

# Canonical Duality Theory for Solving Non-monotone Variational Inequality Problems

Guoshan Liu, David Yang Gao and Shouyang Wang

**Abstract** This paper presents a canonical dual approach for solving a class of non-monotone variational inequality problems. It shows that by using the canonical dual transformation, these challenging problems can be reformulated as a canonical dual problem, which is equivalent to the primal problems in the sense that they have the same set of KKT points. Existence theorem for global optimal solutions is obtained. Based on the canonical duality theory, this dual problem can be solved via well-developed convex programming methods. Applications are illustrated with several examples.

## 1 Problems and Motivation

We are interesting in solving the following non-monotone variational inequality problem:

$$(VI) : \quad \langle F(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle \geq 0, \quad \forall \mathbf{x} \in \mathcal{H}, \quad (1)$$

where the notation  $\langle *, * \rangle$  denotes for inner product,  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a non-monotone operator, defined by

$$F(\mathbf{x}) = Q\mathbf{x} - \mathbf{f} + \nabla W(\mathbf{x}), \quad (2)$$

---

G. Liu (✉)

School of Business, Renmin University of China, Beijing 100872, China  
e-mail: liuguoshan@gmail.com

D.Y. Gao

Faculty of Science and Technology, Federation University Australia,  
Ballarat, VIC 3353, Australia  
e-mail: d.gao@federation.edu.au

S. Wang

Laboratory of Management, Decision and Information Systems  
Chinese Academy of Sciences, Beijing 100084, China  
e-mail: sywang@amss.ac.cn

in which  $Q \in \mathbb{R}^{n \times n}$  is a symmetric matrix,  $\mathbf{f} \in \mathbb{R}^n$  is a given vector, and  $W(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$  is a nonconvex differentiable function. The feasible set  $\mathcal{X}$  in this paper is defined by

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n \mid \phi(\mathbf{x}) \leq 0\},$$

where  $\phi(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^q$  is assumed to be a vector value convex function. In this paper, we assume that  $\text{ri}(\mathcal{X})$  is nonempty, i.e., there exists at least one  $\bar{\mathbf{x}}$  such that  $\phi(\bar{\mathbf{x}}) < 0$ .

The first problem involving a variational inequality is the well known *Signorini problem*, proposed by A. Signorini in 1959 as a frictionless contact problem in linear elastic mechanics and solved by G. Fichera in 1963 (cf. [4]). Mathematical theory of variational inequality was first studied by G. Stampacchia in 1964 [21]. It is known that the Signorini problem is actually equivalent to a variational problem subjected to inequality constraint. By the fact that the total potential energy for linear elasticity is convex, it turns out that extensive mathematical research has been focused mainly on monotone variational inequality problems (see [1, 2, 5, 9, 12, 14, 15, 19]). However, variational problems in large deformation mechanics are usually nonconvex [6, 11, 22]. For example, the total potential energy of the Gao nonlinear beam model [16, 17] is given by

$$J(w) = \int \left[ \frac{1}{2} a (w_{,xx})^2 + \frac{1}{2} \left( \frac{1}{2} (w_{,x})^2 - \lambda \right)^2 - wf \right] dx, \quad (3)$$

where  $a > 0$  is a material constant,  $f(x)$  is a given distributive load, and the unknown function  $w(x)$  is the deformation of the beam. Clearly,  $J(w)$  is nonconvex if the beam is subjected to a compressive load  $\lambda > 0$ . If the beam is supported by an obstacle  $\psi(x)$ , the associated variational inequality problem was formulated in [7, 8]

$$\langle \delta J(w), v - w \rangle \geq 0 \quad \forall v \in \mathcal{V}_a,$$

where  $\mathcal{V}_a$  is a convex set with the inequality constraint  $v(x) \geq \psi(x)$ , and  $\delta J(w)$  is the Gâteaux derivative of  $J(w)$  given by

$$\delta J(w) = a w_{,xxxx} - \frac{3}{2} (w_{,x})^2 w_{,xx} + \lambda w_{,xx} - f$$

which is a non-monotone differential operator. In finite element analysis, if the displacement  $w(x)$  is numerically discretized by a finite vector  $\mathbf{x} \in \mathbb{R}^n$ , the total potential functional  $J(w)$  can be written in a nonconvex function in  $\mathbb{R}^n$  (see [20])

$$P(\mathbf{x}) = W(\mathbf{x}) + \frac{1}{2} \langle \mathbf{x}, Q\mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{f} \rangle$$

where  $W(\mathbf{x})$  is the so-called double-well potential

$$W(\mathbf{x}) = \sum_{k=1}^p \frac{1}{2} \alpha_k \left( \frac{1}{2} \langle \mathbf{x}, B_k \mathbf{x} \rangle + \langle \mathbf{x}, b_k \rangle - d_k \right)^2,$$

in which  $\alpha_k, d_k > 0$  are given constants,  $b_k \in \mathbb{R}^n$ , and  $B_k \in \mathbb{R}^{n \times n}$  is given symmetric matrix for each  $k \in \{1, \dots, p\}$ . This double-well function appears extensively in computational physics, including Landau-Ginzburg equation in phase transitions [13], von Karmen plate [22], nonlinear Schrödinger equation, and much more [10, 11]. Due to the nonconvexity of  $W(\mathbf{x})$ , the operator  $F(\mathbf{x})$  is non-monotone. Traditional direct methods for solving such problems are very difficult.

Duality theory in convex analysis has been well studied in [3]. Application to monotone variational inequality problems was first proposed by Mosco [18]. However, in nonconvex variational problems and non-monotone variational inequalities, these well-developed duality theory and methods usually lead to a so-called duality gap. In order to close this gap, a potentially useful canonical duality theory has been developed in [10].

In this paper, we will demonstrate the application of this method by solving the non-monotone variational inequality problem (VI). In the next section, the canonical dual of the problem is presented, which is equivalent to the primal problem in the sense that they have the same set of KKT points. The extremality conditions for these KKT points are discussed in Sect. 3. In Sect. 4, we discuss the properties of the dual problem and give a sufficient condition for the existence of solution. Applications are illustrated in Sect. 5. Some conclusions are given in the final section.

## 2 Optimization Problem and Its Canonical Dual

It is known that the variational inequality problem (VI) is related to the following optimization problem:

$$(OP) : \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ P(\mathbf{x}) = W(\mathbf{x}) + \frac{1}{2} \langle \mathbf{x}, Q\mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{f} \rangle \right\} \tag{4}$$

*s.t.*  $\phi(\mathbf{x}) \leq 0.$

By introducing Lagrange multipliers  $\lambda \in \mathbb{R}_+^q$  to relax the inequality constraint  $\phi(\mathbf{x}) \leq 0$ , the classical Lagrangian associated with this constrained optimization problem is

$$L(\mathbf{x}, \lambda) = P(\mathbf{x}) + \lambda^\top \phi(\mathbf{x}).$$

Thus, the criticality condition  $\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) = 0$  leads to the equilibrium equation

$$\nabla W(\mathbf{x}) + Q\mathbf{x} - \mathbf{f} + \sum_{s=1}^q \lambda_s \nabla \phi_s(\mathbf{x}) = 0.$$

By the KKT theory, the Lagrange multipliers have to satisfy the following complementarity conditions

$$\lambda^\top \phi(\mathbf{x}) = 0, \quad \phi(\mathbf{x}) \leq 0, \quad \lambda \geq 0.$$

We call the point which satisfies the above two conditions the KKT stationary point of the problem (OP) and (VI). Because we have already assumed that  $\text{ri}(\mathcal{X})$  is nonempty, then the basic constraints qualification must hold for the problem (VI) and (OP), and the following result is obvious.

**Lemma 1.** *If  $\bar{\mathbf{x}}$  solves (VI), then it is a KKT stationary point of (OP) or (VI).*

Due to the nonconvexity of the object function  $P(\mathbf{x})$ , to solve problem (VI) is very difficult. For a given  $\lambda \geq 0$ , the Lagrangian dual function can be defined by

$$P^*(\lambda) = \inf_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \lambda).$$

In the case that  $P(\mathbf{x})$  is convex, we have the well-known saddle duality theorem

$$P(\bar{\mathbf{x}}) = \inf_{\mathbf{x}} \sup_{\lambda \geq 0} L(\mathbf{x}, \lambda) = \sup_{\lambda \geq 0} \inf_{\mathbf{x}} L(\mathbf{x}, \lambda) = P^*(\bar{\lambda}).$$

However, if  $P(\mathbf{x})$  is nonconvex, we have the so-called weak duality

$$\theta = \inf_{\mathbf{x}} \sup_{\lambda \geq 0} L(\mathbf{x}, \lambda) - \sup_{\lambda \geq 0} \inf_{\mathbf{x}} L(\mathbf{x}, \lambda) \geq 0.$$

Very often, this duality gap  $\theta = \infty$ .

Following the standard procedure of the canonical dual transformation, we assume that there exists a *geometrical operator*

$$\xi = \Lambda(\mathbf{x}) = \{\boldsymbol{\varepsilon}(\mathbf{x}), \phi(\mathbf{x})\} : \mathbb{R}^n \rightarrow \mathcal{E}_a \subset \mathbb{R}^p \times \mathbb{R}^q$$

and a canonical function  $\bar{V}(\xi) : \mathcal{E}_a \rightarrow \mathbb{R} \cup \{\infty\}$  such that the nonconvex optimization problem (OP) can be written in the canonical form:

$$\min_{\mathbf{x}} \Pi(\mathbf{x}) = \bar{V}(\Lambda(\mathbf{x})) - U(\mathbf{x}), \quad (5)$$

where  $U(\mathbf{x}) = -\frac{1}{2}\langle \mathbf{x}, Q\mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{f} \rangle$ , and  $\bar{V}(\xi)$  is defined by

$$\bar{V}(\xi) = V(\boldsymbol{\varepsilon}(\mathbf{x})) + \Psi(\phi(\mathbf{x})), \quad (6)$$

in which,  $V(\boldsymbol{\varepsilon}(\mathbf{x})) = W(\mathbf{x})$  and

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

We assume that  $V$  is convex in this paper, and let  $\partial$  denote the sub-gradients set of a convex function such the same as in [3]. Then, we can express the stationary condition for (VI) or (OP) as

$$0 \in \partial \Pi(\mathbf{x}). \tag{7}$$

For any given  $\zeta \in \mathbb{R}^{p+q}$ , the Fenchel sup-conjugate function  $\bar{V}^\sharp$  of the convex function  $\bar{V}$  is given as

$$\bar{V}^\sharp(\zeta) = \sup_{\xi \in \mathcal{E}_a} \{\langle \xi, \zeta \rangle - \bar{V}(\xi)\} = V^\sharp(\sigma) + \Psi^\sharp(\lambda),$$

where

$$\Psi^\sharp(\lambda) = \sup_{\phi \leq 0} \{\langle \phi, \lambda \rangle - \Psi(\phi)\} = \begin{cases} 0 & \text{if } \lambda \geq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

We let

$$\mathcal{S}_a = \text{dom}(\bar{V}^\sharp) = \{\zeta \in \mathbb{R}^{p+q} \mid \bar{V}^\sharp(\zeta) < +\infty\}.$$

By the definition introduced in [10], the pair  $(\xi, \zeta)$  is called an *extended canonical duality pair* on  $\mathcal{E}_a \times \mathcal{S}_a$  if the following duality relations hold on  $\mathcal{E}_a \times \mathcal{S}_a$

$$\zeta \in \partial \bar{V}(\xi) \Leftrightarrow \xi \in \partial \bar{V}^\sharp(\zeta) \Leftrightarrow \langle \xi, \zeta \rangle = \bar{V}(\xi) + \bar{V}^\sharp(\zeta). \tag{8}$$

Thus, for this canonical duality pair,  $W(\mathbf{x}) + \Psi(\phi(\mathbf{x}))$  can be replaced by  $\bar{V}(\Lambda(\mathbf{x})) = \langle \Lambda(\mathbf{x}), \zeta \rangle - \bar{V}^\sharp(\zeta)$ , the so-called *total complementary function*  $\mathcal{E}(\mathbf{x}, \zeta)$  can be defined by

$$\mathcal{E}(\mathbf{x}, \zeta) = \langle \Lambda(\mathbf{x}), \zeta \rangle - \bar{V}^\sharp(\zeta) - U(\mathbf{x}).$$

For a given  $\zeta \in \mathcal{S}_a$ , the canonical dual function can be obtained as

$$P^d(\zeta) = \text{sta}_{\mathbf{x}}\{\mathcal{E}(\mathbf{x}, \zeta) : \mathbf{x} \in \mathbb{R}^n\} = U^\Lambda(\zeta) - \bar{V}^\sharp(\zeta),$$

where  $U^\Lambda(\zeta)$  is called the  $\Lambda$ -conjugate of  $U$ , defined by

$$U^\Lambda(\zeta) = \text{sta}_{\mathbf{x}}\{\langle \Lambda(\mathbf{x}), \zeta \rangle - U(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\},$$

and the notation  $\text{sta}_x\{\dots\}$  stands for solving the stationary point problem given in  $\{\dots\}$  with respect to  $x$ . Let  $\mathcal{S}_c$  denotes the feasible space such that  $U^\Lambda$  is well-defined on

$\mathcal{S}_c$ , then the canonical dual problem ( $\mathcal{P}^d$ ) is eventually proposed as the following

$$(\mathcal{P}^d) : \max\{P^d(\boldsymbol{\zeta}) : \boldsymbol{\zeta} \in \mathcal{S}_c\}. \quad (9)$$

In many applications, the geometrical operator  $\Lambda$  is usually a vector-valued quadratic function:

$$\begin{aligned} \Lambda(\mathbf{x}) &= (\varepsilon(\mathbf{x}), \phi(\mathbf{x})) \\ &= \left\{ \frac{1}{2} \langle \mathbf{x}, \mathbf{B}_k \mathbf{x} \rangle + \langle \mathbf{x}, b_k \rangle - d_k, \frac{1}{2} \langle \mathbf{x}, C_s \mathbf{x} \rangle + \langle \mathbf{x}, c_s \rangle - e_s \right\}, \end{aligned}$$

where  $b_k \in \mathbb{R}^n$  and  $B_k \in \mathbb{R}^{n \times n}$  is a given symmetric matrix for each  $k \in \{1, 2, \dots, p\}$ ;  $c_s \in \mathbb{R}^n$  for each  $s \in \{1, 2, \dots, q\}$ ,  $C_s \in \mathbb{R}^{n \times n}$  is a given positive definite matrix for each  $s \in \{1, 2, \dots, q\}$ ;  $d \in \mathbb{R}^p$  and  $e \in \mathbb{R}^q$ . In this case, the canonical dual function has an explicit form

$$P^d(\boldsymbol{\zeta}) = -\frac{1}{2} \langle \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}), \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle - \langle d, \boldsymbol{\sigma} \rangle - \langle e, \boldsymbol{\lambda} \rangle - \bar{V}^\sharp(\boldsymbol{\zeta}),$$

where  $\boldsymbol{\zeta} = (\boldsymbol{\sigma}, \boldsymbol{\lambda})$  and

$$\mathbf{G}(\boldsymbol{\zeta}) = \mathbf{Q} + \sum_{k=1}^p \mathbf{B}_k \sigma_k + \sum_{s=1}^q C_s \lambda_s, \quad \boldsymbol{\tau}(\boldsymbol{\zeta}) = \mathbf{f} - \sum_{k=1}^p b_k \sigma_k - \sum_{s=1}^q c_s \lambda_s.$$

The notation  $\mathbf{G}^\dagger$  stands for the Moore–Penrose inverse of  $\mathbf{G}$ . We use  $\mathcal{C}_{ol} \mathbf{G}$  to denote the column space of  $\mathbf{G}$ , then the dual feasible space  $\mathcal{S}_c$  can be defined by

$$\mathcal{S}_c = \{\boldsymbol{\zeta} = \{\boldsymbol{\sigma}, \boldsymbol{\lambda}\} \in \mathcal{S}_a \mid \boldsymbol{\tau}(\boldsymbol{\zeta}) \in \mathcal{C}_{ol} \mathbf{G}(\boldsymbol{\zeta})\}.$$

Now, consider the derivative of  $P^d$ , we first have

$$U^A(\boldsymbol{\zeta}) = -\frac{1}{2} \langle \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}), \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle - \langle d, \boldsymbol{\sigma} \rangle - \langle e, \boldsymbol{\lambda} \rangle.$$

It follows that

$$\nabla_{\sigma_k} U^A = \frac{1}{2} \langle \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}), B_k \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle + \langle b_k, \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle - d_k, \quad k = 1, 2, \dots, p \quad (10)$$

and

$$\nabla_{\lambda_s} U^A = \frac{1}{2} \langle \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}), C_s \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle + \langle c_s, \mathbf{G}^\dagger(\boldsymbol{\zeta}) \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle - e_s, \quad s = 1, 2, \dots, q. \quad (11)$$

Therefore, the dual problem associated with (VI) can be given as

$$(DVI) : \quad \langle \nabla \bar{P}^d(\boldsymbol{\zeta}), \boldsymbol{\zeta} - \bar{\boldsymbol{\zeta}} \rangle \geq 0, \quad \forall \boldsymbol{\zeta} \in \mathbb{R}^p \times \mathbb{R}_+^q,$$

where

$$\bar{P}^d(\zeta) = \frac{1}{2} \langle \mathbf{G}^\dagger(\zeta) \boldsymbol{\tau}(\zeta), \boldsymbol{\tau}(\zeta) \rangle + \langle d, \sigma \rangle + \langle e, \lambda \rangle + V^\sharp(\sigma).$$

The stationary conditions for both  $(\mathcal{P}^d)$  and  $(DVI)$  is given as:

$$0 \in \partial \bar{P}^d(\zeta). \tag{12}$$

Similar to Lemma 1, we have the following result.

**Lemma 2.** *If  $\bar{\zeta}$  solves  $(DVI)$ , then it is a stationary point of  $(\mathcal{P}^d)$  or  $(DVI)$ .*

In fact, for any stationary point  $\bar{\mathbf{x}}$  of  $(VI)$ , there is a  $\bar{\zeta} \in \partial \bar{V}(\bar{\xi})$  with  $\bar{\xi} = \Lambda(\bar{\mathbf{x}})$  such that

$$\mathbf{G}(\bar{\zeta})\bar{\mathbf{x}} - \boldsymbol{\tau}(\bar{\zeta}) = 0. \tag{13}$$

On the other hand, if we can solve the dual problem  $(DVI)$  for  $\bar{\zeta}$ , then the solution  $\bar{\mathbf{x}}$  to the primal problem  $(VI)$  should be obtained via the above relation (13).

**Theorem 1.** *If  $\bar{\zeta}$  is a solution of  $(DVI)$ , then  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta})$  is a KKT point of the problem  $(VI)$ . Moreover, if  $\bar{\mathbf{x}}$  is a solution of  $(VI)$  with  $\mathbf{G}(\bar{\zeta})$  is invertible, then  $\bar{\zeta}$  is a KKT point of  $(DVI)$ .*

**Proof.** First, assume that  $\bar{\zeta} \in \mathbb{R}^p \times \mathbb{R}_+^q$  is a solution of  $(DVI)$ , it is obvious that  $\bar{\zeta} \in \mathcal{S}_c \cap \text{dom}(\bar{V}^\sharp)$  and we have that

$$0 \in \partial P^d(\bar{\zeta}) = \nabla U^\Lambda(\bar{\zeta}) - \partial \bar{V}^\sharp(\bar{\zeta}). \tag{14}$$

Let

$$\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta}),$$

then by (10) and (11), we have

$$\nabla U^\Lambda(\bar{\zeta}) = \Lambda(\bar{\mathbf{x}}).$$

By (14), we have  $\bar{\xi} = \Lambda(\bar{\mathbf{x}}) \in \partial \bar{V}^\sharp(\bar{\zeta})$ . which is equivalent to  $\bar{\zeta} \in \partial \bar{V}(\bar{\xi})$ . It follows that

$$\begin{aligned} \partial \Pi(\bar{\mathbf{x}}) &= \partial_{\bar{\xi}} \bar{V}(\bar{\xi}) \nabla \Lambda(\bar{\mathbf{x}}) + Q\bar{\mathbf{x}} - \mathbf{f} \\ &\ni \bar{\zeta} \nabla \Lambda(\bar{\mathbf{x}}) + Q\bar{\mathbf{x}} - \mathbf{f} \\ &= \mathbf{G}(\bar{\zeta})\bar{\mathbf{x}} - \boldsymbol{\tau}(\bar{\zeta}) = 0. \end{aligned}$$

This shows that  $\bar{\mathbf{x}}$  is a KKT of  $(VI)$ .

Now, we assume that  $\bar{\mathbf{x}}$  is a solution of  $(VI)$  with  $\mathbf{G}(\bar{\zeta})$  is invertible. By (13), we have  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta})$ . Therefore,  $\bar{\zeta}$  is a KKT point of  $(DVI)$ .  $\square$

*Remark 1.* In fact, for a stationary point  $\bar{\mathbf{x}}$  of (VI), the associated stationary point  $\bar{\zeta}$  may not be unique if the linear independence constraint qualification (LICQ) does not hold at  $\bar{\mathbf{x}}$  for (VI). This situation is different from the canonical dual for unconstrained optimization problems.

### 3 Global and Local Extremalities

Because the feasible solution set  $\mathcal{K}$  is convex and we assume that the basic constraint qualification always holds at any feasible point, then any local minimizer of (OP) is a solution of (VI). In fact, we can give some sufficient conditions for a stationary point of (OP) to be a local minimizer or a global minimizer. In this section, we assume that  $V$  is twice differentiable. In order to state the results of this section, we need to give a new notation. For any given  $\bar{\mathbf{x}} \in \mathbb{R}^n$ , let  $\mathcal{B}(\bar{\mathbf{x}})$  be a  $p \times n$  matrix, whose  $k$ -th row is given as  $\mathcal{B}_k(\bar{\mathbf{x}}) = \bar{\mathbf{x}}^\top B_k$ ,  $k = 1, 2, \dots, p$ .

**Theorem 2.** *Suppose that  $\bar{\zeta}$  is a solution of (DVI) and  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta})$ . If the matrix  $\mathcal{B}^\top(\bar{\mathbf{x}})\nabla_{\varepsilon\varepsilon}^2 V(\bar{\varepsilon})\mathcal{B}(\bar{\mathbf{x}}) + \mathbf{G}(\bar{\zeta})$  is positive definite, then  $\bar{\mathbf{x}}$  is a local minimizer of the problem (OP).*

**Proof.** Consider the Lagrangian function of the problem (OP):

$$L(\mathbf{x}, \lambda) = W(\mathbf{x}) + \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} - \mathbf{f}^\top \mathbf{x} + \lambda^\top \phi(\mathbf{x}),$$

we know that

$$\nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\lambda}) = \mathbf{G}(\bar{\zeta})\bar{\mathbf{x}} - \boldsymbol{\tau}(\bar{\zeta}) = 0.$$

We also have that

$$\nabla_{\mathbf{xx}}^2 L(\bar{\mathbf{x}}, \bar{\lambda}) = \mathcal{B}^\top(\bar{\mathbf{x}})\nabla_{\varepsilon\varepsilon}^2 V(\bar{\varepsilon})\mathcal{B}(\bar{\mathbf{x}}) + \mathbf{G}(\bar{\zeta}).$$

By assumption of the theorem,  $\nabla_{\mathbf{xx}}^2 L(\bar{\mathbf{x}}, \bar{\lambda})$  is positive definite. Then, the vector  $\bar{\mathbf{x}}$  is a local minimizer of the problem (OP).  $\square$

**Corollary 1.** *Suppose that  $\bar{\zeta}$  is a solution of (DVI) and  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta})$ . If  $\mathbf{G}(\bar{\zeta})$  is positive definite, then  $\bar{\mathbf{x}}$  is a local minimizer of the problem (OP).*

**Proof.** Because  $V$  is convex, we have  $\mathcal{B}^\top(\bar{\mathbf{x}})\nabla_{\varepsilon\varepsilon}^2 V(\bar{\varepsilon})\mathcal{B}(\bar{\mathbf{x}})$  is positive semi-definite for any given  $\bar{\mathbf{x}}$ . By the assumption of this proposition, we know that  $\mathbf{G}(\bar{\zeta})$  is positive definite, then we must have that  $\nabla_{\mathbf{xx}}^2 L(\bar{\mathbf{x}}, \bar{\lambda})$  is positive definite, hence we can conclude that  $\bar{\mathbf{x}}$  is a local minimizer of (OP) by Theorem 2.  $\square$

In fact, the above Corollary can be enhanced and we give a sufficient condition for a global minimizer of (OP).

**Theorem 3.** *Suppose that  $\bar{\zeta}$  is a solution of (DVI) and  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta})$ . If  $\mathbf{G}(\bar{\zeta})$  is positive semi-definite, then  $\bar{\mathbf{x}}$  is a global minimizer of the problem (OP).*



**Proof.** If  $\bar{\mathbf{x}}$  is a stationary point of  $(OP)$ , then we have that

$$\mathbf{G}(\bar{\zeta})\bar{\mathbf{x}} - \boldsymbol{\tau}(\bar{\zeta}) = 0.$$

Now, consider the function  $\Lambda(\mathbf{x})$ , we have

$$\varepsilon_k(\mathbf{x}) - \varepsilon_k(\bar{\mathbf{x}}) = \langle \mathbf{x} - \bar{\mathbf{x}}, B_k\bar{\mathbf{x}} + b_k + \frac{1}{2}B_k(\mathbf{x} - \bar{\mathbf{x}}) \rangle, \quad k = 1, 2, \dots, p$$

and

$$\phi_s(\mathbf{x}) - \phi_s(\bar{\mathbf{x}}) = \langle \mathbf{x} - \bar{\mathbf{x}}, C_s\bar{\mathbf{x}} + c_s + \frac{1}{2}C_s(\mathbf{x} - \bar{\mathbf{x}}) \rangle, \quad s = 1, 2, \dots, q.$$

We also have that

$$U(\mathbf{x}) - U(\bar{\mathbf{x}}) = \langle \mathbf{x} - \bar{\mathbf{x}}, Q\bar{\mathbf{x}} - \mathbf{f} + \frac{1}{2}Q(\mathbf{x} - \bar{\mathbf{x}}) \rangle.$$

Then, we have

$$\begin{aligned} \Pi(\mathbf{x}) - \Pi(\bar{\mathbf{x}}) &\geq \langle \bar{\sigma}, \varepsilon(\mathbf{x}) - \varepsilon(\bar{\mathbf{x}}) \rangle + \langle \bar{\lambda}, \phi(\mathbf{x}) - \phi(\bar{\mathbf{x}}) \rangle + U(\mathbf{x}) - U(\bar{\mathbf{x}}) \\ &= \langle \mathbf{x} - \bar{\mathbf{x}}, \mathbf{G}(\bar{\zeta})\bar{\mathbf{x}} - \boldsymbol{\tau}(\bar{\zeta}) \rangle + \frac{1}{2}\langle \mathbf{x} - \bar{\mathbf{x}}, \mathbf{G}(\bar{\zeta})(\mathbf{x} - \bar{\mathbf{x}}) \rangle \\ &= \frac{1}{2}\langle \mathbf{x} - \bar{\mathbf{x}}, \mathbf{G}(\bar{\zeta})(\mathbf{x} - \bar{\mathbf{x}}) \rangle \\ &\geq 0 \end{aligned}$$

for any  $\mathbf{x} \in \mathbb{R}^n$ . Now, we have proved that  $\bar{\mathbf{x}}$  is a global minimizer of  $(OP)$  and the proof of the theorem is finished.  $\square$

## 4 Existence of the Solution

In order to discuss the existence of solution to the problem, we need the following sets:

$$\mathcal{S}_c^+ = \{\bar{\zeta} \in \mathbb{R}^p \times \mathbb{R}_+^q \mid \gamma_{\mathbf{G}}(\bar{\zeta}) > 0\},$$

$$\bar{\mathcal{S}}_c = \{\bar{\zeta} \in \mathbb{R}^p \times \mathbb{R}_+^q \mid \gamma_{\mathbf{G}}(\bar{\zeta}) = 0\},$$

where  $\gamma_{\mathbf{G}}(\bar{\zeta})$  is the smallest eigenvalue of the matrix  $\mathbf{G}(\bar{\zeta})$  for any given  $\bar{\zeta} \in \mathcal{S}_c$ . We also need to define a  $n \times (p + q)$  matrix  $\bar{\mathcal{E}}(\bar{\mathbf{x}})$  for any given  $\bar{\mathbf{x}}$  with its  $k$ -th column is given as  $\bar{\mathcal{E}}_k(\bar{\mathbf{x}}) = B_k\bar{\mathbf{x}} + b_k$ ,  $k = 1, 2, \dots, p$  and  $(p + s)$ -th column as  $\bar{\mathcal{E}}_{p+s}(\bar{\mathbf{x}}) = C_s\bar{\mathbf{x}} + c_s$ ,  $s = 1, 2, \dots, q$ . The notation  $\|\cdot\|$  can be any norm of a vector in this paper. Then we have the following result.

**Theorem 4.** *The canonical dual function  $P^d$  is concave in  $\mathcal{S}_c^+$ .*

**Proof.** Because  $\bar{V}$  is convex, we need only to prove that  $U^A$  is concave in order to show that  $P^d$  is concave. We have

$$\frac{\partial U^A}{\partial \sigma_k} = \frac{1}{2} \langle \mathbf{G}^\dagger(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma}), B_k \mathbf{G}^\dagger(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma}) \rangle + \langle b_k, \mathbf{G}^\dagger(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma}) \rangle - d_k, \quad k = 1, 2, \dots, p$$

and

$$\frac{\partial U^A}{\partial \lambda_s} = \frac{1}{2} \langle \mathbf{G}^\dagger(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma}), C_s \mathbf{G}^\dagger(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma}) \rangle + \langle c_s, \mathbf{G}^\dagger(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma}) \rangle - e_s, \quad s = 1, 2, \dots, q.$$

Let  $\mathbf{x} = \mathbf{G}^\dagger(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma})$  for any  $\boldsymbol{\varsigma} \in \mathcal{S}_c^+$ , we have

$$\frac{\partial^2 U^A}{\partial \boldsymbol{\varsigma} \partial \boldsymbol{\varsigma}} = -\bar{\boldsymbol{\epsilon}}^\top(\mathbf{x}) \mathbf{G}^\dagger(\boldsymbol{\varsigma}) \bar{\boldsymbol{\epsilon}}(\mathbf{x}).$$

Hence, the Hessian matrix  $\partial^2 U^A / \partial \boldsymbol{\varsigma} \partial \boldsymbol{\varsigma}$  is negative semi-definite and  $U^A$  is concave, then  $P^d$  is concave in  $\mathcal{S}_c^+$ .  $\square$

We denote  $\mathcal{T}_{\mathbf{G}}(\boldsymbol{\varsigma}) = \{\xi \in \mathbb{R}^n \mid \mathbf{G}(\boldsymbol{\varsigma})\xi = \gamma_{\mathbf{G}}(\boldsymbol{\varsigma})\xi \quad \forall \boldsymbol{\varsigma} \in \mathcal{S}_c^+ \cup \bar{\mathcal{S}}_c^+\}$ .

**Theorem 5.** Assume that  $\text{dom}(V^\sharp)$  is closed,  $\text{dom}(V^\sharp) \cap \mathcal{S}_c^+ \neq \emptyset$  and

$$\lim_{\|\boldsymbol{\varsigma}\| \rightarrow \infty} \frac{V^\sharp(\boldsymbol{\varsigma})}{\|\boldsymbol{\varsigma}\|} = +\infty.$$

If  $\langle \boldsymbol{\tau}(\boldsymbol{\varsigma}), \xi \rangle \neq 0$  for any  $\xi \in \mathcal{T}_{\mathbf{G}}(\boldsymbol{\varsigma})$  and any  $\boldsymbol{\varsigma} \in \bar{\mathcal{S}}_c^+$ , then there must be a  $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_c^+$  such that  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\boldsymbol{\varsigma}})\boldsymbol{\tau}(\bar{\boldsymbol{\varsigma}})$  is a global minimizer of the problem (OP).

**Proof.** In order to simplify the proof, we assume that  $\text{dom}(V^\sharp) = \mathbb{R}^{p+q}$ . For any  $\delta > 0$ , we denote a set

$$\Omega(\delta) = \left\{ \boldsymbol{\varsigma} \in \mathcal{S}_c \mid \|\boldsymbol{\varsigma}\| \leq \delta, \quad \gamma(\mathbf{G}(\boldsymbol{\varsigma})) \geq \frac{1}{\delta} \right\}.$$

Let

$$\Gamma(\delta) = \sup_{\boldsymbol{\varsigma} \in \mathcal{S}_c^+ \setminus \Omega(\delta)} P^d(\boldsymbol{\varsigma}),$$

for any  $\delta > 0$ . We will show that

$$\lim_{\delta \rightarrow +\infty} \Gamma(\delta) = -\infty.$$

By contradiction, assume that this conclusion is not true, then there is a sequence  $\{\boldsymbol{\varsigma}_i\}_{i=1,2,\dots} \subseteq \mathcal{S}_c^+$  with  $\boldsymbol{\varsigma}_i \in \mathcal{S}_c^+ \setminus \Omega(i)$  and  $P^d(\boldsymbol{\varsigma}_i) \geq \Gamma(i) - \frac{1}{i}$  for any  $i = 1, 2, \dots$ , such that

$$\lim_{i \rightarrow +\infty} P^d(\mathcal{G}_i) = M, \tag{15}$$

where  $M \in \mathbb{R}$ . If  $\{\mathcal{G}_i\}_{i=1,2,\dots}$  is unbounded, then there is  $\mathcal{K}_1 \subseteq \{1, 2, \dots, \}$  such that

$$\lim_{i \rightarrow +\infty, i \in \mathcal{K}_1} \|\mathcal{G}_i\| = \infty.$$

Then, we have that

$$\frac{1}{2} \langle \mathbf{G}^\dagger(\mathcal{G}_i) \boldsymbol{\tau}(\mathcal{G}_i), \boldsymbol{\tau}(\mathcal{G}_i) \rangle \geq 0, \quad \forall i \in \mathcal{K}_1$$

and

$$\lim_{i \rightarrow +\infty, i \in \mathcal{K}_1} \langle d, \sigma_i \rangle + \langle e, \lambda_i \rangle + \bar{V}^\#(\mathcal{G}_i) = +\infty.$$

It follows that

$$\lim_{i \rightarrow +\infty, i \in \mathcal{K}_1} P^d(\mathcal{G}_i) = -\infty,$$

which contradicts (15).

Now, assume that  $\{\mathcal{G}_i\}_{i=1,2,\dots}$  is bounded. Let  $\xi_i \in \mathcal{T}_{\mathbf{G}}(\mathcal{G}_i)$  with  $\|\xi_i\| = 1$  for  $i = 1, 2, \dots$ . Then there must be a  $\mathcal{K}_2 \subseteq \{1, 2, \dots\}$  such that

$$\begin{aligned} \lim_{i \rightarrow +\infty, i \in \mathcal{K}_2} \gamma_{\mathbf{G}}(\mathcal{G}_i) &= 0 \\ \lim_{i \rightarrow +\infty, i \in \mathcal{K}_2} \mathcal{G}_i &= \bar{\mathcal{G}} \\ \lim_{i \rightarrow +\infty, i \in \mathcal{K}_2} \xi_i &= \bar{\xi}. \end{aligned}$$

Now, we have that

$$\bar{\xi} \in \mathcal{T}_{\mathbf{G}}(\bar{\mathcal{G}}), \quad \bar{\mathcal{G}} \in \bar{\mathcal{S}}_c, \quad \|\bar{\xi}\| = 1.$$

Let

$$\boldsymbol{\tau}_\xi(\mathcal{G}_i) = \langle \boldsymbol{\tau}(\mathcal{G}_i), \xi_i \rangle \xi_i, \quad \boldsymbol{\tau}_\xi^c(\mathcal{G}_i) = \boldsymbol{\tau}(\mathcal{G}_i) - \langle \boldsymbol{\tau}(\mathcal{G}_i), \xi_i \rangle \xi_i,$$

for  $i = 1, 2, \dots$ . Because

$$\xi_i \in \mathcal{T}_{\mathbf{G}}(\mathcal{G}_i), \quad i = 1, 2, \dots,$$

we have

$$\langle \boldsymbol{\tau}_\xi(\mathcal{G}_i), \mathbf{G}(\mathcal{G}_i) \boldsymbol{\tau}_\xi^c(\mathcal{G}_i) \rangle = 0, \quad i \in \mathcal{K}_2.$$

Therefore,

$$\langle \boldsymbol{\tau}_\xi(\mathcal{G}_i), \mathbf{G}^\dagger(\mathcal{G}_i) \boldsymbol{\tau}_\xi^c(\mathcal{G}_i) \rangle = 0, \quad i \in \mathcal{K}_2.$$

Now, we have

$$\begin{aligned}
 U^\Lambda(\mathfrak{G}_i) &= -\frac{1}{2}\langle \mathbf{G}^\dagger(\mathfrak{G}_i)\boldsymbol{\tau}(\mathfrak{G}_i), \boldsymbol{\tau}(\mathfrak{G}_i) \rangle - \langle d, \sigma_i \rangle - \langle e, \lambda_i \rangle \\
 &= -\frac{1}{2}\langle \boldsymbol{\tau}_\xi(\mathfrak{G}_i), \mathbf{G}^\dagger(\mathfrak{G}_i)\boldsymbol{\tau}_\xi(\mathfrak{G}_i) \rangle - \frac{1}{2}\langle \boldsymbol{\tau}_\xi^c(\mathfrak{G}_i), \mathbf{G}^\dagger(\mathfrak{G}_i)\boldsymbol{\tau}_\xi^c(\mathfrak{G}_i) \rangle - \langle d, \sigma_i \rangle - \langle e, \lambda_i \rangle \\
 &\leq -\frac{1}{2\gamma_{\mathbf{G}}(\mathfrak{G}_i)}\langle \boldsymbol{\tau}(\mathfrak{G}_i), \xi_i \rangle^2 - \langle d, \sigma_i \rangle - \langle e, \lambda_i \rangle,
 \end{aligned}$$

for  $i \in \mathcal{H}_2$ . Then, we have

$$\lim_{i \rightarrow +\infty, i \in \mathcal{H}_2} U^\Lambda(\mathfrak{G}_i) = -\infty.$$

Therefore,

$$\lim_{i \rightarrow +\infty, i \in \mathcal{H}_2} P^d(\mathfrak{G}_i) = -\infty,$$

which contradicts (15). Now, we have proved that

$$\lim_{\delta \rightarrow +\infty} \Gamma(\delta) = -\infty.$$

Choose a  $\bar{\mathfrak{G}} \in \mathcal{S}_c^+$ , there must be a  $\bar{\delta} > 0$  such that

$$\Gamma(\bar{\delta}) \leq P^d(\bar{\mathfrak{G}}).$$

Because  $\Omega(\bar{\delta})$  is compact, then there is a  $\tilde{\mathfrak{G}} \in \Omega(\bar{\delta})$  such that

$$P^d(\tilde{\mathfrak{G}}) = \max_{\mathfrak{G} \in \Omega(\bar{\delta})} P^d(\mathfrak{G}).$$

It follows that

$$P^d(\tilde{\mathfrak{G}}) = \max_{\mathfrak{G} \in \mathcal{S}_c^+} P^d(\mathfrak{G}).$$

Then,  $\tilde{\mathfrak{G}}$  is stationary point of  $(DVI)$  with  $\mathbf{G}(\tilde{\mathfrak{G}})$  is positive definite. By Theorem 3,  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\tilde{\mathfrak{G}})\boldsymbol{\tau}(\tilde{\mathfrak{G}})$  is a global minimizer of the problem  $(OP)$ . □

In fact, the condition  $\lim_{\mathfrak{G} \rightarrow \infty} V^\sharp(\mathfrak{G})/\|\mathfrak{G}\| = +\infty$  can be weaken in some cases.

**Theorem 6.** *In case  $b_k = 0, k = 1, 2, \dots, p$  and  $c_s = 0, s = 1, 2, \dots, q$ , assume that  $\text{dom}(V^\sharp)$  is closed,  $\text{dom}(V^\sharp) \cap \mathcal{S}_c^+ \neq \emptyset$  and  $\lim_{\|\mathfrak{G}\| \rightarrow \infty} V^\sharp(\mathfrak{G}) = +\infty$ . If  $\langle \boldsymbol{\tau}(\mathfrak{G}), \xi \rangle \neq 0$  for any  $\xi \in \mathcal{T}_{\mathbf{G}}(\mathfrak{G})$  and any  $\mathfrak{G} \in \tilde{\mathcal{S}}_c^+$ , then there must be a vector  $\bar{\mathfrak{G}} \in \mathcal{S}_c^+$  such that  $\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\mathfrak{G}})\boldsymbol{\tau}(\bar{\mathfrak{G}})$  is a global minimizer of the problem  $(OP)$ .*

**Proof.** It is similar with the Theorem 5. □

### 5 Some Examples

We now present some examples to show how to apply the canonical dual theory for solving real problems. Consider the problem (VI) with

$$F(\mathbf{x}) = \left(\frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) - \mu\right) \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} + \begin{pmatrix} 2\mathbf{x}_1 + \mathbf{x}_2 \\ \mathbf{x}_1 + 2\mathbf{x}_2 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

and

$$\mathcal{K} = \left\{ \mathbf{x} \mid \frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) \leq e \right\}.$$

The primal optimization problem (OP) of (VI) is given as follows:

$$\begin{aligned} \min_{\mathbf{x}} P(\mathbf{x}) &= \frac{1}{2}(\frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) - \mu)^2 + (\mathbf{x}_1^2 + \mathbf{x}_2\mathbf{x}_2 + \mathbf{x}_2^2) - (2\mathbf{x}_1 + 2\mathbf{x}_2) \\ \text{s.t.} \quad &\frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) \leq e. \end{aligned}$$

For the dual problem, we have that

$$\mathbf{G}^\dagger(\boldsymbol{\varsigma}) = \frac{1}{2} \begin{pmatrix} 1, & 1 \\ 1, & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{3+(\sigma+\lambda)}, & 0 \\ 0, & \frac{1}{1+(\sigma+\lambda)} \end{pmatrix} \begin{pmatrix} 1, & 1 \\ 1, & -1 \end{pmatrix}.$$

It follows that

$$\begin{aligned} \nabla U^A(\boldsymbol{\varsigma}) &= \frac{1}{2}(2, 2) \begin{pmatrix} 1, & 1 \\ 1, & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{(3+(\sigma+\lambda))^2}, & 0 \\ 0, & \frac{1}{(1+(\sigma+\lambda))^2} \end{pmatrix} \begin{pmatrix} 1, & 1 \\ 1, & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ &= \frac{8}{(3+(\sigma+\lambda))^2} - \begin{pmatrix} \mu \\ e \end{pmatrix}. \end{aligned}$$

Then the critical condition of the dual problem becomes

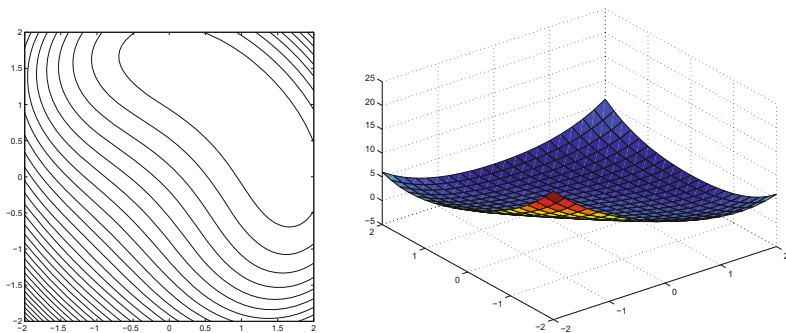
$$\begin{aligned} \frac{4}{(3+(\sigma+\lambda))^2} - \sigma &= \mu \\ \frac{4}{(3+(\sigma+\lambda))^2} - e &\in \partial\Psi^\sharp(\lambda). \end{aligned}$$

We consider three cases with various values of  $\mu$  and  $e$ .

**Example 1** First, we let  $\mu = 2$  and  $e = 2$ , then we can find that  $\bar{\mathbf{x}} = (1, 1)$  is a stationary point of the problem. At this point, we have  $\bar{\xi} = (-1, -1)$  and  $\bar{\varsigma} = (-1, 0)$ . Note that

$$\mathbf{G}(\bar{\varsigma}) = \begin{pmatrix} 1, & 1 \\ 1, & 1 \end{pmatrix}$$

is singular, by Theorem 3.2, we can conclude that  $(1, 1)$  is a solution of (VI) because  $\mathbf{G}(\bar{\varsigma})$  is semi-positive. This result can also be verified graphically in Fig. 1.



**Fig. 1** Contours and graph of  $P(\mathbf{x})$  in Example 1

**Example 2.** We now let  $\mu = 2$  and  $e = 1/2$ . In this case we have

$$\begin{aligned} \nabla P(\mathbf{x}) &= \left(\frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) - 2\right) \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} + \begin{pmatrix} 2\mathbf{x}_1 + \mathbf{x}_2 \\ \mathbf{x}_1 + 2\mathbf{x}_2 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2)\mathbf{x}_1 + \mathbf{x}_2 - 2 \\ \frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2)\mathbf{x}_2 + \mathbf{x}_1 - 2 \end{pmatrix} \\ &\leq \begin{pmatrix} |\mathbf{x}_1| + |\mathbf{x}_2| - 2 \\ |\mathbf{x}_2| + |\mathbf{x}_1| - 2 \end{pmatrix} \\ &< \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{aligned}$$

for any  $\mathbf{x}$  with  $\frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) \leq \frac{1}{2}$ , and the stationary point of the problem is  $\bar{\mathbf{x}} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$  with  $\bar{\boldsymbol{\zeta}} = (-\frac{3}{2}, 2\sqrt{2} - \frac{3}{2})$ , which satisfies the following critical condition of the dual problem

$$\begin{aligned} \frac{4}{(3 + (\sigma + \lambda))^2} - \sigma - 2 &= 0 \\ \frac{4}{(3 + (\sigma + \lambda))^2} - \frac{1}{2} &= 0. \end{aligned}$$

By the fact that

$$\mathbf{G}(\bar{\boldsymbol{\zeta}}) = \begin{pmatrix} 2\sqrt{2} - 1, & 1 \\ 1, & 2\sqrt{2} - 1 \end{pmatrix}$$

is positive definite, we know that  $\mathbf{x} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$  is a solution of (VI) (Fig. 2).

**Example 3.** Finally, we let  $\mu = 4/9$  and  $e = 2$ . In this case, the critical condition of the dual problem becomes

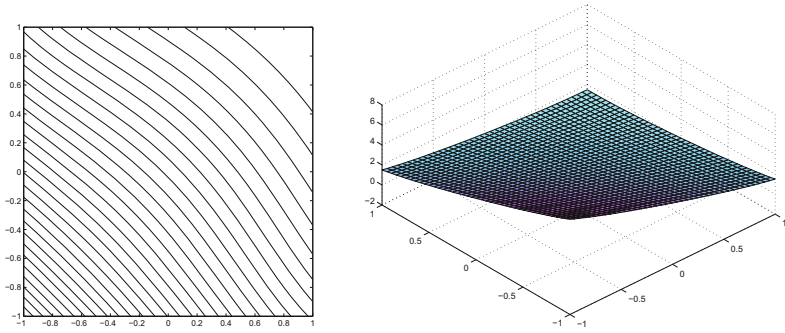


Fig. 2 Contours and graph for Example 2

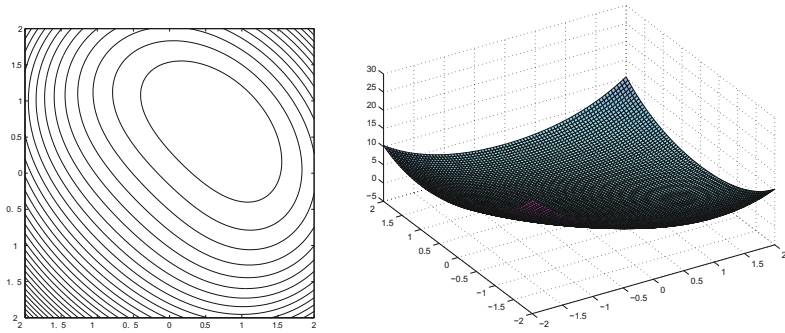


Fig. 3 Contours and graph of Example 3

$$\frac{4}{(3 + (\sigma + \lambda))^2} - \sigma = \frac{4}{9}$$

$$\frac{4}{(3 + (\sigma + \lambda))^2} - 2 \in \partial \Psi^\#(\lambda).$$

It is easily find that  $\bar{\zeta} = (0, 0)$  is a solution of the dual problem with  $\bar{\mathbf{x}} = (\frac{2}{3}, \frac{2}{3})$  a stationary point of the primal problem. Since

$$\mathbf{G}(\bar{\zeta}) = \begin{pmatrix} 2, & 1 \\ 1, & 2 \end{pmatrix}$$

is positive definite, this solution  $\bar{\mathbf{x}} = (\frac{2}{3}, \frac{2}{3})$  is a solution of the primal problem (Fig. 3).

## 6 Conclusions

In this paper, we have proposed the canonical duality theory for solving a class of non-monotone variational inequalities problems. A sufficient condition for a global minimizer of the associated optimization problem (*OP*) is presented. By the fact that the canonical dual problem is equivalent to a convex minimization problem on a convex dual feasible set  $\mathcal{S}_c^+$  with only simple non-negative constraints, which can be solved easily via well-developed methods. Existence of the solution of (*VI*) is also discussed. Examples given in the paper show the various cases that the solution may either exist on the boundary of the feasible space, or a point where  $\mathbf{G}(\bar{\zeta})$  is singular. These facts can help us to understand the difficulties of the primal problem and to develop some effective methods for solving the canonical dual problem in the future.

**Acknowledgements** The research of the first author (Guoshan Liu) was supported by the National Natural Science Foundation of China under its grand # 70771106 and the New Century Excellent Scholarship of Ministry of Education, China. The research of the second author (David Gao) was supported by US Air Force Office of Scientific Research under the grants FA2386-16-1-4082 and FA9550-17-1-0151.

## References

1. Allen, G.: Variational inequalities, complementarity problems, and duality theorems. *J. Math. Anal. Appl.* **58**, 1–10 (1977)
2. Duvaut, G., Lions, J.L.: *Inequalities in Mechanics and Physics*. Springer, Berlin (1976)
3. Ekeland, I., Temam, R.: *Convex Analysis and Variational Problems*. North-Holland, Amsterdam (1976)
4. Fichera, G.: Sul problema elastostatico di Signorini con ambigue condizioni al contorno (On the elastostatic problem of Signorini with ambiguous boundary conditions), *Rendiconti della Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, Matematiche e Naturali*, **8** **34**(2), 138–142 (1963)
5. Friedman, A.: *Variational and Free Boundary Problems*. Springer, New York (1993)
6. Gao, D.Y.: Duality theory in nonlinear buckling analysis for von Karman equations. *Stud. Appl. Math.* **94**, 423–444 (1995)
7. Gao, D.Y.: Nonlinear elastic beam theory with application in contact problems and variational approaches. *Mech. Res. Commun.* **23**(1), 11–17 (1996)
8. Gao, D.Y.: Dual extremum principles in finite deformation theory with applications in post-buckling analysis of nonlinear beam model. *Appl. Mech. Rev. ASME* **50**(11), 64–71 (1997)
9. Gao, D.Y.: Bi-complementarity and duality: a framework in nonlinear equilibria with applications to the contact problem of elastoplastic beam theory. *J. Math. Anal. Appl.* **221**, 672–697 (1998)
10. Gao, D.Y.: *Duality Principles in Nonconvex Systems: Theory, Methods and Applications*. Kluwer Academic Publishers, Dordrecht (2000)
11. Gao, D.Y.: Nonconvex semi-linear problems and canonical dual solutions. In: Gao, D.Y., Ogden, R.W. (eds.) *Advances in Mechanics and Mathematics*, vol. II, pp. 261–312. Kluwer Academic, Dordrecht (2003)
12. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. *Quart. Appl. Math.* **XLVII**(3), 487–504 (1989)



13. Gao, D.Y., Yu, H.F.: Multi-scale modelling and canonical dual finite element method in phase transitions of solids. *Int. J. Solids Struct.* **45**, 3660–3673 (2008)
14. Glowinski, R., Lions, J.L., Trémolières, R.: *Numerical Analysis of Variational Inequalities*. North-Holland, Amsterdam (1981)
15. Isac, G.: *Complementarity Problems*. Lecture Notes in Mathematics. Springer, New York (1992)
16. Kuttler, K.L., Purcell, J., Shillor, M.: Analysis and simulations of a contact problem for a nonlinear dynamic beam with a crack. *Q. J. Mech. Appl. Mech.* (2011). doi:[10.1093/qjmam/hbr018](https://doi.org/10.1093/qjmam/hbr018)
17. MBengue, M.F. Shillor, M.: Regularity result for the problem of vibrations of a nonlinear beam. *Electron. J. Differ. Equ.*, vol. 2008, No. 27, p. 112. <http://ejde.math.txstate.edu> or <http://ejde.math.unt.edu> (2008)
18. Mosco, U.: Dual variational inequalities. *J. Math. Anal. Appl.* **40**, 202–206 (1972)
19. Panagiotopoulos, P.D.: *Inequality Problems in Mechanics and Applications*. Birkhäuser, Boston (1985)
20. Santos, H.A.F.A., Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. *Int. J. Nonlinear Mech.* **47**, 240–247 (2012). doi:[10.1016/j.ijnonlinmec.2011.05.012](https://doi.org/10.1016/j.ijnonlinmec.2011.05.012)
21. Stampacchia, G.: Formes bilinéaires coercitives sur les ensembles convexes. *Comptes rendus hebdomadaires des séances de l'Académie des sci.* **258**, 4413–4416 (1964)
22. Yau, S.T., Gao, D.Y.: Obstacle problem for von Karman equations. *Adv. Appl. Math.* **13**, 123–141 (1992)