J Glob Optim (2009) 45:473–497
DOI 10.1007/s10898-009-9399-x

Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality

David Yang Gao · Ning Ruan · Hanif D. Sherali

Received: 2 November 2007 / Accepted: 6 January 2009 / Published online: 28 January 2009

Abstract This paper presents a canonical duality theory for solving a general nonconvex quadratic minimization problem with nonconvex constraints. By using the *canonical dual transformation* developed by the first author, the nonconvex primal problem can be converted into a canonical dual problem with zero duality gap. A general analytical solution form is obtained. Both global and local extrema of the nonconvex problem can be identified by the triality theory associated with the canonical duality theory. Illustrative applications to quadratic minimization with multiple quadratic constraints, box/integer constraints, and general nonconvex polynomial constraints are discussed, along with insightful connections to classical Lagrangian duality. Criteria for the existence and uniqueness of optimal solutions are presented. Several numerical examples are provided.

Keywords Canonical duality theory · Triality · Lagrangian duality · Global optimization · Integer programming

1 Introduction

In this paper, we address the following general constrained nonlinear programming problem:

$$(\mathcal{P}): \quad \min \left\{ P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{f} \quad : \quad \mathbf{x} \in \mathcal{X}_c \right\}, \tag{1}$$

where $A = \{A_{ij}\} \in \mathbb{R}^{n\times n}$ is a symmetric matrix, $\mathbf{f} \in \mathbb{R}^n$ is a given vector, the feasible space $\mathcal{X}_k \subset \mathbb{R}^n$ is defined as

$$\mathcal{X}_c = \left\{ \mathbf{x} \in \mathcal{X}_a | \;\; \mathbf{g}(\mathbf{x}) \le \mathbf{d} \in \mathbb{R}^m \right\}, \tag{2}$$

D. Y. Gao (✉) · N. Ruan
Department of Mathematics, Virginia Tech, Blacksburg, VA 24061, USA
e-mail: gao@vt.edu

D. Y. Gao · N. Ruan · H. D. Sherali
Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061, USA

where $\mathbf{g}(\mathbf{x}) = \{g_\alpha(\mathbf{x})\} : \mathcal{X}_a \to \mathbb{R}^m$ is a given vector-valued differentiable (not necessary convex) function, $\mathcal{X}_a$ is a convex open set in $\mathbb{R}^n$, and $\mathbf{d} \in \mathbb{R}^m$ is a given vector. (We follow the traditional *tensor notation* used in finite deformation theory [10], where the indices i, j represent components of entities in $\mathbb{R}^n$ or $\mathbb{R}^{n\times n}$, while α, β represent components in other spaces).

The problem $(\mathcal{P})$ involves minimizing a nonconvex quadratic function over a nonconvex feasible space. By introducing a Lagrangian multiplier vector $\boldsymbol{\sigma} \in \mathbb{R}^m_+ = \{\boldsymbol{\sigma} \in \mathbb{R}^m | \ \boldsymbol{\sigma} \geq 0\}$ to relax the inequality constraints in $\mathcal{X}_c$, the classical Lagrangian $L : \mathcal{X}_a \times \mathbb{R}^m_+ \to \mathbb{R}$ is given by

$$L(\mathbf{x}, \boldsymbol{\sigma}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{f} + \boldsymbol{\sigma}^T(\mathbf{g}(\mathbf{x}) - \mathbf{d}). \tag{3}$$

If all the components of $\mathbf{g}(\mathbf{x})$ are convex functions, and $A \succeq 0$, i.e., positive semidefinite (PSD), then Problem $(\mathcal{P})$ has a convex quadratic objective function and convex constraints, and the Lagrangian is a saddle function, i.e., $L(\mathbf{x}, \boldsymbol{\sigma})$ is convex in the primal variables $\mathbf{x}$, concave (linear) in the dual variables (Lagrange multipliers) $\boldsymbol{\sigma}$, and the Lagrangian dual problem can be easily defined by the Fenchel–Moreau–Rockafellar transformation

$$P^*(\boldsymbol{\sigma}) = \inf_{\mathbf{x}\in\mathcal{X}_a} L(\mathbf{x}, \boldsymbol{\sigma}), \tag{4}$$

where, under certain constraint qualifications that insure the existence of a Karush–Kuhn–Tucker (KKT) solution (see [1]), we have the following strong min–max duality relation:

$$\inf_{\mathbf{x}\in\mathcal{X}_c} P(\mathbf{x}) = \sup_{\boldsymbol{\sigma}\in\mathbb{R}^m_+} P^*(\boldsymbol{\sigma}). \tag{5}$$

In this case, the problem can be solved easily by any well-developed convex programming technique.

However, due to the assumed nonconvexity of Problem $(\mathcal{P})$, the Lagrangian $L(\mathbf{x}, \boldsymbol{\sigma})$ is no longer a saddle function and the Fenchel-Young inequality leads to only the following weak duality relation in general:

$$\inf_{\mathbf{x}\in\mathcal{X}_c} P(\mathbf{x}) \geq \sup_{\boldsymbol{\sigma}\in\mathbb{R}^m_+} P^*(\boldsymbol{\sigma}). \tag{6}$$

The slack θ in the inequality (6) is called the *duality gap* in global optimization, where possibly, $\theta = \infty$. This duality gap shows that the well-developed Fenchel–Moreau–Rockafellar duality theory can be used only for solving convex minimization problems. Also, due to the nonconvexity of the objective function and/or constraints, the problem may have multiple local solutions. The identification of a global minimizer has been a fundamentally challenging task in global optimization.

The classical Lagrangian duality theory was originally developed in the context of analytical mechanics [38]. In the realm of linear elasticity, the primal problem $(\mathcal{P})$ is usually a convex variational problem in functional space, while its dual $(\mathcal{P}^*)$ is called the complementary variational problem. In mechanics and mathematical physics, the terminology “complementary” connotes the strong (or perfect) duality, i.e., the so-called *canonical duality* [10], and the strong duality relation (5) is also called the complementary variational principle. However, in finite deformation elasticity, the primal function $P(\mathbf{x})$ is usually nonconvex and the existence of a purely stress-(dual variable)-based complementary variational principle has invoked a lively debate existing for more than 50 years, since E. Reissner (1953) [37,39,41].

In order to close the duality gap inherent in the classical Lagrange duality theory, a so-called *canonical duality theory* has been developed, first in nonconvex analysis [27] and mechanics [7,10], and then in global optimization [12,16,26]. This new theory is composed mainly of a potentially useful *canonical dual transformation*, an associated *complementary-dual principle*, and a *triality theory*, whose components comprise a saddle min–max duality and two pairs of double-min, double-max dualities. The canonical dual transformation can be used to formulate perfect dual problems without a duality gap; the complementary-dual principle shows that the primal problem is equivalent to its canonical dual in the sense that they have the same KKT points; while the triality theory can be used to identify both global and local extrema. The key step in the canonical dual transformation is to introduce a *geometrical operator* $\boldsymbol{\xi} = \Lambda(\mathbf{x})$ and a canonical function $V(\boldsymbol{\xi})$ (whose Legendre or Fenchel conjugate can be uniquely defined) such that the nonconvex constraint can be written in the canonical form $\mathbf{g}(\mathbf{x}) = V(\Lambda(\mathbf{x}))$. Thus, the complementary (i.e., the canonical dual) function can be uniquely formulated via the Legendre-Fenchel transformation. The original idea of this canonical dual transformation comes from the joint work of Gao and Strang [27] on nonconvex/nonsmooth variational problems. They discovered that a necessary condition for a minimizer depends on the Gâteaux derivative Λ_t of the geometrical operator, while a sufficient condition and the so-called *complementary gap function* depends on the complementary operator $\Lambda_c(\mathbf{x}) := \Lambda(\mathbf{x}) - \Lambda_t\mathbf{x}$. This complementary gap function closes the duality gap present in Lagrangian duality and plays an important role in nonconvex analysis and global optimization. Some open problems in one-dimensional nonlinear elasticity have thus been solved recently [22,23]. A comprehensive review on the canonical duality theory and connections between nonconvex mechanics and global optimization appear in Gao and Sherali [26].

The purpose of the present paper is to illustrate the application of the canonical duality theory for solving the foregoing quadratic minimization problem with nonconvex constraints. In the next section, we will show how to use the canonical dual transformation to convert the nonconvex constrained problem into a canonical dual problem, in order to derive related global optimality conditions. Certain particular cases along with numerical examples are used to illustrate the theory in Sect. 3 (quadratic constraints), Sect. 4 (box and integer constraints), and Sect. 5 (polynomial constraints), for using the canonical dual problem to obtain all critical points. We also expose certain insightful connections with classical Lagrangian duality for these special applications. Finally, Sect. 6 provides a summary, some open problems, and conclusions.

2 Canonical dual problem and strong duality theorem

For analytical convenience, we introduce an indicator function of the feasible set $\mathcal{X}_c$:

$$W(\boldsymbol{\epsilon}) = \begin{cases} 0 & \text{if } \boldsymbol{\epsilon} \leq \mathbf{d} \\ +\infty & \text{otherwise} \end{cases} \tag{7}$$

and let

$$U(\mathbf{x}) \equiv -P(\mathbf{x}) = \mathbf{x}^T\mathbf{f} - \frac{1}{2}\mathbf{x}^T A\mathbf{x}.$$

Then the primal problem $(\mathcal{P})$ can be written in the following unconstrained form:

$$\min \{\Pi(\mathbf{x}) = W(\mathbf{g}(\mathbf{x})) - U(\mathbf{x}) : \ \ \mathbf{x} \in \mathcal{X}_a\}. \tag{8}$$

By the Fenchel transformation, the conjugate function $W^\sharp(\boldsymbol{\sigma})$ of $W(\boldsymbol{\epsilon})$ can be defined by

$$W^\sharp(\boldsymbol{\sigma}) = \sup_{\boldsymbol{\epsilon}\in\mathbb{R}^m} \{\boldsymbol{\epsilon}^T\boldsymbol{\sigma} - W(\boldsymbol{\epsilon})\} = \begin{cases} \mathbf{d}^T\boldsymbol{\sigma} & \text{if } \boldsymbol{\sigma} \geq 0 \\ +\infty & \text{otherwise,} \end{cases} \tag{9}$$

which is convex and lower semi-continuous (l.s.c.) on $\mathbb{R}^m$. From convex analysis [2,44], the following relations hold for $(\boldsymbol{\epsilon}, \boldsymbol{\sigma}) \in \mathbb{R}^m \times \mathbb{R}^m$:

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

Replacing $W(\mathbf{g}(\mathbf{x}))$ in $\Pi(\mathbf{x})$ by $\mathbf{g}^T(\mathbf{x})\boldsymbol{\sigma} - W^\sharp(\boldsymbol{\sigma})$ based on the Fenchel–Young equality, the *extended Lagrangian* $\Xi_o : \mathcal{X}_a \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ associated with the problem (8) can be given as

$$\Xi_o(\mathbf{x}, \boldsymbol{\sigma}) = \mathbf{g}^T(\mathbf{x})\boldsymbol{\sigma} - W^\sharp(\boldsymbol{\sigma}) - U(\mathbf{x}). \tag{10}$$

Clearly, we have $\Xi_o(\mathbf{x}, \boldsymbol{\sigma}) = L(\mathbf{x}, \boldsymbol{\sigma}), \quad \forall(\mathbf{x}, \boldsymbol{\sigma}) \in \mathcal{X}_a \times \mathbb{R}^m_+$.

Since $\mathbf{g}(\mathbf{x})$ is a nonconvex function, following the standard procedure of the canonical dual transformation (see [10]), we assume that there exists a Gâteaux differentiable *geometrical operator*

$$\boldsymbol{\xi} = \{\xi^\alpha_\beta\} = \Lambda(\mathbf{x}) : \mathcal{X}_a \subset \mathbb{R}^n \to \mathcal{E}_a \subset \mathbb{R}^{m\times p_\alpha}, \tag{11}$$

and a *canonical function* $V : \mathcal{E}_a \to \mathbb{R}^m$ such that the nonconvex constraint $\mathbf{g}(\mathbf{x})$ can be written in the canonical form:

$$\mathbf{g}(\mathbf{x}) = V(\Lambda(\mathbf{x})). \tag{12}$$

By the definition introduced in [10], a function $V : \mathcal{E}_a \to \mathbb{R}^m$ is called *canonical* if it is Gâteaux differentiable on $\mathcal{E}_a$ such that the duality mapping

$$\boldsymbol{\varsigma} = \{\varsigma^\beta_\alpha\} = \nabla V(\boldsymbol{\xi}) = \left\{\frac{\partial V_\alpha(\boldsymbol{\xi})}{\partial \xi^\alpha_\beta}\right\} : \mathcal{E}_a \to \mathcal{E}^*_a \subset \mathbb{R}^{p_\alpha\times m}, \tag{13}$$

is invertible. We note that the geometric variable $\boldsymbol{\xi} = \{\xi^\alpha_\beta\}$ is an $m \times p_\alpha$ matrix, while its canonical dual variable $\boldsymbol{\varsigma} = \{\varsigma^\beta_\alpha\}$ is a $p_\alpha \times m$ matrix. In finite deformation theory and theoretical physics, these are called second order two-point tensors [10]. For the constrained problem $(\mathcal{P})$ considered in this paper, the dimension p_α of the geometrical variable $\boldsymbol{\xi} = \{\xi^\alpha_\beta\}$ depends on each given constraint $g_\alpha(\mathbf{x}) \leq d_\alpha, \ \alpha = 1, \ldots, m$. Let $I_\alpha = \{\beta |\ \beta \in \{1, \ldots, p_\alpha\}\}$ be an index set and let

$$\langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle = \left\{\sum_{\beta\in I_\alpha} \xi^\alpha_\beta \varsigma^\beta_\alpha\right\} \ : \ \mathcal{E}_a \times \mathcal{E}^*_a \to \mathbb{R}^m$$

denote the *partial bilinear form* on the product space $\mathcal{E}_a \times \mathcal{E}^*_a$. Thus, the Legendre conjugate $V^* : \mathcal{E}^*_a \to \mathbb{R}^m$ of V can be defined by

$$V^*(\boldsymbol{\varsigma}) = \text{sta}\{\langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle - V(\boldsymbol{\xi}) \ : \ \boldsymbol{\xi} \in \mathcal{E}_a\},$$

where the notation sta$\{*\}$ denotes computing the stationary points of $\{*\}$. By the assumption that the duality relation (13) is invertible (i.e., canonical), the Legendre conjugate $V^*(\boldsymbol{\varsigma})$ is uniquely defined on $\mathcal{E}^*_a$ and the inverse duality relation can be written as

$$\boldsymbol{\xi} = \nabla V^*(\boldsymbol{\varsigma}) = \left\{\frac{\partial V^*_\alpha(\boldsymbol{\varsigma})}{\partial \varsigma^\beta_\alpha}\right\}. \tag{14}$$

It is easy to verify that the following equivalent relations hold on $\mathcal{E}_a \times \mathcal{E}_a^*$:

$$\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) \;\Leftrightarrow\; \boldsymbol{\xi} = \nabla V^*(\boldsymbol{\varsigma}) \;\Leftrightarrow\; \langle \boldsymbol{\xi}; \boldsymbol{\varsigma} \rangle = V(\boldsymbol{\xi}) + V^*(\boldsymbol{\varsigma}). \tag{15}$$

In the terminology used in [10], $(\boldsymbol{\xi}, \boldsymbol{\varsigma})$ forms a *canonical duality pair* on $\mathcal{E}_a \times \mathcal{E}_a^*$. Thus, noting (12) and (15) to replace $\mathbf{g}(\mathbf{x})$ in (10) by

$$V(\Lambda(\mathbf{x})) = \langle \Lambda(\mathbf{x}); \boldsymbol{\varsigma} \rangle - V^*(\boldsymbol{\varsigma}),$$

we define the so-called *generalized total complementary function* (or the *extended Lagrangian* [10,27]) $\Xi : \mathcal{X}_a \times \mathbb{R}^m_+ \times \mathcal{E}_a^* \to \mathbb{R}$ as

$$\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \boldsymbol{\sigma}^T [\langle \Lambda(\mathbf{x}); \boldsymbol{\varsigma} \rangle - V^*(\boldsymbol{\varsigma}) - \mathbf{d}] + \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T \mathbf{f}, \tag{16}$$

where $\boldsymbol{\sigma} \in \mathbb{R}^m_+$ is the dual variable vector associated with $\mathbf{g}(\mathbf{x}) \leq \mathbf{d} \in \mathbb{R}^m$. Through this total complementary function, the *canonical dual function* can be defined by

$$P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \text{sta}\,\{\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) \;:\; \mathbf{x} \in \mathcal{X}_a\} = U^\Lambda(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) - \boldsymbol{\sigma}^T (V^*(\boldsymbol{\varsigma}) + \mathbf{d}), \tag{17}$$

where $U^\Lambda(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is the *parametric* Λ*-conjugate function* of the quadratic function $U(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T A\mathbf{x} + \mathbf{x}^T \mathbf{f}$ defined by the following Λ-conjugate transformation [10,12]:

$$U^\Lambda(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \text{sta}\{\boldsymbol{\sigma}^T \langle \Lambda(\mathbf{x}); \boldsymbol{\varsigma} \rangle - U(\mathbf{x}) \;:\; \mathbf{x} \in \mathcal{X}_a\}. \tag{18}$$

Letting $\mathcal{S}_c \subset \mathbb{R}^m_+ \times \mathcal{E}_a^*$ be a canonical dual feasible space on which the canonical dual function $P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is well defined, the canonical dual problem can be posed as follows:

$$(\mathcal{P}^d): \quad \text{sta}\{P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) : (\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c\}. \tag{19}$$

Theorem 1 (Complementary-Dual Principle) *The problem* $(\mathcal{P}^d)$ *is canonically dual to the primal problem* $(\mathcal{P})$ *in the sense that if* $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ *is a critical point of* $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ *over* $(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{X}_a \times \mathbb{R}^m_+ \times \mathcal{E}_a^*$*, then* $\bar{\mathbf{x}}$ *is a KKT point of* $(\mathcal{P})$*,* $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ *is a KKT point of* $(\mathcal{P}^d)$*, and*

$$P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) = P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}). \tag{20}$$

Proof If $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a critical point of Ξ as hypothesized, then we have the following criticality conditions:

$$\nabla_{\mathbf{x}} \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}, \bar{\boldsymbol{\sigma}}) = \bar{\boldsymbol{\sigma}}^T \langle \Lambda_t(\bar{\mathbf{x}}); \bar{\boldsymbol{\varsigma}} \rangle + A\bar{\mathbf{x}} - \mathbf{f} = 0, \tag{21}$$

$$\nabla_{\boldsymbol{\varsigma}} \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}, \bar{\boldsymbol{\sigma}}) = \Lambda(\bar{\mathbf{x}}) - \nabla V^*(\bar{\boldsymbol{\varsigma}}) = 0, \tag{22}$$

where $\Lambda_t(\mathbf{x}) = \nabla \Lambda(\mathbf{x})$ denotes the Gâteaux derivative of Λ, along with the conditions:

$$0 \leq \bar{\boldsymbol{\sigma}} \quad \perp \quad \big(\langle \Lambda(\bar{\mathbf{x}}); \bar{\boldsymbol{\varsigma}} \rangle - V^*(\bar{\boldsymbol{\varsigma}}) - \mathbf{d}\big) \leq 0, \tag{23}$$

where the notation $\perp$ represents the complementarity or orthogonality condition. Since $(\boldsymbol{\xi}, \boldsymbol{\varsigma})$ is a canonical duality pair on $\mathcal{E}_a \times \mathcal{E}_a^*$, the criticality condition (22) is equivalent to $\bar{\boldsymbol{\varsigma}} = \nabla_{\boldsymbol{\xi}} V(\Lambda(\bar{\mathbf{x}})) = \partial V(\boldsymbol{\xi}(\bar{\mathbf{x}}))/\partial \boldsymbol{\xi}$. Substituting this into (21) and using the chain rule to deduce $\nabla \mathbf{g}(\bar{\mathbf{x}}) = \langle \Lambda_t(\bar{\mathbf{x}}); \nabla_{\boldsymbol{\xi}} V(\Lambda(\bar{\mathbf{x}})) \rangle$, we have,

$$A\bar{\mathbf{x}} - \mathbf{f} + \bar{\boldsymbol{\sigma}}^T \nabla \mathbf{g}(\bar{\mathbf{x}}) = \nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}) = 0.$$

This is the criticality condition of the primal problem $(\mathcal{P})$. By the Legendre equality $\langle \Lambda(\bar{\mathbf{x}}); \bar{\boldsymbol{\varsigma}} \rangle - V^*(\bar{\boldsymbol{\varsigma}}) = V(\Lambda(\bar{\mathbf{x}}))$, the condition (23) can be written as

$$0 \leq \bar{\boldsymbol{\sigma}} \perp (\mathbf{g}(\bar{\mathbf{x}}) - \mathbf{d}) \leq 0.$$

This shows that $\bar{\mathbf{x}}$ is a KKT point of the primal problem $(\mathcal{P})$. From the complementarity condition (23), we have,

$$\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) = P(\bar{\mathbf{x}}).$$

On the other hand, by the definition of the canonical dual function, if $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a KKT point as hypothesized, the criticality condition (21) leads to

$$U^{\Lambda}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) = \bar{\boldsymbol{\sigma}}^T \langle \Lambda(\bar{\mathbf{x}}); \bar{\boldsymbol{\varsigma}} \rangle - U(\bar{\mathbf{x}}).$$

Therefore, $\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) = P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ and $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a KKT point of the dual problem $(\mathcal{P}^d)$. □

This theorem shows that there is no duality gap between the primal problem and its canonical dual. In order to identify the global minimizer, we need to study the convexity of the generalized complementary function $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$. Without much loss of generality, we introduce the following assumptions:

$$\begin{array}{l}\text{(i) the geometrical operator } \Lambda(\mathbf{x}) : \mathbb{R}^n \to \mathcal{E}_a \text{ is twice Gâteaux differentiable;} \\ \text{(ii) the canonical function } V : \mathcal{E}_a \to \mathbb{R}^m \text{ is convex.}\end{array} \tag{24}$$

By this assumption, we know that the conjugate function $V^*(\boldsymbol{\varsigma}) : \mathcal{E}_a^* \to \mathbb{R}^m$ is also convex, and for any given $(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c$, the generalized complementary function $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is twice Gâteaux differentiable on $\mathbf{x}$. Let $G_a(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \nabla_{\mathbf{x}}^2 \Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ denote the Hessian matrix of $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ and let

$$\mathcal{S}_c^+ = \{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c \mid \; G_a(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) \succeq 0, \quad \forall \mathbf{x} \in \mathcal{X}_a\} \tag{25}$$

be a subset of $\mathcal{S}_c$. We have the following theorem.

Theorem 2 (Global Optimality Condition) *Suppose that the conditions in (24) hold and that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a critical point of $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$. If $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) \in \mathcal{S}_c^+$, then $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a global maximizer of P^d on $\mathcal{S}_c^+$ and $\bar{\mathbf{x}}$ is a global minimizer of P on $\mathcal{X}_c$, i.e.,*

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_c} P(\mathbf{x}) = \max_{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+} P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}). \tag{26}$$

Proof Actually, this theorem is a special application of the general theory proposed by Gao and Strang in [27]. By the convexity of $V(\boldsymbol{\xi})$, its Legendre conjugate $V^* : \mathcal{E}_a^* \to \mathbb{R}^m$ is also convex. Thus, for any given $\boldsymbol{\sigma} \in \mathbb{R}_+^m$, the linear combination $\boldsymbol{\sigma}^T V^*(\boldsymbol{\varsigma}) : \mathcal{E}_a^* \to \mathbb{R}$ is convex and the generalized complementary function $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is concave in $\boldsymbol{\varsigma}$. By considering $\boldsymbol{\sigma} \in \mathbb{R}_+^m$ as a Lagrange multiplier for the inequality constraint in $\mathcal{X}_c$, the complementary function $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ can be viewed as a concave (linear) function of $\boldsymbol{\sigma} \in \mathbb{R}_+^m$ for any given $(\mathbf{x}, \boldsymbol{\varsigma}) \in \mathcal{X}_a \times \mathcal{E}_a^*$. Therefore, for any given $\mathbf{x} \in \mathbb{R}^n$, we have

$$\max_{\boldsymbol{\sigma} \in \mathbb{R}_+^m} \max_{\boldsymbol{\varsigma} \in \mathcal{E}_a^*} \Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \max_{\boldsymbol{\sigma} \in \mathbb{R}_+^m} L(\mathbf{x}, \boldsymbol{\sigma}) = \begin{cases} P(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{X}_c, \\ \infty & \text{otherwise} . \end{cases}$$

Moreover, if $(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+ \subset \mathbb{R}_+^m \times \mathcal{E}_a^*$, then $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is convex in $\mathbf{x} \in \mathcal{X}_a$ and concave in $\boldsymbol{\varsigma}$ for any given $\boldsymbol{\sigma} \in \mathbb{R}_+^m$. Therefore, if $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) \in \mathcal{X}_a \times \mathcal{S}_c^+$ is a critical point of Ξ, we have

$$\begin{aligned}\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) &= \min_{\mathbf{x} \in \mathcal{X}_a} \max_{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+} \Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \min_{\mathbf{x} \in \mathcal{X}_c} P(\mathbf{x}) \\ &= \max_{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+} \min_{\mathbf{x} \in \mathcal{X}_a} \Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \max_{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+} P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}).\end{aligned}$$

By Theorem 1, we then have (26). □

This theorem provides a sufficient condition for a global minimizer of the nonconvex primal problem. In many applications, the geometrical mapping $\Lambda(\mathbf{x}) : \mathcal{X}_a \to \mathcal{E}_a$ is usually a quadratic operator

$$\Lambda(\mathbf{x}) = \left\{ \frac{1}{2}\mathbf{x}^T \mathbf{B}^\alpha_\beta \mathbf{x} + \mathbf{x}^T \mathbf{C}^\alpha_\beta \right\} : \mathbb{R}^n \to \mathcal{E}_a \subset \mathbb{R}^{m \times p_\alpha}, \tag{27}$$

where $\mathbf{B}^\alpha_\beta = \left\{ B^\alpha_{ij\beta} \right\} = \left\{ B^\alpha_{ji\beta} \right\} \in \mathbb{R}^{n\times n}$, $\mathbf{C}^\alpha_\beta = \left\{ C^\alpha_{i\beta} \right\} \in \mathbb{R}^n$, and the range $\mathcal{E}_a$ depends on both $\mathbf{B}^\alpha_\beta$ and $\mathbf{C}^\alpha_\beta$. In this case, the generalized complementary function has the form:

$$\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \frac{1}{2}\mathbf{x}^T G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{x} - \boldsymbol{\sigma}^T (V^*(\boldsymbol{\varsigma}) + \mathbf{d}) - \mathbf{x}^T \mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}), \tag{28}$$

where

$$G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = A + \sum_{\alpha=1}^{m} \sum_{\beta \in I_\alpha} \sigma_\alpha \mathbf{B}^\alpha_\beta \varsigma^\beta_\alpha \tag{29}$$

is the Hessian matrix of $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$, which does not depend on $\mathbf{x}$, and

$$\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \mathbf{f} - \sum_{\alpha=1}^{m} \sum_{\beta \in I_\alpha} \sigma_\alpha \mathbf{C}^\alpha_\beta \varsigma^\beta_\alpha. \tag{30}$$

The criticality condition (21) in this case is a linear equation of $\mathbf{x}$, i.e.,

$$G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{x} = \mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}), \tag{31}$$

which is also called the *canonical equilibrium equation* [10]. Clearly, for a given $(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$, if $\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is in the column space of $G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$, denoted by $\mathcal{C}_{ol}(G_a)$, the solution of the equation (31) can be written in the form:

$$\mathbf{x} = G_a^+(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}), \tag{32}$$

where G_a^+ is the Moore–Penrose generalized inverse of G_a. Thus, the canonical dual feasible space $\mathcal{S}_c$ can be defined as

$$\mathcal{S}_c = \{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathbb{R}^m_+ \times \mathcal{E}^*_a |\ \mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{C}_{ol}(G_a)\}, \tag{33}$$

and the canonical dual function P^d can be formulated as

$$P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = -\frac{1}{2}\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})^T G_a^+(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) - \boldsymbol{\sigma}^T (V^*(\boldsymbol{\varsigma}) + \mathbf{d}). \tag{34}$$

Since $\Lambda(\mathbf{x})$ is a quadratic operator, its Gâteaux derivative is an affine operator

$$\Lambda_t(\mathbf{x}) = \nabla\Lambda(\mathbf{x}) = \mathbf{x}^T \mathbf{B}^\alpha_\beta + \mathbf{C}^{\alpha T}_\beta.$$

By Gao and Strang [27], the complementary operator $\Lambda_c(\mathbf{x})$ of Λ_t is defined by

$$\Lambda_c(\mathbf{x}) = \Lambda(\mathbf{x}) - \Lambda_t(\mathbf{x})\mathbf{x} = -\frac{1}{2}\mathbf{x}^T \mathbf{B}^\alpha_\beta \mathbf{x}. \tag{35}$$

Thus, the *complementary gap function* [27] can be defined as

$$G(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = \boldsymbol{\sigma}^T \langle -\Lambda_c(\mathbf{x}); \boldsymbol{\varsigma}\rangle + \frac{1}{2}\mathbf{x}^T A\mathbf{x} = \frac{1}{2}\mathbf{x}^T G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{x}. \tag{36}$$

This gap function plays an important role in nonconvex analysis and global optimization. Clearly, $G(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) \geq 0 \; \forall \mathbf{x} \in \mathcal{X}_a$ if $G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \succeq 0$. Let

$$\mathcal{S}_c^+ = \{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c \mid G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \succeq 0\}, \tag{37}$$

$$\mathcal{S}_c^- = \{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c \mid G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \prec 0\}. \tag{38}$$

Then from Theorems 1, 2, and the *triality theory* developed in [10,12,14], we have the following result:

Theorem 3 (Triality Theory) *Suppose that $\Lambda(\mathbf{x})$ is a quadratic operator defined by (27) and the condition (ii) in (24) holds.*

If $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) \in \mathcal{S}_c$ is a critical point of $(\mathcal{P}^d)$, then $\bar{\mathbf{x}} = G_a^+(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})\mathbf{F}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a KKT point of $(\mathcal{P})$ and $P(\bar{\mathbf{x}}) = P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$.

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

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_c} P(\mathbf{x}) = \max_{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+} P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}). \tag{39}$$

If the critical point $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) \in \mathcal{S}_c^-$, then on the neighborhood $\mathcal{X}_o \times \mathcal{S}_o$ of $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}})$, we have either

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) = \min_{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_o} P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) \tag{40}$$

or

$$P(\bar{\mathbf{x}}) = \max_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) = \max_{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_o} P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}). \tag{41}$$

Proof If $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) \in \mathcal{S}_c$ is a critical point of $(\mathcal{P}^d)$, we have

$$\delta_{\boldsymbol{\varsigma}} P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) = \bar{\boldsymbol{\sigma}}^T \left(\Lambda(\bar{\mathbf{x}}) - \nabla V^*(\bar{\boldsymbol{\varsigma}})\right) = 0, \tag{42}$$

$$0 \leq \bar{\boldsymbol{\sigma}} \quad \perp \quad \delta_{\boldsymbol{\sigma}} P^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) = \langle \Lambda(\bar{\mathbf{x}}); \bar{\boldsymbol{\varsigma}} \rangle - V^*(\bar{\boldsymbol{\varsigma}}) - \mathbf{d} \leq 0, \tag{43}$$

where $\bar{\mathbf{x}} = G_a^+(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})\mathbf{F}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$. Equation (42) asserts that if $\bar{\sigma}_\alpha \neq 0$, then the corresponding $\xi_\alpha(\bar{\mathbf{x}}) = \Lambda_\alpha(\bar{\mathbf{x}}) = \nabla_{\varsigma_\alpha} V^*(\bar{\boldsymbol{\varsigma}})$. By the fact that $(\Lambda(\bar{\mathbf{x}}), \bar{\boldsymbol{\varsigma}})$ is a canonical duality pair on $\mathcal{E}_a \times \mathcal{E}_a^*$, from the equivalent relations in (15), we have $\langle \Lambda(\bar{\mathbf{x}}); \bar{\boldsymbol{\varsigma}} \rangle - V^*(\bar{\boldsymbol{\varsigma}}) = V(\Lambda(\bar{\mathbf{x}})) = \mathbf{g}(\bar{\mathbf{x}})$. Therefore, the complementarity condition in (43) leads to $\bar{\sigma}_\alpha(g_\alpha(\bar{\mathbf{x}}) - d_\alpha) = 0$. If $\bar{\sigma}_\alpha \neq 0$, we must have the criticality condition $g_\alpha(\bar{\mathbf{x}}) - d_\alpha = 0$. This shows that if $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a critical point of $P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$, the vector $\bar{\mathbf{x}} = G_a^+(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})\mathbf{F}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a KKT point of $(\mathcal{P})$.

Since the total complementary function $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is a saddle function on $\mathcal{X}_a \times \mathcal{S}_c^+$ and a *super-critical function* on $\mathcal{X}_a \times \mathcal{S}_c^-$ (i.e., Ξ is concave in both $\mathbf{x} \in \mathcal{X}_a$ and $(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^-$), the remainder of the theorem follows from the triality theory developed in [10,12,16]. □

Remark 1 Note that the dual variable $\boldsymbol{\sigma} \in \mathbb{R}_+^m$, i.e., it is restricted by the inequality constraint $\boldsymbol{\sigma} \geq 0 \in \mathbb{R}^m$. From the proof of Theorem 1 (as well as Theorem 3) we know that if $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ is a critical point of Ξ, then $\bar{\mathbf{x}}$ and $\bar{\boldsymbol{\varsigma}}$ satisfy the criticality conditions (21) and (22), respectively, while $\bar{\boldsymbol{\sigma}}$ satisfies the KKT conditions (23). Given a nonconvex quadratic objective function $P(\mathbf{x})$, the global minimizer $\bar{\mathbf{x}}$ is located on the boundary of the feasible set $\mathcal{X}_c$, i.e., there exists at least one $\bar{\sigma}_\alpha > 0$ such that

$$\nabla_{\sigma_\alpha} \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}}) = \langle \Lambda_\alpha(\bar{\mathbf{x}}); \bar{\varsigma}_\alpha \rangle - V_\alpha^*(\bar{\boldsymbol{\varsigma}}) - d_\alpha = g_\alpha(\bar{\mathbf{x}}) - d_\alpha = 0.$$

i.e., the dual variable $\bar{\sigma}_\alpha$ is a critical point of $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$, and the corresponding vector $\bar{\mathbf{x}}$ is KKT point of $P(\mathbf{x})$. By the fact that the canonical dual feasible space $\mathcal{S}_c^+$ is a closed convex set, the critical point $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$ could be located on the boundary of $\mathcal{S}_c^+$. In this case, the matrix $G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ may be singular and the canonical dual problem may have more than one critical point on $\mathcal{S}_c^+$. Particularly, on the open canonical dual feasible space

$$\mathcal{S}_c^{\ddagger} = \{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c \mid G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \succ 0\}, \tag{44}$$

the Hessian matrix $G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is invertible and the canonical dual function can be formulated as

$$P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = -\frac{1}{2}\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})^T G_a^{-1}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) - \boldsymbol{\sigma}^T(V^*(\boldsymbol{\varsigma}) + \mathbf{d}). \tag{45}$$

In this case, the canonical dual problem

$$(\mathcal{P}_+^d): \quad \max\left\{P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}): \quad (\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^{\ddagger}\right\} \tag{46}$$

usually has at most one solution $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$, which is a critical point of $P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$. By Theorem 3, the vector

$$\bar{\mathbf{x}} = G_a^{-1}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})\mathbf{F}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\varsigma}})$$

is a solution to $(\mathcal{P})$, which is located on the boundary of $\mathcal{X}_c$ under the nonconvexity of $P(\mathbf{x})$. The related existence conditions will be discussed in the following section.

The generalized complementary function $\Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma})$ defined in (16) is actually the *second-order Lagrangian* in the canonical duality theory (see Sect. 4.1.2 in [10]). For the quadratic geometrical mapping $\Lambda(\mathbf{x})$, the canonical dual function $P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ can be formulated explicitly, which is also called the pure complementary energy function in finite deformation theory (see [7–9]). The analytical solution of type (32) was first obtained in nonconvex variational problems [7,11] and finite deformation mechanics [9]. Similar results also appear in global optimization [12,16,17]. It is interesting to note that many nonconvex primal problems in nonconvex analysis and global optimization share the same form of the canonical dual function as defined in (34) (see [8,12,15,20,26]). In the case that $\Lambda(\mathbf{x})$ is a higher order nonlinear function, the sequential canonical dual transformation can be used to formulate a higher order complementary function (see Chap. 4 in [10]). □

Theorem 3 shows that by the canonical dual transformation, the nonconvex primal problem $(\mathcal{P})$ can be transformed to a concave maximization dual problem over a convex feasible space $\mathcal{S}_c^+$, which can be solved via well-developed nonlinear programming techniques (cf. [40]).

3 Applications to quadratic constrained problems

We begin by considering the following nonconvex quadratic minimization problem with quadratic inequality constraints, denoted by $(\mathcal{P}_q)$:

$$(\mathcal{P}_q): \quad \begin{cases} \min\left\{P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{f}\right\} \\ s.t. \;\; \frac{1}{2}\mathbf{x}^T\mathbf{B}^\alpha\mathbf{x} + \mathbf{x}^T\mathbf{C}^\alpha \le d_\alpha, \quad \alpha = 1, \ldots, m, \end{cases} \tag{47}$$

where $\mathbf{B}^\alpha = \left\{B_{ij}^\alpha\right\} = \left\{B_{ji}^\alpha\right\} \in \mathbb{R}^{n\times n}$, $\mathbf{C}^\alpha = \left\{C_i^\alpha\right\} \in \mathbb{R}^n, \forall\, \alpha = 1, \ldots, m$, and $\mathbf{d} = \{d_\alpha\} \in \mathbb{R}^m$ is a vector. Due to the nonconvex cost function and nonconvex inequality constraints, this problem is known to be NP-hard [42].

Since the constraint $\mathbf{g}(\mathbf{x})$ is a vector-valued quadratic function defined on $\mathcal{X}_a = \mathbb{R}^n$, we simply let

$$\mathbf{g}(\mathbf{x}) = \Lambda(\mathbf{x}) = \left\{ \frac{1}{2}\mathbf{x}^T\mathbf{B}^\alpha\mathbf{x} + \mathbf{x}^T\mathbf{C}^\alpha \right\} : \mathbb{R}^n \to \mathbb{R}^m. \tag{48}$$

Compared with (27), we have $p_\alpha = 1$ and the canonical function $V(\boldsymbol{\xi}) = \boldsymbol{\xi}$ is a self-mapping. Therefore, the canonical dual variable $\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) = I$ is an identity matrix in $\mathbb{R}^{m\times m}$ and $V^*(\boldsymbol{\varsigma}) = \text{sta}\{\langle\boldsymbol{\xi};\boldsymbol{\varsigma}\rangle - \boldsymbol{\xi} |\ \boldsymbol{\xi} \in \mathbb{R}^m\} = 0$. In this case, the generalized complementary function (28) has a very simple form:

$$\Xi_q(\mathbf{x}, \boldsymbol{\sigma}) = \frac{1}{2}\mathbf{x}^T G_q(\boldsymbol{\sigma})\mathbf{x} - \mathbf{x}^T\mathbf{F}_q(\boldsymbol{\sigma}) - \boldsymbol{\sigma}^T\mathbf{d}, \tag{49}$$

where

$$G_q(\boldsymbol{\sigma}) = A + \sum_{\alpha=1}^m \sigma_\alpha\mathbf{B}^\alpha, \quad \text{and} \quad \mathbf{F}_q(\boldsymbol{\sigma}) = \mathbf{f} - \sum_{\alpha=1}^m \sigma_\alpha\mathbf{C}^\alpha. \tag{50}$$

Therefore, on the dual feasible space

$$\mathcal{S}_q = \left\{\boldsymbol{\sigma} \in \mathbb{R}^m_+ |\ \mathbf{F}_q(\boldsymbol{\sigma}) \in \mathcal{C}_{ol}(G_q)\right\}, \tag{51}$$

the canonical dual function P^d_q can be formulated as

$$P^d_q(\boldsymbol{\sigma}) = -\frac{1}{2}\mathbf{F}_q(\boldsymbol{\sigma})^T G^+_q(\boldsymbol{\sigma})\mathbf{F}_q(\boldsymbol{\sigma}) - \boldsymbol{\sigma}^T\mathbf{d}. \tag{52}$$

In this case, the complementary gap function has a simple form:

$$G(\mathbf{x}, \boldsymbol{\sigma}) = \frac{1}{2}\mathbf{x}^T G_q(\boldsymbol{\sigma})\mathbf{x}, \tag{53}$$

which is nonnegative on $\mathbb{R}^n$ if $G_q(\boldsymbol{\sigma}) \succeq 0$. Let

$$\mathcal{S}^+_q = \{\boldsymbol{\sigma} \in \mathcal{S}_q |\ G_q(\boldsymbol{\sigma}) \succeq 0\}, \quad \text{and} \quad \mathcal{S}^-_q = \{\boldsymbol{\sigma} \in \mathcal{S}_q |\ G_q(\boldsymbol{\sigma}) \prec 0\}. \tag{54}$$

Then the canonical dual problem for this quadratic constrained problem is given by

$$(\mathcal{P}^d_q): \quad \max\{P^d_q(\boldsymbol{\sigma}) : \ \boldsymbol{\sigma} \in \mathcal{S}^+_q\}. \tag{55}$$

From Theorem 3 we have the following result:

Theorem 4 *The problem $(\mathcal{P}^d_q)$ is canonically dual to the primal problem $(\mathcal{P}_q)$ in the sense that for each critical point $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q$ of $(\mathcal{P}^d_q)$, the vector $\bar{\mathbf{x}} = G^+_q(\bar{\boldsymbol{\sigma}})\mathbf{F}_q(\bar{\boldsymbol{\sigma}})$ is a KKT point of $(\mathcal{P}_q)$ and $P(\bar{\mathbf{x}}) = P^d(\bar{\boldsymbol{\sigma}})$.*

Particularly, if the critical point $\bar{\boldsymbol{\sigma}} \in \mathcal{S}^+_q$, then $\bar{\boldsymbol{\sigma}}$ is a global maximizer of $(\mathcal{P}^d_q)$, the vector $\bar{\mathbf{x}}$ is a global minimizer of $(\mathcal{P}_q)$, and

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x}\in\mathcal{X}_c} P(\mathbf{x}) = \max_{\boldsymbol{\sigma}\in\mathcal{S}^+_q} P^d_q(\boldsymbol{\sigma}) = P^d_q(\bar{\boldsymbol{\sigma}}). \tag{56}$$

If $G_q(\bar{\boldsymbol{\sigma}}) \succ 0$, then $\bar{\boldsymbol{\sigma}}$ is a unique global maximizer of $(\mathcal{P}^d_q)$ and the vector $\bar{\mathbf{x}} = G^{-1}_q(\bar{\boldsymbol{\sigma}})\mathbf{F}_q(\bar{\boldsymbol{\sigma}})$ is a unique global minimizer of $(\mathcal{P}_q)$.

However, if the critical point $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^-$, then $\bar{\boldsymbol{\sigma}}$ is a local minimizer of $P_q^d(\boldsymbol{\sigma})$ on the neighborhood $\mathcal{S}_o \subset \mathcal{S}_q^-$ if and only if $\bar{\mathbf{x}}$ is a local minimizer of $P(\mathbf{x})$ on the neighborhood $\mathcal{X}_o \subset \mathcal{X}_c$, i.e.,

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) = \min_{\boldsymbol{\sigma} \in \mathcal{S}_o} P_q^d(\boldsymbol{\sigma}) = P_q^d(\bar{\boldsymbol{\sigma}}). \qquad \square \tag{57}$$

This theorem shows that the Hessian matrix of the complementary gap function $G(\mathbf{x}, \boldsymbol{\sigma})$ provides sufficient and uniqueness conditions for globally minimizing the quadratic constrained problem $(\mathcal{P}_q)$. We note that this theorem is actually a special application of the general canonical duality theory developed in [10,12,16,27]. In order to study the existence theory, we need to introduce the following sets:

$$\partial \mathcal{S}_q = \left\{ \boldsymbol{\sigma} \in \mathbb{R}^m \,|\; \det G_q(\boldsymbol{\sigma}) = 0 \right\}, \tag{58}$$

$$\partial \mathcal{S}_q^+ = \left\{ \boldsymbol{\sigma} \in \mathcal{S}_q \,|\; \det G_q(\boldsymbol{\sigma}) = 0 \right\}. \tag{59}$$

Similar to the general results proposed in [29], we then have the following result:

Theorem 5 *Suppose that for the given matrices A, $\{\mathbf{B}^\alpha\}$, $\{\mathbf{C}^\alpha\}$ and vectors $\mathbf{f}$, $\mathbf{d}$, there exists at least one $\boldsymbol{\sigma}_0 \in \mathcal{S}_q^+$ such that $G_q(\boldsymbol{\sigma}_0) \succeq 0$ and*

$$\lim_{\substack{\|\boldsymbol{\sigma}\| \to \infty \\ \boldsymbol{\sigma} \in \mathcal{S}_q^+}} P_q^d(\boldsymbol{\sigma}) = -\infty. \tag{60}$$

Then the canonical dual problem $(\mathcal{P}_q^d)$ has at least one KKT point $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^+$. If $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^+$ is also a critical point of $P_q^d(\boldsymbol{\sigma})$, then $\bar{\mathbf{x}} = G_q^+(\bar{\boldsymbol{\sigma}})\mathbf{F}_q(\bar{\boldsymbol{\sigma}})$ is a global minimizer for the primal problem $(\mathcal{P}_q)$.

Moreover, if $\partial \mathcal{S}_q \subset \mathbb{R}_+^m$, there exists at least one $\boldsymbol{\sigma}_0 \in \mathcal{S}_q^+$ such that $G_q(\boldsymbol{\sigma}_0) \succ 0$, and

$$\lim_{\substack{\boldsymbol{\sigma} \to \partial \mathcal{S}_q^+ \\ \boldsymbol{\sigma} \in \mathcal{S}_q^+}} P_q^d(\boldsymbol{\sigma}) = -\infty, \tag{61}$$

then the canonical dual problem $(\mathcal{P}_q^d)$ has a unique global maximizer $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^+$ and $\bar{\mathbf{x}} = G_q^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{F}_q(\bar{\boldsymbol{\sigma}})$ is a unique global minimizer for the primal problem $(\mathcal{P}_q)$.

Proof By the fact that the feasible space $\mathcal{S}_q^+$ is a semi-closed convex set whose boundary $\partial \mathcal{S}_q^+$ is a hyper-surface in $\mathbb{R}^m$, if there exists a $\boldsymbol{\sigma}_0$ such that $G_q(\boldsymbol{\sigma}_0) \succeq 0$, then $\mathcal{S}_q^+$ is not empty. Since the canonical dual function $P_q^d(\boldsymbol{\sigma})$ is continuous and concave on $\mathcal{S}_q^+$, which is finite on $\partial \mathcal{S}_q^+$, if the condition (60) holds, then $P_q^d(\boldsymbol{\sigma})$ has at least one maximizer on $\mathcal{S}_q^+$.

Moreover, if $\partial \mathcal{S}_q \subset \mathbb{R}_+^m$, then $\mathcal{S}_q^+ \subset \mathbb{R}_+^m$. If there exists a $\boldsymbol{\sigma}_0$ such that $G_q(\boldsymbol{\sigma}_0) \succ 0$, then $\mathcal{S}_q^+$ is non-empty and has at least one interior point. Under the conditions (60) and (61), the canonical dual function $P_q^d(\boldsymbol{\sigma})$ is strictly concave and coercive on the open convex set $\mathcal{S}_q^+ \backslash \partial \mathcal{S}_q^+$. Therefore, the canonical dual problem $(\mathcal{P}_q)$ has a unique maximizer $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^+$ that is a critical point of $P_q^d(\boldsymbol{\sigma})$. $\square$

Theorem 5 shows that under the conditions (60) and (61), the canonical dual function $P_q^d(\boldsymbol{\sigma})$ has a unique maximizer $\bar{\boldsymbol{\sigma}}$ on the open feasible space

$$\mathcal{S}_q^\ddagger = \{\boldsymbol{\sigma} \in \mathcal{S}_q \,|\, G_q(\boldsymbol{\sigma}) \succ 0\}. \tag{62}$$

In this case, the matrix $G_q(\boldsymbol{\sigma})$ is invertible on $\mathcal{S}_q^\ddagger$ and the canonical dual function P_q^d can be written as

$$P_q^d(\boldsymbol{\sigma}) = -\frac{1}{2}\mathbf{F}_q(\boldsymbol{\sigma})^T G_q^{-1}(\boldsymbol{\sigma})\mathbf{F}_q(\boldsymbol{\sigma}) - \boldsymbol{\sigma}^T\mathbf{d}. \tag{63}$$

Particularly, if $m = 1$ and $\mathbf{C} = \mathbf{0}$, the problem $(\mathcal{P}_q)$ has only one quadratic constraint $g(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T B\mathbf{x} \le d$. Therefore, the canonical dual function has only one variable [18]:

$$P_q^d(\sigma) = -\frac{1}{2}\mathbf{f}^T(A + \sigma B)^{-1}\mathbf{f} - \sigma d, \tag{64}$$

and the criticality condition $\nabla P_q^d(\sigma) = 0$ leads to a nonlinear algebraic equation

$$\frac{1}{2}\mathbf{f}^T(A + \sigma B)^{-1}B(A + \sigma B)^{-1}\mathbf{f} = d, \tag{65}$$

which can be solved easily to obtain all dual solutions. Moreover, if $B = I$ is an identity matrix in $\mathbb{R}^n$, then the constraint $\frac{1}{2}\mathbf{x}^T B\mathbf{x} = \frac{1}{2}\|\mathbf{x}\|^2 \le d$ is an n-dimensional sphere. It was shown in [17,18] that number of solutions depends on the number of negative eigenvalues of A. These canonical dual solutions play an important role in trust region methods.

Remark 2 It is instructive to see connections between (63) and the classical Lagrangian Theory proposed in (3)–(5). Under the assumption that

det $G_q(\boldsymbol{\sigma}) \neq 0$, the necessary optimality condition $\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\sigma}) = 0$ in (3) leads to the unique solution $\mathbf{x} = G_q^{-1}(\boldsymbol{\sigma})\mathbf{F}_q(\boldsymbol{\sigma})$. Substituting this into (4) we get

$$P^*(\boldsymbol{\sigma}) = \inf_{\mathbf{x}\in\mathcal{X}_a} L(\mathbf{x}, \boldsymbol{\sigma}) = \begin{cases} P_q^d(\boldsymbol{\sigma}) & \text{if } \boldsymbol{\sigma} \in \mathcal{S}_q^\ddagger, \\ -\infty & \text{otherwise.} \end{cases}$$

Therefore, under the sufficient (global) optimality condition $\boldsymbol{\sigma} \in \mathcal{S}_q^\ddagger$, the Lagrangian dual function $P^*(\boldsymbol{\sigma})$ is identical to the canonical dual function $P_q^d(\boldsymbol{\sigma})$ as defined in (63). By Theorem 4, we know that if $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^\ddagger$ is a critical point of $P_q^d(\boldsymbol{\sigma})$, then the corresponding $\bar{\mathbf{x}} = G_q^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{F}_q(\bar{\boldsymbol{\sigma}}) \in \mathcal{X}_c$ satisfies $\bar{\boldsymbol{\sigma}} \perp (\mathbf{g}(\bar{\mathbf{x}}) - \mathbf{d})$, and $(\bar{\mathbf{x}}, \bar{\boldsymbol{\sigma}}) \in \mathbb{R}^n \times \mathcal{S}_q^\ddagger$ is a saddle point solution with the following strong duality condition holding:

$$\inf_{\mathbf{x}\in\mathcal{X}_c} P(\mathbf{x}) = \sup_{\boldsymbol{\sigma}\in\mathcal{S}_q^\ddagger} P^*(\boldsymbol{\sigma}).$$

However, in many applications if $\{\mathbf{B}^\alpha\}$ $(\alpha = 1, \ldots, m)$ are either indefinite or zero matrices, the dual feasible space $\mathcal{S}_q^\ddagger$ could be an empty set and in this case,

$$P^*(\boldsymbol{\sigma}) = \inf_{\mathbf{x}\in\mathcal{X}_a} L(\mathbf{x}, \boldsymbol{\sigma}) = -\infty.$$

Therefore, the classical Lagrangian duality theory cannot be used to solve the primal problem (see Example 2 in the next section). But, by the complementary-dual principle (the first part of Theorem 4) we know that the primal problem is equivalent to the following canonical dual minimal stationary problem:

$$\min \operatorname{sta}\{P_q^d(\boldsymbol{\sigma}) : \quad \boldsymbol{\sigma} \in \mathcal{S}_q\}. \tag{66}$$

For each critical point $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q$, the vector $\bar{\mathbf{x}} = G_q^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{F}_q(\bar{\boldsymbol{\sigma}})$ is a KKT point of the primal problem. If $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^-$ is a local minimizer of $P_q^d(\boldsymbol{\sigma})$, by the triality theory we know that $\bar{\mathbf{x}}$ is a local minimizer of $(\mathcal{P}_q)$. □

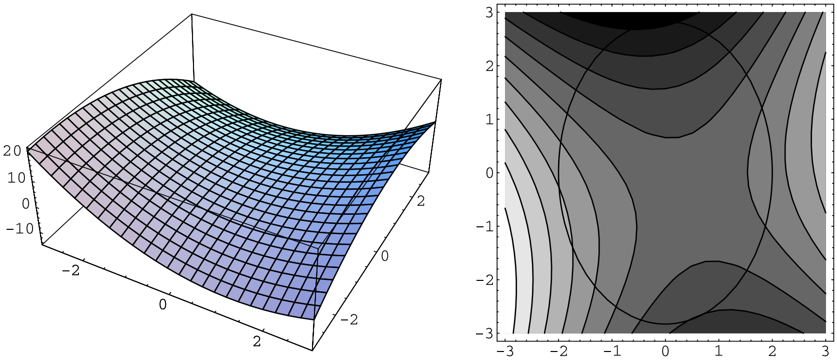


Fig. 1 Graph of $P(\mathbf{x})$ (left); contours of $P(\mathbf{x})$ and boundary of $\mathcal{X}_c$ (right) for Example 1

This remark shows that for nonconvex minimization problems with quadratic constraints, the Lagrangian dual $P^*(\boldsymbol{\sigma})$ is identical to the canonical dual $P^d(\boldsymbol{\sigma})$ only on the canonical dual feasible space $\mathcal{S}_q^\ddagger$. The sufficient global optimality conditions in $\mathcal{S}_q^\ddagger$ close the duality gap that exists in the classical Lagrangian duality theory. The triality theory (Theorem 4) can be used to develop effective canonical dual approaches for solving a wide class of primal problems. The general sufficient global optimality conditions for nonconvex minimization problems was first proposed by Gao and Strang in [27], where $G(\mathbf{x}, \boldsymbol{\sigma})$ is called the complementary gap function associated with the quadratic operator $\Lambda(\mathbf{x})$. They discovered that under the condition $G(\mathbf{x}, \boldsymbol{\sigma}) \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}_a$, the critical point of the total complementary function $\Xi_q(\mathbf{x}, \boldsymbol{\sigma})$ leads to global optimal solutions to the primal problem. We note that similar results and the so-called L-subdifferential condition proposed in the recent papers [34,35] are actually special applications of the general canonical duality theory developed in [10,27].

Example 1 In 2-D space, if we let

$$A = \begin{pmatrix} 3 & 0.5 \\ 0.5 & -2.0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \end{pmatrix}, \quad \text{and} \quad \mathbf{f} = \begin{pmatrix} 1 \\ 1.5 \end{pmatrix},$$

then the matrix A is indefinite, while B is positive definite. Setting $d = 2$, the graph of the primal function $P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{f}$ is a saddle surface (see Fig. 1), and the boundary of the feasible set $\mathcal{X}_c = \{\mathbf{x} \in \mathbb{R}^2 | \ \frac{1}{2}\mathbf{x}^T B\mathbf{x} \leq d\}$ is an ellipse (see Fig. 1). In this case, the canonical dual function (64) can be formulated as

$$P_q^d(\sigma) = -\frac{1}{2}(1 \ \ 1.5)\begin{pmatrix} 3+\sigma & 0.5 \\ 0.5 & -2+0.5\sigma \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1.5 \end{pmatrix} - 2\sigma,$$

which has four critical points (see Fig. 2):

$$\bar{\sigma}_1 = 5.08 > \bar{\sigma}_2 = 3.06 > \bar{\sigma}_3 = -2.46 > \bar{\sigma}_4 = -3.68.$$

Since $G_q(\bar{\sigma}_1) \succ 0$ and $G_q(\bar{\sigma}_4) \prec 0$, the triality theory yields that $\mathbf{x}_1 = (-0.05, 2.83)^T$ is a global minimizer located on the boundary of $\mathcal{X}_c$, and $\mathbf{x}_4 = (-1.95, -0.64)^T$ is a local maximizer. From the graph of $P^d(\sigma)$ we can see that $\mathbf{x}_2 = (0.39, -2.77)^T$ is a local minimizer, and $\mathbf{x}_3 = (2.0, -0.16)^T$ is a local maximizer. We have

$$P(\mathbf{x}_1) = -12.25 < P(\mathbf{x}_2) = -4.24 < P(\mathbf{x}_3) = 4.04 < P(\mathbf{x}_4) = 8.81.$$

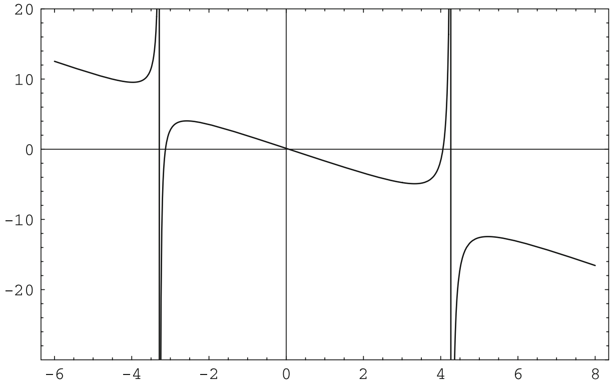


Fig. 2 Graph of $P^d(\sigma)$

4 Applications to box and integer constrained problems

We now turn our attention to the box constrained quadratic minimization problem:

$$(\mathcal{P}_b): \quad \begin{cases} \min \left\{ P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{f} \right\} \\ s.t. \quad -1 \le x_i \le 1, \quad i = 1, \ldots, n. \end{cases} \tag{67}$$

An associated problem is the quadratic integer (Boolean) programming problem

$$(\mathcal{P}_i): \quad \begin{cases} \min \left\{ P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{f} \right\} \\ s.t. \;\; x_i \in \{-1, 1\}, \quad i = 1, \ldots, n. \end{cases} \tag{68}$$

Both problems are known to be NP-hard and have been subjected to considerable study over the past several decades (see [31,33,47,48]).

In order to use the canonical dual transformation, the key step is to rewrite the linear box constraints in the following canonical form:

$$\mathbf{g}(\mathbf{x}) = \Lambda(\mathbf{x}) = \left\{ \frac{1}{2}\mathbf{x}^T\mathbf{B}^\alpha\mathbf{x} \right\} \le \mathbf{d}, \tag{69}$$

where $\mathbf{B}^\alpha = \{B^\alpha_{ij}\} \in \mathbb{R}^{n\times n}$, $\forall\, \alpha = 1, \ldots, n$, i.e.,

$$B^\alpha_{ij} = \begin{cases} 1 & \text{if } i = j = \alpha, \\ 0 & \text{otherwise,} \end{cases}$$

and $\mathbf{d} = \left\{\frac{1}{2}\right\}^n$ is a vector whose components are all $\frac{1}{2}$. Thus, the box constrained problem is actually the most simple case of the problem $(\mathcal{P}_q)$. Particularly, if the inequality constraint in (69) is replaced by

$$\mathbf{g}(\mathbf{x}) = \Lambda(\mathbf{x}) = \left\{ \frac{1}{2}\mathbf{x}^T\mathbf{B}^\alpha\mathbf{x} \right\} = \mathbf{d}, \tag{70}$$

then $\mathbf{x}$ has to be in $\{-1, 1\}^n$. Therefore, the Boolean programming problem $(\mathcal{P}_i)$ is actually a special case of the quadratically constrained quadratic program (where each equality constraint is equivalently represented by two inequalities).

The canonical dual problem in this case has an even simpler form (see [20]):

$$(\mathcal{P}^d_b): \quad \max \left\{ P^d_b(\boldsymbol{\sigma}) = -\frac{1}{2}\mathbf{f}^T G^+_b(\boldsymbol{\sigma})\mathbf{f} - \boldsymbol{\sigma}^T\mathbf{d} : \; \boldsymbol{\sigma} \in \mathcal{S}_b \right\}, \tag{71}$$

where

$$G_b(\boldsymbol{\sigma}) = A + \sum_{\alpha}^{n} \sigma_\alpha \mathbf{B}^\alpha, \tag{72}$$

$$\mathcal{S}_b = \{\boldsymbol{\sigma} \in \mathbb{R}^n_+ | \ \mathbf{f} \in \mathcal{C}_{ol}(G_b)\}. \tag{73}$$

Theorem 6 *If $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_b$ is a critical point of the canonical dual problem $(\mathcal{P}^d_b)$, then $\bar{\mathbf{x}} = G^+_b(\bar{\boldsymbol{\sigma}})\mathbf{f}$ is a KKT point of the box constrained problem $(\mathcal{P}_b)$ and $P(\bar{\mathbf{x}}) = P^d_b(\bar{\boldsymbol{\sigma}})$. If $\bar{\boldsymbol{\sigma}} > 0$, then $\bar{\mathbf{x}} \in \{-1, 1\}^n$ is a KKT point of the Boolean programming problem $(\mathcal{P}_i)$.*

If $G_b(\bar{\boldsymbol{\sigma}}) \succ 0$, then $\bar{\mathbf{x}} = G_b^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{f}$ is a global minimizer of the problem $(\mathcal{P}_b)$.
If $G_b(\bar{\boldsymbol{\sigma}}) \succ 0$ and $\bar{\boldsymbol{\sigma}} > 0$, then $\bar{\mathbf{x}} = G_b^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{f}$ is a global minimizer of the problem $(\mathcal{P}_i)$. □

The proof of this theorem can be found in [20]. This theorem shows that by the canonical dual transformation, the nonconvex integer programming problem $(\mathcal{P}_i)$ can be converted to the following continuous concave maximization dual problem:

$$(\mathcal{P}^d_i): \quad \max \left\{ P^d_b(\boldsymbol{\sigma}) = -\frac{1}{2}\mathbf{f}^T G_b^{-1}(\boldsymbol{\sigma})\mathbf{f} - \boldsymbol{\sigma}^T\mathbf{d} : \ \boldsymbol{\sigma} \in \mathcal{S}^+_i \right\}, \tag{74}$$

where $\mathcal{S}^+_i = \{\boldsymbol{\sigma} \in \mathbb{R}^n_+ | \ \boldsymbol{\sigma} > 0, \quad G_b(\boldsymbol{\sigma}) \succ 0\}$. By the fact that $\mathbf{d} = \{\frac{1}{2}\}^n > \mathbf{0}$ and

$$\lim_{\|\boldsymbol{\sigma}\| \to \infty} P^d_b(\boldsymbol{\sigma}) = -\infty,$$

we know from Theorem 5 that the canonical dual problem $(\mathcal{P}^d_b)$ has at least one KKT point $\bar{\boldsymbol{\sigma}} \in \mathcal{S}^+_b$. If $\bar{\boldsymbol{\sigma}} \in \mathcal{S}^+_i \subset \mathcal{S}^+_b$, then $\bar{\mathbf{x}} = G_b^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{f}$ is a global minimizer for $(\mathcal{P}_i)$. A detailed study on canonical dual approaches for 0–1 programming and multi-integer programming is given in [3,25,49].

Example 2 For a given vector $\mathbf{f} \in \mathbb{R}^n$, we consider the following constrained convex maximization problem:

$$\max\{\|\mathbf{x} + \mathbf{f}\|_2 : \|\mathbf{x}\|_\infty \leq 1\}, \tag{75}$$

which is equivalent to the following concave quadratic minimization problem:

$$\min \left\{ P(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T\mathbf{x} - \mathbf{x}^T\mathbf{f} : |x_i| \leq 1, \quad \forall i = 1, \ldots, n, \quad \mathbf{x}^T\mathbf{x} \leq r \right\}, \tag{76}$$

where $r > n$ to ensure that the additional quadratic constraint $\mathbf{x}^T\mathbf{x} \leq r$ in the feasible space $\mathcal{X}_c = \{\mathbf{x} \in \mathbb{R}^n | \ -1 \leq x_i \leq 1, \quad \forall i = 1, \ldots, n, \quad \mathbf{x}^T\mathbf{x} \leq r\}$ is never active. Therefore, (76) is indeed a box constrained concave minimization problem. It is known that for high dimensional nonconvex constrained optimization problems, to check which constraints are active is fundamentally difficult [42].

If we let $n = 2$, $r = 100$, and $\mathbf{f} = (1, 1)^T$, the optimal solution is $\bar{\mathbf{x}} = (1, 1)^T$ with objective value $P(\bar{\mathbf{x}}) = -3$. To illustrate the difficulty of applying the classical Lagrangian duality theory directly to (76), we first introduce Lagrange multipliers $(\sigma_1, \sigma_2, \sigma_3, \sigma_4)^T \in \mathbb{R}^4_+$ to relax the linear box constraints $-1 \leq x_i \leq 1, \ i = 1, 2$, and $\sigma_5 \geq 0$ to relax the quadratic constraint $\frac{1}{2}\mathbf{x}^T\mathbf{x} \leq 50$. The Lagrangian associated with (76) is

$$L(\mathbf{x}, \boldsymbol{\sigma}) = -\frac{1}{2}(x_1^2 + x_2^2) - (x_1 + x_2) + \sum_{i=1}^{2} \left[\sigma_i(x_i - 1) - \sigma_{i+2}(x_i + 1)\right] + \frac{1}{2}\sigma_5\left(x_1^2 + x_2^2 - 100\right),$$

with the Lagrangian dual function given by

$$P^*(\boldsymbol{\sigma}) = \min_{\mathbf{x}\in\mathbb{R}^2} L(\mathbf{x}, \boldsymbol{\sigma}).$$

When $\sigma_5 < 1$, we get $P^*(\boldsymbol{\sigma}) = -\infty$. When $\sigma_5 = 1$, we obtain

$$\max_{\boldsymbol{\sigma}\geq 0}\{P^*(\boldsymbol{\sigma}) : \sigma_5 = 1\} = -52$$

at the solution $\sigma_0 = (1, 1, 0, 0, 1)$. Finally, for any given $\boldsymbol{\sigma} \in \mathcal{S}_r = \{\boldsymbol{\sigma} \in \mathbb{R}^5_+ | \ \sigma_5 > 1\}$, the Lagrangian dual function can be obtained as

$$P^*(\boldsymbol{\sigma}) = -\frac{1}{2(\sigma_5 - 1)}\left[(1 - \sigma_1 + \sigma_3)^2 + (1 - \sigma_2 + \sigma_4)^2\right] - \sum_{i=1}^{4}\sigma_i - 50\sigma_5.$$

It is easy to check that the solution to the Lagrangian dual problem

$$\sup\{P^*(\boldsymbol{\sigma}) : \boldsymbol{\sigma} \in \mathcal{S}_r\} = -52,$$

realized as $\boldsymbol{\sigma} \to \boldsymbol{\sigma}_o = (1, 1, 0, 0, 1)^T$. Since $\boldsymbol{\sigma}_o$ is not a critical point of $P^*(\boldsymbol{\sigma})$, both Theorems 4 and 6 do not apply for this case and the associated $\mathbf{x}_o$ is not a KKT point of the primal problem (76). Hence, there exists a duality gap between the primal and the Lagrangian dual problem, i.e.,

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x}\in\mathcal{X}_c} P(\mathbf{x}) = -3 > -52 = \max_{\boldsymbol{\sigma}\in\mathbb{R}^5_+} P^*(\boldsymbol{\sigma}) = P^*(\boldsymbol{\sigma}_o).$$

To close this duality gap, we rewrite the constraints in the canonical form

$$\mathbf{g}(\mathbf{x}) = \Lambda(\mathbf{x}) = \left\{\begin{array}{l} \frac{1}{2}x_1^2 \\ \frac{1}{2}x_2^2 \\ \frac{1}{2}(x_1^2 + x_2^2) \end{array}\right\} \leq \left\{\begin{array}{l} \frac{1}{2} \\ \frac{1}{2} \\ 50 \end{array}\right\} = \mathbf{d}$$

as defined in (69) with

$$B_{ij}^{\alpha} = \begin{cases} 1 & \text{if } i = j = \alpha, \\ 0 & \text{otherwise}, \end{cases} \quad i, j, \alpha = 1, 2, \quad B_{ij}^{3} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \quad i, j = 1, 2$$

and $\mathbf{d} = 0.5(1, 1, 100)^T$. In effect, this reformulates (76) for this instance as follows:

$$\min\left\{P(\mathbf{x}) = -\frac{1}{2}(x_1^2 + x_2^2) - (x_1 + x_2) \ : \ \frac{1}{2}x_1^2 \leq \frac{1}{2}, \quad \frac{1}{2}x_2^2 \leq \frac{1}{2}, \quad \frac{1}{2}(x_1^2 + x_2^2) \leq 50\right\}. \tag{77}$$

The associated total complementary function (49) has a simple form:

$$\Xi_q(\mathbf{x}, \boldsymbol{\sigma}) = \frac{1}{2}x_1^2(\sigma_1 + \sigma_3 - 1) + \frac{1}{2}x_2^2(\sigma_2 + \sigma_3 - 1) - (x_1 + x_2) - \frac{1}{2}(\sigma_1 + \sigma_2) - 50\sigma_3.$$

Then, on the canonical dual feasible space $\mathcal{S}_q^+ = \{\boldsymbol{\sigma} \in \mathbb{R}^3_+ | \ \sigma_i + \sigma_3 - 1 > 0, \ \ i = 1, 2\}$ the canonical dual problem $(\mathcal{P}_q^d)$ is

$$\max\left\{ P_q^d(\boldsymbol{\sigma}) = -\frac{1}{2}\left(\frac{1}{\sigma_1+\sigma_3-1} + \frac{1}{\sigma_2+\sigma_3-1}\right) - \frac{1}{2}(\sigma_1+\sigma_2) - 50\sigma_3 \ : \ \boldsymbol{\sigma} \in \mathcal{S}_q^+ \right\}. \tag{78}$$

The optimal solution for this concave maximization problem is $\bar{\boldsymbol{\sigma}} = (2, 2, 0)^T$ with the optimal value $P_q^d(\bar{\boldsymbol{\sigma}}) = -3$. Observe that $\bar{\sigma}_3 = 0$ reflects the fact that the quadratic constraint $\mathbf{x}^T\mathbf{x} \le r$ is inactive. Since $\bar{\boldsymbol{\sigma}} \in \mathcal{S}_q^+$ is a critical point of $P_q^d(\boldsymbol{\sigma})$, therefore, the vector $\bar{\mathbf{x}} = G_q^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{f} = (1, 1)^T$ is a global minimizer of the primal problem with zero duality gap.

Finally, it is insightful to note that *once the primal problem has been reformulated* as (77), the classical Lagrangian dual is then given by $\max\{P^*(\boldsymbol{\sigma}) : \ \boldsymbol{\sigma} \ge 0\}$, where $P^*(\boldsymbol{\sigma}) = \min_{\mathbf{x}}\{\Xi_q(\mathbf{x}, \boldsymbol{\sigma})\}$. It is readily verified that for $\boldsymbol{\sigma} \in \mathcal{S}_q^+$, $P^*(\boldsymbol{\sigma})$ is given by $P_q^d(\boldsymbol{\sigma})$ as in (78), while $P^*(\boldsymbol{\sigma}) = -\infty$ otherwise. Hence, under the sufficient global optimality condition $\boldsymbol{\sigma} \in \mathcal{S}_q^+$, the Lagrangian dual for the *canonically reformulated* problem (77) is precisely given by the canonical dual (78).

Remark 3 This example shows the difficulty of directly applying the classical Lagrangian duality for solving nonconvex minimization problems with linear (including both box and integer) constraints. The classical Lagrangian duality theory was originally developed for linearly constrained convex variational problems in analytical mechanics, where the Lagrange multipliers and the linear constraints possess certain perfect duality (work-conjugate) relations [10,38]. In convex analysis, the classical (saddle) Lagrangian duality theory was generalized for solving the following convex minimization problem [2,44]:

$$\min\{P(\mathbf{x}) = W(\Lambda\mathbf{x}) - F(\mathbf{x})\},$$

where Λ is a linear operator, $W(\mathbf{y})$ is a convex (nonsmooth) function, and $F(\mathbf{x})$ is a concave (or linear) function. By using the Fenchel transformation, the so-called Fenchel–Moreau–Rockafellar dual problem is formulated as

$$\max\{P^d(\boldsymbol{\sigma}) = F^\sharp(\Lambda^*\boldsymbol{\sigma}) - W^\sharp(\boldsymbol{\sigma})\},$$

where Λ^* is an adjoint operator of Λ. In constrained problems, if the function $W(\Lambda\mathbf{x})$ is an indicator of the linear (geometrical) constraint $\Lambda\mathbf{x} \le \mathbf{d}$ as defined by Equation (7), then by the Lagrange multiplier law (see [10], p. 36–37), the Lagrange multiplier $\boldsymbol{\sigma}$ for the primal problem should be a solution for the dual problem. Dually, if the Fenchel conjugate $F^\sharp$ is an indicator of the *balance constraint* $\Lambda^*\boldsymbol{\sigma} = \mathbf{f}$ [10], then the Lagrange multiplier $\mathbf{x}$ for the dual problem should be a solution to the primal problem. The convexity of the primal function $P(\mathbf{x})$ leads to the strong duality $\min P(\mathbf{x}) = \max P^d(\boldsymbol{\sigma})$. In physics and continuum mechanics, the perfect dual function is called the *complementary energy* and the strong duality relation is controlled by certain conservation laws. However, the Lagrange multiplier method has been misused for solving general nonlinear constrained nonconvex problems. The primal problem (76) in Example 2 has both linear and nonlinear (quadratic) constraints, the Lagrange multipliers $\sigma_i, \ i = 1, 2, 3, 4$ are dual to the linear constraints, while σ_5 is dual to the quadratic constraint. Since the linear and nonlinear constraints are different geometrical measures, their corresponding dual variables, i.e., the Lagrange multipliers $\sigma_i, \ i = 1, 2, 3, 4$, and σ_5 are in different metric spaces with different (physical) units. Therefore, the classical Lagrangian dual problem in this case does not make physical sense. The *weak Lagrangian duality theory* leads to various duality gaps and violates the conservation law.

Generally speaking, if the objective function $P(\mathbf{x})$ is nonconvex, the optimal solutions are usually located on the boundary of the feasible set. To determine active constraints is a fundamentally difficult task. For a nonconvex polynomial $P(\mathbf{x})$, the effective domain of the Lagrangian dual $P^*(\boldsymbol{\sigma}) = \sup_\sigma L(\mathbf{x}, \boldsymbol{\sigma})$ is usually empty. Therefore, many extended Lagrangian duality methods have been developed.

As indicated in [12,16,27], the **key step** in the canonical dual transformation is to choose a (nonlinear) geometrical operator $\boldsymbol{\xi} = \Lambda(\mathbf{x})$ such that the duality relation $\boldsymbol{\varsigma} = \boldsymbol{\varsigma}(\boldsymbol{\xi})$ defined by (13) is invertible (one-to-one) and the nonconvex primal problem can be written in the canonical form (12). Detailed discussion on the geometrical operator $\Lambda(\mathbf{x})$ and the canonical dual pairs $(\boldsymbol{\xi}, \boldsymbol{\varsigma})$ are given in [10] (Chap. 6). For most of problems in nonconvex analysis and global optimization, the quadratic geometrical operator $\Lambda(\mathbf{x})$ can be used to formulate canonical dual problems [12,27]. By the fact that the total complementary function is identical to the Lagrangian for the reformulated primal problem (77), the function $\Xi_q(\mathbf{x}, \boldsymbol{\sigma})$ was also called the *nonlinear (or extended) Lagrangian* in [10]. In physics and continuum mechanics, since the terminology *complementarity* represents perfect duality, it is more appropriate to name $\Xi_q(\mathbf{x}, \boldsymbol{\sigma})$ as the total complementary function.

According to the canonical duality theory, when the primal problem is reformulated in the canonical form, the canonical dual function and its feasible space can be formulated precisely via the canonical dual transformation. Both global and local extremality conditions are governed by the triality theory. It is interesting to note that by using the quadratic geometrical measure, different primal problems in differential fields can be transformed to a unified canonical dual form. Actually, the canonical dual function $P_q^d(\boldsymbol{\sigma})$ was first proposed in nonconvex analysis and mechanics [7–9], which leads to analytical solutions for a class of nonconvex variational/boundary value problems [11,22,23].

In continuum mechanics, the quadratic measure $\boldsymbol{\xi} = \Lambda(\mathbf{x})$ corresponds to the Cauchy-Riemann type strain tensor, while its canonical dual $\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi})$ is the well-known second Piola–Kirchhoff stress tensor. For most of hyper-elastic materials, the strain energy $V(\boldsymbol{\xi})$ is a convex function of the canonical strain measure $\boldsymbol{\xi}$ (see [10], Chap. 6). In plasticity of solids, the box constraint $\|\mathbf{x}\|_\infty \le 1$ corresponds to the well-known Tresca yield condition, while the condition $\|\mathbf{x}\|_2 \le \sqrt{n}$ corresponds to the von Mises yield condition (see [10], p. 404). In the famous experiment by Taylor and Quinney (1931), the experimental data for most metals lies between the two criteria; however, the data are generally closer to the von Mises yield condition. By using the quadratic operator $\Lambda(\mathbf{x}) = \|\mathbf{x}\|_2^2$, the canonical dual transformation was first proposed to solve nonlinear finite element programming problems in large scale plastic limit analyses of structures [5,6]. The vector-valued quadratic canonical dual transformation $\Lambda(\mathbf{x}) = \{x_i^2\}$ for solving box and integer constrained problems was presented in [3,20,25,49]. □

5 Nonconvex polynomial constrained problems

We now assume that $\mathbf{g}(\mathbf{x})$ is a general fourth order polynomial constraint given by

$$\mathbf{g}(\mathbf{x}) = \left\{ \sum_{\beta \in I_\alpha} \frac{1}{2} D_\alpha^\beta \left(\frac{1}{2} \mathbf{x}^T \mathbf{B}_\beta^\alpha \mathbf{x} + \mathbf{x}^T \mathbf{C}_\beta^\alpha - E_\beta^\alpha \right)^2 \right\} \le \mathbf{d}, \tag{79}$$

where $\mathbf{B}_\beta^\alpha = \{B_{ij\beta}^\alpha\} \in \mathbb{R}^{n \times n}$ and $\mathbf{C}_\beta^\alpha = \{C_{i\beta}^\alpha\} \in \mathbb{R}^n$ are given as before, I_α is a (finite) index set that depends on each index $\alpha = 1, \ldots, m$, and $\{D_\alpha^\beta\}$ and $\{E_\beta^\alpha\}$ are two given second order tensors. We assume that $D_\alpha^\beta > 0, \quad \forall \alpha \in \{1, \ldots, m\}, \ \beta \in I_\alpha$. This function is also called the fourth order canonical polynomial as discussed in [10,19], which arises in many applications (e.g., in Euclidean distance problems [28,36,45], post-buckling analysis in large deformation beam theory [13], finite element analysis for phase transitions of super-conductivity [29], and sensor network localization [30]).

By introducing a geometrical measure

$$\boldsymbol{\xi} = \Lambda(\mathbf{x}) = \{\xi_\beta^\alpha\} = \left\{ \frac{1}{2}\mathbf{x}^T \mathbf{B}_\beta^\alpha \mathbf{x} + \mathbf{x}^T \mathbf{C}_\beta^\alpha \right\} : \mathbb{R}^n \to \mathcal{E}_a,$$

where the range of $\mathcal{E}_a$ depends on the tensors $\{\mathbf{B}_\beta^\alpha\}$ and $\{\mathbf{C}_\beta^\alpha\}$, the canonical function

$$V(\boldsymbol{\xi}) = \left\{ \sum_{\beta \in I_\alpha} \frac{1}{2} D_\alpha^\beta \left(\xi_\beta^\alpha - E_\beta^\alpha \right)^2 \right\} : \mathcal{E}_a \to \mathbb{R}^m$$

is a quadratic function. Thus, the canonical duality relation

$$\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) = \{D_\alpha^\beta (\xi_\beta^\alpha - E_\beta^\alpha)\} : \mathcal{E}_a \to \mathcal{E}_a^*$$

is a linear mapping, where the range of $\mathcal{E}_a^*$ depends on the tensors $\{D_\alpha^\beta\}$ and $\{E_\beta^\alpha\}$. The Legendre conjugate V^* can be defined uniquely as:

$$V^*(\boldsymbol{\varsigma}) = \left\{ \sum_{\beta \in I_\alpha} \left(\frac{1}{2D_\alpha^\beta} (\varsigma_\alpha^\beta)^2 + E_\beta^\alpha \varsigma_\alpha^\beta \right) \right\}. \tag{80}$$

Substituting this into (34), the canonical dual function has the following form:

$$P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) = -\frac{1}{2}\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})^T G_a^+(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) - \sum_{\alpha=1}^m \sum_{\beta \in I_\alpha} \left(\sigma_\alpha \left(\frac{1}{2D_\alpha^\beta} (\varsigma_\alpha^\beta)^2 + E_\beta^\alpha \varsigma_\alpha^\beta + d_\alpha \right) \right), \tag{81}$$

which is concave on the dual feasible space

$$\mathcal{S}_c^+ = \{(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathbb{R}_+^m \times \mathcal{E}_a^* | \ \mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{C}_{ol}(G_a), \quad G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \succeq 0\}. \tag{82}$$

Remark 4 Similar to Remark 2, it is again insightful to view the connection between the canonical dual (81) and the classical Lagrangian dual defined by (3)–(5). In this case, we have as in (3) that

$$L(\mathbf{x}, \boldsymbol{\sigma}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T \mathbf{f} + \boldsymbol{\sigma}^T (\mathbf{g}(\mathbf{x}) - \mathbf{d}). \tag{83}$$

Clearly, without introducing the canonical dual pair $(\boldsymbol{\xi}, \boldsymbol{\varsigma})$, the Fenchel-Moreau-Rockafellar dual $P^*(\boldsymbol{\sigma}) = \inf_{\mathbf{x} \in \mathcal{X}_a} L(\mathbf{x}, \boldsymbol{\sigma})$ defined by (4) cannot be defined explicitly due to the high order nonlinearity of the constraint $\mathbf{g}(\mathbf{x}) \le \mathbf{d}$. However, by using the canonical dual transformation $\boldsymbol{\xi} = \Lambda(\mathbf{x})$ and the chain rule, the necessary condition for the optimization in (4), i.e., $\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\sigma}) = 0$ leads to

$$A\mathbf{x} - \mathbf{f} + \sum_{\alpha=1}^m \sum_{\beta \in I_\alpha} \left(\sigma_\alpha \varsigma_\alpha^\beta (\mathbf{B}_\beta^\alpha \mathbf{x} + \mathbf{C}_\beta^\alpha) \right) = 0. \tag{84}$$

This is the canonical equilibrium equation $G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{x} = \mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$, where

$$\boldsymbol{\varsigma} \equiv \frac{\partial V(\Lambda(\mathbf{x}))}{\partial \boldsymbol{\xi}} \tag{85}$$

is as defined above. If $G_a(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is invertible for $(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c$ as defined in (33), we get from (84) that $\mathbf{x}$ uniquely satisfies $\mathbf{x} = G_a^{-1}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$. Furthermore, by (84), we get

$$\frac{1}{2}\mathbf{x}^T A\mathbf{x} = \frac{1}{2}\mathbf{x}^T\mathbf{f} - \frac{1}{2}\sum_{\alpha=1}^{m}\sum_{\beta\in I_\alpha} \boldsymbol{\sigma}_\alpha \varsigma_\alpha^\beta[\mathbf{x}^T B_\beta^\alpha \mathbf{x} + \mathbf{x}^T C_\beta^\alpha]. \tag{86}$$

Substituting for $\frac{1}{2}\mathbf{x}^T A\mathbf{x}$ in (83) using (86), and then applying the optimality condition $\mathbf{x} = G_a^{-1}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$, the Lagrangian dual function reduces precisely to $P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ defined in (81) under the global optimality condition $(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+$. Note that $\nabla_{\boldsymbol{\varsigma}} \Xi(\mathbf{x}, \boldsymbol{\sigma}, \boldsymbol{\varsigma}) = 0$ produces the inverse of the identity (85) under the relevant case when $\boldsymbol{\sigma}^\alpha \neq 0, \ \forall \alpha = 1, \ldots, m$, thus validating the foregoing derivation.

In finite deformation theory, the geometrical measure $\boldsymbol{\xi} = \Lambda(\mathbf{x})$ is called the *canonical strain tensor* and the canonical duality relation (85) is called the *constitutive law*. This one-to-one constitutive relation leads to the canonical dual function P^d. The associated Theorem 1 is called the *pure complementary variational principle*, first developed in nonconvex variational analysis [7,8]. This theorem solves an open problem in nonconvex mechanics posed by Hellinger (1914) and Reissner (1953) (see [32,43]) and it is known as the Gao principle [39]. □

We now present some special case applications for (79).

5.1 Quadratic minimization with one nonconvex polynomial constraint

We first assume that the primal problem has only one nonconvex constraint:

$$g(\mathbf{x}) = \frac{1}{2}\left(\frac{1}{2}\mathbf{x}^T B\mathbf{x} + \mathbf{x}^T\mathbf{c} - \eta\right)^2 \leq d, \tag{87}$$

where B is an $n \times n$ matrix, $\mathbf{c} \in \mathbb{R}^n$ is a vector, and $\eta > 0$ is a constant. In physics, this nonconvex fourth order polynomial is known as the *double-well function* and it appears in many applications [10]. In this case, $m = |I_\alpha| = 1$, and

$$G_a(\sigma, \varsigma) = A + \sigma\varsigma B, \quad \mathbf{F}(\sigma, \varsigma) = \mathbf{f} - \sigma\varsigma\mathbf{c}.$$

The canonical dual function is

$$P^d(\sigma, \varsigma) = -\frac{1}{2}\mathbf{F}^T(\sigma, \varsigma)G_a^+(\sigma, \varsigma)\mathbf{F}(\sigma, \varsigma) - \sigma\left(\frac{1}{2}\varsigma^2 + \eta\varsigma + d\right). \tag{88}$$

Example 3 In 2-D space, let B be an identity matrix, $\mathbf{c} = \mathbf{0}$, and let A be a diagonal matrix with $a_{11} = 0.6$, $a_{12} = a_{21} = 0$, and $a_{22} = -0.5$. Setting $\mathbf{f} = (0.2, -0.1)^T$, $d = 1$, and $\eta = 1.5$, the constraint $g(\mathbf{x}) \leq d$ is an annulus (see Fig. 3 (right)). Solving the dual problem, we get

$$\bar{\sigma} = 0.3829201, \quad \text{and} \quad \bar{\varsigma} = 1.4142136.$$

The primal solution

$$\bar{\mathbf{x}} = \begin{pmatrix} 0.1752033 \\ -2.4078477 \end{pmatrix}$$

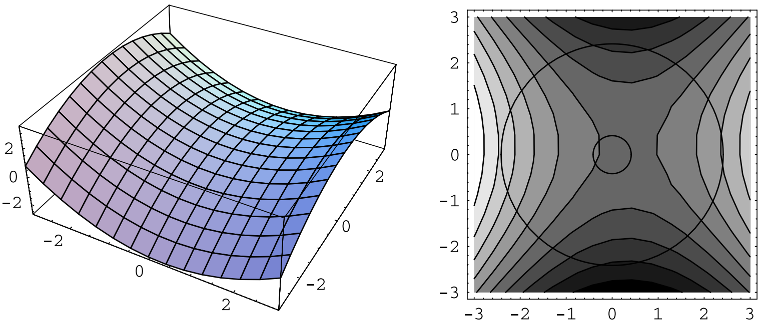


Fig. 3 Graph of $P(\mathbf{x})$ (left); contours of $P(\mathbf{x})$ and boundary of $\mathcal{X}_c$ (right) for Example 3

is located on the boundary of the feasible set $\mathcal{X}_c$ (see Fig. 3) and we have

$$P(\bar{\mathbf{x}}) = -1.7160493 = P^d(\bar{\sigma}, \bar{\varsigma}).$$

5.2 Combined quadratic and nonconvex polynomial constraints

We now consider the problem with the following two constraints:

$$g_1(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{B}^1\mathbf{x} + \mathbf{x}^T\mathbf{C}^1 \le d_1,$$
$$g_2(\mathbf{x}) = \frac{1}{2}\left(\frac{1}{2}\mathbf{x}^T\mathbf{B}^2\mathbf{x} + \mathbf{x}^T\mathbf{C}^2 - \eta\right)^2 \le d_2.$$

In this case, $m = 2$, $I_\alpha = \{1\}$, for $\alpha = 1, 2$, and the geometrical operator

$$\boldsymbol{\xi} = \Lambda(\mathbf{x}) = \left\{\frac{1}{2}\mathbf{x}\mathbf{B}^\alpha\mathbf{x} + \mathbf{x}^T\mathbf{C}^\alpha\right\} : \mathbb{R}^n \to \mathbb{R}^2$$

is a 2-vector. The canonical function $V(\boldsymbol{\xi})$ is a vector-valued function

$$V(\boldsymbol{\xi}) = \left\{\xi_1, \frac{1}{2}(\xi_2 - \eta)^2\right\}.$$

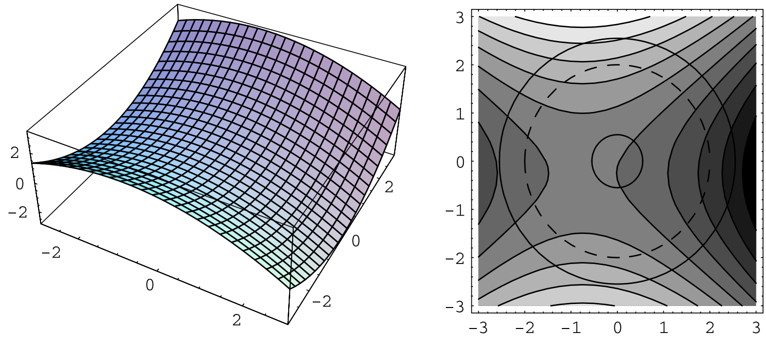
The canonical dual variable is $\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) = \{1, \xi_2 - \eta\}$. Since $\varsigma_1 = 1$, we let $\varsigma_2 = \varsigma$. Thus, the canonical dual function has only three variables $(\sigma_1, \sigma_2, \varsigma) \in \mathbb{R}^3$, i.e.,

$$P^d(\sigma_1, \sigma_2, \varsigma) = -\frac{1}{2}\mathbf{F}(\sigma_1, \sigma_2, \varsigma)^T G_a^+(\sigma_1, \sigma_2, \varsigma)\mathbf{F}(\sigma_1, \sigma_2, \varsigma)$$
$$-\sigma_1 d_1 - \sigma_2\left(\frac{1}{2}\varsigma^2 + \eta\varsigma + d_2\right), \tag{89}$$

where

$$G_a(\sigma_1, \sigma_2, \varsigma) = A + \sigma_1\mathbf{B}^1 + \sigma_2\varsigma\mathbf{B}^2, \quad \mathbf{F}(\sigma_1, \sigma_2, \varsigma) = \mathbf{f} - \sigma_1\mathbf{C}^1 - \sigma_2\varsigma\mathbf{C}^2.$$

Example 4 We consider A to be a 2×2 diagonal matrix, i.e., $a_{11} = -0.4$, $a_{12} = a_{21} = 0$, and $a_{22} = 0.6$. Setting $\mathbf{f} = (0.3, -0.15)^T$, $\mathbf{B}^1 = \mathbf{B}^2 = \mathbf{I}$, $\mathbf{C}^1 = \mathbf{C}^2 = \mathbf{0}$, $d_1 = 2$, $d_2 = 1.2$, and $\eta = 1.7$, the graph of the objective function $P(x_1, x_2)$ is a saddle surface (Fig. 4(left)), the constraint $g_1(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2^2) \le 2$ is a disk of radius 2, while



(a) **(b)**

Fig. 4 (**a**) Graph of $P(\mathbf{x})$; (**b**) Contours of $P(\mathbf{x})$, constraints $g_1(x_1, x_2) \leq d_1$ (disk with radius $R \leq 2$, dashed circle), and $g_2(x_1, x_2) \leq d_2$ (annulus with radius $0.55 \leq R \leq 2.55$) for Example 4

$g_2(x_1, x_2) = \frac{1}{2}\left(\frac{1}{2}(x_1^2 + x_2^2) - 1.7\right)^2 \leq 1.2$ represents an annulus (see Fig. 4 (right)). Solving the dual problem, we get

$$\bar{\sigma}_1 = 0.5503198, \quad \bar{\sigma}_2 = 0, \quad \text{and} \quad \bar{\varsigma} = 0.3159349.$$

The primal solution is therefore

$$\bar{\mathbf{x}} = \begin{pmatrix} 1.9957445 \\ -0.1303985 \end{pmatrix},$$

which is located on the boundary $g_1(\bar{x}_1, \bar{x}_2) = 0$, and

$$P(\bar{\mathbf{x}}) = -1.4097812 = P^d(\bar{\sigma}_1, \bar{\sigma}_2, \bar{\varsigma}).$$

Concrete applications of quadratic minimization with two nonconvex polynomial constraints are given in [28].

6 Concluding remarks and open problems

We have presented a detailed application of the canonical duality theory to the general differentiable nonconvex optimization problem. This problem arises in many real-world applications. Using the canonical dual transformation, a unified canonical dual problem was formulated with zero duality gap, which can be solved by well-developed nonlinear optimization methods. Both global and local optimizers can be identified by the triality theory. Insightful connections of this canonical duality with the classical Lagrangian duality have also been presented for two special applications in order to highlight the main constructs that enable the perfect duality results. Furthermore, these applications show how:

(1) the n-dimensional nonconvex constrained problem $(\mathcal{P}_q)$ in Sect. 3 can be reformulated as an m-dimensional concave maximization dual problem $(\mathcal{P}_q^d)$ over a convex space $\mathcal{S}_q^+$ with $m < n$ ($m = 1$ in Example 1);
(2) the nonconvex discrete integer programming problem $(\mathcal{P}_i)$ in Sect. 4 can be converted to a concave maximization dual problem $(\mathcal{P}_i^d)$ over a convex continuous space $\mathcal{S}_i^+$, which can be solved easily if $\mathcal{S}_i^+$ is non-empty.

Generally speaking, optimal solutions for constrained nonconvex minimization problems are usually KKT points located on the boundary of the feasible sets. Due to the lack of global optimality criteria, it is very difficult for direct methods and the classical Lagrangian relaxations to find global minimizers. Therefore, most of these problems are considered to be NP-hard (see [42,46]). However, by the canonical duality theory, these KKT points can be easily determined by the critical points of the canonical dual problems. The triality theory can be used to develop potentially powerful algorithms for solving these problems.

The canonical duality theory concepts and methodologies presented in this article can be used and generalized to solve many other difficult problems in global optimization, nonconvex analysis and mechanics, network communication, and scientific computations. In general, so long as the geometrical operator Λ is chosen properly, the canonical dual transformation method can be used to formulate perfect dual problems and the triality theory can be used to establish useful theoretical results. If the convex dual feasible space $\mathcal{S}_c^+$ is not empty, the canonical dual $\max\{P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) : (\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c^+\}$ can be solved easily by well-developed convex minimization techniques. As indicated in [21], the primal problem $(\mathcal{P}^d)$ could be NP-hard for the class of problems where the canonical dual function has no critical point in $\mathcal{S}_c^+$. In this case, the primal problem (19) is equivalent to the following canonical dual minimal stationary problem:

$$(\mathcal{P}_g^d): \quad \min \operatorname{sta}\{P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \ : \ (\boldsymbol{\sigma}, \boldsymbol{\varsigma}) \in \mathcal{S}_c\}. \tag{90}$$

Since $P^d(\boldsymbol{\sigma}, \boldsymbol{\varsigma})$ is usually nonconvex on $\mathcal{S}_c$, to solve this minimal stationary problem could be a challenging task and many theoretical issues remain open. For further details and comprehensive applications of the canonical duality theory, we refer the reader to [4,10,21,24,26,45].

Acknowledgements The authors would like to thank the two anonymous referees and an associate editor for their detailed comments and insightful suggestions. In particular, Example 2 was suggested by one of reviewers and Remark 3 is based on the questions raised by this reviewer and the associate editor. This research was supported by the *National Science Foundation* under grants CCF-0514768 and CMMI-0552676.

References

1. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms, 3rd edn. Wiley, New York (2006)
2. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland, Amsterdam (1976)
3. Fang, S.C., Gao, D.Y., Sheu, R.L., Wu, S.Y.: Canonical dual approach for solving 0-1 quadratic programming problems. J. Ind. Manag. Optim. **4**(1), 125–142 (2008)
4. Fang, S.C., Gao, D.Y., Sheu, R.L., Xing, W.X.: Global optimization for a class of fractional programming problems. J. Glob. Optim., (2008) (in press)
5. Gao, D.Y.: Extended bounding theorems of limit analysis. Appl. Math. Mech. **4**, 571–584 (1983)
6. Gao, D.Y.: Panpenalty finite element programming for limit analysis. Comput. Struct. **28**, 749–755 (1988)
7. Gao, D.Y.: Duality, triality and complementary extremun principles in nonconvex parametric variational problems with applications. IMA J. Appl. Math. **61**, 199–235 (1998)
8. Gao, D.Y.: Pure complementary energy principle and triality theory in finite elasticity. Mech. Res. Comm. **26**(1), 31–37 (1999)
9. Gao, D.Y.: General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics. Meccanica **34**, 169–198 (1999)
10. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications. pp. 454 Kluwer Academic Publishers, Dordrecht (2000)
11. Gao, D.Y.: Analytic solution and triality theory for nonconvex and nonsmooth vatiational problems with applications. Nonlinear Anal. **42**(7), 1161–1193 (2000)

12. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
13. Gao, D.Y.: Finite deformation beam models and triality theory in dynamical post-buckling analysis. Int. J. Non-Linear Mech. **5**, 103–131 (2000)
14. Gao, D.Y.: Tri-duality in global optimization. In: Floudas, C.A., Pardalos, P.D. (eds.) Encyclopedia of Optimization, Vol. 1, pp. 485–491. Kluwer Academic Publishers, Dordrecht (2001)
15. Gao, D.Y.: Nonconvex Semi-linear Problems and Canonical Dual Solutions. Advances in Mechanics and Mathematics, Vol. II. pp. 261–312. Kluwer Academic Publishers, Dordrecht (2003)
16. Gao, D.Y.: Perfect duality theory and complete solutions to a class of global optimization problems. Optimization **52**(4–5), 467–493 (2003)
17. Gao, D.Y.: Complete solutions to constrained quadratic optimization problems, Special issue on Duality. J. Glob. Optim. **29**, 377–399 (2004)
18. Gao, D.Y.: Sufficient conditions and perfect duality in nonconvex minimization with inequality constraints. J. Ind. Manag. Optim. **1**, 59–69 (2005)
19. Gao, D.Y.: Complete solutions and extremality criteria to polynomial optimization problems. J. Glob. Optim. **35**, 131–143 (2006)
20. Gao, D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. J. Ind. Maneg. Optim. **3**(2), 293–304 (2007)
21. Gao, D.Y.: Advances in canonical duality theory with applications to global optimization. In: Ierapetriou, M., Bassett, M., Pistikopoulos, S. (eds.) Proceedings of the Fifth International Conference on Foundations of Computer-Aided Process Operations, pp.73–82. Omni Press, Cambridge, MA (2008)
22. Gao, D.Y., Ogden, R.W.: Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem. Zeitschrift für Angewandte Mathematik Und Physik **59**(3), 498–517 (2008)
23. Gao, D.Y., Ogden, R.W.: Multi-solutions to nonconvex variational problems with implications for phase transitions and numerical computation. Q. J. Mech. Appl. Math. **61**(4), 497–522 (2008)
24. Gao, D.Y., Ruan, N.: Solutions and optimality criteria for nonconvex quadratic-exponential minimization problem. Math. Methods Oper. Res. **67**(3), 479–496 (2008)
25. Gao, D.Y., Ruan, N.: On the solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. (2008) (in press)
26. Gao, D.Y., Sherali, H.D.: Canonical Duality Theory: Connections Between Nonconvex Mechanics and Global Optimization, Advances in Applied Mathematics and Global Optimization. pp. 249–316. Springer, New York (2008)
27. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Q. Appl. Math. **47**(3), 487–504 (1989)
28. Gao, D.Y., Yang, W.C.: Complete solutions to minimal distance problem between two nonconvex surfaces. Optimization **57**(5), 705–714 (2008)
29. 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)
30. Gao, D.Y., Ruan, N., Pardalos, P.M.: Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization, to appear. In: Pardalos, P.M., Ye, Y., Boginski, V.L., Commander, C.W. (eds.) Sensors: Theory, Algorithms, and Applications. Springer, Berlin (2008)
31. Hansen, P., Jaumard, B., Ruiz, M., Xiong, J.: Global minimization of indefinite quadratic functions subject to box constraints. Technical report, Technical Report G-91-54, Gread. École Polytechnique, Université McGill, Montreal (1991)
32. Hellinger, E.: Die allgemeine Ansätze der Mechanik der Kontinua. Encyklopädie der Mathematischen Wissenschaften **IV**(Part 4), 602–694 (1914)
33. Horst, R., Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization. Kluwer Academic Publishers, Dordrecht (2000)
34. Jeyakumar, V., Rubinov, A.M., Wu, Z.Y.: Sufficient global optimality conditions for non-convex quadratic minimization problems with box constraints. J. Glob. Optim. **36**(3), 471–481 (2006)
35. Jeyakumar, V., Rubinov, A.M., Wu, Z.Y.: Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions. Math. Program. **110**(3), 521–541 (2007)
36. Johnson, D.E., Cohen, E.: A framework for efficient minimum distance computations. In: Proc. IEEE International Conference on Robotics and Automation, pp. 3678–3684. Leuven, Belgium (1998)
37. Koiter, W.T.: On the complementary energy theorem in nonlinear elasticity theory. In: G. Fichera (ed.) Trends Appl. Pure Math. Mech., Pitman, London (1976)
38. Lagrange, J.L.: Mécanique Analytique. Gauthier-Villars, Paris (1788)
39. Li, S.F., Gupta, A.: On dual configuration forces. J. Elast. **84**, 13–31 (2006)
40. Nash, S.G., Sofer, A.: Linear and Nonlinear Programming. McGraw-Hill, New York (1996)

41. Ogden, R.W.: A note on variational theorems in non-linear elastostatics. Math. Proc. Camb. Phil. Soc **77**, 609–615 (1975)
42. Pardalos, P.M., Vavasis, S.A.: Quadratic programming with one negative eigenvalue is NP-hard. J. Glob. Optim. **21**, 843–855 (1991)
43. Reissner, E.: On a variational theorem for finite elastic deformations. J. Math. Phys., **32**(2–3), 129-135 (1953). (See also Selected Works in Applied Mechanics and Mathematics, Jones and Bartlett Publishers, Boston, MA, 1996)
44. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
45. Ruan, N., Gao, D.Y., Jiao, Y.: Canonical dual least square method for solving general nonlinear systems of quadratic equations. Comput. Optim. Appl. (published online: http://www.springerlink.com/content/c6090221p4g41858/) (2008). DOI:10.1007/s10589-008-9222-5
46. Sahni, S.: Computationally related problems. SIAM J. Comp. **3**, 262–279 (1974)
47. Sherali, H.D., Tuncbilek, C.: A global optimization for polynomial programming problem using a reformulation-linearization technique. J. Glob. Optim. **2**, 101–112 (1992)
48. Sherali, H.D., Tuncbilek, C.: New reformulation-linearization technique based relaxation for univariate and multivariate polynomial programming problems. Oper. Res. Lett. **21**(1), 1–10 (1997)
49. Wang, Z., Fang, S.C., Gao, D.Y., Xing, W.: Global extremal conditions for multi-integer quadratic programming. J. Ind. Manag. Optim. **4**(2), 213–226 (2008)