = 1, 2, \dots, n$ ) so that the inequality constraints in  $\mathcal{X}_{ip}$  can be written in the canonical form  $\boldsymbol{\epsilon} = \{x_i^2 - 1\} \leq 0 \in \mathcal{R}^n$ . It is easy to prove that the canonical duality relations

$$\boldsymbol{\sigma} \in \partial\Psi(\boldsymbol{\epsilon}) \Leftrightarrow \boldsymbol{\epsilon} \in \partial\Psi^\sharp(\boldsymbol{\sigma}) \Leftrightarrow \boldsymbol{\epsilon}^T \boldsymbol{\sigma} = \Psi(\boldsymbol{\epsilon}) + \Psi^\sharp(\boldsymbol{\sigma})$$

are equivalent to

$$\boldsymbol{\epsilon} \leq 0, \boldsymbol{\sigma} \geq 0, \boldsymbol{\epsilon}^T \boldsymbol{\sigma} = 0.$$

Clearly, the complementarity condition leads to the integer constraint

$$\epsilon = \{x_i^2 - 1\} = 0 \in \mathcal{R}^n \quad \text{if } \boldsymbol{\sigma} > 0 \in \mathcal{R}^n.$$

Thus, replacing the dual feasible spaces  $\mathcal{S}_a$  and  $\mathcal{S}_a^+$  by

$$\mathcal{S}_{ip} = \{\mathbf{y}^* = (\varsigma, \boldsymbol{\sigma})^T \in \mathcal{R}^{1+n} \mid \varsigma \geq -\alpha, \boldsymbol{\sigma} > 0, \det G(\varsigma, \boldsymbol{\sigma}) \neq 0\}, \quad (29)$$

$$\mathcal{S}_{ip}^+ = \{(\varsigma, \boldsymbol{\sigma})^T \in \mathcal{S}_{ip} \mid G(\varsigma, \boldsymbol{\sigma}) \text{ is positive definite}\} \quad (30)$$

respectively, the canonical dual problem of  $(\mathcal{P}_{ip})$  can be proposed as

$$(\mathcal{P}_{ip}^d): \quad \text{sta} \left\{ P_{ip}^d(\varsigma, \boldsymbol{\sigma}) = -\frac{1}{2} \mathbf{c}^T [G(\varsigma, \boldsymbol{\sigma})]^{-1} \mathbf{c} - \frac{1}{2} \varsigma^2 - \alpha \varsigma - \sum_{i=1}^n \sigma_i : (\varsigma, \boldsymbol{\sigma})^T \in \mathcal{S}_{ip} \right\}. \quad (31)$$

Similar to Theorem 1 and 2, we have

**Theorem 3.** *Problem  $(\mathcal{P}_{ip}^d)$  is canonically dual to the primal problem  $(\mathcal{P}_{ip})$  in the sense that if  $\bar{\mathbf{y}}^* = (\bar{\varsigma}, \bar{\boldsymbol{\sigma}})^T \in \mathcal{S}_{ip}$  is a KKT point of  $(\mathcal{P}_{ip}^d)$ , then the vector  $\bar{\mathbf{x}} = [G(\bar{\varsigma}, \bar{\boldsymbol{\sigma}})]^{-1} \mathbf{c}$  is a KKT point of  $(\mathcal{P}_{ip})$ , and*

$$P(\bar{\mathbf{x}}) = P^d(\bar{\mathbf{y}}^*). \quad (32)$$

*If  $\bar{\mathbf{y}}^* = (\bar{\varsigma}, \bar{\boldsymbol{\sigma}})^T \in \mathcal{S}_{ip}^+$ , then  $\bar{\mathbf{y}}^*$  is a global maximizer of  $P_{ip}^d$  on  $\mathcal{S}_{ip}^+$ , the vector  $\bar{\mathbf{x}}$  is a global minimizer of  $P$  on  $\mathcal{X}_{ip}$ , and*

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_{ip}} P(\mathbf{x}) = \max_{\mathbf{y}^* \in \mathcal{S}_{ip}^+} P^d(\mathbf{y}^*) = P^d(\bar{\mathbf{y}}^*). \quad (33)$$

Since the canonical dual function  $P_{ip}^d(\mathbf{y}^*)$  is concave on  $\mathcal{S}_{ip}^+$ , this theorem shows that the discrete nonconvex integer minimization problem  $(\mathcal{P}_{ip})$  can be converted into a continuous concave maximization dual problem  $(\mathcal{P}_{ip}^d)$ , which can be solved by any well-developed nonlinear optimization method. Detailed study on the canonical duality theory for solving integer programming was given in [6].

The canonical duality theory can be applied to solve many constrained nonconvex integer programming problems. Particularly, if the function  $W(\mathbf{x})$  in  $(\mathcal{P}_{ip})$  is an indicator of the feasible space  $\mathcal{X}_b := \{\mathbf{x} \in \mathcal{X}_{ip} \mid B\mathbf{x} = \mathbf{b}\}$ , i.e.,

$$W(\mathbf{x}) = \begin{cases} 0 & \text{if } B\mathbf{x} = \mathbf{b}, \\ +\infty & \text{otherwise,} \end{cases}$$

where  $\mathbf{b} \in \mathcal{R}^m$  is a given vector, and the quadratic function  $Q(\mathbf{x})$  can be written in the least square form:

$$Q(\mathbf{x}) = -\frac{1}{2} |\mathbf{A}\mathbf{x} - \mathbf{c}|^2 = -\frac{1}{2} \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{A}^T \mathbf{c} - \frac{1}{2} |\mathbf{c}|^2, \quad (34)$$

where  $\mathbf{A}$  could be any given  $p \times n$  matrix and  $\mathbf{c} \in \mathcal{R}^p$  is a given vector, then the primal problem  $(\mathcal{P}_{ip})$  is the so-called *Boolean least squares problem*:

$$(\mathcal{P}_b): \quad \min_{\mathbf{x} \in \mathcal{X}_{ip}} \left\{ P_b(\mathbf{x}) = -\frac{1}{2} |\mathbf{A}\mathbf{x} - \mathbf{c}|^2 : B\mathbf{x} = \mathbf{b} \right\}. \quad (35)$$

This problem arises from a large number of applications in communication systems such as the channel decoding, MIMO detection, multiuser detection, equalization, resource allocation in wireless systems, etc (see [4, 5]). Since the quadratic function  $Q(\mathbf{x})$  is concave, its minimizers are located on the boundary of the feasible set  $\mathcal{X}_{ip}$ .

Traditional methods for solving this discrete concave minimization problem ( $\mathcal{P}_b$ ) is very difficult.

We assume that  $m < n$  and  $\text{rank } B = m$  so that the problem ( $\mathcal{P}_b$ ) is not over-constrained. Let  $N_B \in \mathcal{R}^{n \times r}$  ( $r = n - m$ ) be the null space of  $B$ , i.e.,  $BN_B \mathbf{x}^o = \mathbf{0} \in \mathcal{R}^m \forall \mathbf{x}^o \in \mathcal{R}^r$ , thus, the general solutions of the constraint equation  $B\mathbf{x} = \mathbf{b}$  is

$$\mathbf{x} = \mathbf{x}_b + N_B \mathbf{x}^o, \quad (36)$$

where  $\mathbf{x}_b \in \mathcal{R}^n$  is a particular solution of  $B\mathbf{x} = \mathbf{b}$ , i.e.,  $B\mathbf{x}_b = \mathbf{b}$ , and  $\mathbf{x}^o \in \mathcal{R}^r$  is a free vector. We let

$$\begin{aligned} G_a(\boldsymbol{\sigma}) &= -A^T A + 2\text{Diag}(\boldsymbol{\sigma}) \\ G_b(\boldsymbol{\sigma}) &= \frac{1}{2}(A^T \mathbf{c} + G_a(\boldsymbol{\sigma})\mathbf{x}_b)^T N_B [N_B^T G_a(\boldsymbol{\sigma}) N_B]^{-1} N_B^T (A^T \mathbf{c} + G_a(\boldsymbol{\sigma})\mathbf{x}_b). \end{aligned}$$

Then, on the dual feasible space

$$\mathcal{S}_b = \{\boldsymbol{\sigma} \in \mathcal{R}^n \mid \boldsymbol{\sigma} > 0, \det G_a(\boldsymbol{\sigma}) \neq 0\}, \quad (37)$$

the canonical dual of the Boolean least squares problem can be formulated as

$$(\mathcal{P}_b^d) : \max_{\boldsymbol{\sigma} \in \mathcal{S}_b} \left\{ P_b^d(\boldsymbol{\sigma}) = \frac{1}{2} \mathbf{x}_b^T G_a(\boldsymbol{\sigma}) \mathbf{x}_b - \frac{1}{2} |\mathbf{c}|^2 + \mathbf{x}_b^T A^T \mathbf{c} - \sum_{i=1}^n \sigma_i - G_b(\boldsymbol{\sigma}) \right\}. \quad (38)$$

**Theorem 4.** *The primal problem ( $\mathcal{P}_b$ ) is perfectly dual to the canonical dual problem ( $\mathcal{P}_b^d$ ) in the sense that if  $\bar{\boldsymbol{\sigma}}$  is a critical point of ( $\mathcal{P}_b^d$ ) and  $\bar{\boldsymbol{\sigma}} > 0$ , then the vector*

$$\bar{\mathbf{x}} = \mathbf{x}_b - N_B [N_B^T G_a(\bar{\boldsymbol{\sigma}}) N_B]^{-1} N_B^T (A^T \mathbf{c} + G_a(\bar{\boldsymbol{\sigma}})\mathbf{x}_b) \quad (39)$$

is a critical point of ( $\mathcal{P}$ ) and

$$P_b(\bar{\mathbf{x}}) = P_b^d(\bar{\boldsymbol{\sigma}}). \quad (40)$$

Moreover, if  $G_a(\bar{\boldsymbol{\sigma}})$  is positive definite, then  $\bar{\boldsymbol{\sigma}}$  is a global maximizer of ( $\mathcal{P}_b^d$ ) on

$$\mathcal{S}_b^+ = \{\boldsymbol{\sigma} \in \mathcal{S}_b \mid G_a(\boldsymbol{\sigma}) \text{ is positive definite}\},$$

and  $\bar{\mathbf{x}}$  is a global minimizer of ( $\mathcal{P}_b$ ) on  $\mathcal{X}_b$ , i.e.

$$P_b(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_b} P_b(\mathbf{x}) = \max_{\boldsymbol{\sigma} \in \mathcal{S}_b^+} P_b^d(\boldsymbol{\sigma}) = P_b^d(\bar{\boldsymbol{\sigma}}). \quad (41)$$

In the case that there is no equilibrium constraint, the primal problem ( $\mathcal{P}_b$ ) is a so-called *lattice-decoding-type* problem:

$$(\mathcal{P}_{bo}) : \min_{\mathbf{x} \in \mathcal{X}_{ip}} \left\{ P_{bo}(\mathbf{x}) = -\frac{1}{2} |A\mathbf{x} - \mathbf{c}|^2 \right\}. \quad (42)$$

In this case, the canonical dual problem has a simple form:

$$(\mathcal{P}_{bo}^d) : \max_{\boldsymbol{\sigma} \in \mathcal{S}_b} \left\{ P_{bo}^d(\boldsymbol{\sigma}) = -\frac{1}{2} \mathbf{c}^T A [G_a(\boldsymbol{\sigma})]^{-1} A^T \mathbf{c} - \sum_{i=1}^n \sigma_i - \frac{1}{2} |\mathbf{c}|^2 \right\}. \quad (43)$$

**Corollary 2.** *If  $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_b$  is a KKT point the canonical dual problem ( $\mathcal{P}_{bo}^d$ ), then  $\bar{\mathbf{x}} = -[G_a(\bar{\boldsymbol{\sigma}})]^{-1} A^T \mathbf{c}$  is a KKT point of the Boolean least squares problem ( $\mathcal{P}_{bo}$ ). If  $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_b^+$ , then  $\bar{\mathbf{x}}$  is a global minimizer of  $P_{bo}(\mathbf{x})$  on  $\mathcal{X}_{ip}$  and*

$$P_{bo}(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_{ip}} P_{bo}(\mathbf{x}) = \max_{\boldsymbol{\sigma} \in \mathcal{S}_b^+} P_{bo}^d(\boldsymbol{\sigma}) = P_{bo}^d(\bar{\boldsymbol{\sigma}}). \quad (44)$$

A detailed study for the canonical duality approach for solving quadratic 0 - 1 integer programming can be found in [6].



#### 4. Applications.

4.1. **Decoupled problem.** In the case that  $A = \text{Diag}(\mathbf{a})$  is a diagonal matrix and  $B$  is a unit matrix in  $\mathcal{R}^n$ , the primal problem  $(P)$  has the decoupled form of

$$\min \left\{ P(\mathbf{x}) = \sum_{i=1}^n \left( \frac{1}{2} a_i x_i^2 - c_i x_i \right) + \frac{1}{2} \left( \frac{1}{2} |\mathbf{x}|^2 - \alpha \right)^2 : x_i^2 \leq \ell_i \right\}. \quad (45)$$

On the the dual feasible set

$$\mathcal{S}_a = \{(\varsigma, \boldsymbol{\sigma}) \in \mathcal{R}^{1+n} \mid \varsigma \geq -\alpha, \sigma_i \geq 0, a_i + \varsigma + 2\sigma_i \neq 0\},$$

the canonical dual function  $P^d(\mathbf{y}^*)$  has a simple form

$$P^d(\varsigma, \boldsymbol{\sigma}) = -\frac{1}{2} \varsigma^2 - \alpha \varsigma - \sum_{i=1}^n \left( \frac{c_i^2}{2(a_i + \varsigma + 2\sigma_i)} + \ell_i \sigma_i \right). \quad (46)$$

For a one-dimensional problem, we chose  $a = -0.5$ ,  $c = 0.1$ ,  $\alpha = 0.3$ . Then, the graph of  $P(x)$  is a double-well function on a given box  $[-\sqrt{\ell}, \sqrt{\ell}]$  (see Fig. 1). If we let  $\ell = 1$ , the canonical dual function  $P^d(x)$  has two critical points  $(\bar{\varsigma}_1, \bar{\sigma}_1) = (0.2, 0.2)$  and  $(\bar{\varsigma}_2, \bar{\sigma}_2) = (0.2, 0.1)$ . Both of them are located on the boundary of the box  $[-1, 1]$ . Since  $(\bar{\varsigma}_1, \bar{\sigma}_1) \in \mathcal{S}_a^+$ , by Theorem 2 we know that  $\bar{x}_1 = \frac{1}{2}c/(a + \bar{\varsigma}_1 + 2\bar{\sigma}_1) = 1$  is a global minimizer of  $P(x)$  on  $[-1, 1]$ , while  $\bar{x}_2 = \frac{1}{2}c/(a + \bar{\varsigma}_2 + 2\bar{\sigma}_2) = -1$  is a local minimizer since  $(\bar{\varsigma}_2, \bar{\sigma}_2) \in \mathcal{S}_a^-$ . It is easy to check that

$$P(\bar{x}_1) = -0.33 = P^d(\bar{\varsigma}_1, \bar{\sigma}_1), \quad P(\bar{x}_2) = -0.13 = P^d(\bar{\varsigma}_2, \bar{\sigma}_2).$$

If we let  $\ell = 4$ , then the canonical dual problem has three KKT points:

$$\bar{\mathbf{y}}_1^* = (0.576, 0), \quad \bar{\mathbf{y}}_2^* = (0.416, 0), \quad \bar{\mathbf{y}}_3^* = (-0.292, 0).$$

By Theorem 2 we know that  $\bar{x}_1 = 1.32$  is a global minimizer,  $\bar{x}_2 = -1.197$  is a local minimizer, and  $\bar{x}_3 = -0.126$  is a local maximizer (see Fig. 1).

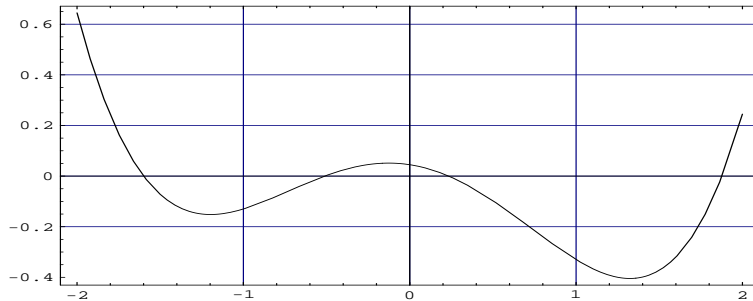


FIGURE 1. Nonconvex minimization with box constraints: (1) in  $[-1, 1]$  both global minimizer  $\bar{x}_1 = 1$  and local minimizer  $\bar{x}_2 = -1$  are located on the boundary; (2) in  $[-2, 2]$ , the global minimizer  $\bar{x}_1 = 1.32$  and local minimizer  $\bar{x}_2 = -1.197$  are critical points

**4.2. Two-dimensional box constrained problem.** We now consider the following box constrained concave minimization problem

$$\min P_{bo}(x_1, x_2) = -\frac{1}{2}|A\mathbf{x} - \mathbf{c}|^2, \quad (47)$$

$$s.t. \quad x_i^2 \leq 1, \quad i = 1, 2, \quad (48)$$

where  $A = \{a_{ij}\}$  is an arbitrarily given  $2 \times 2$  matrix. If we choose  $a_{11} = -1.0$ ,  $a_{12} = 0$ ,  $a_{21} = -1$ ,  $a_{22} = -2$ , and  $\mathbf{c} = (3, 2)^T$ , the dual function

$$P_{bo}^d(\boldsymbol{\sigma}) = -\frac{1}{2}\mathbf{c}^T A[G_a(\boldsymbol{\sigma})]^{-1}A^T\mathbf{c} - \sum_{i=1}^2 \sigma_i - \frac{1}{2}|\mathbf{c}|^2$$

has four critical points:

$$\boldsymbol{\sigma}_1 = (4.5, 5.0), \quad \boldsymbol{\sigma}_2 = (2.5, -1), \quad \boldsymbol{\sigma}_3 = (-2.5, 3), \quad \boldsymbol{\sigma}_4 = (-0.5, 1).$$

The corresponding primal solutions  $\mathbf{x}_k = -[G_a(\boldsymbol{\sigma}_k)]^{-1}A^T\mathbf{c}$  ( $k = 1, 2, 3, 4$ ) are

$$\mathbf{x}_1 = (1, 1), \quad \mathbf{x}_2 = (1, -1), \quad \mathbf{x}_3 = (-1, 1), \quad \mathbf{x}_4 = (-1, -1).$$

It is easy to check that we have only one critical point  $\boldsymbol{\sigma}_1 \in \mathcal{S}_b^+$ . By Corollary 2, we know that  $\mathbf{x}_1$  is a global minimizer (see Fig. 2). It is easy to verify that  $P(\mathbf{x}_k) = P^d(\boldsymbol{\sigma}_k)$   $k = 1, 2, 3, 4$  and

$$P_{bo}(\mathbf{x}_1) = -20.5 < P_{bo}(\mathbf{x}_2) = -8.5 < P_{bo}(\mathbf{x}_3) = -6.5 < P_{bo}(\mathbf{x}_4) = -2.5$$

We note that if the inequality constraints in (48) are replaced by  $x_i^2 = 1$ ,  $i = 1, 2$ , then the problem (47) is a Boolean least square problem. Actually, since all the four dual solutions  $\boldsymbol{\sigma}_k \neq 0$  ( $k = 1, 2, 3, 4$ ), it turns out that the primal solutions  $\mathbf{x}_k$  are all integer vectors.

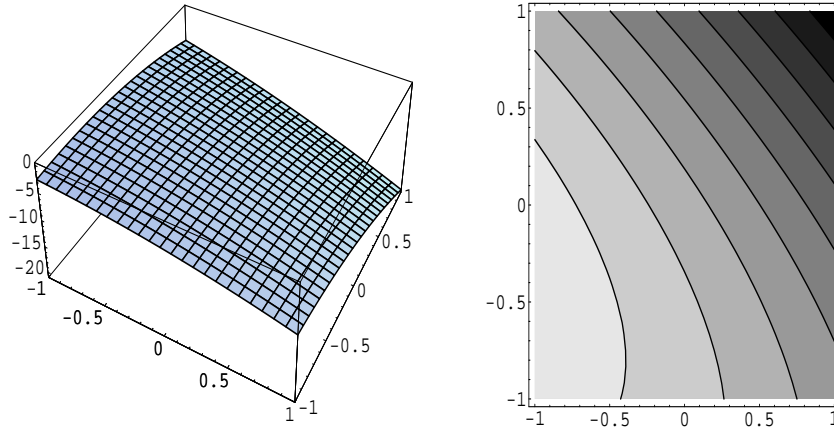


FIGURE 2. Graph of the least square function  $P_{bo}$  and its contour

## 5. Conclusion remarks and a conjecture on NP-Hard problems.

Applications of the canonical duality theory have been demonstrated by solving the box constrained nonconvex optimization problems  $(\mathcal{P})$ ,  $(\mathcal{P}_{ip})$ , and their special cases. Theorems proved in this paper show that by the canonical dual transformation, nonconvex primal problems can be converted into canonical dual problems. By the fact that the canonical dual function  $P^d(\mathbf{y}^*)$  is concave on the dual feasible space  $\mathcal{S}_a^+$ , if  $\mathcal{S}_a^+$  is non-empty, i.e., the dual problem  $(\mathcal{P}^d)$  has at least one KKT point in  $\mathcal{S}_a^+$ , the canonical dual problem

$$\max_{\mathbf{y}^* \in \mathcal{S}_a^+} P^d(\mathbf{y}^*) \quad (49)$$

can be solved by well-developed deterministic optimization methods. Since the primal problem  $(\mathcal{P})$  could be NP-hard, a conjecture can be proposed as the following.

**Conjecture.** *The box constrained nonconvex minimization problem  $(\mathcal{P})$  is NP-Hard if its canonical dual  $(\mathcal{P}^d)$  has no KKT point in  $\mathcal{S}_a^+$ .*

Generally speaking, the function  $W(\mathbf{x})$  in the primal problem  $(\mathcal{P})$  could be any given nonconvex/nonsmooth function as long as it can be written in the canonical form, i.e., there exists a nonlinear operator  $\mathbf{y} = \Lambda(\mathbf{x})$  and a canonical function  $U(\mathbf{y})$  such that  $W(\mathbf{x}) = U(\Lambda(\mathbf{x}))$ . Thus, the theorems proposed in this paper can be generalized for solving more complicated problems in global optimization.

In nonlinear analysis and mechanics, the canonical duality theory has been successfully applied for solving many nonconvex/nonsmooth variational/boundary value problems [11, 20]. In finite deformation elasticity, the canonical duality theory is the so-called *Gao principle* [27]. Detailed discussion on canonical duality theory and its extensive applications can be found in the recent review articles [15, 21].

**Acknowledgements.** The research of this work was supported by National Science Foundation grant CCF-0514768. The author would like to thank the anonymous referees whose valuable comments and remarks improved the paper.

## REFERENCES

- [1] Akrotirianakis, I.G. and Floudas, C.A., *Computational experience with a new class of convex underestimators: Box-constrained NLP problems*, J. Global Optimization, **29** (2004), 249-264.
- [2] Akrotirianakis, I.G. and Floudas, C.A., *A new class of improved convex underestimators for twice continuously differentiable constrained NLPs*, J. Global Optimization, **30** (2004), 367-390.
- [3] Al-Khayyal, F.H. and Falk, J.E., *Jointly constrained biconvex programming*, Mathematics of Operations Research, **8** (1983), 523.
- [4] Chiang, M., *Geometrical Programming for Communication Systems*, nox Publishers Inc., PO Box 1024, Hanover, MA, USA, 2005.
- [5] Chiang, M., *Nonconvex optimization in communication systems*, in *Advances in Mechanics and Mathematics*, Vol. **III**, (eds: D.Y. Gao and H.D. Sherali), Springer, 2006.
- [6] Fang, S.-C., Gao, D.Y., Sheu, R.-L., and Wu, S.Y., *Canonical duality and complete solutions to nonconvex integer programming problems*, to appear in *Optimization*.
- [7] Floudas, C.A., "Deterministic Optimization. Theory, Methods, and Applications," Kluwer Academic Publishers, 2000.
- [8] Floudas, C.A., Akrotirianakis, I.G., Caratzoulas, S., Meyer, C.A., and Kallrath, J., *Global optimization in the 21th century: Advances and challenges*, Computers & Chemical Engineering, **29** (2005), 1185-1202.

- [9] Floudas C.A. and V. Visweswaran, *A primal-relaxed dual global optimization approach*, Journal of Optimization, Theory, and its Applications, **78** (1993), 187-225.
- [10] Floudas, C.A. and V. Visweswaran, *Quadratic optimization*, in “Handbook of Global Optimization” (eds: R. Horst and P.M. Pardalos), Kluwer Academic Publishers, Dordrecht/Boston/London, 1995, 217-270.
- [11] Gao, D.Y., *Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications*, Nonlinear Analysis, **42** (2000), 1161-1193.
- [12] Gao, D.Y., “Duality Principles in Nonconvex Systems: Theory, Methods and Applications,” Kluwer Academic Publishers, Dordrecht/Boston/London, 2000.
- [13] Gao, D.Y., *Canonical dual transformation method and generalized triality theory in nonsmooth global optimization*, J. Global Optimization, **17** (2000), 127-160.
- [14] Gao, D.Y., *Perfect duality theory and complete solutions to a class of global optimization problems*, Optimisation, **52** (2003), 467-493.
- [15] Gao, D.Y., *Nonconvex semi-linear problems and canonical duality solutions*, in “Advances in Mechanics and Mathematics”, Vol. **II**, Kluwer Academic Publishers, Dordrecht/Boston/London, 2003, 261-312.
- [16] Gao, D.Y., *Canonical duality theory and solutions to constrained nonconvex quadratic programming*, Journal of Global Optimization, **29** (2004), 377-399.
- [17] Gao, D.Y., *Sufficient conditions and perfect duality in nonconvex minimization with inequality constraints*, J. Industrial and Management Optimization, **1** (2005), 59-69.
- [18] Gao, D.Y., “Duality-Mathematics” (eds: J.G. Webster), Wiley Encyclopedia of Electrical and Electronics Engineering, Vol. **6**, (Electronical Edition), 2006.
- [19] Gao, D.Y., *Complete solutions and extremality criteria to polynomial optimization problems*, J. Global Optimization, **35** (2006), 131-143.
- [20] Gao, D.Y. and Ogden, R.W., *Complete solutions to the azimuthal shear problem in finite elasticity*, to be submitted.
- [21] Gao, D.Y. and Serali, H.D., “Canonical Duality Theory: Connection between Nonconvex Mechanics and Global Optimization,” Advances in Mechanics and Mathematics, Vol. **III**, Springer, 2006.
- [22] Gao, D.Y. and Strang, G., *Geometric nonlinearity: Potential energy, complementary energy, and the gap function*, Quart. Appl. Math., **47** (1989), 487-504.
- [23] Grippo, L. and Lucidi, S., *A differentiable exact penalty function for bound constrained quadratic programming problems*, Optimization, **22** (1991), 557-578.
- [24] Han, C.G., Pardalos, P.M., and Ye, Y., *An interior point algorithm for large-scale quadratic problems with box constraints*, in “Springer-Verlag Lecture Notes in Control and Information” (eds: A. Bensoussan and J.L. Lions), Vol. **144**, 1990, 413-422.
- [25] Hansen, P., Jaumard, B., Ruiz, M., and Xiong, J., *Global minimization of indefinite quadratic functions subject to box constraints*, Technical Report G-91-54, GERAD, École Polytechnique, Université McGill, Montreal, (1991).
- [26] Horst, R., Panos M. Pardalos, Nguyen Van Thoai, “Introduction to Global Optimization,” Kluwer Academic Publishers, 2000.
- [27] Li, S.F. and Gupta, A., *On dual configuration forces*, J. of Elasticity, **84** (2006), 13-31.
- [28] Murty, K.G. and Kabadi, S.N., *Some NP-hard problems in quadratic and nonlinear programming*, Math. Programming, **39** (1987), 117-129.
- [29] Pardalos, P.M. and Schnitger, G., *Checking local optimality in constrained quadratic and nonlinear programming*, Operations Research Letters, **7** (1988), 33-35.
- [30] Serali, H.D. and Alameddine, A., *A new reformulation linearization technique for bilinear programming problems*, J. of Global Optimization, **2** (1992), 379.

Received August 2006; 1st revision October 2006; 2nd revision January 2007.

E-mail address: gao@vt.edu