

Canonical Primal–Dual Method for Solving Nonconvex Minimization Problems

Changzhi Wu and David Yang Gao

Abstract A new primal–dual algorithm is presented for solving a class of nonconvex minimization problems. This algorithm is based on canonical duality theory such that the original nonconvex minimization problem is first reformulated as a convex–concave saddle point optimization problem, which is then solved by a quadratically perturbed primal–dual method. Numerical examples are illustrated. Comparing with the existing results, the proposed algorithm can achieve better performance.

1 Problems and Motivations

The nonconvex minimization problem to be studied is proposed as the following:

$$(\mathcal{P}_o) : \min \left\{ P(\mathbf{x}) = W(\mathbf{x}) + \frac{1}{2} \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{f} \rangle \mid \mathbf{x} \in \mathcal{X}_a \right\}, \quad (1)$$

where $\mathbf{x} = \{x_i\} \in \mathbb{R}^n$ is a decision vector, $\mathbf{A} = \{A_{ij}\} \in \mathbb{R}^{n \times n}$ is a given real symmetrical matrix, $\mathbf{f} = \{f_i\} \in \mathbb{R}^n$ is a given vector, $\langle *, * \rangle$ denotes a bilinear form in $\mathbb{R}^n \times \mathbb{R}^n$; the feasible space \mathcal{X}_a is an open convex subset of \mathbb{R}^n such that on which the nonconvex function $W : \mathcal{X}_a \rightarrow \mathbb{R}$ is well-defined.

Due to the nonconvexity, Problem (\mathcal{P}_o) may admit many local minima and local maxima [4]. It is not an easy task to identify or numerically compute its global minimizer. Therefore, many numerical methods have been developed in literature, including the extended Gauss–Newton method (see [22]), the proximal method (see [21]), as well as the popular semi-definite programming (SDP) relaxation (see [18]). Generally speaking, Gauss–Newton type methods are local-based such that only

C. Wu

Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia

D.Y. Gao (✉)

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

local optimal solutions can be expected. To find global optimal solution often relies on the branch-and-bound [2] as well as the moment matrix-based SDP relaxation [20, 36]. However, these methods are computationally expensive which can be used for solving mainly small or medium size problems. The main goal of this paper is to develop an efficient algorithm for solving the nonconvex problem (\mathcal{P}_o).

Generally speaking, a powerful algorithm should be based on a precise theory. Canonical duality theory is a newly developed, powerful methodological theory that has been used successfully for solving a large class of global optimization problems in both continuous and discrete systems [4, 6, 8]. The main feature of this theory is that, which depends on the objective function $W(\mathbf{x})$, the nonconvex/nonsmooth/discrete primal problems can be transformed into a unified concave maximization problem over a convex continuous space, which can be solved easily using well-developed convex optimization techniques (see review articles [6, 8] for details). This powerful theory was developed from Gao and Strang's original work [7] where the nonconvex function $W(\mathbf{x})$ is the so-called *stored energy*, which is required, by the concept (see [24], p. 8), to be an objective function. In mathematical physics, a real-valued function $W(\mathbf{x})$ is said to be *objective* if $W(\mathbf{Q}\mathbf{x}) = W(\mathbf{x})$ for all rotation matrix \mathbf{Q} such that $\mathbf{Q}^{-1} = \mathbf{Q}^T$ and $\det \mathbf{Q} = 1$ (see Chap. 6 in [4]), i.e., an objective function $W(\mathbf{x})$ should be an invariant under certain coordinate transformations. In continuum mechanics, the objectivity is also referred as the *frame-indifference* (see [1, 24]). Therefore, instead of the decision variables directly, an objective function usually depends on certain measure (norm) of \mathbf{x} , say, the Euclidean norm $\|\mathbf{x}\|$ as we have $\|\mathbf{Q}\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{Q}^T \mathbf{Q} \mathbf{x} = \|\mathbf{x}\|^2$. In this paper, we shall need only the following weak assumptions for the nonconvex function $W(\mathbf{x})$ in (\mathcal{P}_o).

Assumption 1

(A1). There exists a *geometrical operator* $\Lambda(\mathbf{x}) : \mathcal{X}_a \rightarrow \mathcal{V}_a \subset \mathbb{R}^m$ and a strictly convex differentiable function $V : \mathcal{V}_a \subset \mathbb{R}^m \rightarrow \mathbb{R}$ such that

$$W(\mathbf{x}) = V(\Lambda(\mathbf{x})) \quad \forall \mathbf{x} \in \mathcal{X}_a. \quad (2)$$

(A2). The geometrical operator $\Lambda(\mathbf{x})$ is a vector-valued quadratic mapping in the form of

$$\Lambda(\mathbf{x}) = \left\{ \frac{1}{2} \langle \mathbf{x}, \mathbf{A}_1 \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{b}_1 \rangle, \dots, \frac{1}{2} \langle \mathbf{x}, \mathbf{A}_m \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{b}_m \rangle \right\}, \quad (3)$$

where $\mathbf{A}_i, i = 1, \dots, m$, are symmetrical matrices with appropriate dimensions and $\mathbf{b}_i, i = 1, \dots, m$, are given vectors such that the range \mathcal{V}_a is a closed convex set in \mathbb{R}^m .

Actually, Assumption (A1) is the so-called *canonical transformation*. Particularly, if $\mathbf{A}_i \geq 0, \mathbf{b}_i = 0 \quad \forall i = 1, \dots, m$, then $\Lambda(\mathbf{x})$ is an objective (Cauchy–Riemann type) measure (see [4]). Based on this assumption, the proposed nonconvex problem (\mathcal{P}_o) can be reformulated in the following canonical form:

$$(\mathcal{P}) : \min \left\{ P(\mathbf{x}) = V(\Lambda(\mathbf{x})) + \frac{1}{2} \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{f} \rangle : \mathbf{x} \in \mathcal{X}_a \right\}. \quad (4)$$

The canonical primal problem (\mathcal{P}) arises naturally from a wide range of applications in engineering and sciences. For instance, the canonical function $V(\boldsymbol{\xi})$ is simply a quadratic function of $\boldsymbol{\xi} = \Lambda(\mathbf{x})$ in the least squares methods for solving systems of quadratic equations $\Lambda(\mathbf{x}) = \mathbf{d} \in \mathbb{R}^m$ (see [32]), chaotic dynamical systems [31], wireless sensor network localization [11], general Euclidean distance geometry [26], and computational biology [38]. In computational physics and networks optimization, the position variable \mathbf{x} is usually a matrix (second-order tensor) and the geometrical operator $\boldsymbol{\xi} = \Lambda(\mathbf{x})$ is a positive semi-definite (discredited Cauchy–Riemann measure) tensor (see [11]), the convex function $V(\boldsymbol{\xi})$ is then an objective function, which is the instance studied by Gao and Strang [4, 7]. Particularly, if $W(\mathbf{x})$ is a quadratic function, the canonical dual problem is equivalent to a SDP problem (see [11]). By the facts that the geometrical operator defined in Assumption (A2) is a general quadratic mapping, the nonconvex function $W(\mathbf{x})$ studied in this paper is not necessary to be “objective”, which certainly has extensive applications in complex systems.

The rest of this paper is divided into six sections. The canonical dual problem is formulated in the next section, where some existing difficulties are addressed. The associated canonical min-max duality theory is discussed in Sect. 3. A proximal point method is proposed in Sect. 4 to solve this canonical min-max problem. Section 5 presents some numerical experiments. Applications to sensor network optimization are illustrated in Sect. 6. The paper is ended by some concluding remarks.

2 Canonical Duality Theory

By Assumption (A1), the canonical function $V(\cdot)$ is strictly convex and differentiable on \mathcal{V}_a , therefore, the canonical dual mapping $\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) : \mathcal{V}_a \rightarrow \mathcal{V}_a^* \subset \mathbb{R}^m$ is one-to-one onto the convex set $\mathcal{V}_a^* \subset \mathbb{R}^m$ such that the following canonical duality relations hold on $\mathcal{V}_a \times \mathcal{V}_a^*$

$$\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) \Leftrightarrow \boldsymbol{\xi} = \nabla V^*(\boldsymbol{\varsigma}) \Leftrightarrow V(\boldsymbol{\xi}) + V^*(\boldsymbol{\varsigma}) = \langle \boldsymbol{\xi}; \boldsymbol{\varsigma} \rangle, \quad (5)$$

where $\langle *; * \rangle$ is a bilinear form on $\mathbb{R}^m \times \mathbb{R}^m$, and $V^*(\boldsymbol{\varsigma})$ is the Legendre conjugate of $V(\boldsymbol{\xi})$ defined by

$$V^*(\boldsymbol{\varsigma}) = \max \{ \langle \boldsymbol{\xi}; \boldsymbol{\varsigma} \rangle - V(\boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \mathcal{V}_a \}. \quad (6)$$

By convex analysis, we have

$$V(\Lambda(\mathbf{x})) = \max \{ \langle \Lambda(\mathbf{x}); \boldsymbol{\varsigma} \rangle - V^*(\boldsymbol{\varsigma}) \mid \boldsymbol{\varsigma} \in \mathcal{V}_a^* \}. \quad (7)$$

Substituting (7) into (4), Problem (\mathcal{P}) can be equivalently written as

$$\min_{\mathbf{x}} \max_{\boldsymbol{\zeta}} \{ \Xi(\mathbf{x}, \boldsymbol{\zeta}) \mid (\mathbf{x}, \boldsymbol{\zeta}) \in \mathcal{X}_a \times \mathcal{V}_a^* \}, \quad (8)$$

where $\Xi : \mathcal{X}_a \times \mathcal{V}_a^* \rightarrow \mathbb{R}$ is the *total complementary function* defined by

$$\begin{aligned} \Xi(\mathbf{x}, \boldsymbol{\zeta}) &= \langle \Lambda(\mathbf{x}); \boldsymbol{\zeta} \rangle - V^*(\boldsymbol{\zeta}) + \frac{1}{2} \langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{f} \rangle \\ &= \frac{1}{2} \langle \mathbf{x}, \mathbf{G}(\boldsymbol{\zeta})\mathbf{x} \rangle - V^*(\boldsymbol{\zeta}) - \langle \mathbf{x}, \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle, \end{aligned} \quad (9)$$

in which

$$\mathbf{G}(\boldsymbol{\zeta}) = \mathbf{A} + \sum_{k=1}^m \zeta_k \mathbf{A}_k, \quad (10)$$

and

$$\boldsymbol{\tau}(\boldsymbol{\zeta}) = \mathbf{f} + \sum_{k=1}^m \zeta_k \mathbf{b}_k. \quad (11)$$

For a given $\boldsymbol{\zeta} \in \mathcal{V}_a^*$, the stationary condition $\nabla_{\mathbf{x}} \Xi(\mathbf{x}, \boldsymbol{\zeta}) = 0$ leads to the following *canonical equilibrium equation*

$$\mathbf{G}(\boldsymbol{\zeta})\mathbf{x} = \boldsymbol{\tau}(\boldsymbol{\zeta}). \quad (12)$$

Let

$$\mathcal{S}_a = \{ \boldsymbol{\zeta} \in \mathcal{V}_a^* \mid \exists \mathbf{x} \in \mathcal{X}_a, \text{ such that } \mathbf{G}(\boldsymbol{\zeta})\mathbf{x} = \boldsymbol{\tau}(\boldsymbol{\zeta}) \}$$

be the dual feasible space, in which, the canonical dual function is defined by

$$P^d(\boldsymbol{\zeta}) = \text{sta} \{ \Xi(\mathbf{x}, \boldsymbol{\zeta}) \mid \mathbf{x} \in \mathcal{X}_a \} = -\frac{1}{2} \langle \mathbf{G}^\dagger(\boldsymbol{\zeta})\boldsymbol{\tau}(\boldsymbol{\zeta}), \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle - V^*(\boldsymbol{\zeta}), \quad (13)$$

where $\text{sta} \{ f(\mathbf{x}) \mid \mathbf{x} \in \mathcal{X}_a \}$ stands for finding stationary points of $f(\mathbf{x})$ on \mathcal{X}_a , and \mathbf{G}^\dagger represents the generalized inverse of \mathbf{G} . Particularly, let

$$\mathcal{S}_a^+ = \{ \boldsymbol{\zeta} \in \mathcal{V}_a^* \mid \mathbf{G}(\boldsymbol{\zeta}) \geq 0 \}, \quad (14)$$

where $\mathbf{G}(\boldsymbol{\zeta}) \geq 0$ means that the matrix $\mathbf{G}(\boldsymbol{\zeta})$ is positive semi-definite. Clearly, \mathcal{S}_a^+ is a convex set of \mathcal{S}_a and the total complementary function $\Xi(\mathbf{x}, \boldsymbol{\zeta})$ is convex-concave on $\mathcal{X}_a \times \mathcal{S}_a^+$, by which, the canonical dual problem can be proposed as the following:

$$(\mathcal{P}^d) : \quad \max \{ P^d(\boldsymbol{\zeta}) \mid \boldsymbol{\zeta} \in \mathcal{S}_a^+ \}. \quad (15)$$

The following result is due to the canonical duality theory.

Theorem 1 (Gao [6]). *Problem (\mathcal{P}^d) is canonically dual to (\mathcal{P}) in the sense that if $\bar{\zeta}$ is a stationary solution to (\mathcal{P}^d) , then the vector*

$$\bar{\mathbf{x}} = \mathbf{G}^\dagger(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta}) \tag{16}$$

is a stationary point to (\mathcal{P}) and $P(\bar{\mathbf{x}}) = P^d(\bar{\zeta})$.

Moreover, if $\bar{\zeta} \in \mathcal{S}_a^+$, then $\bar{\mathbf{x}}$ is a global minimizer of (\mathcal{P}) if and only if $\bar{\zeta}$ is a global maximizer of (\mathcal{P}^d) , i.e.,

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_a} P(\mathbf{x}) \Leftrightarrow \max_{\zeta \in \mathcal{S}_a^+} P^d(\zeta) = P^d(\bar{\zeta}). \tag{17}$$

This theorem shows that if the canonical dual problem (\mathcal{P}^d) has a stationary solution on \mathcal{S}_a^+ , then the nonconvex primal problem (\mathcal{P}) is equivalent to a concave maximization dual problem (\mathcal{P}^d) without duality gap. If we further assume that $\mathcal{X}_a = \mathbb{R}^n$ and the optimal solution $\bar{\zeta}$ to Problem (\mathcal{P}^d) is an interior point of \mathcal{S}_a^+ , i.e., $\mathbf{G}(\bar{\zeta}) \succ 0$, then the optimal solution $\bar{\mathbf{x}}$ of Problem (\mathcal{P}) can be obtained uniquely by $\bar{\mathbf{x}} = \mathbf{G}^{-1}(\bar{\zeta})\boldsymbol{\tau}(\bar{\zeta})$ (see [10]).

However, our experiences show that for a class of “difficult” global optimization problems, the canonical dual problem has no stationary solution in \mathcal{S}_a^+ such that $\mathbf{G}(\bar{\zeta}) \succ 0$. In this paper, we propose a computational scheme to solve the case in which the solution is located on the boundary of \mathcal{S}_a^+ . To continue, we need an additional mild assumption:

(A3) There exists an optimal solution $\bar{\mathbf{x}}$ of Problem (\mathcal{P}) such that $\mathbf{G}(\bar{\zeta}) \succeq 0$, where $\bar{\zeta} = \nabla V(\boldsymbol{\xi})|_{\boldsymbol{\xi}=\Lambda(\bar{\mathbf{x}})}$.

In fact, Assumption (A3) is easily satisfied by many real-world problems. To see this, let us first examine the following examples.

Example 1. Suppose that \mathcal{X}_a is a bounded convex polytope subset of \mathbb{R}^n . Since \mathcal{X}_a contains only linear constraints, both \mathcal{V}_a and \mathcal{S}_a are also close and bounded. Let χ be the smallest eigenvalue of $\sum_{k=1}^m \varsigma_k \mathbf{A}_k$, where $\boldsymbol{\zeta} = [\varsigma_1, \dots, \varsigma_m]^T \in \mathcal{S}_a$. Since \mathcal{S}_a is bounded, $\chi > -\infty$. Let $\bar{\chi}$ be the smallest eigenvalue of \mathbf{A} . If $\bar{\chi} + \chi \geq 0$, then Assumption (A3) is satisfied.¹

This example shows that if the quadratic function $\frac{1}{2}(\mathbf{x}, \mathbf{A}\mathbf{x})$ is sufficiently convex, the nonconvexity of $V(\Lambda(\mathbf{x}))$ becomes insignificant. Thus, the combination of them is still convex. However, this is a special case in nonconvex systems. The following example has a wide applications in network optimization.

¹In fact, Problem (\mathcal{P}) is convex under the condition $\bar{\chi} + \chi \geq 0$. The proof of this result is similar to that of Proposition 1 given in [16].

Example 2. Euclidean distance optimization problem:

$$\min \left\{ \sum_{i,j} \left(\|\mathbf{x}_i - \mathbf{x}_j\|^2 - d_{i,j}^2 \right)^2 + \sum_k \left(\|\mathbf{x}_k - \mathbf{a}_k\|^2 - d_k^2 \right)^2 \mid \mathbf{x}_i \in \mathbb{R}^d \forall i = 1, \dots, n \right\}, \quad (18)$$

where \mathbf{x}_i is the location vector in Euclidean space \mathbb{R}^d , d_{ij} and d_k are given distance values, the vectors $\{\mathbf{a}_k\}$ are pre-fixed locations. Problem (18) has many applications, such as wireless sensor network localization and molecular design, etc. For this nonconvex problem, we can choose $\Lambda(\mathbf{x})$ to be the collection of all $\Lambda_{ij}(\mathbf{x}) = \|\mathbf{x}_i - \mathbf{x}_j\|^2$ and $\Lambda_k(\mathbf{x}) = \|\mathbf{x}_k - \mathbf{a}_k\|^2$. In this case, $V(\boldsymbol{\xi}) = \sum_{i,j} (\xi_{ij} - d_{ij}^2)^2 + \sum_k (\xi_k - d_k^2)^2$. If (18) has the optimal function value of 0, then $\xi_{ij} = d_{ij}^2$ and $\xi_k = d_k^2$, where $\boldsymbol{\xi} = \Lambda(\bar{\mathbf{x}})$ and $\bar{\mathbf{x}}$ is an optimal solution of problem (18). It is easy to check that the dual variable $\bar{\boldsymbol{\zeta}} = 0$. Thus, $\det \mathbf{G}(\bar{\boldsymbol{\zeta}}) = 0$. Therefore, Assumption (A3) holds.

This example shows that Assumption (A3) is satisfied in the least squares method for solving a large class of nonlinear systems [31, 32]. It is known that for the conventional SDP relaxation methods, the solution of problem (18) can be exactly recovered if and only if the SDP solution of Problem (18) is a relative interior and the optimal function value of problem (18) is 0 [29]. If the problem (18) has more than one solution, the conventional SDP relaxation does not produce any solution. The goal of this paper is to overcome this difficulty by proposing a canonical primal–dual iterative scheme.

3 Saddle Point Problem

Based on Assumption (A1–A3), the primal problem (\mathcal{P}) is relaxed to the following canonical saddle point problem:

$$(\mathcal{S}_p) : \min_{\mathbf{x}} \max_{\boldsymbol{\zeta}} \left\{ \Xi(\mathbf{x}, \boldsymbol{\zeta}) = \frac{1}{2} \langle \mathbf{x}, \mathbf{G}(\boldsymbol{\zeta}) \mathbf{x} \rangle - V^*(\boldsymbol{\zeta}) - \langle \mathbf{x}, \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle \mid (\mathbf{x}, \boldsymbol{\zeta}) \in \mathcal{X}_a \times \mathcal{S}_a^+ \right\}. \quad (19)$$

Suppose that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p). If $\det(\mathbf{G}(\bar{\boldsymbol{\zeta}})) \neq 0$, we call Problem (\mathcal{S}_p) is non-degenerate. Otherwise, we call it degenerate.

3.1 Non-degenerate Problem (\mathcal{S}_p)

Theorem 2. *Suppose that Problem (\mathcal{S}_p) is non-degenerate. Then, $\bar{\mathbf{x}}$ is a unique solution of Problem (\mathcal{P}) if and only if $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a solution of Problem (\mathcal{S}_p).*

Proof. Suppose that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is the solution of Problem (\mathcal{S}_p) . Since Problem (\mathcal{S}_p) is non-degenerate, $\mathbf{G}(\bar{\boldsymbol{\zeta}}) \succ 0$, i.e., $\bar{\boldsymbol{\zeta}} \in \text{int}\mathcal{S}_a^+$. Thus, $\nabla_{\boldsymbol{\zeta}}\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) = \bar{\boldsymbol{\zeta}} - \Lambda(\bar{\mathbf{x}}) = 0$. For any $\mathbf{x} \in \mathcal{X}_a$, we have

$$\min_{\mathbf{x} \in \mathcal{X}_a} P(\mathbf{x}) = \min_{\mathbf{x} \in \mathcal{X}_a} \max_{\boldsymbol{\zeta} \in \mathcal{Y}_a^*} \Xi(\mathbf{x}, \boldsymbol{\zeta}) = \min_{\mathbf{x} \in \mathcal{X}_a} \max_{\boldsymbol{\zeta} \in \mathcal{S}_a^+} \Xi(\mathbf{x}, \boldsymbol{\zeta}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) = P(\bar{\mathbf{x}}).$$

Thus, $\bar{\mathbf{x}}$ is the optimal solution of Problem (\mathcal{P}) .

On the other hand, we suppose that $\bar{\mathbf{x}}$ is the optimal solution of Problem (\mathcal{P}) . Let $\bar{\boldsymbol{\zeta}} = \nabla V(\Lambda(\bar{\mathbf{x}}))$. Then,

$$P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) = \max_{\boldsymbol{\zeta} \in \mathbb{R}^m} \Xi(\bar{\mathbf{x}}, \boldsymbol{\zeta}).$$

Since $V(\cdot)$ is strictly convex, we have

$$\Xi(\bar{\mathbf{x}}, \boldsymbol{\zeta}) \leq \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) \quad \forall \boldsymbol{\zeta} \in \mathcal{Y}_a^* \subset \mathbb{R}^m. \quad (20)$$

The equality holds in (20) if and only if $\boldsymbol{\zeta} = \bar{\boldsymbol{\zeta}}$ since $\Xi(\bar{\mathbf{x}}, \boldsymbol{\zeta})$ is strictly concave in terms of $\boldsymbol{\zeta}$. Suppose that $(\mathbf{x}_1, \boldsymbol{\zeta}_1)$ is also a saddle point of Problem (\mathcal{S}_p) . By a similar induction as above, we can show that \mathbf{x}_1 is an optimal solution of Problem (\mathcal{P}) . Furthermore, $P(\mathbf{x}_1) = \Xi(\mathbf{x}_1, \boldsymbol{\zeta}_1)$. Since $\mathbf{x}_1 \in \mathcal{X}_a$, we have

$$P(\mathbf{x}_1) = \Xi(\mathbf{x}_1, \boldsymbol{\zeta}_1) \leq \Xi(\bar{\mathbf{x}}, \boldsymbol{\zeta}_1) \leq \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) = P(\bar{\mathbf{x}}).$$

The first equality holds only when $\mathbf{x}_1 = \bar{\mathbf{x}}$ since $\mathbf{G}(\boldsymbol{\zeta}_1) \succ 0$. The second inequality becomes equality if and only if $\boldsymbol{\zeta}_1 = \bar{\boldsymbol{\zeta}}$ since $V(\cdot)$ is strictly convex. By the fact that $\bar{\mathbf{x}}$ is an optimal solution of Problem (\mathcal{P}) and $\mathbf{x}_1 \in \mathcal{X}_a$, $P(\mathbf{x}_1) = P(\bar{\mathbf{x}})$, $\mathbf{x}_1 = \bar{\mathbf{x}}$ and $\boldsymbol{\zeta}_1 = \bar{\boldsymbol{\zeta}}$. Thus, $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is the solution of Problem (\mathcal{S}_p) . We complete the proof. ■

If $\mathcal{X}_a = \mathbb{R}^n$, the saddle point Problem (\mathcal{S}_p) can be further recast as a convex semi-definite programming problem.

Proposition 1. *Suppose that Problem (\mathcal{S}_p) is non-degenerate and $\mathcal{X}_a = \mathbb{R}^n$. Let $\bar{\boldsymbol{\zeta}}$ be the solution of the following convex SDP problem:*

$$(SDP) : \quad \min \{V^*(\boldsymbol{\zeta}) + g\} \quad \text{s.t.} \quad \begin{bmatrix} \mathbf{G}(\boldsymbol{\zeta}) & \boldsymbol{\tau}(\boldsymbol{\zeta}) \\ \boldsymbol{\tau}^T(\boldsymbol{\zeta}) & 2g \end{bmatrix} \succeq 0. \quad (21)$$

Then, the SDP problem defined by (21) has a unique solution $(\bar{g}, \bar{\boldsymbol{\zeta}})$ such that $\mathbf{G}(\bar{\boldsymbol{\zeta}}) \succ 0$. Furthermore, $\bar{\mathbf{x}} = \mathbf{G}^{-1}(\bar{\boldsymbol{\zeta}})\boldsymbol{\tau}(\bar{\boldsymbol{\zeta}})$ is the unique solution of Problem (\mathcal{P}) .

Proof. By Schur complement lemma [15], the SDP problem (21) has a unique solution $(\bar{g}, \bar{\boldsymbol{\zeta}})$ such that $\mathbf{G}(\bar{\boldsymbol{\zeta}}) \succ 0$ if and only if the following convex minimization problem

$$\min \left\{ V^*(\boldsymbol{\zeta}) + \frac{1}{2} \langle \mathbf{G}^{-1}(\boldsymbol{\zeta})\boldsymbol{\tau}(\boldsymbol{\zeta}), \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle \mid \mathbf{G}(\boldsymbol{\zeta}) \succeq 0 \right\} \quad (22)$$

has a unique solution $\bar{\boldsymbol{\zeta}}$ such that $\mathbf{G}(\bar{\boldsymbol{\zeta}}) \succ 0$. Since $\mathcal{X}_a = \mathbb{R}^n$, the convex minimization problem (22) is equivalent to Problem (\mathcal{S}_p) by Theorem 3.1 in [10]. ■

Remark 1. Theorem 2 is actually a special case of the general result obtained by Gao and Strang in finite deformation theory [7]. Indeed, if we let $\bar{W}(\mathbf{x}) = W(\mathbf{x}) + \frac{1}{2}\langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle$ and $\Lambda(\mathbf{x}) = \{\Lambda(\mathbf{x}), \frac{1}{2}\langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle\}$, then, the Gao–Strang complementary gap function is simply defined as

$$G(\mathbf{x}, \boldsymbol{\zeta}) = \frac{1}{2}\langle \mathbf{x}, \mathbf{G}(\boldsymbol{\zeta})\mathbf{x} \rangle.$$

Clearly, this gap function is strictly positive for any nonzero $\mathbf{x} \in \mathcal{X}_a$ if and only if $\mathbf{G}(\boldsymbol{\zeta}) \succ 0$. Then by Theorem 2 in [7] we know that the primal problem has a unique solution if the problem (\mathcal{S}_p) is non-degenerate. By Theorem 2 and Proposition 1 we know that the nonconvex problem (\mathcal{P}) can be solved easily either by solving a sequence of strict convex–concave saddle point problems, or via solving a convex semi-definite programming problem if Problem (\mathcal{S}_p) is non-degenerate. By the fact that $g = \frac{1}{2}\langle \mathbf{G}^{-1}(\boldsymbol{\zeta})\boldsymbol{\tau}(\boldsymbol{\zeta}), \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle$ is actually the pure complementary gap function (see Eq. (19) in [6]), the convex SDP problem (21) is indeed a special case of the canonical dual problem (\mathcal{P}^d) defined by (15). Moreover, the canonical duality theory can also be used to find the biggest local extrema of the nonconvex problem (\mathcal{P}) (see [10]).

3.2 Degenerate Problem (\mathcal{S}_p) and Linear Perturbation

If Problem (\mathcal{S}_p) is degenerate, i.e., $\mathbf{G}(\bar{\boldsymbol{\zeta}}) \succeq 0$ and $\det(\mathbf{G}(\bar{\boldsymbol{\zeta}})) = 0$ or $\bar{\boldsymbol{\zeta}} \in \partial\mathcal{S}_a^+$, it has multiple saddle points. The following theorem reveals the relations between Problem (\mathcal{P}) and Problem (\mathcal{S}_p) .

Theorem 3. *Suppose that Problem (\mathcal{S}_p) is degenerate.*

- 1) *If $\bar{\mathbf{x}}$ is a solution of Problem (\mathcal{P}) and $\bar{\boldsymbol{\zeta}} = \nabla V(\Lambda(\bar{\mathbf{x}}))$, then $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) .*
- 2) *If $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) , then $\bar{\mathbf{x}}$ is a solution of Problem (\mathcal{P}) .*
- 3) *If $(\mathbf{x}_1, \boldsymbol{\zeta}_1)$ and $(\mathbf{x}_2, \boldsymbol{\zeta}_2)$ are two saddle points of Problem (\mathcal{S}_p) , then $\boldsymbol{\zeta}_1 = \boldsymbol{\zeta}_2$.*

Proof. 1). Since $\bar{\mathbf{x}}$ is a solution of Problem (\mathcal{P}) and $\bar{\boldsymbol{\zeta}} = \nabla V(\Lambda(\bar{\mathbf{x}})) \in \mathcal{S}_a^+$ (by Assumption (A3)),

$$\Xi(\bar{\mathbf{x}}, \boldsymbol{\zeta}) \leq \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}), \quad \forall \boldsymbol{\zeta} \in \mathcal{S}_a^+.$$

Furthermore,

$$\langle \nabla P(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}_a. \tag{23}$$

Substituting $\nabla P(\bar{\mathbf{x}}) = \mathbf{G}(\bar{\boldsymbol{\zeta}})\bar{\mathbf{x}} - \boldsymbol{\tau}(\bar{\boldsymbol{\zeta}}) = \nabla_{\mathbf{x}}\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ into (23), we obtain

$$\langle \nabla_{\mathbf{x}}\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}), \mathbf{x} - \bar{\mathbf{x}} \rangle \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}_a.$$

Thus,

$$\min_{\mathbf{x} \in \mathcal{X}_a} \Xi(\mathbf{x}, \bar{\boldsymbol{\zeta}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}).$$

Therefore,

$$\Xi(\bar{\mathbf{x}}, \boldsymbol{\zeta}) \leq \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) \leq \Xi(\mathbf{x}, \bar{\boldsymbol{\zeta}}), \quad \forall (\mathbf{x}, \boldsymbol{\zeta}) \in \mathcal{X}_a \times \mathcal{S}_a^+.$$

This implies that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) .

2). Suppose that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) and $\nabla_{\boldsymbol{\zeta}} \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) = 0$. Then,

$$P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) \leq \Xi(\mathbf{x}, \bar{\boldsymbol{\zeta}}), \quad \forall (\mathbf{x}, \boldsymbol{\zeta}) \in \mathcal{X}_a \times \mathcal{S}_a^+.$$

On the other hand,

$$\Xi(\mathbf{x}, \bar{\boldsymbol{\zeta}}) = \langle \Lambda(\mathbf{x}); \bar{\boldsymbol{\zeta}} \rangle - V^*(\bar{\boldsymbol{\zeta}}) - U(\mathbf{x}) \leq V(\Lambda(\mathbf{x})) - U(\mathbf{x}) = P(\mathbf{x}).$$

Combining the above two inequalities, $\bar{\mathbf{x}}$ is a solution Problem (\mathcal{P}) .

3). This result follows directly from the strict convexity of both $V(\cdot)$ and $V^*(\cdot)$. The proof is completed. \blacksquare

Theorem 3 shows that the nonconvex minimization Problem (\mathcal{P}) is equivalent to the canonical saddle min-max Problem (\mathcal{S}_p) . What we should emphasize is that the solutions set of Problem (\mathcal{P}) is in general nonconvex, while the set of saddle points of Problem (\mathcal{S}_p) is convex. For example, let us consider the following optimization problem:

$$\min \left\{ \frac{1}{2} ((x_1 + x_2)^2 - 1)^2 + \frac{1}{2} ((x_1 - x_2)^2 - 1)^2 \mid (x_1, x_2) \in \mathbb{R}^2 \right\}. \quad (24)$$

Let $\xi = \Lambda(\mathbf{x}) = [(x_1 + x_2)^2 - 1, (x_1 - x_2)^2 - 1]^T$. Then,

$$\mathbf{G}(\boldsymbol{\zeta}) = \begin{bmatrix} \zeta_1 + \zeta_2 & \zeta_1 - \zeta_2 \\ \zeta_1 - \zeta_2 & \zeta_1 + \zeta_2 \end{bmatrix},$$

$V^*(\boldsymbol{\zeta}) = \frac{1}{2} \boldsymbol{\zeta}^T \boldsymbol{\zeta}$. Thus, $\mathbf{G}(\boldsymbol{\zeta}) \geq 0 \Leftrightarrow \zeta_1 \geq 0$ and $\zeta_2 \geq 0$. Clearly, $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) if and only if $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is the solution of the following variational inequality:

$$\mathbf{G}(\bar{\boldsymbol{\zeta}}) \bar{\mathbf{x}} = 0, \quad (25)$$

$$\langle \nabla V^*(\bar{\boldsymbol{\zeta}}) - \Lambda(\bar{\mathbf{x}}); \boldsymbol{\zeta} - \bar{\boldsymbol{\zeta}} \rangle \geq 0, \quad \forall \boldsymbol{\zeta} \geq 0. \quad (26)$$

It is easy to verify that the optimization problem (24) has four solutions $(1, 0)$, $(0, 1)$, $(-1, 0)$ and $(0, -1)$. Clearly, its solution set is nonconvex. On the other hand, by the statement 3) in Theorem 3, we have $\bar{\boldsymbol{\zeta}} = 0$. Thus, $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) if and only if $\bar{\boldsymbol{\zeta}} = 0$ and $\bar{\mathbf{x}}$ satisfies

$$\begin{aligned} (x_1 + x_2)^2 &\leq 1, \\ (x_1 - x_2)^2 &\leq 1. \end{aligned}$$

Denote $\Omega = \text{convhull}\{(1, 0), (0, 1), (-1, 0), (0, -1)\}$, where convhull means convex hull. Therefore, the saddle point set of Problem (\mathcal{S}_p) is $\Omega \times 0$ which is a convex set. This example also shows that the solutions of Problem (\mathcal{P}) are the vertex points of the saddle points set of Problem (\mathcal{S}_p) .

Now we turn our attention to the saddle point problem (\mathcal{S}_p) . For some simple optimization problems, we can simply use linear perturbation method to solve it. To illustrate it, let us consider a simple optimization problem given as below:

$$(\mathcal{P}_1) : \min_{\mathbf{x}} P_1(\mathbf{x}) = \frac{1}{2} \left(\frac{1}{2} \mathbf{x}^T \mathbf{A}_1 \mathbf{x} - b_1 \right)^2 + \frac{1}{2} \left(\frac{1}{2} \mathbf{x}^T \mathbf{A}_2 \mathbf{x} - b_2 \right)^2 - \langle \mathbf{x}, \mathbf{f} \rangle. \tag{27}$$

Proposition 2. *Suppose that there exists $(\varsigma_1, \varsigma_2)$ such that $\varsigma_1 \mathbf{A}_1 + \varsigma_2 \mathbf{A}_2 \succ 0$. If the saddle point $(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}})$ of the associated Problem (\mathcal{S}_{p1}) is on the boundary of \mathcal{S}_a^+ , then for any given $\varepsilon > 0$, there exists a $\Delta \mathbf{f} \in \mathbb{R}^n$ such that $\|\Delta \mathbf{f}\| \leq \varepsilon$ and the perturbed saddle point Problem (\mathcal{S}_{p1})*

$$(\mathcal{S}_{p1}) : \min_{\mathbf{x}} \max_{\boldsymbol{\varsigma}} \left\{ \frac{1}{2} \langle \mathbf{x}, \mathbf{G}(\boldsymbol{\varsigma}) \mathbf{x} \rangle - \frac{1}{2} \boldsymbol{\varsigma}^T \boldsymbol{\varsigma} - \langle \mathbf{x}, \mathbf{f} + \Delta \mathbf{f} \rangle : (\mathbf{x}, \boldsymbol{\varsigma}) \in \mathbb{R}^n \times \mathcal{S}_a^+ \right\}$$

has a unique saddle point $(\bar{\mathbf{x}}_p, \bar{\boldsymbol{\varsigma}}_p)$ such that $\mathbf{G}(\bar{\boldsymbol{\varsigma}}_p) \succ 0$. Furthermore, $\bar{\mathbf{x}}_p$ is the unique solution of

$$(\mathcal{P}_1^{ptb}) : \min_{\mathbf{x}} P_1(\mathbf{x}) = \frac{1}{2} \left(\frac{1}{2} \mathbf{x}^T \mathbf{A}_1 \mathbf{x} - b_1 \right)^2 + \frac{1}{2} \left(\frac{1}{2} \mathbf{x}^T \mathbf{A}_2 \mathbf{x} - b_2 \right)^2 - \langle \mathbf{x}, \mathbf{f} + \Delta \mathbf{f} \rangle,$$

where $\mathbf{G}(\boldsymbol{\varsigma}) = \varsigma_1 \mathbf{A}_1 + \varsigma_2 \mathbf{A}_2$.

Proof. Since $\mathbf{x} \in \mathbb{R}^n$, Problem (\mathcal{S}_{p1}) is equivalent to the following optimization problem:

$$\begin{aligned} \max_{\boldsymbol{\varsigma}} \quad & -V^*(\boldsymbol{\varsigma}) - \frac{1}{2} (\mathbf{f} + \Delta \mathbf{f})^T (\boldsymbol{\varsigma}) \mathbf{G}^{-1}(\boldsymbol{\varsigma}) (\mathbf{f} + \Delta \mathbf{f}) \\ \text{s.t.} \quad & \mathbf{G}(\boldsymbol{\varsigma}) \succeq 0. \end{aligned} \tag{28}$$

By the assumption that there exists $(\varsigma_1, \varsigma_2)$ such that $\varsigma_1 \mathbf{A}_1 + \varsigma_2 \mathbf{A}_2 \succ 0$, \mathbf{A}_1 and \mathbf{A}_2 are simultaneously diagonalizable via congruence. More specifically, there exists an invertible matrix \mathbf{C} such that

$$\begin{aligned} \mathbf{C}^T \mathbf{A}_1 \mathbf{C} &= \text{diag}(a_1^1, \dots, a_n^1), \\ \mathbf{C}^T \mathbf{A}_2 \mathbf{C} &= \text{diag}(a_1^2, \dots, a_n^2). \end{aligned}$$

Under this condition, it is easy to show that for any given $\varepsilon > 0$, there exists a $\Delta \mathbf{f} \in \mathbb{R}^n$ such that $\|\Delta \mathbf{f}\| \leq \varepsilon$ and

$$\lim_{\boldsymbol{\zeta} \rightarrow \partial \mathcal{S}_a^+} \frac{1}{2} (\mathbf{f} + \Delta \mathbf{f})^T (\boldsymbol{\zeta}) \mathbf{G}^{-1} (\boldsymbol{\zeta}) (\mathbf{f} + \Delta \mathbf{f}) = +\infty.$$

Thus, the solution of the optimization problem (28) cannot be located in the boundary of \mathcal{S}_a^+ for this $\Delta \mathbf{f}$. The results follow readily. We complete the proof. ■

From Proposition 1 we know that if the solution $\bar{\mathbf{x}}$ of Problem (\mathcal{P}_1) satisfies $\mathbf{G}(\bar{\boldsymbol{\zeta}}) \succ 0$, then it can be obtained by simply solving the concave maximization dual problem (\mathcal{P}^d). Otherwise, Proposition 2 shows that this solution can be obtained under a small perturbation. Thus, the nonconvex optimization problem (\mathcal{P}_1) can be completely solved by either the convex SDP or the canonical duality. However, for general optimization problems, the linear perturbation method may not produce an interior saddle point of Problem (\mathcal{S}_p). To overcome this difficulty, we shall introduce a nonlinear perturbation method in the next section.

4 Quadratic Perturbation Method

We now focus on solving the degenerated Problem (\mathcal{S}_p). Clearly, Problem (\mathcal{S}_p) is strictly concave with respect to $\boldsymbol{\zeta}$. However, if Problem (\mathcal{S}_p) is degenerate, i.e., $\bar{\boldsymbol{\zeta}} \in \partial \mathcal{S}_a^+$, then Problem (\mathcal{S}_p) is convex but not strictly in terms of \mathbf{x} . In this case, Problem (\mathcal{S}_p) has multiple solutions. To stabilize such kind of optimization problems, nonlinear perturbation methods can be used (see [9]). Thus, using the quadratic perturbation method to Problem (\mathcal{S}_p), a regularized saddle point problem can be proposed as

$$\min_{\mathbf{x}} \max_{\boldsymbol{\zeta} \in \mathcal{S}_a^+} \Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta}) = \Xi(\mathbf{x}, \boldsymbol{\zeta}) + \frac{\rho_k}{2} \|\mathbf{x} - \mathbf{x}_k\|^2, \tag{29}$$

where both \mathbf{x}_k and $\rho_k, k = 1, 2, \dots$, are given. In practical computation, the canonical dual feasible space \mathcal{S}_a^+ can also be relaxed as

$$\mathcal{S}_{\mu_k}^+ = \{\boldsymbol{\zeta} \in \mathcal{V}_a^* \subset \mathbb{R}^m \mid \mathbf{G}(\boldsymbol{\zeta}) + \mu_k \mathbf{I} \succeq 0\},$$

where $\mu_k < \rho_k$. Note that

$$\Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta}) = \frac{1}{2} \langle \mathbf{x}, (\mathbf{G}(\boldsymbol{\zeta}) + \rho_k \mathbf{I}) \mathbf{x} \rangle - V^*(\boldsymbol{\zeta}) - \langle \mathbf{x}, \rho_k \mathbf{x}_k + \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle + \frac{\rho_k}{2} \langle \mathbf{x}_k, \mathbf{x}_k \rangle.$$

Thus, $\Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta})$ is strictly convex–concave in $\mathbb{R}^n \times \mathcal{S}_{\mu_k}^+$ and

$$\min_{\mathbf{x}} \max_{\boldsymbol{\zeta} \in \mathcal{S}_{\mu_k}^+} \Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta}) = \max_{\boldsymbol{\zeta} \in \mathcal{S}_{\mu_k}^+} \min_{\mathbf{x}} \Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta}).$$

For each given $\boldsymbol{\zeta} \in \mathcal{S}_{\mu_k}^+$, denote

$$\mathbf{x}(\boldsymbol{\zeta}) = \arg \min_{\mathbf{x}} \Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta}).$$

Then, $\mathbf{x}(\boldsymbol{\zeta}) = (\mathbf{G}(\boldsymbol{\zeta}) + \rho_k I)^{-1}(\rho_k \mathbf{x}_k + \boldsymbol{\tau}(\boldsymbol{\zeta}))$. Substituting this $\mathbf{x}(\boldsymbol{\zeta})$ into $\Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta})$, we obtain the perturbed canonical dual function

$$P_{\rho_k}^d(\boldsymbol{\zeta}) = -\frac{1}{2} \langle (\mathbf{G}(\boldsymbol{\zeta}) + \rho_k I)^{-1}(\rho_k \mathbf{x}_k + \boldsymbol{\tau}(\boldsymbol{\zeta})), \rho_k \mathbf{x}_k + \boldsymbol{\tau}(\boldsymbol{\zeta}) \rangle - V^*(\boldsymbol{\zeta}) + \frac{\rho_k}{2} \langle \mathbf{x}_k, \mathbf{x}_k \rangle.$$

Now our canonical primal–dual algorithm can be proposed as follows.

Algorithm 1

Step 1 Initialization \mathbf{x}_0, ρ_0, N and the error tolerance ε . Set $k = 0$.

Step 2 Set $\boldsymbol{\zeta}_{k+1} = \arg \max_{\boldsymbol{\zeta} \in \mathcal{S}_{\mu_k}^+} P_{\rho_k}^d(\boldsymbol{\zeta})$ and $\mathbf{x}_{k+1} = (\mathbf{G}(\boldsymbol{\zeta}_{k+1}) + \rho_k I)^{-1}(\rho_k \mathbf{x}_k + \boldsymbol{\tau}(\boldsymbol{\zeta}_{k+1}))$.

Step 3 If $\|\boldsymbol{\zeta}_{k+1} - \boldsymbol{\zeta}_k\| \leq \varepsilon$, stop. Otherwise, set $k = k + 1$ and go to Step 2.

Theorem 4. Suppose that

- 1) $\bar{\rho} \geq \rho_k > 0, \sigma_k = \sum_{i=1}^k \rho_i \rightarrow +\infty, \rho_k \downarrow 0, \mu_k \downarrow 0$ and $0 < \mu_k < \rho_k$;
- 2) For any given $\mathbf{x}, \lim_{\|\boldsymbol{\zeta}_k\| \rightarrow \infty} \Xi(\mathbf{x}, \boldsymbol{\zeta}_k) = -\infty$;
- 3) The sequence $\{\mathbf{x}_k\}$ is a bounded;

Then, there exists a $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) \in \mathbb{R}^n \times \mathcal{S}_a^+$ such that $\{\mathbf{x}_k, \boldsymbol{\zeta}_k\} \rightarrow (\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$. Furthermore, $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) .

Proof. Note that $0 < \mu_k < \rho_k$, the perturbed total complementary function $\Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta})$ is strictly convex–concave with respect to $(\mathbf{x}, \boldsymbol{\zeta})$ in $\mathbb{R}^n \times \mathcal{S}_{\mu_k}^+$. Since $(\mathbf{x}_k, \boldsymbol{\zeta}_k)$ is generated by Algorithm 1, we have

$$(\mathbf{x}_k, \boldsymbol{\zeta}_k) = \arg \min_{\mathbf{x}} \max_{\boldsymbol{\zeta} \in \mathcal{S}_{\mu_k}^+} \Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta}) = \arg \min_{\mathbf{x}} \max_{\boldsymbol{\zeta} \in \mathcal{S}_{\mu_k}^+} \left\{ \Xi(\mathbf{x}, \boldsymbol{\zeta}) + \frac{\rho_{k-1}}{2} \|\mathbf{x} - \mathbf{x}_{k-1}\|^2 \right\}. \quad (30)$$

That is

$$\Xi_{\rho_k}(\mathbf{x}_k, \boldsymbol{\zeta}) \leq \Xi_{\rho_k}(\mathbf{x}_k, \boldsymbol{\zeta}_k) \leq \Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\zeta}_k), \quad \forall (\mathbf{x}, \boldsymbol{\zeta}) \in \mathbb{R}^n \times \mathcal{S}_{\mu_k}^+.$$

By the fact that $\mu_k \downarrow 0$ and $\mathcal{S}_{\mu_k}^+ = \{\boldsymbol{\zeta} \in \mathcal{V}_a^* \mid \mathbf{G}(\boldsymbol{\zeta}) + \mu_k I \geq 0\}$, we have $\mathcal{S}_{\mu_k}^+ \supseteq \mathcal{S}_{\mu_{k+1}}^+$ and $\bigcap_k \mathcal{S}_{\mu_k}^+ = \mathcal{S}_a^+$.

To continue, we suppose that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ is a saddle point of Problem (\mathcal{S}_p) , i.e.,

$$\Xi(\bar{\mathbf{x}}, \boldsymbol{\zeta}) \leq \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}}) \leq \Xi(\mathbf{x}, \bar{\boldsymbol{\zeta}}), \quad \forall (\mathbf{x}, \boldsymbol{\zeta}) \in \mathbb{R}^n \times \mathcal{S}_a^+.$$

Now we adopt the following steps to prove our results.

1) The sequence $\{\mathbf{x}_k\}$ is convergent, i.e., there exists a $\bar{\mathbf{x}}$ such that $\mathbf{x}_k \rightarrow \bar{\mathbf{x}}$. From (30), we have

$$\Xi_{\rho_{k-1}}(\mathbf{x}_k, \boldsymbol{\varsigma}_k) = \Xi(\mathbf{x}_k, \boldsymbol{\varsigma}_k) + \frac{\rho_{k-1}}{2} \|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2 \leq \Xi_{\rho_{k-1}}(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_k) = \Xi(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_k). \quad (31)$$

Clearly,

$$\Xi(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_k) + \frac{\rho_{k-2}}{2} \|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\|^2 = \Xi_{\rho_{k-2}}(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_k). \quad (32)$$

Since $\boldsymbol{\varsigma}_k \in \mathcal{S}_{\mu_k}^+ \subset \mathcal{S}_{\mu_{k-1}}^+$ and $(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_{k-1})$ is the saddle point of $\Xi_{\rho_{k-1}}(\mathbf{x}, \boldsymbol{\varsigma})$ in $\mathbb{R}^n \times \mathcal{S}_{\mu_{k-1}}^+$, we obtain

$$\Xi_{\rho_{k-2}}(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_k) \leq \Xi_{\rho_{k-2}}(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_{k-1}) = \Xi(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_{k-1}) + \frac{\rho_{k-2}}{2} \|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\|^2. \quad (33)$$

Combining (32) and (33), we obtain

$$\Xi(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_k) \leq \Xi(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_{k-1}).$$

Thus,

$$\Xi(\mathbf{x}_k, \boldsymbol{\varsigma}_k) + \frac{\rho_{k-1}}{2} \|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2 \leq \Xi(\mathbf{x}_{k-1}, \boldsymbol{\varsigma}_{k-1}). \quad (34)$$

Repeating the above process, we get

$$\Xi(\mathbf{x}_k, \boldsymbol{\varsigma}_k) + \sum_{i=1}^{k-1} \frac{\rho_{i-1}}{2} \|\mathbf{x}_i - \mathbf{x}_{i-1}\|^2 \leq \Xi(\mathbf{x}_1, \boldsymbol{\varsigma}_1). \quad (35)$$

On the other hand,

$$\begin{aligned} \Xi_{\rho_{k-1}}(\mathbf{x}_k, \boldsymbol{\varsigma}_k) &= \Xi(\mathbf{x}_k, \boldsymbol{\varsigma}_k) + \frac{\rho_{k-1}}{2} \|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2 \\ &\geq \Xi_{\rho_{k-1}}(\mathbf{x}_k, \bar{\boldsymbol{\varsigma}}) = \Xi(\mathbf{x}_k, \bar{\boldsymbol{\varsigma}}) + \frac{\rho_{k-1}}{2} \|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2 \\ &\geq \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}) + \frac{\rho_{k-1}}{2} \|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2. \end{aligned} \quad (36)$$

Substituting (36) into (35) gives rise to

$$\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}) + \sum_{i=1}^{k-2} \frac{\rho_{i-1}}{2} \|\mathbf{x}_i - \mathbf{x}_{i-1}\|^2 \leq \Xi(\mathbf{x}_1, \boldsymbol{\varsigma}_1), \quad \forall k \in \mathbb{N}.$$

Since $\{\mathbf{x}_k\}$ is a bounded sequence, $\sigma_k \rightarrow +\infty$ and $\rho_k \downarrow 0$, the sequence \mathbf{x}_k is convergent, i.e., there exists a $\bar{\mathbf{x}}$ such that $\mathbf{x}_k \rightarrow \bar{\mathbf{x}}$.

- 2) The sequence $\{\varsigma_k\}$ is convergent. We first show that ς_k is a bounded sequence. In a similar argument to the inequality (34), we can show that

$$\Xi(\mathbf{x}_{k+1}, \varsigma_{k+1}) \geq \Xi(\mathbf{x}_{k+1}, \bar{\varsigma}) \geq \Xi(\bar{\mathbf{x}}, \bar{\varsigma}).$$

On the other hand,

$$\begin{aligned} \Xi_{\rho_k}(\mathbf{x}_{k+1}, \varsigma_{k+1}) &= \Xi(\mathbf{x}_{k+1}, \varsigma_{k+1}) + \frac{\rho_k}{2} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|^2 \\ &\leq \Xi_{\rho_k}(\bar{\mathbf{x}}, \varsigma_{k+1}) = \Xi(\bar{\mathbf{x}}, \varsigma_{k+1}) + \frac{\rho_k}{2} \|\bar{\mathbf{x}} - \mathbf{x}_k\|^2. \end{aligned}$$

Summing the above inequalities together yields that

$$\begin{aligned} \Xi(\bar{\mathbf{x}}, \bar{\varsigma}) - \frac{\bar{\rho}}{2} \|\bar{\mathbf{x}} - \mathbf{x}_k\|^2 &\leq \Xi(\bar{\mathbf{x}}, \bar{\varsigma}) - \frac{\rho_k}{2} \|\bar{\mathbf{x}} - \mathbf{x}_k\|^2 \\ &\leq \Xi(\mathbf{x}_{k+1}, \bar{\varsigma}) - \frac{\rho_k}{2} \|\bar{\mathbf{x}} - \mathbf{x}_k\|^2 \leq \Xi(\bar{\mathbf{x}}, \varsigma_{k+1}). \end{aligned}$$

By Assumption (2) and $\mathbf{x}^k \rightarrow \bar{\mathbf{x}}$, we know that ς_k is a bounded sequence.

Now we suppose that there are two subsequences $\{\varsigma_k^1\}$ and $\{\varsigma_k^2\}$ of $\{\varsigma_k\}$ such that $\{\varsigma_k^1\} \rightarrow \varsigma^1$ and $\{\varsigma_k^2\} \rightarrow \varsigma^2$. Denote $\{\mathbf{x}_k^1\}$ and $\{\mathbf{x}_k^2\}$ are two subsequences of $\{\mathbf{x}_k\}$ associated with $\{\varsigma_k^1\}$ and $\{\varsigma_k^2\}$. Clearly, $\varsigma^1, \varsigma^2 \in \mathcal{S}_a^+$. Note that

$$\begin{aligned} \Xi(\mathbf{x}_{k+1}^1, \varsigma^2) + \frac{\rho_k^1}{2} \|\mathbf{x}_{k+1}^1 - \mathbf{x}_k^1\|^2 &= \Xi_{\rho_k^1}(\mathbf{x}_{k+1}^1, \varsigma^2) \\ &\leq \Xi_{\rho_k^1}(\mathbf{x}_{k+1}^1, \varsigma_{k+1}^1) = \Xi(\mathbf{x}_{k+1}^1, \varsigma_{k+1}^1) + \frac{\rho_k^1}{2} \|\mathbf{x}_{k+1}^1 - \mathbf{x}_k^1\|^2. \end{aligned} \quad (37)$$

Thus,

$$\Xi(\mathbf{x}_{k+1}^1, \varsigma^2) \leq \Xi(\mathbf{x}_{k+1}^1, \varsigma_{k+1}^1).$$

Taking limit on both sides of the above inequality yields to

$$\Xi(\bar{\mathbf{x}}, \varsigma^2) \leq \Xi(\bar{\mathbf{x}}, \varsigma^1).$$

In a similar way, we can show that

$$\Xi(\bar{\mathbf{x}}, \varsigma^1) \leq \Xi(\bar{\mathbf{x}}, \varsigma^2).$$

Therefore,

$$\Xi(\bar{\mathbf{x}}, \varsigma^1) = \Xi(\bar{\mathbf{x}}, \varsigma^2)$$

which implies that $\varsigma^1 = \varsigma^2$. Hence, $\{\varsigma_k\}$ is a convergent sequence.

- 3) We show that if $\{\mathbf{x}_k, \boldsymbol{\varsigma}_k\} \rightarrow (\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}})$, then $(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}})$ is a saddle point of Problem (\mathcal{S}_p) . In a similar argument to 2), it is easy to show that for any $\boldsymbol{\varsigma} \in \mathcal{S}_a^+$, we have

$$\Xi(\bar{\mathbf{x}}, \boldsymbol{\varsigma}) \leq \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}).$$

So we only need to show that for any \mathbf{x} ,

$$\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}) \leq \Xi(\mathbf{x}, \bar{\boldsymbol{\varsigma}}). \tag{38}$$

Indeed, by the fact that

$$\Xi_{\rho_k}(\mathbf{x}_{k+1}, \boldsymbol{\varsigma}_{k+1}) \leq \Xi_{\rho_k}(\mathbf{x}, \boldsymbol{\varsigma}_{k+1}), \quad \forall \mathbf{x}.$$

Passing limit to the above inequality yields to the inequality (38). We complete the proof. ■

In Theorem 4, there are three assumptions. Assumption (1) is on the selection of the parameters and Assumption (2) is always satisfied for strictly convex functions. Assumption (3) is important to ensure the convergence of Algorithm 1. In fact, from our numerical experiments, we found that \mathbf{x}_k might become unbound for certain cases. Therefore, a modified algorithm for solving Problem (\mathcal{P}) is suggested as the following.

Algorithm 2

- Step 1* Adopt Algorithm 1 to solve Problem (\mathcal{S}_p) . Denote the obtained solution as $(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}})$.
- Step 2* If $\|\Lambda(\bar{\mathbf{x}}) - \nabla V^*(\bar{\boldsymbol{\varsigma}})\| \leq \varepsilon$, output $\bar{\mathbf{x}}$ is a global minimizer of Problem (\mathcal{P}) , where ε is the tolerance. Otherwise, a gradient-based optimization method is used to refine Problem (\mathcal{P}) with initial condition $\bar{\mathbf{x}}$.

Remark 2. Since Problem (\mathcal{S}_p) is a convex–concave saddle point problem, many exact and inexact proximal point methods can be adapted [12, 14, 30]. In fact, solving Problem (\mathcal{S}_p) is an easy task since it is essentially a convex optimization problem. However, to obtain a solution of Problem (\mathcal{P}) from the solution set of Problem (\mathcal{S}_p) is a difficult task since the identification of degenerate indices in the nonlinear complementarity problem is hard [37]. Unlike the classical proximal point methods, our proposed Algorithm 1 is based on a sequence of exterior point approximation. In this case, the gradient operator $[\nabla_{\mathbf{x}} \Xi(\mathbf{x}, \boldsymbol{\varsigma}), -\nabla_{\boldsymbol{\varsigma}} \Xi(\mathbf{x}, \boldsymbol{\varsigma})]$ in $\mathbb{R}^n \times \mathcal{S}_a^+$ is not a monotone operator, but $[\nabla_{\mathbf{x}} \Xi(\mathbf{x}, \boldsymbol{\varsigma}) + \mu_k I, -\nabla_{\boldsymbol{\varsigma}} \Xi(\mathbf{x}, \boldsymbol{\varsigma})]$ is monotone in $\mathbb{R}^n \times \mathcal{S}_a^+$. By the fact that $\bigcap_k \mathcal{S}_{\mu_k}^+ = \mathcal{S}_a^+$, our algorithm generates a convergent sequence and its clustering point is a saddle point of Problem (\mathcal{S}_p) under certain conditions. Since $[\nabla_{\mathbf{x}} \Xi(\mathbf{x}, \boldsymbol{\varsigma}), -\nabla_{\boldsymbol{\varsigma}} \Xi(\mathbf{x}, \boldsymbol{\varsigma})]$ in $\mathbb{R}^n \times \mathcal{S}_a^+$ is not monotone for each subproblem, it is natural to approximate an optimal solution of Problem (\mathcal{P}) under the perturbation of the regularized term $\frac{1}{2}\rho_k \|\mathbf{x} - \mathbf{x}_k\|^2$. This illustrates why our perturbed (exterior penalty-type) algorithm usually produces an optimal solution of Problem (\mathcal{P}) , while the existing proximal point methods based on the interior point algorithm do not.

Remark 3. In our proof of Theorem 4, we require that $\rho_k \rightarrow 0$. For classical proximal point methods, this condition was not required. In fact, this condition is adopted for simple proof that of clustering point $(\bar{\mathbf{x}}, \bar{\boldsymbol{\zeta}})$ of the sequence $\{\mathbf{x}_k, \boldsymbol{\zeta}_k\}$ being a saddle point of Problem (\mathcal{P}) . Our simulations show that $\rho_k \rightarrow 0$ can be relaxed. Indeed, in our test simulations, we found that the convergence for the case of ρ_k being chosen as a proper constant parameter is faster than that one of $\rho_k \rightarrow 0$.

5 Numerical Experiments

This section presents some numerical results by proposed canonical primal–dual method. In our simulations, the involved SDP is solved by YALMIP [23] and SeDuMi [34].

Example 5.1. Let us first consider the optimization problem (24). Taking $\rho_k = \frac{1}{k}$ and $\mu_k = 0.1\rho_k$, the initial condition is randomly generated. Table 1 reports the results obtained by our method.

From Table 1, we can see that all the four solutions $(0, 1)$, $(1, 0)$, $(0, -1)$, and $(-1, 0)$ can be detected by our algorithm with different (randomly generated) initial conditions. The corresponding $\mathbf{G}(\bar{\boldsymbol{\zeta}}) \approx 0$, as we shown in Proposition 2, can also be solved by perturbation method under any given tolerance. However, the following optimization problem

$$\min_{\mathbf{x}} P(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^m (\mathbf{x}^T \mathbf{A}_i \mathbf{x} - d_i)^2 \tag{39}$$

cannot be solved by perturbation method in general, where $\mathbf{A}_i, i = 1, \dots, m$, are randomly generated semi-definite matrix and $d_i, i = 1, \dots, m$, are chosen such that the optimal function value of $P(\mathbf{x})$ is 0. In fact, $\mathbf{G}(\bar{\boldsymbol{\zeta}}) = 0$ since the optimal cost function value of the optimization problem (39) is 0. Suppose that m is not too small

Table 1 Numerical results for optimization problem (24)

Initial condition	$\bar{\mathbf{x}}$	$\bar{\boldsymbol{\zeta}}$	$P(\bar{\mathbf{x}}) = \frac{1}{2} \ \bar{\boldsymbol{\zeta}} - \Lambda(\bar{\mathbf{x}})\ ^2$
$\begin{pmatrix} 0.81472369 \\ 0.90579194 \end{pmatrix}$	$\begin{pmatrix} -1.12001364 \times 10^{-14} \\ 1.00004756 \end{pmatrix}$	$\begin{pmatrix} -3.48372378 \\ -3.48372376 \end{pmatrix} \times 10^{-9}$	$0.93735607 \times 10^{-8}$
$\begin{pmatrix} 0.60684258 \\ 0.48598247 \end{pmatrix}$	$\begin{pmatrix} 1.00004756 \\ 5.39453096 \times 10^{-14} \end{pmatrix}$	$\begin{pmatrix} -3.48358490 \\ -3.48358548 \end{pmatrix} \times 10^{-9}$	$0.93735508 \times 10^{-8}$
$\begin{pmatrix} -0.61543234 \\ -0.79193703 \end{pmatrix}$	$\begin{pmatrix} 0.56709252 \times 10^{-14} \\ -1.00004840 \end{pmatrix}$	$\begin{pmatrix} -3.48379359 \\ -3.48379378 \end{pmatrix} \times 10^{-9}$	$0.93735627 \times 10^{-8}$
$\begin{pmatrix} -0.92181297 \\ -0.73820724 \end{pmatrix}$	$\begin{pmatrix} -1.00004756 \\ 0.12834042 \times 10^{-13} \end{pmatrix}$	$\begin{pmatrix} -3.48370090 \\ -3.48370051 \end{pmatrix} \times 10^{-9}$	$0.93735602 \times 10^{-8}$

Table 2 Numerical results for optimization problem (39) after 50 iterations

(n, m)	$P(\bar{x})$ with $\rho_k = 1/k$ and $\mu_k = 0.1\rho_k$	$P(\bar{x})$ with $\rho_k = 0.1$ and $\mu_k = 0.1\rho_k$
(20, 25)	$4.67244827 \times 10^{-6}$	$4.44146192 \times 10^{-8}$
(30, 35)	$2.10227829 \times 10^{-5}$	$0.80404292 \times 10^{-5}$
(40, 50)	0.00154861	$2.34887665 \times 10^{-5}$
(50, 60)	0.00951209	0.00032821

(for example $m \geq 20$), for any given small perturbation Δf , the corresponding saddle point problem (\mathcal{S}_p) has no solution (\bar{x}, \bar{z}) such that $\mathbf{G}(\bar{z}) \succ 0$ by our numerical experiences. Thus, the linear perturbation method cannot be applied. Now we use our proposed algorithm to solve (39) with different ρ_k and μ_k . In about 80% cases, our method can capture a solution of Problem (\mathcal{S}). The corresponding numerical results are reported in Table 2.

During our numerical computation, we observe that for very few steps (for example, less than 20 iterations), the numerical solution by our method is very close to one solution of Problem (\mathcal{S}). In fact, for all the cases in Table 2, if we set $\varepsilon = 10^{-4}$, then all the obtained results are satisfied with $\max_i |\bar{x}_i^* - x_i^{true}| \leq \varepsilon, i = 1, \dots, n$, where $\mathbf{x}^{true} = [x_1^{true}, \dots, x_n^{true}]^T$ is one of exact optimal solutions of Problem (\mathcal{S}). However, it suffers from slow convergence. Table 2 shows it clearly for the last two cases. If a gradient-based optimization method is applied, then the optimal function value is $P(\bar{x}) \approx 10^{-8}$ for all cases in Table 2.

It is obvious that Problem (39) has at least two solutions because of its symmetry, i.e., if \bar{x} is its solution, so is $-\bar{x}$. Thus, classical SDP-based relaxation methods in [18, 33, 36] cannot produce an exact solution. However, our method can produce one at the expense of iterative computation of a sequence of SDPs in most cases.

6 Applications to Sensor Networks

In this section, we apply our proposed method for sensor network localization problems.

Consider N sensors and M anchors, both located in the d -dimensional Euclidean space \mathbb{R}^d , where d is 2 or 3. Let the locations of M anchor points be given as $a_1, a_2, \dots, a_M \in \mathbb{R}^d$. The locations of N sensor points $x_1, x_2, \dots, x_N \in \mathbb{R}^d$ are to be determined. Let N_x be a subset of $\{(i, j) : 1 \leq i < j \leq N\}$ in which the distance between the i th and the j th sensor point is given as d_{ij} and N_a be a subset of $\{(i, k) : 1 \leq i \leq N, 1 \leq k \leq M\}$ in which the distance between the i th sensor point and the k th anchor point is given as e_{ik} . Then, a sensor network localization problem is to find vector $x_i \in \mathbb{R}^d$ for all $i = 1, 2, \dots, N$, such that

$$\|x_i - x_j\|^2 = d_{ij}^2, \quad \forall (i, j) \in N_x, \tag{40}$$

$$\|x_i - a_k\|^2 = e_{ik}^2, \quad \forall (i, k) \in N_a. \tag{41}$$

When the given distances d_{ij} , $(i, j) \in N_x$, and e_{ik} , $(i, k) \in N_a$, contain noise, the equalities (40) and (41) may become infeasible. Thus, instead of solving (40) and (41), we formulate it as a nonconvex optimization as given below:

$$\min_{x_1, \dots, x_N} \sum_{(i,j) \in N_x} (\|x_i - x_j\|^2 - d_{ij}^2)^2 + \sum_{(i,k) \in N_a} (\|x_i - a_k\|^2 - e_{ik}^2)^2. \quad (42)$$

Denote $\mathbf{x} = [x_1^T, \dots, x_N^T]^T \in \mathbb{R}^{dN}$. Then, (42) can be rewritten as

$$\min_{\mathbf{x}} \left\{ P(\mathbf{x}) = \sum_{ij \in N_x} (\mathbf{x}^T \mathbf{A}_{ij} \mathbf{x} - d_{ij}^2)^2 + \sum_{ik \in N_a} (\mathbf{x}^T \mathbf{B}_{ii} \mathbf{x} - 2\mathbf{f}_{ik}^T \mathbf{x} - (e_{ik}^2 - \mathbf{f}_{ik}^T \mathbf{f}_{ik}))^2 \right\}, \quad (43)$$

where $\mathbf{A}_{ij} = (\mathbf{E}_i - \mathbf{E}_j)(\mathbf{E}_i - \mathbf{E}_j)^T$, $\mathbf{B}_{ii} = \mathbf{E}_i \mathbf{E}_i^T$,

$$\mathbf{E}_i = \begin{pmatrix} 0_{d \times d} \\ \dots \\ 0_{d \times d} \\ I_{d \times d} \leftarrow i \\ 0_{d \times d} \\ \dots \\ 0_{d \times d} \end{pmatrix} \text{ and } \mathbf{f}_{ik} = \begin{pmatrix} 0_d \\ \dots \\ 0_d \\ a_k \leftarrow i \\ 0_d \\ \dots \\ 0_d \end{pmatrix}.$$

As in [17, 33], the root mean square distance

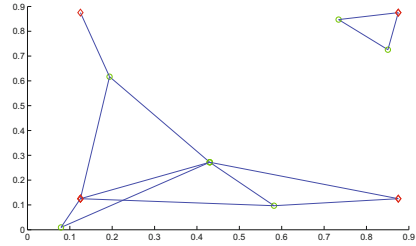
$$RMSD = \left(\frac{1}{N} \sum_{i=1}^N \|\widehat{x}_i - x_i^*\|_2^2 \right)$$

is adopted to measure the accuracy of the locations of the sensor i , $i = 1, \dots, N$, where \widehat{x}_i and x_i^* are the estimated position and true positions, respectively, $i = 1, \dots, N$. The software package SFSDP [17] is applied for generating test problems and comparison. During our simulation, all of sensors are placed in $[0, 1] \times [0, 1]$ randomly and four anchors are fixed at $(0.125, 0.125)$, $(0.125, 0.875)$, $(0.875, 0.125)$, and $(0.875, 0.875)$, respectively.

For the conventional SDP relaxation methods, the computed sensor locations match its true locations if and only if the corresponding sensor network is uniquely localizable [33, 36]. Thus, if the localized sensor network has multiple solutions, the conventional SDP relaxation methods [17, 33] fail to produce a good solution of the optimization problem defined by (43). Let us consider the following network with multiple solutions:

Example 6.1 Consider a sensor network containing six sensors and four anchors depicted in Fig. 1. From Fig. 1, we can see that the sensors x_2^* , x_3^* , and x_5^* have two positions.

Fig. 1 Network topology of six sensors and four anchors



More specifically, x_2 can be either $(0.0791, 0.0091)$ or $(0.0091, 0.1709)$, x_3, x_5 can be either the pair of $[(0.7342, 0.8470), (0.8506, 0.7257)]$ or the pair of $[(1.0158, 0.9030), (0.8994, 1.0243)]$. Let x^* , \check{x} , and \hat{x} be the true sensor locations, sensor locations computed by the SDP method ([18]), and sensor locations computed by Algorithm 1, respectively. The results are depicted in Fig. 2a, c. The true sensor locations (denoted by circles) and the computed locations (denoted by stars) are connected by solid lines. From the two figures, we can clearly see that our method produce better estimations than the SDP relaxation method in [18]. However, we need to solve a sequence of SDPs, but in [18], only one SDP is involved (Table 3). To achieve a higher accuracy, we apply the gradient-based optimization method in SFSDP to refine the solutions obtained by our method and that obtained by SDP method in [18]. After refinement, RMSD obtained by SFSDP is 4.91×10^{-5} and 2.07×10^{-8} is obtained by our method. The refined results are depicted in Fig. 2b, d. From Fig. 2b, we observe that there are still big errors for the sensor 3 and sensor 5 obtained by the refinement of SDP method in [18]. Figure 2d shows that our method produces one of the exact solutions of the optimization problem defined by (43). Thus, our method achieves better performance no matter before or after refinement.

In practical circumstances, the exact distances d_{ij} and e_{ik} are unavailable because of the presence of noise during the measurement. To model such a case, we perturb the distances as

$$\hat{d}_{ij} = \max\{(1 + \xi_{ij}), 0.1\}d_{ij} \quad ((i, j) \in N_x), \tag{44}$$

$$\hat{e}_{ik} = \max\{(1 + \xi_{ik}), 0.1\}e_{ik} \quad ((i, k) \in N_a), \tag{45}$$

where ξ_{ij}, ξ_{ik} are random variables and chosen from the standard normal distribution $N(0, \sigma)$, where σ is the noisy parameter. By substituting (44) and (45) into (43), the corresponding optimization problem involved in noisy distance is obtained.

Example 6.2 Consider a sensor network localization problem with 20 sensors and 4 anchors. Let the radio range be 0.3 and the noisy parameter be 0.001, respectively. A sensor network generated randomly by these parameters is depicted in Fig. 3.

From Fig. 3, we can verify that for this sensor network, it has a unique solution.

We apply Algorithm 2 and the SDP method in [18] in conjunction with a gradient-based refinement method to solve it. The computed results are listed in Table 4. The

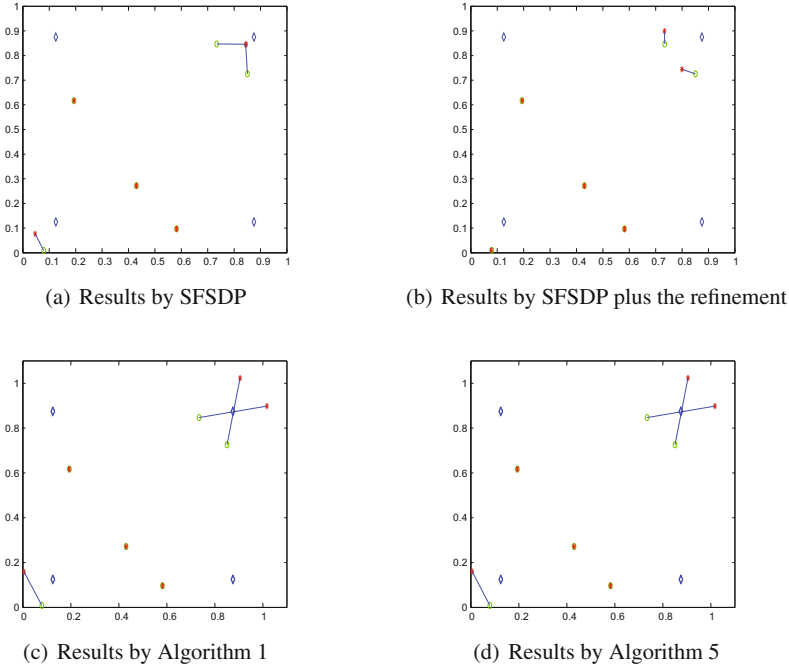


Fig. 2 Computed locations information of six sensors and four anchors

Table 3 Numerical results for six sensors and four anchors

	True solutions		Solutions by SDP in [18]		Solutions by Algorithm 1
x_1^*	(0.5818, 0.0968)	\check{x}_1	(0.5818, 0.0961)	\hat{x}_1	(0.5818, 0.0967)
x_2^*	(0.0791, 0.0091) (0.0091, 0.1709)	\check{x}_2	(0.0775, 0.0100)	\hat{x}_2	(0.0056, 0.1599)
x_3^*	(0.7342, 0.8470) (1.0158, 0.9030)	\check{x}_3	(0.7334, 0.8985)	\hat{x}_3	(1.0167, 0.8980)
x_4^*	(0.1936, 0.6169)	\check{x}_4	(0.1946, 0.6170)	\hat{x}_4	(0.1937, 0.6169)
x_5^*	(0.8506, 0.7257) (0.8994, 1.0243)	\check{x}_5	(0.7995, 0.7439)	\hat{x}_5	(0.9047, 1.0234)
x_6^*	(0.4301, 0.2720)	\check{x}_6	(0.4300, 0.2713)	\hat{x}_6	(0.4299, 0.2718)

RMSD computed by SFSDP in conjunction with a gradient-based refinement method is 9.95×10^{-2} while that computed by our method is 4.1041×10^{-7} . The computed results by Algorithm 2 and by SDP in conjunction with a gradient-based refinement method in [18] are depicted in Fig. 4. From Fig. 4 and the values of RMSD, we know that our method achieves better performance than that by SFSDP in conjunction with a gradient-based refinement method. This is because if the distances are inexact, the SDP-based methods in [18] are not ensured to produce a good solution. However, our

Fig. 3 Network topology of 20 sensors and 4 anchors

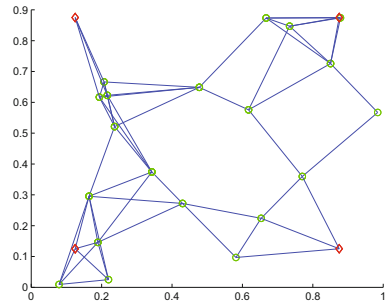


Table 4 Numerical results for 20 sensors and 4 anchors

	True solutions		Solutions by SDP + refinement in [18]		Solutions by Algorithm 2
x_1^*	(0.5818, 0.0968)	\check{x}_1	(0.6203, 0.2107)	\hat{x}_1	(0.5815, 0.0963)
x_2^*	(0.0791, 0.0091)	\check{x}_2	(0.1379, 0.0015)	\hat{x}_2	(0.0795, 0.0091)
x_3^*	(0.7342, 0.8470)	\check{x}_3	(0.7369, 0.8030)	\hat{x}_3	(0.7343, 0.8475)
x_4^*	(0.1936, 0.6169)	\check{x}_4	(0.2384, 0.6406)	\hat{x}_4	(0.1939, 0.6168)
x_5^*	(0.8506, 0.7257)	\check{x}_5	(0.8610, 0.7040)	\hat{x}_5	(0.8503, 0.7258)
x_6^*	(0.4301, 0.2720)	\check{x}_6	(0.4319, 0.2943)	\hat{x}_6	(0.4301, 0.2719)
x_7^*	(0.9846, 0.5671)	\check{x}_7	(0.7621, 0.5022)	\hat{x}_7	(0.9833, 0.5670)
x_8^*	(0.3429, 0.3741)	\check{x}_8	(0.3399, 0.3793)	\hat{x}_8	(0.3430, 0.3739)
x_9^*	(0.2070, 0.6663)	\check{x}_9	(0.2612, 0.6874)	\hat{x}_9	(0.2067, 0.6662)
x_{10}^*	(0.6176, 0.5756)	\check{x}_{10}	(0.6612, 0.5025)	\hat{x}_{10}	(0.6172, 0.5762)
x_{11}^*	(0.1644, 0.2955)	\check{x}_{11}	(0.1643, 0.3085)	\hat{x}_{11}	(0.1643, 0.2956)
x_{12}^*	(0.6533, 0.2237)	\check{x}_{12}	(0.6984, 0.3363)	\hat{x}_{12}	(0.6530, 0.2229)
x_{13}^*	(0.6673, 0.8736)	\check{x}_{13}	(0.6683, 0.8336)	\hat{x}_{13}	(0.6676, 0.8746)
x_{14}^*	(0.2161, 0.6226)	\check{x}_{14}	(0.2607, 0.6429)	\hat{x}_{14}	(0.2165, 0.6226)
x_{15}^*	(0.7701, 0.3595)	\check{x}_{15}	(0.6232, 0.2186)	\hat{x}_{15}	(0.7691, 0.3595)
x_{16}^*	(0.1894, 0.1458)	\check{x}_{16}	(0.1637, 0.1663)	\hat{x}_{16}	(0.1893, 0.1460)
x_{17}^*	(0.8786, 0.8741)	\check{x}_{17}	(0.8746, 0.8626)	\hat{x}_{17}	(0.8789, 0.8743)
x_{18}^*	(0.4776, 0.6487)	\check{x}_{18}	(0.5169, 0.5805)	\hat{x}_{18}	(0.4777, 0.6502)
x_{19}^*	(0.2370, 0.5215)	\check{x}_{19}	(0.2477, 0.5368)	\hat{x}_{19}	(0.2378, 0.5215)
x_{20}^*	(0.2197, 0.0249)	\check{x}_{20}	(0.0236, 0.0836)	\hat{x}_{20}	(0.2202, 0.0253)

method is based on the global solution of the optimization problem defined by (43). Thus, the inexact measurements do not deteriorate the performance of our method.

Example 6.3 Consider a sensor network localization problem with 50 sensors, 4 anchors, and noisy perturbation being 0.001. The corresponding connections between sensors and sensors and sensors and anchors are depicted in Fig. 5.

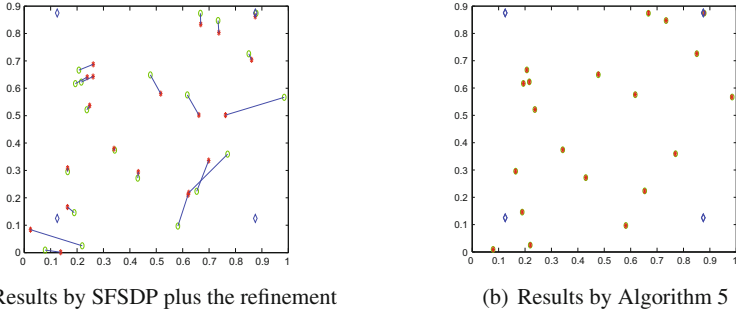


Fig. 4 Computed locations information of 20 sensors and 4 anchors

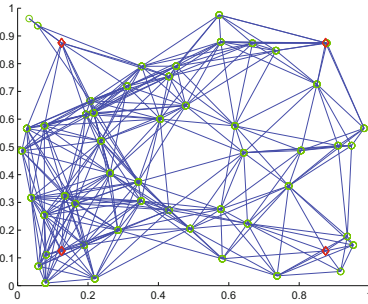


Fig. 5 Network topology of 50 sensors and 4 anchors

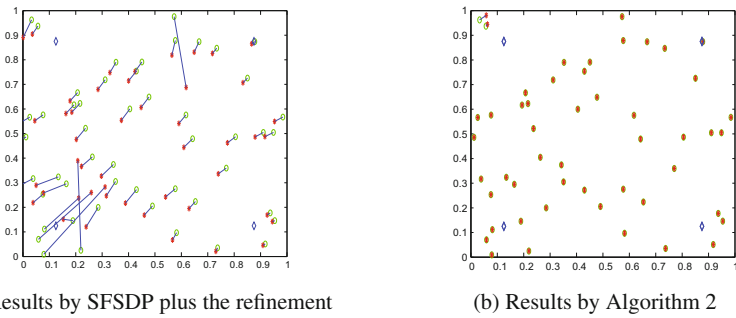


Fig. 6 Computed locations information of 50 sensors and 4 anchors

The computed results by Algorithm 2 and by SFSDP in conjunction with a gradient-based refinement method are depicted in Fig. 6. The RMSD computed by SFSDP in conjunction with a gradient-based refinement method is 1.07×10^{-1} , while that by our method is 1.9956×10^{-5} . Both Fig. 6 and the values of RMSD show that our method achieves better performance.

7 Conclusion

This paper presented an effective method and algorithms for solving a class of non-convex optimization problems. Using the canonical duality theory, the original non-convex optimization problem is first relaxed to a convex–concave saddle point optimization problem. Depending on the singularity of the matrix \mathbf{G} , this relaxed saddle point problem is classified in two cases: degenerate or non-degenerate. For the non-degenerate case, the solution of the primal problem can be recovered exactly through solving a convex SDP problem. Otherwise, a quadratic perturbed primal–dual scheme is proposed to solve the corresponding degenerate saddle point problem. We proved that, under certain conditions, the sequence generated by our proposed scheme converges to a solution of the corresponding saddle point problem. If this saddle point satisfies the condition of $\|\Lambda(\bar{\mathbf{x}}) - \nabla V^*(\bar{\boldsymbol{\zeta}})\| \leq \varepsilon$ within a given error tolerance, then the solution of the primal problem is also recovered exactly. Otherwise, $\bar{\mathbf{x}}$ is taken as a starting point and a gradient-based optimization method is applied to refine the primal solution. Numerical simulations show that our method can achieve better performance than the conventional SDP-based relaxation methods.

Acknowledgements The research was supported by US Air Force Office of Scientific Research under the grants AFOSR FA9550-17-1-0151 and AOARD FOST-16-265. Numerical computation was performed by research student Mr. Chaojie Li at Federation University.

References

1. Ball, J.M.: Some open problems in elasticity. *Geometry, Mechanics, and Dynamics*, pp. 3–59. Springer, New York (2002)
2. Birgin, E.G., Floudas, C.A., Martinez, J.M.: Global minimization using an Augmented Lagrangian method with variable lower-level constraints. *Math. Program. Ser. A* **125**, 139–162 (2010)
3. Gallier, J.: The Schur complement and symmetric positive semidefinite (and definite) matrices. www.cis.upenn.edu/jean/schurcomp.pdf
4. Gao, D.Y.: *Duality Principles in Nonconvex Systems: Theory Methods and Applications*. Springer, New York (2000)
5. Gao, D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. *J. Ind. Manag. Optim.* **3**(2), 293–304 (2007)
6. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solutions for global optimization. *Comput. Chem. Eng.* **33**, 1964–1972 (2009)
7. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. *Q. Appl. Math.* **47**(3), 487–504 (1989)
8. Gao, D.Y., Sherali, H.D.: Canonical duality: connection between nonconvex mechanics and global optimization. *Advances in Applied Mathematics and Global Optimization*, pp. 249–316. Springer, Berlin (2009)
9. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. *J. Global Optim.* **47**, 463–484 (2010)
10. Gao, D.Y., Wu, C.Z.: On the triality theory for a quartic polynomial optimization problem. *J. Ind. Manag. Optim.* **8**(1), 229–242 (2012)

11. 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. In: Pardalos, P.M., Ye, Y.Y., Boginski, V., Commander, C. (eds.) *Sensors: Theory, Algorithms and Applications*, vol. 61, pp. 37–54. Springer, Berlin (2012)
12. Guler, O.: New proximal point algorithms for convex minimization. *SIAM J. Optim.* **2**, 649–664 (1992)
13. He, B.S., Liao, L.Z.: Improvement of some projection methods for monotone nonlinear variational inequalities. *J. Optim. Theory Appl.* **112**, 111–128 (2002)
14. He, B.S., Yuan, X.M.: An accelerated inexact proximal point algorithm for convex minimization. *J. Optim. Theory Appl.* **154**, 536–548 (2012)
15. Horn, R.A., Johnson, C.R.: *Matrix Analysis*. Cambridge University Press, Cambridge (1985)
16. Kaplan, A., Tichatschke, R.: Proximal point methods and nonconvex optimization. *J. Global Optim.* **13**, 389–406 (1998)
17. Kim, S., Kojima, M., Waki, H., Yamashita, M.: User Manual for SFSDDP: a Sparse versions of Full Semidefinite programming relaxation for sensor network localization problems. *Research Reports on Mathematical and Computer Science, SERIES B* (2009)
18. Kim, S., Kojima, M., Waki, H.: Exploiting sparsity in SDP relaxation for sensor network localization. *SIAM J. Optim.* **1**, 192–215 (2009)
19. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. *Ekonomika i Matematicheskie* **12**, 747–756 (1976)
20. Lasserre, J.B.: Global optimization with polynomials and the problems of moments. *SIAM J. Optim.* **11**, 796–817 (2001)
21. Lewis, A.S., Wright, S.J.: A proximal method for composite minimization. [arXiv:0812.0423v1](https://arxiv.org/abs/0812.0423v1)
22. Li, C., Wang, X.: On convergence of the Gauss–Newton method for convex composite optimization. *Math. Program.* **91**, 349–356 (2002)
23. Löberg, J.: YALMIP: a toolbox for modeling and optimization in Matlab. In: *Proceedings of the International Symposium on CACSD, Taipei, Taiwan*, pp. 284–89 (2004)
24. Marsden, J.E., Hughes, T.J.R.: *Mathematical Foundations of Elasticity*. Prentice-Hall, New Jersey (1983)
25. More, J.J.: Generalizations of the trust region problem, Technical Report MCS-P349-0193. Argonne National Labs, Argonne, IL (1993)
26. More, J., Wu, Z.: Distance geometry optimization for protein structures. *J. Global Optim.* **15**, 219–234 (1999)
27. Nesterov, Y.: Dual extrapolation and its applications to solving variational inequalities and related problems. *Math. Program. Ser. B* **109**, 319–344 (2007)
28. Nesterov, Y.: Primal-dual subgradient methods for convex problems. *Math. Program. Ser. B* **120**, 221–259 (2009)
29. Pong, T.K., Tseng, P.: (Robust) Edge-based semidefinite programming relaxation of sensor network localization. *Math. Program.* **130**(2), 321–358 (2011)
30. Rockafellar, R.T.: Monotone operators and the proximal point algorithms. *SIAM J. Cont. Optim.* **14**, 887–898 (1976)
31. Ruan, N., Gao, D.Y.: Canonical duality approach for nonlinear dynamical systems. *IMA J. Appl. Math.* **79**(2), 313–325 (2014)
32. Ruan, N., Gao, D.Y., Jiao, Y.: Canonical dual least square method for solving general nonlinear systems of quadratic equations. *Comput. Optim. Appl.* **47**, 335–347 (2010)
33. So, A.M., Ye, Y.: Theory of semidefinite programming for sensor network localization. *Math. Program. Ser. B* **109**, 367–384 (2007)
34. Sturm, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim. Methods Softw.* **12**, 625–633 (1999)
35. Taskar, B., Julien, S.L., Jordan, M.I.: Structured prediction, dual extragradient and bregman projections. *J. Mach. Learn. Res.* **7**, 1627–1653 (2006)
36. Wang, Z., Zheng, S., Ye, Y., Boyd, S.: Further relaxations of the semidefinite programming approach to sensor network localization. *SIAM J. Optim.* **19**, 655–673 (2008)

37. Yamashita, N., Dan, H., Fukushima, M.: On the identification of degenerate indices in the nonlinear complementarity problem with the proximal point algorithm. *Math. Program.* **99**, 377–397 (2004)
38. Zhang, J., Gao, D.Y., Yearwood, J.: A novel canonical dual computational approach for prion AGAAAAGA amyloid fibril molecular modeling. *J. Theor. Biol.* **284**, 149–157 (2011). doi:[10.1016/j.jtbi.2011.06.024](https://doi.org/10.1016/j.jtbi.2011.06.024)