Advances in Mechanics and Mathematics 37

David Yang Gao
Vittorio Latorre
Ning Ruan *Editors*

Canonical Duality Theory

Unified Methodology for Multidisciplinary Study

Springer

Advances in Mechanics and Mathematics

Volume 37

More information about this series at http://www.springer.com/series/5613

David Yang Gao · Vittorio Latorre
Ning Ruan
Editors

Canonical Duality Theory

Unified Methodology for Multidisciplinary Study

Springer

Editors
David Yang Gao
Faculty of Science and Technology
Federation University Australia
Mt. Helen, VIC
Australia

Ning Ruan
Faculty of Science and Technology
Federation University Australia
Mt. Helen, VIC
Australia

Vittorio Latorre
Department of Computer, Control,
and Managment Engineering
Sapienza University of Rome
Rome, Roma
Italy

ISSN 1571-8689 ISSN 1876-9896 (electronic)
Advances in Mechanics and Mathematics
ISBN 978-3-319-58016-6 ISBN 978-3-319-58017-3 (eBook)
DOI 10.1007/978-3-319-58017-3

Library of Congress Control Number: 2017940806

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Duality is a beautiful concept that pervades all natural phenomena. In classical mechanics, each potential energy principle is linked with a unique complementary energy principle via the traditional Legendre transformation. The popular Hellinger–Reissner principle is actually a special saddle Lagrangian duality theory in convex analysis and optimization, which lays a foundation for hybrid/mixed finite element methods in computational mechanics and primal-dual interior point methods in mathematical programming. However, this one-to-one duality is broken in nonconvex systems due to a so-called duality gap produced by the modern Fenchel–Moreau transformation. In finite elasticity, the existence of a pure stress-based complementary energy principle was a well-known open problem, existing for several decades. In global optimization and computer science, many nonconvex problems are considered as NP-hard (Non-deterministic Polynomial-time hard) due to the lack of global optimality criteria. Unfortunately, this well-known difficulty is not fully recognized in computational mechanics due to the significant gap between engineering mechanics and global optimization. Indeed, engineers and scientists are mistakenly attempting to use traditional finite element methods and commercial software for solving nonconvex mechanics problems.

Canonical duality theory is a breakthrough methodological theory that can be used not only for modeling complex systems within a unified framework, but also for solving a large class of challenging problems in multidisciplinary fields of engineering, mathematics, and sciences. The concept of canonical (i.e. one-to-one) duality is from the traditional Chinese Yin-Yang philosophy. Niels Bohr realized its value in quantum mechanics. Based on this philosophy, a unified canonical duality framework in mathematical physics was first proposed in the work by Gao and Strang in 1989. This framework reveals an intrinsic duality in nonconvex systems and lays a foundation for the canonical duality theory. The canonical duality theory was developed originally from nonconvex mechanics (1989–2000) and then generalized to global optimization (2000–2010). This theory is composed mainly of (1) a canonical dual transformation, which can be used to formulate perfect dual problems without duality gap; (2) a complementary-dual principle, which solved the open problem in finite elasticity and provides a unified analytical solution form for general nonconvex/nonsmooth/discrete problems; (3) a triality theory, which can be used to identify both global and local optimality conditions and to develop powerful algorithms for solving challenging problems in complex systems. During the past 10 years, the canonical duality theory has been applied successfully for solving a wide class of real-world problems in chaotic dynamics, computational biology, filter design, information technology, logistics and transportation, machine learning, network communication, nonlinear partial differential equations (PDEs) in finite deformation theory, operations research, post-bifurcation, phase transitions in solid mechanics, and materials science, as well as modeling of complex systems, etc.

The original motivation of this book was a colloquium talk presented by David Yang Gao at UC Berkeley in 2013. This volume provides a comprehensive review of the canonical duality theory, its methodology, and algorithms for solving challenging problems in complex systems with applications in nonconvex analysis, variational inequalities, large deformation problems, global optimization, and computational mechanics. It is the authors' hope that by reading this book, the readers should be able to see the beauty and unity of the canonical duality theory and its potential applications in multidisciplinary fields.

The research projects on the canonical duality theory have been continuously supported by US National Science Foundation and US Air Force Office of Scientific Research (AFOSR) under the grants FA9550-09-1-0285, FA9550-10-1-0487, FA2386-16-1-4082 and FA9550-17-1-0151. The authors sincerely thank the program managers, Drs. Juan Zhang, Jay Myung, James H. Lawton, and Kristopher Ahlers at AFOSR, for their professional managements and constant support. The authors wish to express their sincere appreciation to the contributors of this book for their collaborations. Special thanks go to Marc Strauss and Dimana Tzvetkova at Springer for their enthusiasm and professional help for this book.

Ballarat, Australia David Yang Gao
Rome, Italy Vittorio Latorre
Ballarat, Australia Ning Ruan
October 2016

Contents

Canonical Duality-Triality Theory: Bridge Between Nonconvex Analysis/Mechanics and Global Optimization in Complex System

David Yang Gao, Ning Ruan and Vittorio Latorre

Abstract Canonical duality-triality is a breakthrough methodological theory, which can be used not only for modeling complex systems within a unified framework, but also for solving a wide class of challenging problems from real-world applications. This paper presents a brief review on this theory, its philosophical origin, physics foundation, and mathematical statements in both finite- and infinite-dimensional spaces. Particular emphasis is placed on its role for bridging the gap between nonconvex analysis/mechanics and global optimization. Special attentions are paid on unified understanding the fundamental difficulties in large deformation mechanics, bifurcation/chaos in nonlinear science, and the NP-hard problems in global optimization, as well as the theorems, methods, and algorithms for solving these challenging problems. Misunderstandings and confusion on some basic concepts, such as objectivity, nonlinearity, Lagrangian, and generalized convexities are discussed and classified. Breakthrough from recent challenges and conceptual mistakes by M. Voisei, C. Zălinescu and his coworker are addressed. The paper is ended with some open problems and future works in global optimization and nonconvex mechanics.

D.Y. Gao (✉) · N. Ruan
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia
e-mail: d.gao@federation.edu.au

N. Ruan
e-mail: n.ruan@federation.edu.au

V. Latorre
Department of Computer, Control and Management Engineering University of Rome,
Rome, Italy
e-mail: latorre@dis.uniroma1.it

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics
and Mathematics 37, DOI 10.1007/978-3-319-58017-3_1

1 Introduction

Duality is one of the oldest and most beautiful ideas in human knowledge. It has a simple origin from the oriental philosophy of *yin-yang principle* tracing back 5000 years ago [41]. According to *I Ching*,[1] the fundamental law of the nature is the *Dao*, the duality of one yin and one yang, which gives two opposite or complementary points of view of looking at the same object. In quantum mechanics, the wave–particle duality is a typical example to fully describe the behavior of quantum scale objects. Mathematically, duality represents certain translation of concepts, theorems, or mathematical structures in a one-to-one fashion, i.e., if the dual of A is B, then the dual of B is A (cf. [5, 19, 46, 115]). This one-to-one complementary relation is called the *canonical duality*. It is emphasized recently by Sir Michael Atiyah that duality in mathematics is not a theorem, but a "principle" [5]. Therefore, any duality gap is not allowed. This fact is well-known in mathematics and physics, but not in optimization due to the existing gap between these fields. To bridge this gap, a canonical duality-triality theory has been developed originally from nonconvex mechanics [49] with extensive applications in engineering, mathematics, and sciences, especially in the multidisciplinary fields of nonconvex mechanics and global optimization [57, 65, 76].

1.1 Nonconvex Analysis/Mechanics and Difficulties

Mathematical theory of duality for convex problems has been well-established. In linear elasticity, it is well-known that each potential energy principle is associated with a unique complementary energy principle through Legendre transformation. This one-to-one duality is guaranteed by convexity of the stored energy. The well-known Hellinger–Reissner principle is actually a special Lagrangian saddle min–max duality theory in convex analysis, which lays a foundation for mixed/hybrid finite element methods with successful applications in structural limit analysis [28, 29]. However, the one-to-one duality is broken in nonconvex systems. In large deformation theory, the stored energy is generally nonconvex and its Legendre conjugate cannot be uniquely determined. It turns out that the existence of a pure stress-based complementary-dual energy principle (no duality gap) was a well-known open problem over a half century and subjected to extensive discussions by many leading experts including Levison [105], Koiter [95], Oden and Reddy [118], Lee and Shield [104], Stumpf [132], etc.

[1] Also known as the *Book of Changes, Zhouyi* and *Yijing,* is the world oldest and most sophisticated system of wisdom divination, the fundamental source of most of the eastern philosophy, medicine, and spirituality. Traditionally, it was believed that the principles of the I Ching originated with the mythical King Fu Xi during the 3rd and 2nd millennia BCE.

Nonconvex phenomena arise naturally in large classes of engineering applications. Many real-life problems in modern mechanics and complex systems require consideration of nonconvex effects for their accurate modeling. For example, in modeling of hysteresis, phase transitions, shape-memory alloys, and superconducting materials, the free energy functions are usually nonconvex due to certain internal variables [66, 71, 72]. In large deformation analysis, thin-walled structure can buckle even before the stress reaches its elastic limit [37, 38, 78]. Mathematically speaking, many fundamentally difficult problems in engineering and the sciences are mainly due to the nonconvexity of their modeling. In static systems, the nonconvexity usually leads to multi-solutions in the related governing equations. Each of these solutions represents certain possible phase or buckled state in large deformed solids. These local solutions are very sensitive to the internal parameters and external force. In dynamical systems, the so-called chaotic behavior is mainly due to nonconvexity of the objective functions [56]. Numerical methods (such as FEM, FDM, etc.) for solving nonconvex minimal potential variational problems usually end up with nonconvex optimization problems [40, 51, 83, 92, 129]. Due to the lack of global optimality criteria, finding global optimal solutions is fundamentally difficult, or even impossible by traditional numerical methods and optimization techniques. For example, it was discovered by Gao and Ogden [71, 72] that for certain given external loads, both the global and local minimizers are nonsmooth and cannot be determined by any Newton-type numerical methods. In fact, many nonconvex problems are considered as NP-hard (Nondeterministic Polynomial-time hard) in global optimization and computer science [65, 76]. Unfortunately, these well-known difficulties are not fully recognized in computational mechanics due to the significant gap between engineering mechanics and global optimization. Indeed, engineers and scientists are mistakenly attempting to use traditional finite element methods and commercial software for solving nonconvex mechanics problems. In order to identify the fundamental difficulty of the nonconvexity from the traditional definition of nonlinearity, the terminology of *Nonconvex Mechanics* was formally proposed by Gao, Ogden and Stavroulakis in 1999 [84]. The *Handbook of Nonconvex Analysis* by Gao and Motreanu [70] presents recent advances in the field.

1.2 Global Optimization and Challenges

In parallel with the nonconvex mechanics, global optimization is a multidisciplinary research field developed mainly from nonconvex/combinatorial optimization and computational science during the last 90s. In general, the global optimization problem is formulated in terms of finding the absolutely best set of solutions for the following constrained optimization problem

$$\min f(x), \quad \text{s.t. } h_i(x) = 0, \quad g_j(x) \le 0 \;\; \forall i \in I_m, \quad j \in I_p, \tag{1}$$

where $f(x)$ is the so-called "objective function",[2] $h_i(x)$ and $g_j(x)$ are constraint functions, $I_m = \{1, \ldots, m\}$ and $I_p = \{1, \ldots, p\}$ are index sets. It must be emphasized that, different from the basic concept of *objectivity* in continuum physics, the objective function extensively used in mathematical optimization is allowed to be any arbitrarily given function, even the linear function. Clearly, this mathematical model is artificial. Although it enables one to "model" a very wide range of problems, it comes at a price: even very special kinds of nonconvex/discrete optimization problems are considered to be NP-hard. This dilemma is due to the gap between mathematical optimization and mathematical physics. In science, the concept of objectivity is often attributed with the property of scientific measurements that can be measured independently of the observer. Therefore, a function in mathematical physics is called objective only if it depends on certain measure of its variables (see Definition 6.1.2, [49] and the next section). Generally speaking, a useful mathematical model must obey certain fundamental law of nature. Without detailed information on these arbitrarily given functions, it is impossible to have a general theory for finding global extrema of the general nonconvex problem (1). This could be the reason why there was no breakthrough in nonlinear programming during the past 60 years.

In addition to the nonconvexity, many global optimization problems in engineering design and operations research explicitly require integer or binary decision variables. For example, in topology optimization of engineering structures, the design variable of material density $\rho(\mathbf{x}) = \{0, 1\}$ is a discrete selection field, i.e., by selection it has to take the value, 1, and by deselection it has to take the value, 0 (see [8]). By the fact that the deformation variable is a continuous field, which should be determined in each iteration for topological structure, therefore, the finite element method for solving topology optimization problems ends up with a coupled mixed integer nonlinear programming problem. Discrete problems are frequently encountered in modeling real-world systems for a wide spectrum of applications in decision science, management optimization, industrial and systems engineering. Imposing such integer constraints on the variables makes the global optimization problems much more difficult to solve. It is well-known in computational science and global optimization that even the most simple quadratic minimization problem with boolean constraint

$$\min \left\{ \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{x}^T\mathbf{f} |\ \ \mathbf{x} \in \{0, 1\}^n \right\} \tag{2}$$

is considered to be NP-hard (Nondeterministic Polynomial-time hard) [75]. Indeed, this integer minimization problem has 2^n local solutions. Due to the lack of global optimality criterion, traditional direct approaches, such as the popular branch and bound methods, can only handle very small size problems. Actually, it was proved by Pardalos and Vavasis [121, 135] that instead of the integer constraint, the continuous quadratic minimization with box constraints $\mathbf{x} \in [0, 1]^n$ is NP-hard as long as the matrix $\mathbf{Q}$ has one negative eigenvalue.

[2] This terminology is used mainly in English literature. The function $f(x)$ is called the target function in Chinese and Japanese literatures, the goal function in Russian and German literatures.

During the last 20 years, the field of global optimization has been developed dramatically to across almost every branch of sciences, engineering, and complex systems [26, 27, 123]. By the fact that the mathematical model (1) is too general to have a mathematical theory for identifying global extrema, the main task in global optimization is to study algorithmic methods for numerically solving the optimal solutions. These methods can be categorized into two main groups: deterministic and stochastic. *Stochastic methods* are based on an element of random choice. Because of this, one has to sacrifice the possibility of an absolute guarantee of success within a finite amount of computation. *Deterministic methods,* such as the cutting plane, branch and bound methods, can find global optimal solutions, but not in polynomial time. Therefore, this type of methods can be used only for solving very small-sized problems. Indeed, global optimization problems with 200 variables are referred to as "medium scale", problems with 1,000 variables as "large scale", and the so-called "extralarge scale" is only around 4,000 variables [10]. In topology optimization, the variables could be easily 100 times more than this extralarge scale in global optimization. Therefore, to develop a unified deterministic theory for efficiently solving general global optimization problems is fundamentally important, not only in mathematical optimization, but also in general nonconvex analysis and mechanics.

2 Canonical Duality-Triality Theory

The canonical duality-triality theory comprises mainly three parts:

(i) a *canonical dual transformation*, (ii) a *complementary-dual principle,* and (iii) a *triality theory*.

The canonical dual transformation is a versatile methodology which can be used to model complex systems within a unified framework and to formulate perfect dual problems without a duality gap. The complementary-dual principle presents a unified analytic solution form for general problems in continuous and discrete systems. The triality theory reveals an intrinsic duality pattern in multiscale systems, which can be used to identify both global and local extrema, and to develop deterministic algorithms for effectively solving a wide class of nonconvex/nonsmooth/discrete optimization/variational problems.

2.1 General Modeling and Objectivity

A useful methodological theory should have solid foundations not only in physics, but also in mathematics, even in philosophy and aesthetics. The canonical duality theory was developed from Gao and Strang's original work for solving the following general nonconvex/nonsmooth variational problem [77]:

$$\min\{\Pi(\chi) = W(D\chi) - F(\chi) \mid \chi \in \mathscr{X}_c\}, \tag{3}$$

where $F(\boldsymbol{\chi})$ is the external energy, which must be linear on its domain $\mathscr{X}_a$; the linear operator $D : \mathscr{X}_a \to \mathscr{W}_a$ assigns each configuration $\boldsymbol{\chi}$ to an internal variable $\boldsymbol{\varepsilon} = D\boldsymbol{\chi}$ and, correspondingly, $W : \mathscr{W}_a \to \mathbb{R}$ is called the internal (or stored) energy. The feasible set $\mathscr{X}_c = \{\boldsymbol{\chi} \in \mathscr{X}_a | \ D\boldsymbol{\chi} \in \mathscr{W}_a\}$ is the *kinetically admissible space*.

By Riesz representation theorem, the external energy can be written as $F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle$, where $\bar{\boldsymbol{\chi}}^* \in \mathscr{X}^*$ is a given input (or source). The bilinear form $\langle \boldsymbol{\chi}, \boldsymbol{\chi}^* \rangle : \mathscr{X} \times \mathscr{X}^* \to \mathbb{R}$ puts $\mathscr{X}$ and $\mathscr{X}^*$ in duality. Therefore, the variation (or Gâteaux derivative) of $F(\boldsymbol{\chi})$ leads to the *action–reaction duality*: $\bar{\boldsymbol{\chi}}^* = \partial F(\boldsymbol{\chi})$. Dually, the internal energy must be an *objective function* on its domain $\mathscr{W}_a$ such that the intrinsic physical behavior of the system can be described by the *constitutive duality*: $\boldsymbol{\sigma} = \partial W(\boldsymbol{\varepsilon})$.

Objectivity is a basic concept in mathematical modeling [17, 91, 111, 120], but is still subjected to seriously study in continuum physics [109, 116, 117]. The mathematical definition was given in Gao's book (Definition 6.1.2 [49]).

Definition 1 (Objectivity and Isotropy). Let $\mathscr{R}$ be a proper orthogonal group, i.e., $\mathbf{R} \in \mathscr{R}$ if and only if $\mathbf{R}^T = \mathbf{R}^{-1}$, $\det \mathbf{R} = 1$. A set $\mathscr{W}_a$ is said to be objective if

$$\mathbf{R}\boldsymbol{\varepsilon} \in \mathscr{W}_a \ \ \forall \boldsymbol{\varepsilon} \in \mathscr{W}_a, \ \forall \mathbf{R} \in \mathscr{R}.$$

A real-valued function $W : \mathscr{W}_a \to \mathbb{R}$ is said to be objective if

$$W(\mathbf{R}\boldsymbol{\varepsilon}) = W(\boldsymbol{\varepsilon}) \ \ \forall \boldsymbol{\varepsilon} \in \mathscr{W}_a, \ \forall \mathbf{R} \in \mathscr{R}. \tag{4}$$

A set $\mathscr{W}_a$ is said to be isotropic if $\boldsymbol{\varepsilon}\mathbf{R} \in \mathscr{W}_a \ \ \forall \boldsymbol{\varepsilon} \in \mathscr{W}_a, \ \forall \mathbf{R} \in \mathscr{R}$.

A real-valued function $W : \mathscr{W}_a \to \mathbb{R}$ is said to be isotropic if

$$W(\boldsymbol{\varepsilon}\mathbf{R}) = W(\boldsymbol{\varepsilon}) \ \ \forall \boldsymbol{\varepsilon} \in \mathscr{W}_a, \ \forall \mathbf{R} \in \mathscr{R}. \tag{5}$$

Geometrically speaking, an objective function does not depend on the rotation, but only on certain measure of its variable. The isotropy means that the function $W(\boldsymbol{\varepsilon})$ possesses a certain symmetry. In continuum physics, the right Cauchy–Green tensor[3] $\mathbf{C}(\mathbf{F}) = \mathbf{F}^T\mathbf{F}$ is an objective strain measure, while the left Cauchy–Green tensor $\mathbf{c} = \mathbf{F}\mathbf{F}^T$ is an isotropic strain measure. In Euclidean space $\mathscr{W}_a \subset \mathbb{R}^n$, the simplest objective function is the ℓ_2-norm $\|\boldsymbol{\varepsilon}\|$ in $\mathbb{R}^n$ as we have $\|\mathbf{R}\boldsymbol{\varepsilon}\|^2 = \boldsymbol{\varepsilon}^T\mathbf{R}^T\mathbf{R}\boldsymbol{\varepsilon} = \|\boldsymbol{\varepsilon}\|^2 \ \ \forall \mathbf{R} \in \mathscr{R}$. In this case, the objectivity is equivalent to isotropy and, in Lagrangian mechanics, the kinetic energy is required to be isotropic [98].

Physically, an objective function does not depend on observers [117], which is essential for any real-world mathematical modeling. In continuum physics, objectivity implies that the equilibrium condition of angular momentum (symmetry of the

[3]Tensor is a geometrical object which is defined as a multidimensional array satisfying a transformation law (see [120]). A tensor must be independent of a particular choice of coordinate system (frame-indifference). But this terminology has been misused in optimization literature, where, any multidimensional array of data is called tensor (see [6]).

Cauchy stress tensor $\boldsymbol{\sigma} = \partial W(\boldsymbol{\varepsilon})$, Sect. 6.1 [49]) holds. It is emphasized by P. Ciarlet that the objectivity is not an assumption, but an axiom [17]. Indeed, the objectivity is also known as the *axiom of material frame invariance*, which lays a foundation for the canonical duality theory.

As an objective function, the internal energy $W(\boldsymbol{\varepsilon})$ does not depends on each particular problem. Dually, the external energy $F(\boldsymbol{\chi})$ can be called the *subjective function,* which depends on each given problem, such as the inputs, boundary conditions and geometrical constraints in $\mathscr{X}_a$. Together, $\Pi(\boldsymbol{\chi}) = W(D\boldsymbol{\chi}) - F(\boldsymbol{\chi})$ is called the total potential energy and the minimal potential principle leads to the general optimization problem (3).

For dynamical problems, the liner operator $D = \{\partial_t, \partial_x\}$ and $W(D\boldsymbol{\chi}) = T(\partial_t\boldsymbol{\chi}) - V(\partial_x\boldsymbol{\chi})$, where $T(\mathbf{v})$ is the kinetic energy and $V(\mathbf{e})$ can be viewed as stored potential energy, then

$$\Pi(\boldsymbol{\chi}) = T(\partial_t\boldsymbol{\chi}) - V(\partial_x\boldsymbol{\chi}) - F(\boldsymbol{\chi})$$

is the total action in dynamical systems.

The necessary condition $\delta\Pi(\boldsymbol{\chi}) = 0$ for the solution of the minimization problem (3) leads to a general equilibrium equation:

$$A(\boldsymbol{\chi}) = D^*\partial_{\boldsymbol{\varepsilon}} W(D\boldsymbol{\chi}) = \bar{\boldsymbol{\chi}}^*. \tag{6}$$

This abstract form of equilibrium equation covers extensive real-world applications ranging from traditional mathematical physics, modern economics, ecology, game theory, information technology, network optimization, operations research, and much more [49, 76, 131]. Particularly, if $W(\boldsymbol{\varepsilon})$ is quadratic such that $\partial^2 W(\boldsymbol{\varepsilon}) = H$, then the operator $A : \mathscr{X}_c \to \mathscr{X}^*$ is linear and can be written in the triality form: $A = D^*HD$, which appears extensively in mathematical physics, optimization, and linear systems [49, 119, 131]. Clearly, any convex quadratic function $W(\boldsymbol{\varepsilon})$ is objective due to the Cholesky decomposition $A = \Lambda^*\Lambda \succeq 0$.

Example 1 (Manufacturing/Production Systems). In management science, the configuration variable is a vector $\boldsymbol{\chi} \in \mathbb{R}^n$, which could represent the products of a manufacture company. Its dual variable $\bar{\boldsymbol{\chi}}^* \in \mathbb{R}^n$ can be considered as market price (or demands). Therefore, the external energy $F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle = \boldsymbol{\chi}^T \bar{\boldsymbol{\chi}}^*$ in this example is the total income of the company. The products are produced by workers $\boldsymbol{\varepsilon} \in \mathbb{R}^m$. Due to the cooperation, we have $\boldsymbol{\varepsilon} = D\boldsymbol{\chi}$ and $D \in \mathbb{R}^{m\times n}$ is a matrix. Workers are paid by salary $\boldsymbol{\sigma} = \partial W(\boldsymbol{\varepsilon})$, therefore, the internal energy $W(\boldsymbol{\varepsilon})$ in this example is the cost, which should be an objective function. Thus, $\Pi(\boldsymbol{\chi}) = W(D\boldsymbol{\chi}) - F(\boldsymbol{\chi})$ is the total cost or target *and the minimization problem* min $\Pi(\boldsymbol{\chi})$ *leads to the equilibrium equation*

$$D^T\partial_{\boldsymbol{\varepsilon}} W(D\boldsymbol{\chi}) = \bar{\boldsymbol{\chi}}^*,$$

which is an algebraic equation in $\mathbb{R}^n$. The weak form of this equilibrium equation is $\langle \boldsymbol{\chi}, D^T\boldsymbol{\sigma} \rangle = \langle D\boldsymbol{\chi}; \boldsymbol{\sigma} \rangle = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle$, which is the well-known D'Alembert's principle *or the* principle of virtual work *in Lagrangian mechanics. The cost function $W(\boldsymbol{\varepsilon})$*

could be convex for a very small company, but usually nonconvex for big companies to allow some people having the same salaries.

Example 2 (Lagrange Mechanics). In analytical mechanics, the configuration $\boldsymbol{\chi} \in \mathscr{X}_a \subset \mathscr{C}^1[I;\mathbb{R}^n]$ *is a continuous vector-valued function of time* $t \in I \subset \mathbb{R}$. *Its components* $\{\chi_i\}$ $(i = 1,\dots,n)$ *are known as the* Lagrangian coordinates.[4] *Its dual variable* $\bar{\boldsymbol{\chi}}^*$ *is the action vector function in* $\mathbb{R}^n$, *say* $\mathbf{f}(t)$. *The external energy* $F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle = \int_I \boldsymbol{\chi}(t) \cdot \mathbf{f}(t)\, dt$. *While the internal energy* $W(D\boldsymbol{\chi})$ *is the so-called action:*

$$W(D\boldsymbol{\chi}) = \int_I L(\boldsymbol{\chi}, \dot{\boldsymbol{\chi}})\, dt, \quad L = T(\dot{\boldsymbol{\chi}}) - V(\boldsymbol{\chi})$$

where T *is the kinetic energy density,* V *is the potential density, and* $L = T - V$ *is the standard* Lagrangian density. *In this case, the linear operator* $D\boldsymbol{\chi} = \{\partial_t, 1\}\boldsymbol{\chi} = \{\dot{\boldsymbol{\chi}},\ \boldsymbol{\chi}\}$ *is a vector-valued mapping. The kinetic energy* T *must be an objective function of the velocity* $\mathbf{v}_k = \dot{\mathbf{x}}_k(\boldsymbol{\chi})$ *(or isotropic since* $\mathbf{v}_k$ *is a vector) of each particle* $\mathbf{x}_k = \mathbf{x}_k(\boldsymbol{\chi}) \in \mathbb{R}^3 \ \ \forall k \in I_m$, *while the potential density* V *depends on each problem. Together,* $\Pi(\boldsymbol{\chi}) = W(D\boldsymbol{\chi}) - F(\boldsymbol{\chi})$ *is called* total action. *Its stationary condition leads to the* Euler–Lagrange equation*:*

$$D^*\partial W(D\boldsymbol{\chi}) = -\partial_t \frac{\partial T(\dot{\boldsymbol{\chi}})}{\partial \dot{\boldsymbol{\chi}}} - \nabla V(\boldsymbol{\chi}) = \mathbf{f}. \tag{7}$$

For Newton mechanics, $T(\mathbf{v}) = \frac{1}{2}\sum_{k\in I_m} m_k \|\mathbf{v}_k\|^2$ *is quadratic, where* $\|\mathbf{v}_k\|$ *represents the Euclidean norm (speed) of the* k*-th particle in* $\mathbb{R}^3$. *For Einstein's special relativity theory,* $T(\mathbf{v}) = -m_0 c\sqrt{c^2 - \|\mathbf{v}\|^2}$ *is convex (see Sect. 2.1.2, [49]), where* $m_0 > 0$ *is the mass of a particle at rest,* c *is the speed of light. Therefore, the total action* $\Pi(\boldsymbol{\chi})$ *is convex only if* $V(\boldsymbol{\chi})$ *is linear. In this case, the solution of the Euler–Lagrange equation (7) minimizes the total action. The total action is nonconvex as long as the potential density* $V(\boldsymbol{\chi})$ *is nonlinear. In this case, the system may have periodic solution if* $V(\boldsymbol{\chi})$ *is convex and the well-known* least action principle is indeed a misnomer *(see Chap. 2, [49]). The system may have chaotic solution if the potential density* $V(\boldsymbol{\chi})$ *is nonconvex [50, 57]. Unfortunately, these important facts are not well-realized in both classical mechanics and modern nonlinear dynamical systems. The recent review article [67] presents a unified understanding bifurcation, chaos, and NP-hard problems in complex systems.*

In nonlinear analysis, the linear operator D is a partial differential operator, say $D = \{\partial_t,\ \partial_x\}$, and the abstract equilibrium equation (6) is a nonlinear partial differential equation. For convex $W(\boldsymbol{\varepsilon})$, the solution of this equilibrium equation is also a solution to the minimization problem (3). However, for nonconvex $W(\boldsymbol{\varepsilon})$, the solution of (6) is only a stationary point of $\Pi(\boldsymbol{\chi})$. In order to study stability and regularity of the local solutions in nonconvex problems, many generalized definitions,

[4] It is an unfortunate truth that many people do not know the relation between the Lagrangian space $\mathbb{R}^n$ they work in and the Minkowski (physical) space $\mathbb{R}^3 \times \mathbb{R}$ they live in.

such as quasi-, poly- and rank-one convexities have been introduced and subjected to extensively study for more than fifty years [7]. But all these generalized convexities provide only local extremality conditions, which lead to many "outstanding open problems" in nonlinear analysis [7]. However, by the canonical duality-triality theory, we can have clear understandings on these challenges.

2.2 Canonical Transformation and Classification of Nonlinearities

According to the canonical duality, the linear measure $\boldsymbol{\varepsilon} = D\boldsymbol{\chi}$ can't be used directly for studying constitutive law due to the objectivity. Also, the linear operator cannot change the nonconvexity of $W(D\boldsymbol{\chi})$. Indeed, it is well-known that the deformation gradient $\mathbf{F} = \nabla\boldsymbol{\chi}$ is not considered as a strain measure in nonlinear elasticity. The most commonly used strain measure is the right Cauchy–Green strain tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$, which is, clearly, an objective function since $\mathbf{C}(\mathbf{F}) = \mathbf{C}(\mathbf{QF})$. According to P. Ciarlet (Theorem 4.2-1, [16]), the stored energy $W(\mathbf{F})$ of a hyperelastic material is objective if and only if there exists a function $\tilde{W}$ such that $W(\mathbf{F}) = \tilde{W}(\mathbf{C})$. Based on this fact in continuum physics, the canonical transformation is naturally introduced.

Definition 2 (Canonical Function and Canonical Transformation).

A real-valued function $\Phi : \mathscr{E}_a \to \mathbb{R}$ is called canonical if the duality mapping $\partial\Phi : \mathscr{E}_a \to \mathscr{E}_a^*$ is one-to-one and onto.

For a given nonconvex function $W : \mathscr{W}_a \to \mathbb{R}$, if there exists a geometrically admissible mapping $\Lambda : \mathscr{W}_a \to \mathscr{E}_a$ and a canonical function $\Phi : \mathscr{E}_a \to \mathbb{R}$ such that

$$W(\boldsymbol{\varepsilon}) = \Phi(\Lambda(\boldsymbol{\varepsilon})), \tag{8}$$

then, the transformation (8) is called the canonical transformation and $\boldsymbol{\xi} = \Lambda(\boldsymbol{\varepsilon})$ is called the canonical measure.

By this definition, the one-to-one duality relation $\boldsymbol{\xi}^* = \partial\Phi(\boldsymbol{\xi}) : \mathscr{E}_a \to \mathscr{E}_a^*$ implies that the canonical function $\Phi(\boldsymbol{\xi})$ is differentiable and its conjugate function $\Phi^* : \mathscr{E}_a^* \to \mathbb{R}$ can be uniquely defined by the Legendre transformation [49]

$$\Phi^*(\boldsymbol{\xi}^*) = \{\langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle - \Phi(\boldsymbol{\xi}) | \ \boldsymbol{\xi}^* = \partial\Phi(\boldsymbol{\xi})\}, \tag{9}$$

where $\langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle$ represents the bilinear form on $\mathscr{E}$ and its dual space $\mathscr{E}^*$. In this case, $\Phi : \mathscr{E}_a \to \mathbb{R}$ is a canonical function if and only if the following canonical duality relations hold on $\mathscr{E}_a \times \mathscr{E}_a^*$:

$$\boldsymbol{\xi}^* = \partial\Phi(\boldsymbol{\xi}) \quad \Leftrightarrow \quad \boldsymbol{\xi} = \partial\Phi^*(\boldsymbol{\xi}^*) \quad \Leftrightarrow \quad \Phi(\boldsymbol{\xi}) + \Phi^*(\boldsymbol{\xi}^*) = \langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle. \tag{10}$$

A canonical function $\Phi(\boldsymbol{\xi})$ can also be nonsmooth but should be convex such that its conjugate can be well-defined by Fenchel transformation

$$\Phi^{\sharp}(\boldsymbol{\xi}^*) = \sup\{\langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle - \Phi(\boldsymbol{\xi}) | \ \boldsymbol{\xi} \in \mathscr{E}_a\}. \tag{11}$$

In this case, $\partial \Phi(\boldsymbol{\xi}) \subset \mathscr{E}_a^*$ is understood as the subdifferential and the canonical duality relations (10) should be written in the generalized form

$$\boldsymbol{\xi}^* \in \partial \Phi(\boldsymbol{\xi}) \quad \Leftrightarrow \quad \boldsymbol{\xi} \in \partial \Phi^{\sharp}(\boldsymbol{\xi}^*) \quad \Leftrightarrow \quad \Phi(\boldsymbol{\xi}) + \Phi^{\sharp}(\boldsymbol{\xi}^*) = \langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle. \tag{12}$$

This generalized canonical duality plays an important role in unified understanding Lagrangian duality and KKT theory for constrained optimization problems (see [85, 101] and Sect. 5.4).

In analysis, nonlinear PDEs are classified as semilinear, quasi-linear, and fully nonlinear three categories based on the degree of the nonlinearity [25]. A *semilinear PDE* is a differential equation that is nonlinear in the unknown function but linear in all its partial derivatives. A *quasi-linear PDE* is one that is nonlinear in (at least) one of the lower order derivatives but linear in the highest order derivative(s) of the unknown function. *Fully nonlinear PDEs* are referred to as the class of nonlinear PDEs which are nonlinear in the highest order derivatives of the unknown function. However, this classification is not essential as we know that the main difficulty is nonconvexity, instead of nonlinearity since these nonlinear PDEs could be related to certain convex variational problems, which can be solved easily by numerical methods.

The concepts of geometrical and physical nonlinearities are well-known in continuum physics [30–36, 143], but not in abstract analysis and optimization. This leads to many confusions. Based on the canonical transformation, we can have the following classification.

Definition 3 (Geometrical, Physical, and Complete Nonlinearities).

The general problem (3) is called geometrically nonlinear (resp. linear) if the geometrical operator $\Lambda(\boldsymbol{\varepsilon})$ is nonlinear (resp. linear);

The problem (3) is called physically nonlinear (resp. linear) if the constitutive relation $\boldsymbol{\xi}^* = \partial \Phi(\boldsymbol{\xi})$ is nonlinear (resp. linear);

The general problem (3) is called completely nonlinear if it is both geometrically and physically nonlinear.

According to this clarification, the minimization problem (3) is geometrically linear as long as the stored energy $W(\boldsymbol{\varepsilon})$ is convex. In this case, $\Lambda(D\boldsymbol{\chi}) = D\boldsymbol{\chi}$ and $\Phi(\Lambda(\boldsymbol{\varepsilon})) = W(\boldsymbol{\varepsilon})$. Thus, a physically nonlinear but geometrically linear problem could be equivalent to a fully nonlinear PDE, which can be solved easily by well-developed convex optimization techniques. Therefore, the main difficulty in complex systems is the geometrical nonlinearity. This is the reason why only this nonlinearity was emphasized in the title of Gao–Strang's paper [77]. The complete nonlinearity is also called fully nonlinearity in engineering mechanics. Hope this new classification will clear out this confusion. By the canonical transformation, the completely nonlinear minimization problem (3) can be equivalently written in the following canonical form

$$(\mathscr{P}): \quad \min\{\Pi(\boldsymbol{\chi}) = \Phi(\Lambda(D\boldsymbol{\chi})) - F(\boldsymbol{\chi}) |\ \boldsymbol{\chi} \in \mathscr{X}_c\}. \tag{13}$$

In order to solving this nonconvex problem, we need to find its canonical dual form.

2.3 Complementary-Dual Principle

For geometrically linear problems, the stored energy $W(\boldsymbol{\varepsilon})$ is convex and the complementary energy $W^*(\boldsymbol{\sigma})$ can be uniquely defined on $\mathscr{W}_a^*$ by Legendre transformation. Therefore, by using equality $W(\boldsymbol{\varepsilon}) = \langle \boldsymbol{\varepsilon}; \boldsymbol{\sigma}\rangle - W^*(\boldsymbol{\sigma})$, the total potential $\Pi(\boldsymbol{\chi})$ can be equivalently written in the classical Lagrangian form $L : \mathscr{X}_a \times \mathscr{W}_a^* \to \mathbb{R}$

$$L(\boldsymbol{\chi}, \boldsymbol{\sigma}) = \langle D\boldsymbol{\chi}; \boldsymbol{\sigma}\rangle - W^*(\boldsymbol{\sigma}) - F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, D^*\boldsymbol{\sigma} - \bar{\boldsymbol{\chi}}^*\rangle - W^*(\boldsymbol{\sigma}), \tag{14}$$

where, $\boldsymbol{\chi}$ can be viewed as a Lagrange multiplier for the equilibrium equation $D^*\boldsymbol{\sigma} = \bar{\boldsymbol{\chi}}^*$. In linear elasticity, $L(\boldsymbol{\chi}, \boldsymbol{\sigma})$ is the well-known Hellinger–Reissner complementary energy. Let $\mathscr{S}_c = \{\boldsymbol{\sigma} \in \mathscr{W}_a^* |\ D^*\boldsymbol{\sigma} = \bar{\boldsymbol{\chi}}^*\}$ be the so-called *statically admissible space*. Then the Lagrangian dual of the general problem (3) is given by

$$\max\{\Pi^*(\boldsymbol{\sigma}) = -W^*(\boldsymbol{\sigma}) |\ \boldsymbol{\sigma} \in \mathscr{S}_c\}, \tag{15}$$

and the following Lagrangian min–max duality is well-known:

$$\min_{\boldsymbol{\chi}\in\mathscr{X}_c} \Pi(\boldsymbol{\chi}) = \min_{\boldsymbol{\chi}\in\mathscr{X}_a} \max_{\boldsymbol{\sigma}\in\mathscr{W}_a^*} L(\boldsymbol{\chi}, \boldsymbol{\sigma}) = \max_{\boldsymbol{\sigma}\in\mathscr{W}_a^*} \min_{\boldsymbol{\chi}\in\mathscr{X}_a} L(\boldsymbol{\chi}, \boldsymbol{\sigma}) = \max_{\boldsymbol{\sigma}\in\mathscr{S}_c} \Pi^*(\boldsymbol{\sigma}). \tag{16}$$

In continuum mechanics, this one-to-one duality is called *complementary-dual variational principle* [119]. In finite elasticity, the Lagrangian dual is also known as the *Levison–Zubov principle*. However, this principle holds only for convex problems. If the stored energy $W(\boldsymbol{\varepsilon})$ is nonconvex, its complementary energy can't be determined uniquely by the Legendre transformation. Although its Fenchel conjugate $W^\sharp : \mathscr{W}_a^* \to \mathbb{R} \cup \{+\infty\}$ can be uniquely defined, the Fenchel–Moreau dual problem

$$\max\{\Pi^\sharp(\boldsymbol{\sigma}) = -W^\sharp(\boldsymbol{\sigma}) |\ \ \boldsymbol{\sigma} \in \mathscr{S}_c\} \tag{17}$$

is not considered as a complementary-dual problem due to Fenchel–Young inequality:

$$\min\{\Pi(\boldsymbol{\chi}) |\ \boldsymbol{\chi} \in \mathscr{X}_c\} \geq \max\{\Pi^\sharp(\boldsymbol{\sigma}) |\ \ \boldsymbol{\sigma} \in \mathscr{S}_c\}, \tag{18}$$

and $\theta = \min \Pi(\boldsymbol{\chi}) - \max \Pi^\sharp(\boldsymbol{\sigma}) \neq 0$ is the so-called *duality gap*. This duality gap is intrinsic to all type of Lagrangian duality problems since the nonconvexity of

$W(D\boldsymbol{\chi})$ cannot be changed by any linear operator. It turns out that the existence of a pure stress based complementary-dual principle has been a well-known debet in finite elasticity for more than forty years [107].

Remark 1 (Lagrange Multiplier Law). Strictly speaking, the Lagrange multiplier method can be used mainly for equilibrium constraint in $\mathscr{S}_c$ and the Lagrange multiplier must be the solution to the primal problem (see Sect. 1.5.2 [49]). The equilibrium equation $D^*\boldsymbol{\sigma} = \bar{\boldsymbol{\chi}}^*$ must be an invariant under certain coordinates transformation, say the law of angular momentum conservation, which is guaranteed by the objectivity of the stored energy $W(D\boldsymbol{\chi})$ in continuum mechanics (see Definition 6.1.2, [49]), or by the isotropy of the kinetic energy $T(\dot{\boldsymbol{\chi}})$ in Lagrangian mechanics [98]. Specifically, the equilibrium equation for Newton's mechanics is an invariant under the Galilean transformation; while for Einstein's special relativity theory, the equilibrium equation $D^*\boldsymbol{\sigma} = \bar{\boldsymbol{\chi}}^*$ is an invariant under the Lorentz transformation. For linear equilibrium equation, the quadratic $W(\boldsymbol{\varepsilon})$ is naturally an objective function for convex systems. Unfortunately, since the concept of the objectivity is misused in mathematical optimization, the Lagrange multiplier method has been mistakenly used for solving general nonconvex problems, which produces many different duality gaps.

In order to recover the duality gap in nonconvex problems, we use the canonical transformation $W(D\boldsymbol{\chi}) = \Phi(\Lambda(D\chi))$ such that the nonconvex total potential $\Pi(\boldsymbol{\chi})$ can be reformulated as the total complementary energy $\Xi : \mathscr{X}_a \times \mathscr{E}_a^* \to \mathbb{R}$

$$\Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*) = \langle \Lambda(D\boldsymbol{\chi}); \boldsymbol{\xi}^* \rangle - \Phi^*(\boldsymbol{\xi}^*) - F(\boldsymbol{\chi}), \tag{19}$$

which was first introduced by Gao and Strang in 1989 [77]. The stationary condition $\delta\Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*) = 0$ leads to the following canonical equations:

$$\Lambda(D\boldsymbol{\chi}) = \partial\Phi^*(\boldsymbol{\xi}^*), \tag{20}$$

$$D^*\Lambda_t(D\boldsymbol{\chi})\boldsymbol{\xi}^* = \partial F(\boldsymbol{\chi}), \tag{21}$$

where $\Lambda_t(\boldsymbol{\varepsilon}) = \partial\Lambda(\boldsymbol{\varepsilon})$ is a generalized Gâteaux derivative of $\Lambda(\boldsymbol{\varepsilon})$. By the canonical duality, (20) is equivalent to $\boldsymbol{\xi}^* = \partial_{\xi}\Phi(\Lambda(D\boldsymbol{\chi}))$. Therefore, the canonical equilibrium equation (21) is the general equilibrium equation (6).

By using the Gao–Strang complementary function, the canonical dual of $\Pi(\boldsymbol{\chi})$ can be obtained as

$$\Pi^d(\boldsymbol{\xi}^*) = \text{sta}\{\Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*) |\ \boldsymbol{\chi} \in \mathscr{X}_a\} = F^{\Lambda}(\boldsymbol{\xi}^*) - \Phi^*(\boldsymbol{\xi}^*), \tag{22}$$

where $F^{\Lambda}(\boldsymbol{\xi}^*)$ is the Λ-transformation defined by [51]

$$F^{\Lambda}(\boldsymbol{\xi}^*) = \text{sta}\{\langle \Lambda(D\boldsymbol{\chi}); \boldsymbol{\xi}^* \rangle - F(\boldsymbol{\chi}) |\ \boldsymbol{\chi} \in \mathscr{X}_a\}. \tag{23}$$

Clearly, the stationary condition in this Λ-transformation is the canonical equilibrium equation (21). Let $\mathscr{S}_c \subset \mathscr{E}_a^*$ be a feasible set, on which $F^\Lambda(\boldsymbol{\xi}^*)$ is well-defined. Then we have the following result.

Theorem 1 (**Complementary-Dual Principle** [45, 47]) *If* $(\bar{\boldsymbol{\chi}}, \bar{\boldsymbol{\xi}}^*) \in \mathscr{X}_a \times \mathscr{E}_a^*$ *is a stationary point of* $\Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*)$*, then* $\bar{\boldsymbol{\chi}}$ *is a stationary point of* $\Pi(\boldsymbol{\chi})$ *on* $\mathscr{X}_c$*, while* $\bar{\boldsymbol{\xi}}^*$ *is a stationary point of* $\Pi^d(\boldsymbol{\xi}^*)$ *on* $\mathscr{S}_c$*, and*

$$\Pi(\bar{\boldsymbol{\chi}}) = \Xi(\bar{\boldsymbol{\chi}}, \bar{\boldsymbol{\xi}}^*) = \Pi^d(\bar{\boldsymbol{\xi}}^*). \tag{24}$$

This theorem shows that there is no duality gap between $\Pi(\boldsymbol{\chi})$ and $\Pi^d(\boldsymbol{\xi}^*)$. In many real-world applications, the geometrical operator $\Lambda(\boldsymbol{\varepsilon})$ is usually quadratic such that the total complementary function $\Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*)$ can be written as

$$\Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*) = \frac{1}{2}\langle \boldsymbol{\chi}, \mathbf{G}(\boldsymbol{\xi}^*)\boldsymbol{\chi}\rangle - \Phi^*(\boldsymbol{\xi}^*) - \langle \boldsymbol{\chi}, \mathbf{F}(\boldsymbol{\xi}^*)\rangle \tag{25}$$

where $\mathbf{G}(\boldsymbol{\xi}^*) = \nabla^2_{\boldsymbol{\chi}} \Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*)$ and $\mathbf{F}(\boldsymbol{\xi}^*)$ depends on the linear terms in $\Lambda(D\boldsymbol{\chi})$ and the input $\bar{\boldsymbol{\chi}}^*$. The first term in $\Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*)$

$$G_{ap}(\boldsymbol{\chi}, \boldsymbol{\xi}^*) = \frac{1}{2}\langle \boldsymbol{\chi}, \mathbf{G}(\boldsymbol{\xi}^*)\boldsymbol{\chi}\rangle \tag{26}$$

is the so-called *complementary gap function* introduced by Gao and Strang in [77]. In this case, the canonical equilibrium equation $\nabla_{\boldsymbol{\chi}} \Xi(\boldsymbol{\chi}, \boldsymbol{\xi}^*) = \mathbf{G}(\boldsymbol{\xi}^*)\boldsymbol{\chi} - \mathbf{F}(\boldsymbol{\xi}^*) = 0$ is linear in $\boldsymbol{\chi}$ and the canonical dual Π^d can be explicitly formulated as

$$\Pi^d(\boldsymbol{\xi}^*) = -G^*_{ap}(\boldsymbol{\xi}^*) - \Phi^*(\boldsymbol{\xi}^*), \tag{27}$$

where $G^*_{ap}(\boldsymbol{\xi}^*) = \frac{1}{2}\langle \mathbf{G}^{-1}(\boldsymbol{\xi}^*)\mathbf{F}(\boldsymbol{\xi}^*), \mathbf{F}(\boldsymbol{\xi}^*)\rangle$ is called *pure complementary gap function*. Comparing this canonical dual with the Lagrangian dual $\Pi^*(\boldsymbol{\sigma}) = -W^*(\boldsymbol{\sigma})$ in (15) we can find that in addition to replace W^* by the canonical dual Φ^*, the first term in Π^d is identical to the Gao–Strang complementary gap function, which recovers the duality gap in Lagrangian duality theory and plays an important role in triality theory.

Theorem 2 (**Analytical Solution Form**) *If* $\bar{\boldsymbol{\xi}}^* \in \mathscr{S}_c$ *is a stationary point of* $\Pi^d(\boldsymbol{\xi}^*)$*, then*

$$\bar{\boldsymbol{\chi}} = \mathbf{G}^{-1}(\bar{\boldsymbol{\xi}}^*)\mathbf{F}(\bar{\boldsymbol{\xi}}^*) \tag{28}$$

is a stationary point of $\Pi(\boldsymbol{\chi})$ *on* $\mathscr{X}_c$ *and* $\Pi(\bar{\boldsymbol{\chi}}) = \Pi^d(\bar{\boldsymbol{\xi}}^*)$.

This theorem shows that the primal solution is analytically depends on its canonical dual solution. Clearly, the canonical dual of a nonconvex primal problem is also nonconvex and may have multiple stationary points. By the canonical duality, each of these stationary solutions is corresponding to a primal solution via (28). Their extremality is governed by Gao and Strang's complementary gap function.

2.4 Triality Theory

In order to identify extremality of these stationary solutions, we need to assume that the canonical function $\Phi : \mathscr{E}_a \to \mathbb{R}$ is convex and let

$$\mathscr{S}_c^+ = \{\boldsymbol{\xi}^* \in \mathscr{S}_c | \ \mathbf{G}(\boldsymbol{\xi}^*) \succ 0\}, \quad \mathscr{S}_c^- = \{\boldsymbol{\xi}^* \in \mathscr{S}_c | \ \mathbf{G}(\boldsymbol{\xi}^*) \prec 0\}. \tag{29}$$

Clearly, for any given $\boldsymbol{\chi} \in \mathscr{X}_a$ and $\boldsymbol{\chi} \neq 0$, we have

$$G_{ap}(\boldsymbol{\chi}, \boldsymbol{\xi}^*) > 0 \ \Leftrightarrow \ \boldsymbol{\xi}^* \in \mathscr{S}_c^+, \ \ G_{ap}(\boldsymbol{\chi}, \boldsymbol{\xi}^*) < 0 \ \Leftrightarrow \ \boldsymbol{\xi}^* \in \mathscr{S}_c^-.$$

Theorem 3 (**Triality Theorem**) *Suppose $\bar{\boldsymbol{\xi}}^*$ is a stationary point of $\Pi^d(\boldsymbol{\xi}^*)$ and $\bar{\boldsymbol{\chi}} = \mathbf{G}^{-1}(\bar{\boldsymbol{\xi}}^*)\bar{\boldsymbol{\xi}}^*$. If $\bar{\boldsymbol{\xi}}^* \in \mathscr{S}_c^+$, we have*

$$\Pi(\bar{\boldsymbol{\chi}}) = \min_{\boldsymbol{\chi} \in \mathscr{X}_c} \Pi(\boldsymbol{\chi}) \ \Leftrightarrow \ \max_{\boldsymbol{\xi}^* \in \mathscr{S}_c^+} \Pi^d(\boldsymbol{\xi}^*) = \Pi^d(\bar{\boldsymbol{\xi}}^*); \tag{30}$$

If $\bar{\boldsymbol{\xi}}^* \in \mathscr{S}_c^-$, then on a neighborhood[5] *$\mathscr{X}_o \times \mathscr{S}_o \subset \mathscr{X}_c \times \mathscr{S}_c^-$ of $(\bar{\boldsymbol{\chi}}, \bar{\boldsymbol{\xi}}^*)$, we have either*

$$\Pi(\bar{\boldsymbol{\chi}}) = \max_{\boldsymbol{\chi} \in \mathscr{X}_o} \Pi(\boldsymbol{\chi}) \ \Leftrightarrow \ \max_{\boldsymbol{\xi}^* \in \mathscr{S}_o} \Pi^d(\boldsymbol{\xi}^*) = \Pi^d(\bar{\boldsymbol{\xi}}^*), \tag{31}$$

or (only if $\dim \bar{\boldsymbol{\chi}} = \dim \bar{\boldsymbol{\xi}}^*$*)*

$$\Pi(\bar{\boldsymbol{\chi}}) = \min_{\boldsymbol{\chi} \in \mathscr{X}_o} \Pi(\boldsymbol{\chi}) \ \Leftrightarrow \ \min_{\boldsymbol{\xi}^* \in \mathscr{S}_o} \Pi^d(\boldsymbol{\xi}^*) = \Pi^d(\bar{\boldsymbol{\xi}}^*). \tag{32}$$

The first statement (30) is called *canonical min–max duality*. Its weak form was discovered by Gao and Strang in 1989 [77]. This duality can be used to identify global minimizer of the nonconvex problem (3). According this statement, the nonconvex problem (3) is equivalent to the following canonical dual problem, denoted by $(\mathscr{P}^d)$:

$$(\mathscr{P}^d): \quad \max\{\Pi^d(\boldsymbol{\xi}^*) | \ \boldsymbol{\xi}^* \in \mathscr{S}_c^+\}. \tag{33}$$

This is a concave maximization problem which can be solved easily by well-developed convex analysis and optimization techniques. The second statement (31) is the *canonical double-max duality* and (32) is the *canonical double-min duality*. These two statements can be used to identify the biggest local maximizer and local minimizer of the primal problem, respectively.

The triality theory was first discovered by Gao 1996 in post-buckling analysis of a large deformed beam [42, 52]. The generalization to global optimization was made in 2000 [51]. It was realized in 2003 that the double-min duality (32) holds

---

[5]The neighborhood $\mathscr{X}_o$ of $\bar{\boldsymbol{\chi}}$ means that on which, $\bar{\boldsymbol{\chi}}$ is the only stationary point.

under certain additional condition [57, 58]. Recently, it is proved that this additional condition is simply dim $\bar{\boldsymbol{\chi}} = \dim \bar{\boldsymbol{\xi}}^*$ to have the strong canonical double-min duality (32), otherwise, this double-min duality holds weakly in subspaces of $\mathscr{X}_o \times \mathscr{S}_o$ [79, 80, 112, 113].

Example 3 *To explain the theory, let us consider a very simple nonconvex optimization in* $\mathbb{R}^n$*:*

$$\min \left\{ \Pi(\mathbf{x}) = \frac{1}{2}\alpha \left(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda \right)^2 - \mathbf{x}^T \mathbf{f} \quad \forall \mathbf{x} \in \mathbb{R}^n \right\}, \tag{34}$$

where $\alpha, \lambda > 0$ *are given parameters. The criticality condition* $\nabla \Pi(\mathbf{x}) = 0$ *leads to a nonlinear algebraic equation system in* $\mathbb{R}^n$

$$\alpha(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda)\mathbf{x} = \mathbf{f}. \tag{35}$$

Clearly, to solve this nonlinear algebraic equation directly is difficult. Also traditional convex optimization theory cannot be used to identify global minimizer. However, by the canonical dual transformation, this problem can be solved completely and easily. To do so, we let $\xi = \Lambda(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2 \in \mathbb{R}$*, which is an objective measure. Then, the nonconvex function* $W(\mathbf{x}) = \frac{1}{2}\alpha(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda)^2$ *can be written in canonical form* $\Phi(\xi) = \frac{1}{2}\alpha(\xi - \lambda)^2$*. Its Legendre conjugate is given by* $\Phi^*(\varsigma) = \frac{1}{2}\alpha^{-1}\varsigma^2 + \lambda\varsigma$*, which is strictly convex. Thus, the total complementary function for this nonconvex optimization problem is*

$$\Xi(\mathbf{x}, \varsigma) = \frac{1}{2}\|\mathbf{x}\|^2\varsigma - \frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma - \mathbf{x}^T\mathbf{f}. \tag{36}$$

For a fixed $\varsigma \in \mathbb{R}$*, the criticality condition* $\nabla_{\mathbf{x}}\Xi(\mathbf{x}) = 0$ *leads to*

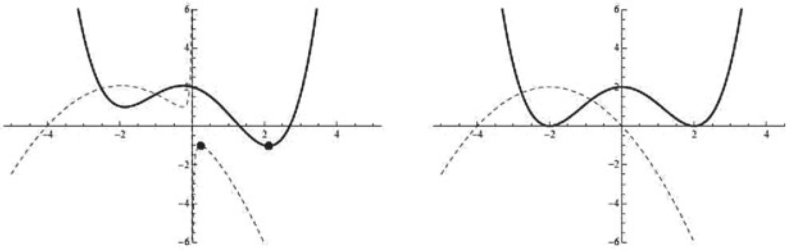
$$\varsigma\mathbf{x} - \mathbf{f} = 0. \tag{37}$$

For each $\varsigma \neq 0$*, the Eq. (37) gives* $\mathbf{x} = \mathbf{f}/\varsigma$ *in vector form. Substituting this into the total complementary function* Ξ*, the canonical dual function can be easily obtained as*

$$\Pi^d(\varsigma) = \{\Xi(\mathbf{x}, \varsigma) | \nabla_{\mathbf{x}}\Xi(\mathbf{x}, \varsigma) = 0\} = -\frac{\mathbf{f}^T\mathbf{f}}{2\varsigma} - \frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma, \quad \forall \varsigma \neq 0. \tag{38}$$

The critical point of this canonical function is obtained by solving the following dual algebraic equation

$$2(\alpha^{-1}\varsigma + \lambda)\varsigma^2 = \mathbf{f}^T\mathbf{f}. \tag{39}$$



(a) $f = 0.5$ (b) $f = 0.$

Fig. 1 Graphs of $\Pi(\mathbf{x})$ (*solid*) and $\Pi^d(\varsigma)$ (*dashed*)

For any given parameters α, λ *and the vector* $\mathbf{f} \in \mathbb{R}^n$*, this cubic algebraic equation has at most three real roots satisfying* $\varsigma_1 \geq 0 \geq \varsigma_2 \geq \varsigma_3$*, and each of these roots leads to a critical point of the nonconvex function* $P(\mathbf{x})$*, i.e.,* $\mathbf{x}_i = \mathbf{f}/\varsigma_i$, $i = 1, 2, 3$. *By the fact that* $\varsigma_1 \in \mathscr{S}_c^+ = \{\varsigma \in \mathbb{R} \mid \varsigma > 0\}$, $\varsigma_{2,3} \in \mathscr{S}_c^- = \{\varsigma \in \mathbb{R} \mid \varsigma < 0\}$*, then Theorem 3 tells us that* $\mathbf{x}_1$ *is a global minimizer of* $\Pi(\mathbf{x})$, $\mathbf{x}_3$ *is a local maximizer of* $\Pi(\mathbf{x})$*, while* $\mathbf{x}_2$ *is a local minimizer if* $n = 1$ *(see Fig. 1). If we choose* $n = 1, \quad \alpha = 1$, $\lambda = 2$*, and* $f = \frac{1}{2}$*, the primal function and canonical dual function are shown in Fig. 1a, where,* $x_1 = 2.11491$ *is global minimizer of* $\Pi(\mathbf{x})$, $\varsigma_1 = 0.236417$ *is global maximizer of* $\Pi^d(\boldsymbol{\varsigma})$*, and* $\Pi(x_1) = -1.02951 = \Pi^d(\varsigma_1)$ *(see the two black dots). Also it is easy to verify that* x_2 *is a local minimizer, while* x_3 *is a local maximizer.*

If we let $\mathbf{f} = 0$*, the graph of* $\Pi(\mathbf{x})$ *is symmetric (i.e., the so-called double-well potential or the Mexican hat for* $n = 2$ *[57]) with infinite number of global minimizers satisfying* $\|\mathbf{x}\|^2 = 2\lambda$*. In this case, the canonical dual* $\Pi^d(\varsigma) = -\frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma$ *is strictly concave with only one critical point (local maximizer)* $\varsigma_3 = -\alpha\lambda \in \mathscr{S}_c^-$ *(for* $\alpha, \lambda > 0$*). The corresponding solution* $\mathbf{x}_3 = \mathbf{f}/\varsigma_3 = 0$ *is a local maximizer. By the canonical dual equation (39) we have* $\varsigma_1 = \varsigma_2 = 0$ *located on the boundary of* $\mathscr{S}_c^+$*, which corresponding to the two global minimizers* $x_{1,2} = \pm\sqrt{2\lambda}$ *for* $n = 1$*, see Fig. 1b. This is similar to the post-buckling of large deformed beam. Due to symmetry* ($f = 0$)*, the nonconvex function* $\Pi(\mathbf{x})$ *has two possible buckled solutions* $\mathbf{x}_{1,2} = (\pm\sqrt{2\lambda}, 0)$ *with the axial load* $\lambda = \frac{1}{2}(b^2 - a^2)$*. While the local maximizer* $\mathbf{x}_3 = \{0, 0\}$ *is corresponding to the unbuckled state.*

This simple example shows a fundament issue in global optimization, i.e., the optimal solutions of a nonconvex problem depends sensitively on the linear term (input or perturbation) $\mathbf{f}$*. Geometrically speaking, the objective function* $W(D\mathbf{x})$ *in* $\Pi(\mathbf{x})$ *possesses certain symmetry. If there is no linear term (subjective function) in* $\Pi(\mathbf{x})$*, the nonconvex problem usually has more than one global minimizer due to the symmetry. Traditional direct approaches and the popular SDP method are usually failed to deal with this situation. By the canonical duality theory, we understand that in this case the canonical dual function* $\Pi^d(\varsigma)$ *has no critical point in* $\mathscr{S}_c^+$*. Therefore, the input* $\mathbf{f}$ *breaks the symmetry so that* $\Pi^d(\varsigma)$ *has a unique stationary point in* $\mathscr{S}_c^+$ *which can be obtained easily. This idea was originally from Gao's work (1996) on post-buckling analysis of large deformed beam [39], where the triality theorem was* *first proposed [42]. The potential energy of this beam model is a double-well function, similar to this example, without lateral force or imperfection, the beam could have two buckling states (corresponding to two minimizers) and one unbuckled state (local maximizer). Later on (2008) in the Gao and Ogden work on analytical solutions in phase transformation [71], they further discovered that the nonconvex system has no phase transition unless the force distribution $f(x)$ vanished at certain points. They also discovered that if force field $f(x)$ changes dramatically, all the Newton-type direct approaches failed even to find any local minimizer. The linear perturbation method has been used successfully for solving global optimization problems [14, 112, 128, 140].*

3 Applications for Modeling of Complex Systems

By the fact that the canonical duality is a fundamental law governing natural phenomena and the objectivity is a basic condition for mathematical models, the canonical duality-triality theory can be used for modeling real-world problems within a unified framework.

3.1 Mixed Integer Nonlinear Programming

The most general and challenging problem in global optimization could be the mixed integer nonlinear program (MINP), which is a minimization problem generally formulated as (see [90])

$$\min\{f(\mathbf{x},\mathbf{y})|\ \ g_i(\mathbf{x},\mathbf{y}) \leq 0 \ \ \forall i \in I_m, \ \ \mathbf{x} \in \mathbb{R}^n, \ \ \mathbf{y} \in \mathbb{Z}^p\} \tag{40}$$

where $\mathbb{Z}^p$ is an integer set, the "objective function" $f(\mathbf{x},\mathbf{y})$ and constraints $g_i(\mathbf{x},\mathbf{y})$ for $i \in I_m$ are arbitrary functions [11]. Certainly, this artificial model is virtually applicable to any problem in operations research, but it is impossible to develop a general theory and powerful algorithm without detailed information given on these functions. As we know that the objectivity is a fundamental concept in mathematical modeling. Unfortunately, this concept has been mistakenly used with other functions, such as target, cost, energy, and utility functions, etc.[6]

Based on the Gao–Strang model (3), we let $\boldsymbol{\chi} = (\mathbf{x},\mathbf{y})$, $\mathbf{D}\boldsymbol{\chi} = (\mathbf{D}_x\mathbf{x},\ \mathbf{D}_y\mathbf{y})$, and $\bar{\boldsymbol{\chi}}^* = (\mathbf{b},\mathbf{t})$. Then the general MINP problem (40) can be remodeled in the following form

$$\min\{\Pi(\mathbf{x},\mathbf{y}) = W(\mathbf{D}_x\mathbf{x},\mathbf{D}_y\mathbf{y}) - \mathbf{x}^T\mathbf{b} - \mathbf{y}^T\mathbf{t} \ |\ \ (\mathbf{x},\mathbf{y}) \in \mathscr{X}_c \times \mathscr{Y}_c, \ \ \mathbf{x} \in \mathbb{Z}^p\}, \tag{41}$$

[6]See http://en.wikipedia.org/wiki/Mathematical_optimization.

where the feasible sets are, correspondingly,

$$\mathscr{X}_c = \{\mathbf{x} \in \mathscr{X}_a \subset \mathbb{R}^n \mid \mathbf{D}_x\mathbf{x} \in \mathscr{U}_a\}, \quad \mathscr{Y}_c = \{\mathbf{y} \in \mathscr{Y}_a \subset \mathbb{R}^p \mid \mathbf{D}_y\mathbf{y} \in \mathscr{V}_a\}.$$

In $\mathscr{X}_a, \mathscr{Y}_a$, certain linear constraints are given, while in $\mathscr{U}_a$, $\mathscr{V}_a$, general nonlinear (constitutive) constraints are prescribed such that the nonconvex (objective) function $W : \mathscr{U}_a \times \mathscr{V}_a \to \mathbb{R}$ can be written in the canonical form $W(\mathbf{D}\boldsymbol{\chi}) = \Phi_\chi(\boldsymbol{\Lambda}(\boldsymbol{\chi}))$ for certain geometrical operator $\boldsymbol{\Lambda}(\boldsymbol{\chi})$. By the fact that any integer set $\mathbb{Z}^p$ is equivalent to a Boolean set [139], we simply let $\mathbb{Z}^p = \{0, 1\}^p$. This constitutive constraint can be relaxed by the canonical transformation [64, 75]

$$\boldsymbol{\varepsilon} = \boldsymbol{\Lambda}_x(\mathbf{x}) = \mathbf{x} \circ (\mathbf{x} - \mathbf{1}) = \{x_i^2 - x_i\}^p, \tag{42}$$

and the canonical function $\Phi_x(\boldsymbol{\varepsilon}) = \{0 \text{ if } \boldsymbol{\varepsilon} = \mathbf{0}, \ \infty \text{ otherwise}\}$. Therefore, the canonical form for the MINP problem is

$$\min\{\Pi(\mathbf{x}, \mathbf{y}) = \Phi_\chi(\boldsymbol{\Lambda}(\mathbf{x}, \mathbf{y})) + \Phi_x(\boldsymbol{\Lambda}_x(\mathbf{x})) - \mathbf{x}^T\mathbf{b} - \mathbf{y}^T\mathbf{t} \mid (\mathbf{x}, \mathbf{y}) \in \mathscr{X}_c \times \mathscr{Y}_c\}. \tag{43}$$

This canonical form covers many real-world applications, including the so-called fixed cost problem [86]. By the fact that the canonical function $\Phi_x(\boldsymbol{\varepsilon})$ is convex, semi-continuous, the canonical duality relation should be replaced by the subdifferential form $\boldsymbol{\sigma} \in \partial\Phi_x(\boldsymbol{\varepsilon})$, which is equivalent to

$$\boldsymbol{\sigma}^T\boldsymbol{\varepsilon} = 0 \quad \Leftrightarrow \quad \boldsymbol{\varepsilon} = \mathbf{0} \ \ \forall \boldsymbol{\sigma} \neq \mathbf{0}. \tag{44}$$

Thus, the integer constraint $\boldsymbol{\varepsilon} = \boldsymbol{\Lambda}_x(\mathbf{x}) = \{x_i(x_i - 1)\} = \mathbf{0}$ can be relaxed by the canonical dual constraint $\boldsymbol{\sigma} \neq \mathbf{0}$ in continuous space.

The canonical duality-triality theory has be used successfully for solving mixed integer programming problems [75, 86]. Particularly, for the quadratic integer programming problem (2), i.e.,

$$\min\left\{\Pi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{x}^T\mathbf{f} \mid \ \mathbf{x} \in \{0, 1\}^n\right\},$$

the canonical dual is [23, 64]

$$\max\left\{\Pi^d(\boldsymbol{\sigma}) = -\frac{1}{2}(\mathbf{f} + \boldsymbol{\sigma})^T\mathbf{G}^{-1}(\boldsymbol{\sigma})(\mathbf{f} + \boldsymbol{\sigma}) \mid \ \boldsymbol{\sigma} \in \mathscr{S}_c^+\right\} \tag{45}$$

where $\mathbf{G}(\boldsymbol{\sigma}) = \mathbf{Q} + 2\text{Diag}\,(\boldsymbol{\sigma})$. This is a concave maximization problem over the convex set in continuous space

$$\mathscr{S}_c^+ = \{\boldsymbol{\sigma} \in \mathbb{R}^n \mid \boldsymbol{\sigma} \neq \mathbf{0}, \ \ \mathbf{G}(\boldsymbol{\sigma}) \succ 0\},$$

which can be solved easily if $\mathscr{S}_c^+ \neq \emptyset$. Otherwise, the integer programming problem (2) could be NP-hard, which is a conjecture proposed in [64]. In this case, a second canonical dual problem has been proposed in [65, 88]

$$\min \left\{ \Pi^g(\boldsymbol{\sigma}) = -\frac{1}{2}\boldsymbol{\sigma}^T \mathbf{Q}^{-1}\boldsymbol{\sigma} - \sum_{i=1}^{n} |f_i - \sigma_i| \ \mid \boldsymbol{\sigma} \in \mathbb{R}^n \right\}. \tag{46}$$

This is a unconstrained nonsmooth minimization problem, which can be solved by some deterministic methods, such as DIRECT method [88].

Remark 2 (Subjective Function and NP-hard Problems). The subjective function $F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle$ *in the general model* $\Pi(\boldsymbol{\chi}) = W(D\boldsymbol{\chi}) - F(\boldsymbol{\chi})$ *plays an important role in global optimization problems. It was proved in [65] that for quadratic integer programming problem (2), if the source term* $\mathbf{f}$ *is bigger enough, the solution is simply* $\{x_i\} = \{0 \text{ if } f_i < 0,\ 1 \text{ if } f_i > 0\}$ *(Theorem 8, [65]). If a system has no input, by Newton's law, it has either trivial solution or infinite number solutions. For example, the well-known max-cut problem*

$$\max \left\{ \Pi(\mathbf{x}) = \frac{1}{4}\sum_{i,j=1}^{n+1} \omega_{ij}(1 - x_i x_j) \mid x_i \in \{-1, 1\} \forall i = 1, \ldots, n \right\} \tag{47}$$

is a special case of quadratic integer programming problem without the linear term. The integer condition is a physical (constitutive) constraint. Since there is no geometrical constraint, the graph is not fixed and any rigid motion is possible. Due to the symmetry $\omega_{ij} = \omega_{ji} > 0$*, the global solution is not unique. The canonical dual feasible space* $\mathscr{S}_c^+$ *in this example is empty and the problem is considered as NP-complete even if* $\omega_{ij} = 1$ *for all edges* $i, j = 1, \ldots, n$ *[94]. However, by adding a linear perturbation term, this problem can be solved efficiently by the canonical duality theory [140].*

3.2 Unified Model in Mathematical Physics

In analysis and mathematical physics, the configuration variable $\boldsymbol{\chi}(t, \mathbf{x})$ is a continuous field function $\boldsymbol{\chi} : [0, T] \times \Omega \subset \mathbb{R} \times \mathbb{R}^d \to \omega \subset \mathbb{R}^p$ (which is a hypersurface if $d + 1 = p$ in differential geometry). The linear operator $D = (\partial_t, \partial_x)$ is a partial differential operator and the stored energy $W(D\boldsymbol{\chi}) = T(\partial_t \boldsymbol{\chi}) - U(\partial_x \boldsymbol{\chi})$ with $T(\mathbf{v})$ as the kinetic energy and $U(\boldsymbol{\varepsilon})$ as deformation energy. Since $\mathbf{v} = \partial_t \boldsymbol{\chi}$ is a vector, the objectivity for kinetic energy $T(\mathbf{v})$ is also known as isotropy. But $\boldsymbol{\varepsilon} = \partial_x \boldsymbol{\chi}$ is a tensor, the deformation energy $U(\boldsymbol{\varepsilon})$ should be an objective function. In this case, the Gao and Strang model (3) is

$$\min \left\{ \Pi(\boldsymbol{\chi}) = T(\partial_t \boldsymbol{\chi}) - U(\partial_x \boldsymbol{\chi}) - \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle \mid \boldsymbol{\chi} \in \mathscr{X}_c \right\}. \tag{48}$$

The stationary condition $\delta\Pi(\boldsymbol{\chi}) = 0$ leads to a general nonlinear partial differential equation

$$\partial_t^* \partial_{\mathbf{v}} T(\partial_t \boldsymbol{\chi}) - \partial_x^* \partial_{\boldsymbol{\varepsilon}} U(\partial_x \boldsymbol{\chi}) = \bar{\boldsymbol{\chi}}^*. \tag{49}$$

The nonlinearity of this equation mainly depends on T and U. For Newtonian mechanics, $T(\mathbf{v})$ is quadratic. By the objectivity, the deformation energy $U(\boldsymbol{\varepsilon})$ can also be split into quadratic part and a nonlinear part such that $W(D\boldsymbol{\chi}) = \frac{1}{2}\langle \boldsymbol{\chi}, \mathbf{Q}\boldsymbol{\chi}\rangle + V(\mathbf{D}\boldsymbol{\chi})$, where $\mathbf{Q} : \mathscr{X}_c \to \mathscr{X}^*$ is a self-adjoint operator, $\mathbf{D}$ is a linear operator, and $V(\boldsymbol{\varepsilon})$ is a nonlinear objective functional. The most simple example is a fourth-order polynomial

$$V(\boldsymbol{\varepsilon}) = \int_\Omega \frac{1}{2}\left(\frac{1}{2}\|\boldsymbol{\varepsilon}\|^2 - \lambda\right)^2 \mathrm{d}\Omega, \tag{50}$$

which is nonconvex for $\lambda > 0$. This nonconvex functional appears extensively in mathematical physics. In fluid mechanics and thermodynamics, $V(\boldsymbol{\varepsilon})$ is the well-known *van de Waals double-well energy*. It is also known as the *sombrero potential* in cosmic string theory [18], or the *Mexican hat* in *Higgs mechanism* [20] and quantum field theory [93]. For this most simple nonconvex potential, the general model (3) can be written as

$$\mathbf{Q}\boldsymbol{\chi} + \mathbf{D}^*\left[\left(\frac{1}{2}\|\mathbf{D}\boldsymbol{\chi}\|^2 - \lambda\right)\mathbf{D}\boldsymbol{\chi}\right] = \bar{\boldsymbol{\chi}}^*. \tag{51}$$

This model covers many well-known equations.

(1) **Duffing equation** ($\mathbf{Q} = -\partial_t^2$ and $\mathbf{D} = \mathbf{I}$ is an identical operator):

$$\chi_{tt} + \left(\frac{1}{2}\chi^2 - \lambda\right)\chi = f(t) \tag{52}$$

(2) **Landau–Ginzburg equation** ($\mathbf{Q} = -\Delta,\ \mathbf{D} = \mathbf{I}$):

$$-\Delta\boldsymbol{\chi} + \left(\frac{1}{2}\|\boldsymbol{\chi}\|^2 - \lambda\right)\boldsymbol{\chi} = \mathbf{f} \tag{53}$$

(3) **Cahn–Hillar equation** ($\mathbf{Q} = -\Delta + \mathrm{curlcurl},\quad \mathbf{D} = \mathbf{I}$):

$$-\Delta\boldsymbol{\chi} + \mathrm{curlcurl}\boldsymbol{\chi} + \left(\frac{1}{2}\|\boldsymbol{\chi}\|^2 - \lambda\right)\boldsymbol{\chi} = \mathbf{f}. \tag{54}$$

(4) **Nonlinear Gorden equation** ($\mathbf{Q} = -\partial_{tt} + \Delta,\quad \mathbf{D} = \mathbf{I}$):

$$-\chi_{tt} + \Delta\boldsymbol{\chi} + \left(\frac{1}{2}\|\boldsymbol{\chi}\|^2 - \lambda\right)\boldsymbol{\chi} = \mathbf{f}. \tag{55}$$

(5) **Nonlinear Gao beam** ($\mathbf{Q} = \rho\partial_{tt} + K\partial_{xxxx}, \quad \mathbf{D} = \partial_x$):

$$\rho\chi_{tt} + K\chi_{xxxx} - \left[\left(\frac{1}{2}\chi_x^2 - \lambda\right)\chi_x\right]_x = f, \tag{56}$$

where $\lambda \in \mathbb{R}$ is an axial force and $f(t, x)$ is the lateral load.

According to the nonlinear classification discussed in Sect. 2.2, the general equation (51) is semilinear as long as $\mathbf{D} = \mathbf{I}$. While the nonlinear Gao beam is quasi-linear. However, if $\lambda > 0$, all these PDEs equations are geometrically nonlinear but physically linear since by the canonical transformation

$$\boldsymbol{\xi} = \Lambda(\boldsymbol{\varepsilon}) = \frac{1}{2}\|\boldsymbol{\varepsilon}\|^2 - \lambda, \quad V(\boldsymbol{\varepsilon}) = \Phi(\Lambda(\boldsymbol{\varepsilon})) = \int_\Omega \frac{1}{2}\boldsymbol{\xi}^2 \, d\Omega,$$

the canonical duality relation $\boldsymbol{\xi}^* = \partial\Phi(\boldsymbol{\xi}) = \boldsymbol{\xi}$ is linear.

The geometrical nonlinearity represents large deformation in continuum physics, or far from the equilibrium state in complex systems, which is necessary for nonconvexity but not sufficient. The nonconvexity of a geometrically nonlinear problem depends on external force and internal parameters. For example, the total potential of the nonlinear Gao beam is nonconvex only if the compressive load $\lambda > \lambda_c$, the Euler buckling load, i.e., the first eigenvalue of $K\chi_{xxxx}$ [39, 48]. In this case, the two minimizers represent the two buckled states, while the local maximizer represents the unbuckled (unstable) state. For dynamical loading, these two local minimizers are very sensitive to the driving force and initial conditions this nonconvex beam model could produce chaotic vibration. The so-called strange attractor is actually a local minimizer [56, 57]. Particularly, if the variable $\chi(t, x)$ can be separate variable as $\chi = q(t)\sin(\theta x)$, this nonlinear beam model is equivalent to the Duffing equation, which is well-known in chaotic dynamics. Figure 2 shows clearly that for the same given initial data, the same Runge–Kutta iteration but with different solvers in MATLAB produces very different "trajectories" in phase space q-p ($p = q_{,t}$). Therefore, this nonlinear beam model is important for understanding many challenging problems in both mathematics and engineering applications and has been subjected to extensive study recently [1, 2, 9, 12, 56, 96, 106, 110].

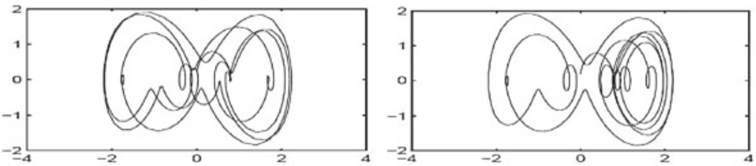


Fig. 2 Chaotical trajectories of the nonlinear Gao beam computed by "ode23" (*left*) and "ode15s" (*right*) in MATLAB

The canonical duality theory has been successfully for modeling real-world problems in nonconvex/nonsmooth dynamical systems [55, 67], differential geometry [81], contact mechanics [44], post-buckling structures [48], multiscale phase transitions of solids [60, 83], and general mathematical physics (see Chap. 4, [49]).

4 Applications in Large Deformation Mechanics

For mixed boundary value problems, the input $\bar{\boldsymbol{\chi}}^*$ is the body force $\mathbf{f}$ in the domain $\Omega \subset \mathbb{R}^d$ and surface traction $\mathbf{t}$ on the boundary $\Gamma_t \subset \partial\Omega$. The external energy

$$F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle = \int_\Omega \boldsymbol{\chi} \cdot \mathbf{f} \, \mathrm{d}\Omega + \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} \, \mathrm{d}\Gamma \tag{57}$$

is a linear functional defined on $\mathscr{X}_a = \{\boldsymbol{\chi} \in \mathscr{C}^1[\Omega; \mathbb{R}^p]| \;\; \boldsymbol{\chi} = 0 \text{ on } \Gamma_\chi\}$. For a hyperelastic material deformation problem, we have $\dim \Omega = d = p = 3$. The stored energy $W(\mathbf{F})$ is usually a nonconvex functional of the deformation gradient tensor $\mathbf{F} = \nabla \boldsymbol{\chi}$

$$W(\mathbf{F}) = \int_\Omega U(\mathbf{F}) \, \mathrm{d}\Omega, \tag{58}$$

where $U(\mathbf{F})$ is the stored energy density defined on $\mathscr{W}_a = \mathbb{M}^3_+ = \{\mathbf{F} = \{F^i_\alpha\} \in \mathbb{R}^{3\times 3}| \;\; \det \mathbf{F} > 0\}$. Thus, on the kinetically admissible space

$$\mathscr{X}_c = \{\boldsymbol{\chi} \in \mathscr{C}^1[\Omega; \mathbb{R}^d]| \;\; \det(\nabla \boldsymbol{\chi}) > 0, \;\; \boldsymbol{\chi} = 0 \text{ on } \Gamma_\chi\},$$

the general model (3) is a typical nonconvex variational problem

$$\min_{\boldsymbol{\chi} \in \mathscr{X}_c} \left\{ \Pi(\boldsymbol{\chi}) = \int_\Omega U(\nabla \boldsymbol{\chi}) \, \mathrm{d}\Omega - \int_\Omega \boldsymbol{\chi} \cdot \mathbf{f} \, \mathrm{d}\Omega - \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} \, \mathrm{d}\Gamma \right\}. \tag{59}$$

The linear operator $D = \text{grad} : \mathscr{X}_a \to \mathbb{M}^3_+$ in this problem is a gradient. The stationary condition $\delta\Pi(\boldsymbol{\chi}) = 0$ leads to a mixed boundary value problem (BVP)

$$(BVP): \quad A(\boldsymbol{\chi}) = \nabla^* \partial_{\mathbf{F}} W(\nabla \boldsymbol{\chi}) = \begin{cases} -\nabla \cdot \nabla_{\mathbf{F}} U(\nabla \boldsymbol{\chi}) = \mathbf{f} & \text{in } \Omega, \\ \mathbf{n} \cdot \nabla_{\mathbf{F}} U(\nabla \boldsymbol{\chi}) = \mathbf{t} & \text{on } \Gamma_t. \end{cases} \tag{60}$$

According to the definition of nonlinear PDEs, the first equilibrium equation (60) is fully nonlinear as long as $\partial U(\mathbf{F})$ is nonlinear. However, it is geometrically linear if $U(\mathbf{F})$ is convex. It is completely nonlinear only if $U(\mathbf{F})$ is nonconvex. Therefore, the definition of fully nonlinearity in PDEs cannot be used to identify difficulty of the nonlinear problems.

It is well-known in finite deformation theory that the convexity of the stored energy density $U(\mathbf{F})$ contradicts the most immediate physical experience (see Theorem 4.8-1, [16]). Indeed, even its domain $\mathbb{M}^3_+$ is not a convex subset of $\mathbb{R}^{3\times 3}$ (Theorem 4.7-4, [16]). Therefore, the solution to the (BVP) is only a stationary point of the total potential $\Pi(\boldsymbol{\chi})$. In order to identify minimizer of the problem, many generalized convexities have been suggested and the following results are well-known[7] (see [49]):

$$U(\mathbf{F}) \text{ is convex } \Rightarrow \text{ polyconvex } \Rightarrow \text{ quasi-convex } \Rightarrow \text{rank-one convex.} \quad (61)$$

If $U \in \mathscr{C}^2(\mathbb{M}^3_+)$, then the rank-one convexity is equivalent to the Legendre–Hadamard (L.H.) condition:

$$\sum_{i,j=1}^{3}\sum_{\alpha,\beta=1}^{3} \frac{\partial^2 U(\mathbf{F})}{\partial F^i_\alpha \partial F^j_\beta} a_i a_j b^\alpha b^\beta \geq 0 \;\; \forall \mathbf{a} = \{a_i\} \in \mathbb{R}^3, \;\; \forall \mathbf{b} = \{b^\alpha\} \in \mathbb{R}^3. \quad (62)$$

The Legendre–Hadamard condition in finite elasticity is also referred to as the *ellipticity condition*, i.e., if the L.H. condition holds, the partial differential operator $A(\boldsymbol{\chi})$ in (60) is considered to be elliptic. For one-dimensional problems $\Omega \subset \mathbb{R}$, all these convexities are equivalent and the rank-one convexity is the well-known convexity in vector space. We should emphasize that these generalized convexities and L.H. condition are local criteria not global. As long as the total potential $\Pi(\boldsymbol{\chi})$ is locally nonconvex in certain domain of Ω, the boundary value problem (60) could have multiple solutions $\boldsymbol{\chi}(\mathbf{x})$ at each material point $\mathbf{x} \in \Omega$ and the total potential $\Pi(\boldsymbol{\chi})$ could have infinitely number of local minimizers (see [71]). This is the main difference between nonconvex analysis and nonlinear PDEs, which is a key point to understand NP-hard problems in computer science and global optimization. Unfortunately, this difference is not fully understood in both fields. It turns out that extensive efforts have been devoted for solving nonconvex variational problems directly. It was discovered by Gao and Ogden in 2008 that even for one-dimensional problems, the L.H. condition can only identify local minimizers, and a geometrically nonlinear ODE could have infinite number solutions, both local and global minimal solutions could be nonsmooth and cannot be determined by any Newton type of numerical methods [71].

By the objectivity of the stored energy density $U(\mathbf{F})$, it is reasonable to assume a canonical function $V(\mathbf{C})$ such that the following canonical transformation holds:

$$W(\mathbf{F}) = \Phi(\Lambda(\mathbf{F})) = \int_\Omega V(\mathbf{F}^T\mathbf{F})\, \mathrm{d}\Omega. \quad (63)$$

[7]The quasiconvexity used in variational calculus and continuum physics has an entirely different meaning from that used in optimization, where a function $f: \mathbb{R}^n \to \mathbb{R}$ is called quasiconvex if its level set $\mathscr{L}_\alpha[f] = \{x \in \mathbb{R}^n |\ f(x) \leq \alpha\}$ is convex. For example, the nonconvex function $f(x) = \sqrt{|x|}$ is quasiconvex.

In this transformation, the geometrical nonlinear operator $\Lambda(\mathbf{F}) = \mathbf{F}^T\mathbf{F}$ is quadratic (objective) and $\mathbf{C} = \mathbf{F}^T\mathbf{F} \in \mathbb{S}^+ = \{\mathbf{C} = \{C_{\alpha\beta}\} \in \mathbb{R}^{3\times 3}|\ \ \mathbf{C} = \mathbf{C}^T,\ \ \mathbf{C} \succ 0\}$ is the well-known right Cauchy–Green strain tensor. Its canonical dual $\mathbf{S} = \partial\Phi(\mathbf{C}) = \nabla V(\mathbf{C}) \in \mathbb{S}$ is a *second Piola–Kirchhoff type stress tensor*.[8] In terms of the canonical strain measure $\mathbf{C}(\mathbf{F})$, the kinetically admissible space $\mathscr{X}_c = \{\boldsymbol{\chi} \in \mathscr{C}^1[\Omega, \mathbb{R}^3]|\ \ \mathbf{C}(\nabla\boldsymbol{\chi}) \in \mathbb{S}^+,\ \ \boldsymbol{\chi} = 0 \text{ on } \Gamma_\chi\}$ is convex and the nonconvex variational problem (59) can be written in the canonical form

$$\min\left\{\Pi(\boldsymbol{\chi}) = \Phi(\mathbf{C}(\nabla\boldsymbol{\chi})) - \langle\boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^*\rangle|\ \ \boldsymbol{\chi} \in \mathscr{X}_c\right\}. \tag{64}$$

By the Legendre transformation $V^*(\mathbf{S}) = \{\mathbf{C} : \mathbf{S} - V(\mathbf{C})|\ \mathbf{S} = \nabla V(\mathbf{C})\}$, the total complementary functional $\Xi(\boldsymbol{\chi}, \mathbf{S})$ has the following form:

$$\Xi(\boldsymbol{\chi}, \mathbf{S}) = \int_\Omega \left[\mathbf{C}(\nabla\boldsymbol{\chi}) : \mathbf{S} - V^*(\mathbf{S}) - \boldsymbol{\chi}\cdot\mathbf{f}\right] \mathrm{d}\Omega - \int_{\Gamma_t} \boldsymbol{\chi}\cdot\mathbf{t}\,\mathrm{d}\Gamma. \tag{65}$$

By the fact that the linear operator $D = \text{grad}$ is a differential operator, it is difficult to find its inverse operator. In order to obtain the canonical dual $\Pi^d(\mathbf{T})$, we need to introduce the following *statically admissible space*

$$\mathscr{T}_c = \{\boldsymbol{\tau} \in \mathscr{C}^1[\Omega; \mathbb{R}^{3\times 3}]|\ -\nabla\cdot\boldsymbol{\tau} = \mathbf{f} \text{ in } \Omega,\ \ \mathbf{n}\cdot\boldsymbol{\tau} = \mathbf{t} \text{ on } \Gamma_t\}.$$

Clearly, for any given $\boldsymbol{\chi} \in \mathscr{X}_a = \{\boldsymbol{\chi} \in \mathscr{C}^1[\Omega; \mathbb{R}^3]|\ \det(\nabla\boldsymbol{\chi}) > 0,\ \ \boldsymbol{\chi} = 0 \text{ on } \Gamma_\chi\}$, the external energy $F(\boldsymbol{\chi})$ can be written equivalently as

$$F_{\boldsymbol{\tau}}(\boldsymbol{\chi}) = \int_\Omega \boldsymbol{\chi}\cdot(-\nabla\cdot\boldsymbol{\tau})\,\mathrm{d}\Omega + \int_{\Gamma_t} \boldsymbol{\chi}\cdot\mathbf{t}\,\mathrm{d}\Gamma = \int_\Omega (\nabla\boldsymbol{\chi}) : \boldsymbol{\tau}\,\mathrm{d}\Omega \quad \forall\boldsymbol{\tau} \in \mathscr{T}_c \tag{66}$$

Thus, for any given $\boldsymbol{\tau} \in \mathscr{T}_c$, the Λ-conjugate of $F(\boldsymbol{\chi})$ can be obtained

$$F_{\boldsymbol{\tau}}^\Lambda(\mathbf{S}) = \text{sta}\{\langle\mathbf{C}(\nabla\boldsymbol{\chi}); \mathbf{S}\rangle - F_{\boldsymbol{\tau}}(\boldsymbol{\chi})|\ \boldsymbol{\chi} \in \mathscr{X}_a\} = -\int_\Omega \frac{1}{4}\text{tr}(\boldsymbol{\tau}\cdot\mathbf{S}^{-1}\cdot\boldsymbol{\tau}^T)\,\mathrm{d}\Omega. \tag{67}$$

Its domain should be

$$\mathscr{S}_c = \{\mathbf{S} \in \mathscr{E}_a^*|\ \det(\boldsymbol{\tau}\cdot\mathbf{S}^{-1}) > 0\}. \tag{68}$$

Therefore, the pure complementary energy can be obtained as

$$\Pi^d(\mathbf{S}; \boldsymbol{\tau}) = -\int_\Omega \left[\frac{1}{4}\text{tr}(\boldsymbol{\tau}\cdot\mathbf{S}^{-1}\cdot\boldsymbol{\tau}^T) + V^*(\mathbf{S})\right] \mathrm{d}\Omega, \tag{69}$$

---

[8]The second Piola–Kirchhoff stress tensor is defined by $\mathbf{T} = \partial\Phi(\mathbf{E})$, where $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$ is the Green–St. Venant strain tensor. Therefore, we have $\mathbf{S} = 2\mathbf{T}$.

which depends on not only the canonical stress $\mathbf{S} \in \mathscr{S}_c$, but also the statically admissible field $\boldsymbol{\tau} \in \mathscr{T}_c$. Let

$$\mathscr{S}_c^+ = \{\mathbf{S} \in \mathscr{S}_c | \ \mathbf{S} \succ 0\}, \quad \mathscr{S}_c^- = \{\mathbf{S} \in \mathscr{S}_c | \ \mathbf{S} \prec 0\}. \tag{70}$$

Theorem 4 (**Pure Complementary Energy Principle, Gao** [43, 45, 49])
If $(\bar{\mathbf{S}}, \bar{\boldsymbol{\tau}}) \in \mathscr{S}_c \times \mathscr{T}_c$ *is a stationary points of* $\Pi^d(\mathbf{S}; \boldsymbol{\tau})$*, then the deformation* $\bar{\boldsymbol{\chi}} \in \mathscr{X}_a$ *such that* $\nabla \bar{\boldsymbol{\chi}} = \bar{\boldsymbol{\tau}} \cdot \bar{\mathbf{S}}^{-1}$ *is a critical point of* $\Pi(\boldsymbol{\chi})$ *and* $\Pi(\bar{\boldsymbol{\chi}}) = \Pi^d(\bar{\mathbf{S}}; \bar{\boldsymbol{\tau}})$. *The critical point* $\bar{\boldsymbol{\chi}}(\mathbf{x})$ *is a global minimizer of* $\Pi(\boldsymbol{\chi})$ *if* $\bar{\mathbf{S}}(\mathbf{x}) \in \mathscr{S}_c^+ \ \ \forall \mathbf{x} \in \Omega$.

Moreover, if the compatibility condition $\nabla \times (\bar{\boldsymbol{\tau}} \cdot \bar{\mathbf{S}}^{-1}) = 0$ *holds, the deformation defined by*

$$\bar{\boldsymbol{\chi}}(\mathbf{x}) = \frac{1}{2} \int_{\mathbf{x}_0}^{\mathbf{x}} \bar{\boldsymbol{\tau}} \cdot \bar{\mathbf{S}}^{-1} d\mathbf{x} \tag{71}$$

along any path from $\mathbf{x}_0 \in \Gamma_\chi$ *to* $\mathbf{x} \in \Omega$ *is a solution to the boundary value problem (60).*

Proof. Using Lagrange multiplier $\boldsymbol{\chi} \in \mathscr{X}_a$ to relax the equilibrium conditions in $\mathscr{T}_c$, we have

$$\Theta(\mathbf{S}; \boldsymbol{\tau}, \boldsymbol{\chi}) = -\int_\Omega \left[\frac{1}{4} \mathrm{tr}(\boldsymbol{\tau} \cdot \mathbf{S}^{-1} \cdot \boldsymbol{\tau}^T) + V^*(\mathbf{S}) \right] \mathrm{d}\Omega - \int_\Omega \boldsymbol{\chi} \cdot (\nabla \cdot \boldsymbol{\tau} + \mathbf{f}) \, \mathrm{d}\Omega + \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} \, \mathrm{d}\Gamma.$$

Its stationary condition leads to

$$2\nabla \boldsymbol{\chi} = \boldsymbol{\tau} \cdot \mathbf{S}^{-1} \tag{72}$$

$$4\mathbf{S} \cdot (\nabla V^*(\mathbf{S})) \cdot \mathbf{S} = \boldsymbol{\tau}^T \cdot \boldsymbol{\tau} \tag{73}$$

and the equilibrium equations in $\mathscr{T}_c$. From (72) we have $\boldsymbol{\tau} = 2(\nabla \boldsymbol{\chi}) \cdot \mathbf{S}$. Substituting this into (73) we have $(\nabla \boldsymbol{\chi})^T (\nabla \boldsymbol{\chi}) = \nabla V^*(\mathbf{S})$, which is equivalent to $\mathbf{S} = \nabla V(\mathbf{C}(\nabla \boldsymbol{\chi}))$ due to the canonical duality. Thus, from the canonical transformation, we have

$$\boldsymbol{\tau} = 2(\nabla \boldsymbol{\chi}) \cdot (\nabla_{\mathbf{C}} V(\mathbf{C}(\nabla \boldsymbol{\chi}))) = \nabla_{\mathbf{F}} U(\nabla \boldsymbol{\chi}) \tag{74}$$

due to the chain role. This shows that the integral (71) is indeed a stationary point of $\Pi(\boldsymbol{\chi})$ since $\boldsymbol{\tau} \in \mathscr{T}_c$.

By the fact that $\mathbf{C} = \Lambda(\mathbf{F})$ is a quadratic operator, the Gao–Strang gap function is

$$G_{ap}(\boldsymbol{\chi}, \mathbf{S}) = \int_\Omega \mathrm{tr}[(\nabla \boldsymbol{\chi}) \cdot \mathbf{S} \cdot (\nabla \boldsymbol{\chi})] \, \mathrm{d}\Omega.$$

Clearly, $G_{ap}(\boldsymbol{\chi}, \mathbf{S})$ is non negative for any given $\boldsymbol{\chi} \in \mathscr{X}_a$ if and only if $\mathbf{S}(\mathbf{x}) \in \mathscr{S}_c^+ \ \ \forall \mathbf{x} \in \Omega$. Replace $\nabla \boldsymbol{\chi} = \frac{1}{2} \boldsymbol{\tau} \cdot \mathbf{S}^{-1}$, this gap function reads

$$G_{ap}(\boldsymbol{\chi}(\mathbf{S},\boldsymbol{\tau}),\mathbf{S}) = \int_{\Omega} \frac{1}{4}\mathrm{tr}[\boldsymbol{\tau}\cdot\mathbf{S}^{-1}\cdot\boldsymbol{\tau}^T]\,\mathrm{d}\Omega,$$

which is convex for any $\boldsymbol{\tau} \in \mathscr{T}_c$ if and only if $\mathbf{S}(\mathbf{x}) \in \mathscr{S}_c^+ \;\; \forall \mathbf{x} \in \Omega$. Therefore, the canonical dual $\Pi^d(\mathbf{S};\boldsymbol{\tau})$ is concave on $\mathscr{S}_a^+ \times \mathscr{T}_c$. By the canonical min–max duality, $\bar{\boldsymbol{\chi}}$ is a unique global minimizer if $\bar{\mathbf{S}}((\mathbf{x}) \succ 0 \;\; \forall \mathbf{x} \in \Omega$.

The compatibility condition $\nabla \times (\boldsymbol{\tau}\cdot\mathbf{S}) = 0$ is necessary for an analytical solution to the mixed boundary value problem (60) due to the fact that curl grad $\boldsymbol{\chi} = 0$. □

The pure complementary energy principle was first proposed by Gao (1997) in post-buckling problems of a large deformed beam [42]. Generalization to 3-D finite deformation theory and nonconvex analysis were given during 1998–2000 [43, 45, 47, 49, 50]. The Eq. (73) is called the *canonical dual algebraic equation* first obtained in 1998 [43]. This equation shows that by the canonical dual transformation, the nonlinear partial differential equation can be equivalently reformed as an algebraic equation. The Eq. (74) show that the statically admissible field $\boldsymbol{\tau} = \nabla U(\mathbf{F})$ is actually the *first Piola–Kirchhoff stress*. For one-dimensional problems, $\boldsymbol{\tau} \in \mathscr{T}_c$ can be easily obtained by the given input. For geometrically nonlinear problems, $\nabla V^*(\mathbf{S})$ is linear and (73) can be solved analytically to obtain a complete set of analytical solutions [49, 50, 66, 71, 72]. By the triality theory, the positive solution $\mathbf{S} \in \mathscr{S}_c^+$ produces a global minimal solution $\bar{\boldsymbol{\chi}}$, while the negative $\mathbf{S} \in \mathscr{S}_c^-$ can be used to identify local extremal solutions. To see this, let us consider the Hessian of the stored energy $U(\mathbf{F}) = V(\mathbf{C}(\mathbf{F}))$. By chain rule, we have

$$\frac{\partial^2 U(\mathbf{F})}{\partial F^i_\alpha \partial F^j_\beta} = 2\delta^{ij} S_{\alpha\beta} + 4 \sum_{\theta,\nu=1}^{3} F^i_\theta H_{\theta\alpha\beta\nu} F^j_\nu, \tag{75}$$

where $\mathbf{H} = \{H_{\theta\alpha\beta\nu}\} = \nabla^2 V(\mathbf{C}) \succ 0$ due to the convexity of the canonical function $V(\mathbf{C})$. Clearly, if $\mathbf{S} \succeq 0$, the L.H. condition holds and the associated $\bar{\boldsymbol{\chi}}$ is a global minimal solution. By the fact that $2\mathbf{F} = \boldsymbol{\tau}\mathbf{S}^{-1}$, we know that $\nabla^2 U(\mathbf{F})$ could be either positive or negative definite even if $\mathbf{S} \prec 0$. Therefore, depending the eigenvalues of $\mathbf{S} \prec 0$, the L.H. condition could also hold at a local minimizer of $\Pi(\boldsymbol{\chi})$ [66]. This shows that the triality theory can be used to identify both global and local extremal solutions, while the L.H. condition is only a necessary condition for a local minimal solution. It is known that an elliptic equation is corresponding to a convex variational problem. Therefore, it is a question if the Legendre–Hadamard condition can still be called as the ellipticity condition in finite elasticity and nonconvex analysis. By the fact that the well-known open problem left by Reissner et al. [125] has been solved by Theorem 4, the pure complementary energy principle is known as the Gao principle in literature (see [107]).

The canonical transformation $W(\mathbf{F}) = \Phi(\Lambda(\mathbf{F}))$ is not unique since the geometrical operator $\Lambda(\mathbf{F})$ can be chosen differently to have different canonical strain measures. For example, the well-known *Hill-Seth strain family*

$$\mathbf{E}^{(\eta)} = \Lambda(\mathbf{F}) = \frac{1}{2\eta}[(\mathbf{F}^T \cdot \mathbf{F})^\eta - \mathbf{I}] \tag{76}$$

is a geometrically admissible objective strain measure for any given $\eta \in \mathbb{R}$ (see Definition 6.3.1, [49]). Particularly, $\mathbf{E}^{(1)}$ is the well-known *Green–St. Venant strain tensor* $\mathbf{E}$. For *St. Venant–Kirchhoff materials*, the stored strain density is quadratic: $V(\mathbf{E}) = \frac{1}{2}\mathbf{E} : \mathbf{H} : \mathbf{E}$, where $\mathbf{H}$ is the Hooke tensor. Clearly, $V(\mathbf{E})$ is convex but

$$U(\mathbf{F}) = V(\mathbf{E}(\mathbf{F})) = \frac{1}{8}(\mathbf{F}^T \cdot \mathbf{F} - \mathbf{I}) : \mathbf{H} : (\mathbf{F}^T \cdot \mathbf{F} - \mathbf{I})$$

is a (nonconvex) double-well type function of $\mathbf{F}$, which is not even rank-one convex [124]. The canonical duality is linear $\mathbf{T} = \nabla V(\mathbf{E}) = \mathbf{H} : \mathbf{E}$ and the generalized total complementary energy $\Xi(\boldsymbol{\chi}, \mathbf{T})$ is the well-known Hellinger–Reissner complementary energy

$$\Xi(\boldsymbol{\chi}, \mathbf{T}) = \int_\Omega \left[\mathbf{E}(\nabla \boldsymbol{\chi}) : \mathbf{T} - \frac{1}{2}\mathbf{T} : \mathbf{H}^{-1} : \mathbf{T} - \boldsymbol{\chi} \cdot \mathbf{f} \right] \mathrm{d}\Omega - \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} \, \mathrm{d}\Gamma. \tag{77}$$

In this case, the primal problem (59) is a geometrically nonlinear variational problem, and its canonical dual functional is

$$\Pi^d(\mathbf{T}; \boldsymbol{\tau}) = -\int_\Omega \frac{1}{2}\left[\mathrm{tr}(\boldsymbol{\tau} \cdot \mathbf{T}^{-1} \cdot \boldsymbol{\tau}^T + \mathbf{T}) + \mathbf{T} : \mathbf{H}^{-1} : \mathbf{T} \right] \mathrm{d}\Omega. \tag{78}$$

The canonical dual algebraic equation (73) is then a cubic tensor equation

$$2\,\mathbf{T} \cdot (\mathbf{H}^{-1} : \mathbf{T} + \mathbf{I}) \cdot \mathbf{T} = \boldsymbol{\tau}^T \cdot \boldsymbol{\tau} \tag{79}$$

For a given statically admissible stress field $\boldsymbol{\tau} \in \mathscr{T}_c$, this tensor equation could have at most 27 solutions $\mathbf{T}(\mathbf{x})$ at each material point $\mathbf{x} \in \Omega$, but only one $\mathbf{T}(\mathbf{x}) \succ 0$, which leads to a global minimal solution [68].

For many real-world problems, the statically admissible stress $\boldsymbol{\tau} \in \mathscr{T}_c$ can be uniquely obtained and the canonical dual algebraic equation (79) can be solved to obtain all possible stress solutions. The canonical duality-triality theory has been used successfully for solving a class of nonconvex variational/boundary value problems [50, 69, 71], pure azimuthal shear [72] and anti-plane shear problems [66].

5 Applications to Computational Mechanics and Global Optimization

Numericalization for solving the nonconvex variational problem (3) leads to a global optimization problem in a finite dimensional space $\mathscr{X} = \mathscr{X}^*$. In complex systems, the decision variable $\boldsymbol{\chi}$ could be either vector or matrix. In operations research, such as logistic and supply chain management sciences, $\boldsymbol{\chi}$ can be even a high-order matrix $\boldsymbol{\chi} = \{\chi_{ij...k}\}$. Correspondingly, the linear operator $D : \mathscr{X}_a \to \mathscr{W}_a$ is a matrix or high-order tensor. In general global optimization problems, the internal energy $W(D\boldsymbol{\chi})$ is not necessary to be an objective function. As long as the canonical transformation $W(D(\boldsymbol{\chi}) = \Phi(\Lambda(D\boldsymbol{\chi}))$ holds, the canonical duality-triality theory can be used for solving a large class of nonconvex/discrete optimization problems.

5.1 Canonical Dual Finite Element Method

It was shown in [40] that by using independent finite element interpolations for displacement and generalized stress:

$$\boldsymbol{\chi}(\mathbf{x}) = \mathbf{N}_u(\mathbf{x})\mathbf{q}^e, \quad \mathbf{S}(\mathbf{x}) = \mathbf{N}_\varsigma(\mathbf{x})\mathbf{p}^e \quad \forall \mathbf{x} \in \Omega^e \subset \Omega, \tag{80}$$

the total complementary functional $\Xi(\boldsymbol{\chi}, \mathbf{S})$ defined by (65) can be discretized as a function in finite-dimensional space

$$\Xi(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\mathbf{q}^T \mathbf{G}(\mathbf{p})\mathbf{q} - \Phi^*(\mathbf{p}) - \mathbf{q}^T \mathbf{f}, \tag{81}$$

where $\mathbf{f}$ is the generalized force and $\mathbf{G}(\mathbf{p})$ is the Hessian matrix of the discretized Gao–Strang complementary gap function. In this case, the pure complementary energy can be formulated explicitly as [40]

$$\Pi^d(\mathbf{p}) = -\frac{1}{2}\mathbf{f}^T \mathbf{G}^+(\mathbf{p})\mathbf{f} - \Phi^*(\mathbf{p}), \tag{82}$$

where $\mathbf{G}^+$ represents a generalized inverse of $\mathbf{G}$. Let

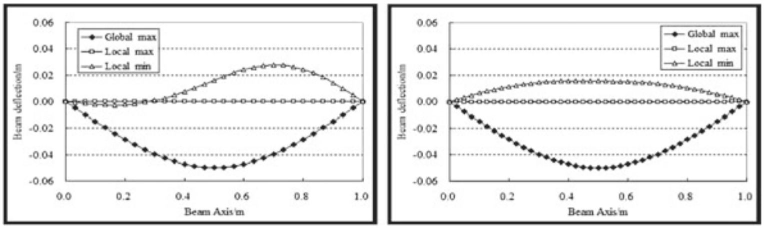
$$\mathscr{S}_c^+ = \{\mathbf{p} \in \mathbb{R}^m |\ \mathbf{G}(\mathbf{p}) \succeq 0\}, \quad \mathscr{S}_c^- = \{\mathbf{p} \in \mathbb{R}^m |\ \mathbf{G}(\mathbf{p}) \prec 0\}.$$

By the fact that $\Pi^d(\mathbf{p})$ is concave on the convex set $\mathscr{S}^+$, the canonical dual FE programming problem

$$\max\{\Pi^d(\mathbf{p}) |\ \mathbf{p} \in \mathscr{S}_c^+\} \tag{83}$$

can be solved easily (if $\mathscr{S}_c^+ \neq \emptyset$) to obtain the global maximizer $\bar{\mathbf{p}}$. By the triality theory, we know that $\bar{\mathbf{q}} = \mathbf{G}^+(\bar{\mathbf{p}})\mathbf{f}$ is a global minimizer of the nonconvex potential $\Pi(\mathbf{q})$. On the other hand, if $\dim \mathbf{q} = \dim \mathbf{p}$, the biggest local min and local max of $\Pi(\mathbf{q})$ can be obtained respectively by [79]

$$\min\{\Pi^d(\mathbf{p}) |\ \mathbf{p} \in \mathscr{S}_c^-\}, \quad \max\{\Pi^d(\mathbf{p}) |\ \mathbf{p} \in \mathscr{S}_c^-\}.$$



(a) 30 elements. (b) 40 elements.

Fig. 3 Canonical dual FEM solutions for post-buckled nonlinear beam: Global minimal solution, i.e., stable buckled state (*doted*); local min, i.e., unstable buckled state (*triangle*); and local max, i.e., unbuckled state (*squared*)

The canonical dual FEM has been used successfully in phase transitions of solids [83] and in post-buckling analysis for the large deformation beam model (2) to obtain all three possible solutions [129] (see Fig. 3). It was discovered that the local minimum is very sensitive to the lateral load and the size of the finite element meshes (see Fig. 3). This method can be used for solving general nonconvex mechanics problems.

5.2 *Global Optimal Solutions for Discrete Nonlinear Dynamical Systems*

General nonlinear dynamical systems can be modeled as a nonlinear initial value problem

$$\boldsymbol{\chi}'(t) = \mathbf{F}(t, \boldsymbol{\chi}(t)) \quad t \in [0, T], \quad \boldsymbol{\chi}(0) = \boldsymbol{\chi}_0, \tag{84}$$

where $T > 0$, $\mathbf{F} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a given vector-valued function. Generally speaking, if a nonlinear equation has multiple solutions at each time t in a subset of its domain $[0, T]$, then the associated initial-valued problem should have infinite number of solutions since the unknown $\boldsymbol{\chi}(t)$ is a continuous function. With time step size $h = T/n$, a discretization of the configuration $\boldsymbol{\chi}(t)$ is $\mathbf{X} = (\boldsymbol{\chi}_1, \cdots, \boldsymbol{\chi}_n) \in \mathscr{X}_a \subset \mathbb{R}^{d\times n}$. By the finite difference method and trapezoidal rule,[9] the initial value problem (84) can be written approximately as

$$\boldsymbol{\chi}_k = \boldsymbol{\chi}_{k-1} + \frac{1}{2} h[\mathbf{F}_{k-1} + \mathbf{F}_k], \quad k = 1, \cdots, n \tag{85}$$

where $\mathbf{F}_k = \mathbf{F}(t_k, \boldsymbol{\chi}_k)$. This is a nonlinear algebraic system. Clearly, if $\boldsymbol{\chi}_k$ in $\mathbf{F}_k$ is replaced by the Euler linear iteration $\boldsymbol{\chi}_k = \boldsymbol{\chi}_{k-1} + h\mathbf{F}_{k-1}$, then (85) is the generalized

[9] Clearly, we can adopt high-order rule for approximation of $F(t, Y)$ at the $k-1$ step, which will should be subjected to study in the future.

Euler method. The popular Runge–Kutta method is also a generalized linear iteration. It is well-known that any linear iteration can only produce one of the infinite number solutions to a nonconvex system and such a numerical "solution" is very sensitive to the step size and numerical errors. This is the reason why different numerical solvers produce totally different results, i.e., the so-called chaotic solutions. Rather than the traditional linear iteration from an initial value, we use the least squares method such that the nonlinear algebraic system (85) can be equivalently written as

$$\min_{\mathbf{X}\in\mathscr{X}_a} \left\{ \Pi(\mathbf{X}) = \frac{1}{2}\sum_{k=1}^{n} \left\| \boldsymbol{\chi}_k - \boldsymbol{\chi}_{k-1} - \frac{1}{2}h(\mathbf{F}_{k-1} + \mathbf{F}_k) \right\|^2 \right\}. \tag{86}$$

Clearly, for any given nonlinear function $\mathbf{F}(t, \boldsymbol{\chi}(t))$, this is a global optimization problem, which could have multiple minimizers at each $\boldsymbol{\chi}_k$. Particularly, if $\mathbf{F}(t, \boldsymbol{\chi})$ is quadratic, then $\Pi(\mathbf{X})$ is a double-well typed fourth-order polynomial function, and is considered to be NP-hard in global optimization even for $d = 1$ (one-dimensional systems) [4, 130]. However, by simply using the quadratic geometrical operator $\boldsymbol{\xi}_k = \Lambda(\boldsymbol{\chi}_k) = \mathbf{F}(t_k, \boldsymbol{\chi}_k)$, the nonconvex leas squares problem (86) can be solved by the canonical duality-triality theory to obtain global optimal solution. Applications have been given to population growth problems [126] and chaotic dynamics [102].

5.3 Unconstrained Nonconvex Minimization

The general model (3) for unconstrained global optimization can be written in the following form

$$\min \left\{ \Pi(\boldsymbol{\chi}) = W(\mathbf{D}\boldsymbol{\chi}) + \frac{1}{2}\langle \boldsymbol{\chi}, \mathbf{A}\boldsymbol{\chi}\rangle - \langle \boldsymbol{\chi}, \mathbf{f}\rangle | \;\; \boldsymbol{\chi} \in \mathscr{X}_a \right\}, \tag{87}$$

where $\mathbf{D} : \mathscr{X}_a \to \mathscr{W}_a$ and $\mathbf{A} = \mathbf{A}^T$ are two given operators and $\mathbf{f} \in \mathscr{X}_a$ is a given input. For the nonconvex function $W(\boldsymbol{\varepsilon})$, we assume that the canonical transformation $W(\mathbf{D}\boldsymbol{\chi}) = \Phi(\Lambda(\boldsymbol{\chi}))$ holds for a quadratic operator

$$\Lambda(\boldsymbol{\chi}) = \left\{ \frac{1}{2}\boldsymbol{\chi}^T \mathbf{H}_{\alpha\beta}\boldsymbol{\chi} \right\} : \mathscr{X}_a \to \mathscr{E}_a \subset \mathbb{R}^{m\times m}, \tag{88}$$

where $\mathbf{H}_{\alpha\beta} = \mathbf{H}_{\alpha\beta}^T \;\; \forall \alpha, \beta \in I_m = \{1, \ldots, m\}$ is a linear operator such that $\mathscr{E}_a$ is either a vector ($\beta = 1$) or tensor ($\alpha, \beta > 1$) space. By the convexity of the canonical function $\Phi : \mathscr{E}_a \to \mathbb{R}$, the canonical duality $\mathbf{S} = \partial\Phi(\boldsymbol{\xi}) \in \mathscr{E}_a^* \subset \mathbb{R}^{m\times m}$ is invertible and the total complementary function $\Xi : \mathscr{X}_a \times \mathscr{E}_a^* \to \mathbb{R}$ reads

$$\Xi(\boldsymbol{\chi}, \mathbf{S}) = \frac{1}{2}\langle \boldsymbol{\chi}, \mathbf{G}(\mathbf{S})\boldsymbol{\chi}\rangle - \Phi^*(\mathbf{S}) - \langle \boldsymbol{\chi}, \mathbf{f}\rangle \tag{89}$$

where $\mathbf{G}(\mathbf{S}) = \mathbf{A} + \sum_{\alpha,\beta \in I_m} \mathbf{H}_{\alpha\beta} S_{\alpha\beta}$. Thus, on $\mathscr{S}_c^+ = \{\mathbf{S} \in \mathscr{E}_a^* | \ \mathbf{G}(\mathbf{S}) \succ 0\}$, the canonical dual problem (33) for the unconstrained global optimization reads

$$\max \left\{ \Pi^d(\mathbf{S}) = -\frac{1}{2}\langle \mathbf{G}^{-1}(\mathbf{S})\mathbf{f}, \mathbf{f} \rangle - \Phi^*(\mathbf{S}) \ | \ \mathbf{S} \in \mathscr{S}_c^+ \right\}. \tag{90}$$

The canonical duality-triality theory has been used successfully for solving the following nonconvex problems.

(1) **Euclidian Distance Geometry Problem**

$$W(\mathbf{D}\boldsymbol{\chi}) = \sum_{i,j=1}^{n} \omega_{ij} \left[\| \boldsymbol{\chi}_i - \boldsymbol{\chi}_j \|^2 - d_{ij} \right]^2, \tag{91}$$

where the decision variable $\boldsymbol{\chi}_i \in \mathbb{R}^d$ is a position (location) vector, $\omega_{ij}, \ d_{ij} > 0 \ \forall i, \ j = 1, \ldots, n, \ i \neq j$ are given weight and distance parameters, respectively. The linear operator $\mathbf{D}\boldsymbol{\chi} = \{\boldsymbol{\chi}_i - \boldsymbol{\chi}_j\}$ in this problem is similar to the finite difference in numerical analysis. Such a problem appears frequently in computational biology [144], chaotic dynamics [108, 126], numerical algebra [128], sensor localization [99, 127], network communication [87], transportation optimization, as well as finite element analysis of structural mechanics [12, 83], etc. These problems are considered to be NP-hard even the Euclidian dimension $d = 1$ [4]. However, by the combination of the canonical duality-triality theory and perturbation methods, these problems can be solved efficiently (see [127]).

(2) **Sum of Fractional Functions**

$$W(\mathbf{D}\boldsymbol{\chi}) = \sum_{i \in I_m} \frac{G_i(\mathbf{D}_g \boldsymbol{\chi})}{H_i(\mathbf{D}_h \boldsymbol{\chi})} \tag{92}$$

where G_i and $H_i > 0 \ \ \forall \ i \in I_m$ are given functions, $\mathbf{D}_g$ and $\mathbf{D}_h$ are linear operators.

(3) **Exponential-Sum-Polynomials**

$$W(\mathbf{D}\boldsymbol{\chi}) = \sum_{i \in I_m} \exp\left(\frac{1}{2} \boldsymbol{\chi}^T \mathbf{B}_i \boldsymbol{\chi} - \alpha_i \right) + \sum_{j \in I_p} \frac{1}{2} \left(\frac{1}{2} \boldsymbol{\chi}^T \mathbf{C}_j \boldsymbol{\chi} - \beta_j \right)^2, \tag{93}$$

where $\mathbf{B}_i$ and $\mathbf{C}_j$ are given symmetrical matrices in $\mathbb{R}^{n \times n}$, α_i, β_j are given parameters.

(4) **Log-Sum-Exp Functions**

$$W(\mathbf{D}\boldsymbol{\chi}) = \frac{1}{\beta} \log \left[1 + \sum_{i \in I_p} \exp\left(\beta \left(\frac{1}{2} \boldsymbol{\chi}^T \mathbf{B}_i \boldsymbol{\chi} + d \right) \right) \right], \tag{94}$$

where $\beta > 0$, $\mathbf{B}_i = \mathbf{B}_i^T$, and $d \in \mathbb{R}$ are given.

All these functions appear extensively in modeling real-world problems, such as computational biology [144], biomechanics, phase transitions [71], filter design [142], location/transportation and networks optimization [87, 127], communication and information theory (see [100]), etc. By using the canonical duality-triality theory, these problems can be solved nicely (see [15, 24, 59, 61, 74, 103, 113, 145]).

5.4 Constrained Global Optimization

Recall the standard mathematical model in global optimization (1)

$$\min f(\mathbf{x}), \quad \text{s.t. } h_i(\mathbf{x}) = 0, \quad g_j(\mathbf{x}) \le 0 \;\; \forall i \in I_m, \quad j \in I_p, \tag{95}$$

where f, g_i and h_j are differentiable, real-valued functions on a subset of $\mathbb{R}^n$ for all $i \in I_m$ and $j \in I_p$. For notational convenience, we use vector forms for constraints

$$\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))\,, \quad \mathbf{h}(\mathbf{x}) = \big(h_1(\mathbf{x}), \ldots, h_p(\mathbf{x})\big)\,.$$

Therefore, the feasible space can be defined as

$$\mathscr{X}_c := \{\mathbf{x} \in \mathbb{R}^n | \mathbf{g}(\mathbf{x}) \le 0, \quad \mathbf{h}(\mathbf{x}) = 0\}.$$

Lagrange multiplier method was originally proposed by J-L Lagrange from analytical mechanics in 1811 [97]. During the past two hundred years, this method and the associated Lagrangian duality theory have been well-developed with extensively applications to many fields of physics, mathematics, and engineering sciences. Strictly speaking, the Lagrange multiplier method can be used only for equilibrium constraints. For inequality constraints, the well-known KKT conditions are involved. Here we show that both the classical Lagrange multiplier method and the KKT theory can be unified by the canonical duality theory.

For convex constrained problem, i.e., $f(\mathbf{x})$, $\mathbf{g}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$ are convex, the standard canonical dual transformation can be used. We can choose the geometrical operator $\boldsymbol{\xi}_0 = \boldsymbol{\Lambda}_0(\mathbf{x}) = \{\mathbf{g}(\mathbf{x}), \mathbf{h}(\mathbf{x})\} : \mathbb{R}^n \to \mathbb{R}^{m+p}$ and let

$$\Phi_0(\boldsymbol{\xi}_0) = \Psi_g(\mathbf{g}) + \Psi_h(\mathbf{h}),$$

where

$$\Psi_g(\mathbf{g}) = \{0 \;\text{ if } \mathbf{g} \le 0, \;\; +\infty \;\text{ otherwise}\}, \quad \Psi_h(\mathbf{h}) = \{0 \;\text{ if } \mathbf{h} = 0, \;\; +\infty \;\text{ otherwise}\},$$

are the so-called indicator functions for the inequality and equality constraints. Then the convex constrained problem (95) can be written in the following canonical form

$$\min \{\Pi(\mathbf{x}) = f(\mathbf{x}) + \Phi_0(\boldsymbol{\Lambda}_0(\mathbf{x}))| \;\; \forall \mathbf{x} \in \mathbb{R}\}\,. \tag{96}$$

By the fact that the canonical function $\Phi_0(\boldsymbol{\xi}_0)$ is convex and lower semi-continuous, the canonical duality relations (10) should be replaced by the following subdifferential forms [51]:

$$\boldsymbol{\xi}_0^* \in \partial\Phi_0(\boldsymbol{\xi}_0) \quad \Leftrightarrow \quad \boldsymbol{\xi}_0 \in \partial\Phi_0^*(\boldsymbol{\xi}_0^*) \quad \Leftrightarrow \quad \Phi_0(\boldsymbol{\xi}_0) + \Phi_0^*(\boldsymbol{\xi}_0^*) = \boldsymbol{\xi}_0^T\boldsymbol{\xi}_0^*, \tag{97}$$

where $\Phi_0^*(\boldsymbol{\xi}_0^*) = \Psi_g^*(\boldsymbol{\lambda}) + \Psi_h^*(\boldsymbol{\mu})$ is the Fenchel conjugate of $\Phi_0(\boldsymbol{\xi}_0)$ and $\boldsymbol{\xi}_0^* = (\boldsymbol{\lambda}, \boldsymbol{\mu})$. By the Fenchel transformation, we have

$$\Psi_g^*(\boldsymbol{\lambda}) = \sup_{\mathbf{g}\in\mathbb{R}^m}\{\mathbf{g}^T\boldsymbol{\lambda} - \Psi_g(\mathbf{g})\} = \begin{cases} 0 & \text{if } \boldsymbol{\lambda} \geq 0 \\ +\infty & \text{otherwise,} \end{cases}$$

$$\Psi_h^*(\boldsymbol{\mu}) = \sup_{\mathbf{h}\in\mathbb{R}^p}\{\mathbf{h}^T\boldsymbol{\mu} - \Psi_h(\mathbf{h})\} = 0 \quad \forall\, \boldsymbol{\mu} \in \mathbb{R}^p.$$

It is easy to verify that for the indicator $\Psi_g(\mathbf{g})$, the canonical duality leads to

$$\begin{aligned} \boldsymbol{\lambda} \in \partial\Psi_g(\mathbf{g}) \quad & \Longrightarrow \boldsymbol{\lambda} \geq 0 \\ \mathbf{g} \in \partial\Psi_g^*(\boldsymbol{\lambda}) \quad & \Longrightarrow \mathbf{g} \leq 0 \\ \boldsymbol{\lambda}^T\mathbf{g} = \Psi_g(\mathbf{g}) + \Psi_g^*(\boldsymbol{\lambda}) & \Longrightarrow \boldsymbol{\lambda}^T\mathbf{g} = 0, \end{aligned} \tag{98}$$

which are the KKT conditions for the inequality constrains $\mathbf{g}(\mathbf{x}) \leq 0$. While for $\Psi_h(\mathbf{h})$, the canonical duality lead to

$$\begin{aligned} \boldsymbol{\mu} \in \partial\Psi_h(\mathbf{h}) \quad & \Longrightarrow \boldsymbol{\mu} \in \mathbb{R}^p \\ \mathbf{h} \in \partial\Psi_h^*(\boldsymbol{\mu}) \quad & \Longrightarrow \mathbf{h} = 0 \\ \boldsymbol{\mu}^T\mathbf{h} = \Psi_h(\mathbf{h}) + \Psi_h^*(\boldsymbol{\mu}) & \Longrightarrow \boldsymbol{\mu}^T\mathbf{h} = 0. \end{aligned} \tag{99}$$

From the second and third conditions in the (99), it is clear that in order to enforce the constrain $\mathbf{h}(\mathbf{x}) = 0$, the dual variable $\boldsymbol{\mu} = \{\mu_i\}$ must be not zero $\forall i \in I_p$. This is a special complementarity condition for equality constrains, generally not mentioned in many textbooks. However, the implicit constraint $\boldsymbol{\mu} \neq 0$ is important in nonconvex optimization.

By using the Fenchel–Young equality $\Phi_0(\boldsymbol{\xi}_0) = \boldsymbol{\xi}_0^T\boldsymbol{\xi}_0^* - \Phi_0^*(\boldsymbol{\xi}_0^*)$ to replace Φ_0 $(\boldsymbol{\Lambda}_0(\mathbf{x}))$ in (96), the total complementarity function can be obtained in the following form:

$$\Xi_0(\mathbf{x}, \boldsymbol{\xi}_0^*) = f(\mathbf{x}) + [\boldsymbol{\lambda}^T\mathbf{g}(\mathbf{x}) - \Psi_g^*(\boldsymbol{\lambda})] + [\boldsymbol{\mu}^T\mathbf{h}(\mathbf{x}) - \Psi_g^*(\boldsymbol{\mu})]. \tag{100}$$

Let $\boldsymbol{\sigma}_0 = (\boldsymbol{\lambda}, \boldsymbol{\mu})$. The dual feasible spaces should be defined as

$$\mathscr{S}_0 = \{\boldsymbol{\sigma}_0 = (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^{m\times p} |\ \ \lambda_i \geq 0 \ \ \forall i \in I_m, \quad \mu_j \neq 0 \ \ \forall j \in I_p\}.$$

Thus, on the feasible space $\mathbb{R}^n \times \mathscr{S}_0$, the total complementary function (100) can be simplified as

$$\Xi_0(\mathbf{x}, \boldsymbol{\sigma}_0) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{x}) = \mathscr{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}), \tag{101}$$

which is the classical Lagrangian and we have

$$P(\mathbf{x}) = \sup \{\Xi_0(\mathbf{x}, \boldsymbol{\sigma}_0) | \ \forall \boldsymbol{\sigma}_0 \in \mathscr{S}_0\} .$$

This shows that the canonical duality theory is an extension of the Lagrangian theory (indeed, the total complementary function was called the extended Lagrangian in [49]).

For nonconvex constrained problems, the so-called *sequential canonical transformation* (see Chap. 4, [49])

$$\boldsymbol{\Lambda}_0(\boldsymbol{\Lambda}_1(\ldots(\boldsymbol{\Lambda}_k(\mathbf{x}))\ldots))$$

can be used for target function and constraints to obtain high-order canonical dual problem. Applications have been given to the high-order polynomial optimization [62, 89], nonconvex analysis [49], neural network [100], and nonconvex constrained problems [82, 85, 101, 114, 146].

5.5 SDP Relaxation and Canonical Primal-Dual Algorithms

Recall the primal problem $(\mathscr{P})$ (13)

$$(\mathscr{P}): \quad \min\{\Pi(\boldsymbol{\chi}) = \Phi(\Lambda(D\boldsymbol{\chi})) - \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle | \ \boldsymbol{\chi} \in \mathscr{X}_c\}$$

and its canonical dual $(\mathscr{P}^d)$ (33)

$$(\mathscr{P}^d): \quad \max\left\{\Pi^d(\mathbf{S}) = -G^*_{ap}(\mathbf{S}) - \Phi^*(\mathbf{S}) \mid \mathbf{S} \in \mathscr{S}_c^+\right\},$$

where $G^*_{ap}(\mathbf{S}) = \frac{1}{2}\langle \mathbf{G}^{-1}(\mathbf{S})\mathbf{F}(\mathbf{S}), \mathbf{F}(\mathbf{S})\rangle$ is the pure gap function. By the fact that $(\mathscr{P}^d)$ is a concave maximization on a convex domain $\mathscr{S}_c^+$, this canonical dual can be solved easily if $\Pi^d(\boldsymbol{\xi}^*)$ has a stationary point in $\mathscr{S}_c^+$. For many challenging (NP-hard) problems, the stationary points $\Pi^d(\mathbf{S})$ are usually located on the boundary of $\mathscr{S}_c^+ = \{\mathbf{S} \in \mathscr{S}_c | \ \mathbf{G}(\mathbf{S}) \succ 0\}$. In this case, the matrix $\mathbf{G}(\mathbf{S})$ is singular and the canonical dual problem could have multiple solutions. Two methods can be suggested for solving this challenging case.

(1) SDP Relaxation. By using the Schur complement Lemma, the canonical dual problem $(\mathscr{P}^d)$ can be relaxed as [145]

$$(\mathscr{P}^r): \quad \min \Phi^*(\mathbf{S}) \ \text{ s.t. } \begin{pmatrix} \mathbf{G}(\mathbf{S}) & \mathbf{F}(\mathbf{S}) \\ \mathbf{F}^T(\mathbf{S}) & 2G_{ap}(\mathbf{S}) \end{pmatrix} \succeq 0, \quad \forall \mathbf{S} \in \mathscr{S}_c. \tag{102}$$

Since $\Phi^*(\mathbf{S})$ is convex and the feasible space is closed, this relaxed canonical dual problem has at least one solution $\bar{\mathbf{S}}$. The associated $\bar{\boldsymbol{\chi}} = \mathbf{G}(\bar{\mathbf{S}})^{-1}\mathbf{F}(\bar{\mathbf{S}})$ is a solution to $(\mathscr{P})$ only if $\bar{\mathbf{S}}$ is a stationary point of $\Pi^d(\mathbf{S})$. Particularly, if $\Phi^*(\mathbf{S}) = \langle \mathbf{Q}; \mathbf{S}\rangle$ is linear, $\mathbf{F} = 0$, $\mathbf{G}(\mathbf{S}) = \mathbf{S}$, and

$$\mathscr{S}_c = \{\mathbf{S} \in \mathbb{S}_n |\ \langle \mathbf{A}_i; \mathbf{S}\rangle = b_i\ \forall i \in I_m\}$$

is a linear manifold, where $\mathbb{S}_n = \{\mathbf{S} \in \mathbb{R}^{n\times n} | \ \mathbf{S} = \mathbf{S}^T\}$ is a symmetrical $n \times n$-matrix space, $\mathbf{Q},\ \mathbf{A}_i \in \mathbb{S}_n\ \ i \in I_m$ are given matrices and $\mathbf{b} = \{b_i\} \in \mathbb{R}^m$ is a given vector, then by the notation $\mathbf{Q} \bullet \mathbf{S} = \langle \mathbf{Q}; \mathbf{S}\rangle = \mathrm{tr}(\mathbf{Q}\cdot\mathbf{C}) = \mathbf{Q} : \mathbf{C}$, the relaxed canonical dual problem can be written as

$$\min \mathbf{Q} \bullet \mathbf{S} \quad \text{s.t.}\ \ \mathbf{S} \succeq 0, \quad \mathbf{A}_i \bullet \mathbf{S} = b_i, \quad \forall i \in I_m, \tag{103}$$

which is a typical Semi-Definite Programming (SDP) problem in optimization [134]. This shows that the popular SDP problem is indeed a special case of the canonical duality-triality theory for solving the general global optimization problem (3). The SDP method and algorithms have been well-studied in global optimization. But this method provides only a lower bound approach for the global minimal solution to $(\mathscr{P})$ if its canonical dual has no stationary point in $\mathscr{S}_c^+$. Also, in many real-world applications, the local solutions are also important. Therefore, a second method is needed.

(2) Quadratic perturbation and canonical primal-dual algorithm. By introducing a quadratic perturbation, the total complementary function (25) can be written as

$$\begin{aligned}\Xi_{\delta_k}(\boldsymbol{\chi}, \mathbf{S}) &= \Xi(\boldsymbol{\chi}, \mathbf{S}) + \frac{1}{2}\delta_k \|\boldsymbol{\chi} - \boldsymbol{\chi}_k\|^2 \\ &= \frac{1}{2}\langle \boldsymbol{\chi}, \mathbf{G}_{\delta_k}(\mathbf{S})\boldsymbol{\chi}\rangle - \Phi^*(\mathbf{S}) - \langle \boldsymbol{\chi}, \mathbf{F}_{\delta_k}(\mathbf{S})\rangle + \frac{1}{2}\delta_k\langle \boldsymbol{\chi}_k, \boldsymbol{\chi}_k\rangle,\end{aligned}$$

where $\delta_k > 0$, $\boldsymbol{\chi}_k\ \ k \in I_p$ are perturbation parameters, $\mathbf{G}_{\delta_k}(\mathbf{S}) = \mathbf{G}(\mathbf{S}) + \delta_k \mathbf{I}$, $\mathbf{F}_{\delta_k}(\mathbf{S}) = \mathbf{F}(\mathbf{S}) + \delta_k \boldsymbol{\chi}_k$. Thus, the original canonical dual feasible space $\mathscr{S}_c^+$ can be enlarged to $\mathscr{S}_{\delta_k}^+ = \{\mathbf{S} \in \mathscr{S}_c |\ \mathbf{G}_{\delta_k}(\boldsymbol{\varsigma}) \succ 0\}$. Using the perturbed total complementary function Ξ_{δ_k}, the perturbed canonical dual problem can be proposed

$$(\mathscr{P}_k^d): \quad \max\left\{\min\{\Xi_{\delta_k}(\boldsymbol{\chi}, \mathbf{S}) |\ \ \boldsymbol{\chi} \in \mathscr{X}_a\} |\ \ \mathbf{S} \in \mathscr{S}_{\delta_k}^+\right\} \tag{104}$$

Based on this perturbed canonical dual problem, a canonical primal-dual algorithm has been developed [141, 145].

Canonical Primal-Dual Algorithm. Given initial data $\delta_0 > 0$, $\boldsymbol{\chi}_0 \in \mathscr{X}_a$, and error allowance $omega > 0$. Let $k = 1$.

(1) Solve the perturbed canonical dual problem $(\mathscr{P}_k^d)$ to obtain $\mathbf{S}_k \in \mathscr{S}_{\delta_k}^+$.
(2) Computer $\bar{\boldsymbol{\chi}}_k = [\mathbf{G}_{\delta_k}(\mathbf{S}_k)]^{-1}\mathbf{F}_{\delta_k}(\mathbf{S}_k)$ and let

$$\boldsymbol{\chi}_k = \boldsymbol{\chi}_{k-1} + \beta_k(\bar{\boldsymbol{\chi}}_k - \boldsymbol{\chi}_{k-1}), \quad \beta_k \in [0, 1].$$

(3) If $|\Pi(\boldsymbol{\chi}_k) - \Pi(\boldsymbol{\chi}_{k-1})| \leq \omega$, then stop, $\boldsymbol{\chi}_k$ is the optimal solution to $(\mathscr{P})$. Otherwise, let $k = k + 1$, go back to 1).

In this algorithm, $\{\beta_k\}$ are given the parameters, which change the search directions. Clearly, if $\beta_k = 1$, we have $\boldsymbol{\chi}_k = \bar{\boldsymbol{\chi}}_k$. This algorithm has been used successfully for solving a class of benchmark problems and sensor network optimization problems [127, 145].

Let $\mathscr{S}^-_{\delta_k} = \{\mathbf{S} \in \mathscr{S}_c | \; \mathbf{G}_{\delta_k}(\mathbf{S}) \prec 0\}$. The combination of this algorithm with the double-min and double-max dualities in the triality theory can be used for finding local optimal solutions [12].

6 Challenges and Breakthrough

In the history of sciences, a ground breaking theory usually has to pass through serious arguments and challenges. This is duality nature and certainly true for the canonical duality-triality theory, which has benefited from recent challenges by M. Voisei, C. Zălinescu and his former student R. Strugariu in a set of 11 papers. These papers fall naturally into three interesting groups.

6.1 Group 1: Bi-level Duality

One paper in this group by Voisei and Zălinescu [138] challenges Gao and Yang's work for solving the following minimal distance between two surfaces [82]

$$\min \left\{ \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|^2 \;|\; g(\mathbf{x}) = 0, \quad h(\mathbf{y}) = 0, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \right\}, \tag{105}$$

where $g(\mathbf{x})$ is convex, while $h(\mathbf{y})$ is a nonconvex function. By the canonical transformation $h(\mathbf{y}) = V(\Lambda(\mathbf{y})) - \mathbf{y}^T\mathbf{f}$, the Gao–Strang complementary function was written in the form of $\Xi(\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}_0, \varsigma)$, where $\boldsymbol{\sigma}_0 = \{\lambda, \mu)$ is the first level canonical dual variable, i.e., the Lagrange multiplier for $\{g(\mathbf{x}) = 0, \; h(\mathbf{y}) = 0\}$, while ς is the second level canonical dual variable for the nonconvex constraint (see Eq. (11) in [82]). Using one counterexample

$$g(\mathbf{x}) = \frac{1}{2}(\|\mathbf{x}\|^2 - 1), \quad h(\mathbf{y}) = \frac{1}{2}\left(\frac{1}{2}\|\mathbf{y} - \mathbf{c}\|^2 - 1\right)^2 - \mathbf{f}^T(\mathbf{y} - \mathbf{c}), \tag{106}$$

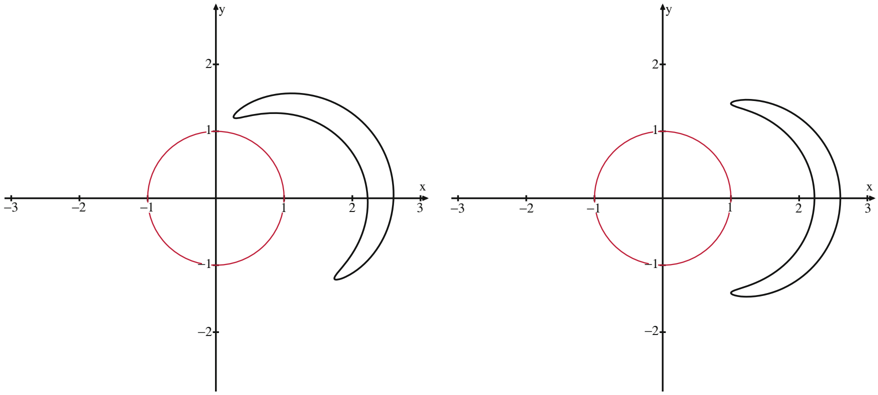


Fig. 4 Perturbations for breaking symmetry with $k = 64$ (*left*) and $k = 10^5$ (*right*)

with $n = 2$ and $\mathbf{c} = (1, 0)$, $\mathbf{f} = (\frac{\sqrt{6}}{96}, 0)$, Voisei and Zălinescu proved that "the main results in Gao and Yang [82] are false" and they concluded: "The consideration of the function Ξ is useless, at least for the problem studied in [82]."

This paper raises up two issues on different levels.

The first issue is elementary: there is indeed a mistake in Gao and Yang's work, i.e., instead of $(\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}_0, \varsigma)$ used in [82], the variables in the total complementary function Ξ should be the vectors $\boldsymbol{\chi} = (\mathbf{x}, \mathbf{y})$ and $(\boldsymbol{\sigma}_0, \varsigma)$ since $\Xi(\boldsymbol{\chi}, \boldsymbol{\sigma}_0, \varsigma)$ is convex in $\mathbf{x}$ and $\mathbf{y}$ but may not in $\boldsymbol{\chi}$. This mistake has been easily corrected in [114]. Therefore, the duality on this level is: opposite to Voisei and Zălinescu's conclusion, the consideration of the Gao–Strang total complementary function Ξ is indeed quite useful for solving the challenging problem (105) [114].

The second issue is crucial. The "counterexample" (106) has two global minimal solutions due to the symmetry (see Fig. 4). Similar to Example 1, the canonical dual problem (33) $\max\{\Pi^d(\boldsymbol{\sigma}_0, \varsigma) | (\boldsymbol{\sigma}_0, \varsigma) \in \mathcal{S}_c^+\}$ has two stationary points on the boundary of $\mathcal{S}_c^+$ (cf. Fig. 1b). Such case has been discussed by Gao in integer programming problem [64]. It was first realized that many so-called NP-hard problems in global optimization usually have multiple global minimal solutions and a conjecture was proposed in [64], i.e., a global optimization problem is NP-hard if its canonical dual has no stationary point in $\mathcal{S}_c^+$. In order to solve such challenging problems, different perturbation methods have been suggested with successful applications in global optimization [75, 127, 128, 140], including a recent paper on solving hard case of a trust region subproblem [14]. For this problem, by simply using linear perturbation $\mathbf{f}_k = (\frac{\sqrt{6}}{96}, \frac{1}{k})$ with $|k| \gg 1$, both global minimal solutions can be easily obtained by the canonical duality-triality theory [114] (see Figs. 1a and 4). Therefore, the duality on this level is: Voisei and Zălinescu's "counterexample" does not contradict the canonical duality-triality theory even in this crucial case.

Actually, by the general model (3), the nonconvex hypersurface $h(\mathbf{y})$ in this paper can be written as $h(\mathbf{y}) = W(D\mathbf{y}) - F(\mathbf{y})$, where the double-well function $W(D\mathbf{y})$

is objective (also isotropic), which represents the modeling with symmetry; while the linear term $F(\mathbf{y})$ is a subjective function, which breaks the symmetry and leads to a particular problem. By the fact that nothing in this world is perfect, therefore, any real-world problem must have certain input or defects. This simple truth lays a foundation for the perturbation method and the triality theory for solving challenging problems. However, this fact is not well-recognized in mathematical optimization and computational science,[10] it turns out that many challenges and NP-hard problems are artificially proposed.

6.2 Group 2: Conceptual Duality

Of four papers in this group, two were published in pure math journals [133, 136] and two were rejected by applied math journals (*ZAMP* and *Q.J. Mech. Appl. Math*). The paper by Voisei and Zălinescu [136] challenges Gao and Strang's original work [77] on solving the general nonconvex variational problem (3) in finite deformation theory. As we discussed in Sect. 2.2 that the stored energy $W(\boldsymbol{\varepsilon})$ must be objective and can't be linear, the deformation operator Λ should be geometrically admissible in order to have the canonical transformation $W(\boldsymbol{\varepsilon}) = \Phi(\Lambda(\boldsymbol{\varepsilon}))$, and the external energy $F(\boldsymbol{\chi})$ must be linear such that $\bar{\boldsymbol{\chi}}^* = \partial F(\boldsymbol{\chi})$ is the given input. Oppositely, by listing total six counterexamples, Voisei and Zălinescu choose a piecewise linear function $g(u,v) = \{u \text{ (if } v = u^2) \text{ ; } 0 \text{ (otherwise)}\}$ as $\Phi(\boldsymbol{\xi})$, a parametric function $f(t) = (t, t^2)$ as the geometrically nonlinear operator $\Lambda(t)$ (see Example 3.1 in [136]), and quadratic functions as $F(\boldsymbol{\chi})$ (see Examples 3.2, 3.4, 3.5, and 3.6 in [136]). While in the rest counterexample (Example 3.3 in [136]), they simply let the external energy $F(u) = 0$ and $\Lambda(u) = u^2 - u$.

Clearly, the piecewise linear function listed by Voisei and Zălinescu is not objective and can't be the stored energy for any real material. Also, both $\Lambda(t)$ and $\Lambda(u)$ are simply not strain measures. Such conceptual mistakes are repeatedly made in their recent papers, say in the paper by Strugariu, Voisei, and Zălinescu (Example 3.3 in [133]), they let $(x(t), y(t)) = \Lambda(t) = (\frac{1}{2}t^2, t)$ be the geometrical mapping $\boldsymbol{\xi}(t) = \Lambda(t)$ and, in their notation, $f(x,y) = xy^3(x^2 + (x - y^4)^2)^{-1}$ as the stored energy $\Phi(\boldsymbol{\xi})$.

For quadratic $F(\boldsymbol{\chi})$, the input $\bar{\boldsymbol{\chi}}^* = \partial F(\boldsymbol{\chi})$ depends linearly on the output $\boldsymbol{\chi}$, which is called the *follower force*. In this case, the system is not conservative and the traditional variational methods do not apply. In order to study such nonconservative minimization problems, a so-called rate variational method and duality principle were proposed by Gao and Onat [73]. While for $F(\boldsymbol{\chi}) = 0$, the minimization $\min\{\Pi(\boldsymbol{\chi}) = W(D\boldsymbol{\chi})\}$ is not a problem but a modeling, which has either trivial solution $\boldsymbol{\chi} = 0$ or multiple solutions $\boldsymbol{\chi} = constant$ due to certain symmetry of the

[10] Indeed, one authors' paper [127] was first submitted to a computational optimization journal and received such a reviewer's comment: "the authors applied a perturbation, which changed the problem mathematically, ... and I suggest an immediate rejection."

mathematical modeling. This is a key mistake happened very often in global optimization, which leads to many man-made NP-hard problems as we discussed in the previous subsection.

The concept of a Lagrangian was introduced by J.L. Lagrange in analytic mechanics 1788, which has a standard notation in physics as (see [98])

$$L(\boldsymbol{\chi}) = T(\dot{\boldsymbol{\chi}}) - V(\boldsymbol{\chi}), \tag{107}$$

where T is the kinetic energy and V is the potential energy. By the Legendre transformation $T^*(\mathbf{p}) = \langle \dot{\boldsymbol{\chi}}, \mathbf{p} \rangle - T(\dot{\boldsymbol{\chi}})$, the Lagrangian is also written as

$$L(\boldsymbol{\chi}, \mathbf{p}) = \langle \dot{\boldsymbol{\chi}}, \mathbf{p} \rangle - T^*(\mathbf{p}) - V(\boldsymbol{\chi}). \tag{108}$$

It is commonly known that for problems with linear potential $V(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle$, the Lagrangian $L(\boldsymbol{\chi})$ is convex and $L(\boldsymbol{\chi}, \mathbf{p})$ is a saddle point functional which leads to a well-known min–max duality in convex systems.

But for problems with convex potential $V(\boldsymbol{\chi})$, the Lagrangian $L(\boldsymbol{\chi})$ is a d.c. function (difference of convex functions) and $L(\boldsymbol{\chi}, \mathbf{p})$ is not a saddle functional any more. In this case, the Hamiltonian $H(\boldsymbol{\chi}, \mathbf{p}) = \langle \dot{\boldsymbol{\chi}}, \mathbf{p} \rangle - L(\boldsymbol{\chi}, \mathbf{p}) = T^*(\mathbf{p}) + V(\boldsymbol{\chi})$ is convex. Therefore, a *bi-duality* (i.e., the combination of the double-min and double-max dualities) was proposed in convex Hamilton systems (see Chap. 2 [49]). However, in the paper by Strugariu, Voisei, and Zălinescu [133], the function

$$L(x, y) = \langle a, x \rangle \langle b, y \rangle - \frac{1}{2}\alpha \|x\|^2 - \frac{1}{2}\beta \|y\|^2$$

is defined as the "Lagrangian", by which, they produced several "counterexamples" for the bi-duality in convex Hamilton systems. In this "Lagrangian", if we consider $V(x) = \frac{1}{2}\alpha\|x\|^2$ as a potential energy and $T^*(y) = \frac{1}{2}\beta\|y\|^2$ as the complementary kinetic energy, but the term $\langle a, x\rangle\langle b, y\rangle$ is not the bilinear form $\langle Dx; y\rangle$ required in Lagrange mechanics, where D is a differential operator such that Dx and y form a (constitutive) duality pair. This term does not make any sense in Lagrangian mechanics [98] and duality theory [22]. Therefore, the "Lagrangian" used by Strugariu, Voisei, and Zălinescu for producing counterexamples of the bi-duality theory is not the Lagrangian used in Gao's book [49], i.e., the standard Lagrangian in classical mechanics [98, 122], convex analysis [22], and modern physics [20, 93]. Actually, the bi-duality theory in finite dimensional space is a corollary of the so-called *Iso-Index Theorem* and the proof was given in Gao's book (see Theorem 5.3.6 and Corollary 5.3.1 [49]).

Papers in this group show a big gap between mathematical physics/analysis and optimization. As V.I. Arnold said [3]: "In the middle of the twentieth century it was attempted to divide physics and mathematics. The consequences turned out to be catastrophic."

6.3 Group 3: Anti-triality

Six papers are in this group on the triality theory. By listing simple counterexamples (cf. e.g., [137]), Voisei and Zălinescu claimed: "a correction of this theory is impossible without falling into trivial".[11] However, even some of these counterexamples are correct, they are not new. This type of counterexamples was first discovered by Gao in 2003 [57, 58], i.e., the double-min duality holds under certain additional constraints (see Remark on page 288 [57] and Remark 1 on page 481 [58]). But neither [57] nor [58] was cited by Voisei and Zălinescu in their papers.

As mentioned in Sect. 2.4, the triality was proposed originally from post-buckling analysis [42] in "either-or" format since the double-max duality is always true but the double-min duality was proved only in one-dimensional nonconvex analysis [49]. As a corollary of an *Iso-Index Theorem* given in Gao's book, the double-min and double-max duality statements were first proved for nonconvex optimization problems in finite dimensional space (see Theorem 5.3.6 in [49]). Recently, this double-min duality has been proved for polynomial optimization [79, 112, 113], and for general global optimization problems [15, 80]. The "certain additional constraints" are simply the dimensions of the primal problem and its canonical dual should be the same in order to have strong double-min duality. Otherwise, this double-min duality holds weakly in subspaces with elegant symmetrical forms. Therefore, the triality theory now has been proved in global optimization, which should play important roles for solving NP-hard problems in complex systems.

7 Concluding Remarks and Open Problems

In this article we have discussed the existing gaps between nonconvex analysis/mechanics and global optimization. Common misunderstandings and confusions on some basic concepts have been addressed and clarified, including the objectivity, nonlinearity, and Lagrangian. the canonical duality is a fundamental law in nature, the canonical duality-triality theory is indeed powerful for unified understanding complicated phenomena and solving challenging problems. So far, this theory can be summarized for having the following functions:

1. To correctly model complex phenomena in multiscale systems within a unified framework [49, 57, 83].
2. To solve a large class of nonconvex/nonsmooth/discrete global optimization problems for obtaining both global and local optimal solutions.
3. To reformulate certain nonlinear partial differential equations in algebraic forms with possibility to obtain all possible analytical solutions [50, 66, 71, 72].

---

[11] This sentence is deleted by Voisei and Zălinescu in their revision of [137] after they were informed by referees that their counterexamples are not new and the triality theory has been proved.

4. To understand and identify certain NP-hard problems, i.e., the general global optimization problems are not NP-hard if they can be solved by the canonical duality-triality theory [64, 75, 127].
5. To understand and solve nonlinear (chaotic) dynamic systems by obtaining global stable solutions [108, 126].
6. To check and verify correctness of existing modeling and theories.

Based on the canonical duality-triality theory, a unified understanding for bifurcation, chaos, and NP-hard problems is given in a recent review article [67]. There are still many open problems existing in the canonical duality-triality theory. Here we list a few of them.

1. Sufficient condition for the existence of the canonical dual solutions on $\mathscr{S}_c^+$.
2. NP-Harness conjecture: A global optimization problem is NP-hard if its canonical dual $\Pi^d(\boldsymbol{\xi}^*)$ has no stationary point on the closed domain $\bar{\mathscr{S}}_c^+ = \{\boldsymbol{\xi}^* \in \mathscr{S}_a | \ \mathbf{G}(\boldsymbol{\xi}^* \succeq 0\}$.
3. Extremality conditions for stationary points of $\Pi^d(\boldsymbol{\xi}^*)$ on the domain such that $\mathbf{G}(\boldsymbol{\xi}^*)$ is indefinite in order to identify all local extrema.
4. Bi-duality and triality theory for d-dimensional ($d > 1$) nonconvex analysis problems.

The following research topics are challenging:

1. Canonical duality-triality theory for solving bi-level optimization problems.

2. Using least-squares method and canonical duality theory for solving 3-dimensional chaotic dynamical problems, such as Lorenz system and Navier–Stokes equation, etc.

3. Perturbation methods for solving NP-hard integer programming problems, such as quadratic Knapsack problem, TSP, and mixed integer nonlinear programming problems.

4. Unilateral post-buckling problem of the Gao nonlinear beam

$$\min_{\chi \in \mathscr{X}_a} \left\{ \Pi(\chi) = \int_0^L \left[\frac{1}{2} E I \chi_{xx}^2 + \frac{1}{12} \alpha E \chi_x^4 - \frac{1}{2} \lambda E \chi_x^2 - f \chi \right] \mathrm{d}x \ \mid \ \chi(x) \geq 0 \right\}. \tag{109}$$

Due to the axial compressive load $\lambda > 0$, the downward lateral load $f(x)$ and the unilateral constraint $\chi(x) \geq 0 \ \ \forall x \in [0, L]$, the solution of this nonconvex variational problem is a local minimizer of $\Pi(\chi)$ which can be obtained numerically by the canonical dual finite element methods [12, 129] if λ and f are not big enough such that $\chi(x) > 0 \ \ \forall x \in [0, L]$. However, if the buckling state $\chi(x) = 0$ happens at any $x \in [0, L]$, the problem could be NP-hard. The open problems include:

(1) under what conditions for the external loads $\lambda > 0$ and $f(x)$, the problem has a solution $\chi(x) > 0 \ \ \forall x \in [0, L]$?

(2) how to solve the unilateral buckling problem when $\chi(x) = 0$ holds for certain $x \in [0, L]$?

Acknowledgements This paper is based on a series of plenary lectures presented at international conferences of mathematics, mechanics and global optimization during 2012–2014. Invitations from organizers are sincerely acknowledged. The research was supported by a grant (AFOSR FA9550-10-1-0487) from the US Air Force Office of Scientific Research.

References

1. Ahn, J., Kuttler, K.L., Shillor, M.: Dynamic contact of two Gao beams. Electron. J. Differ. Equ. **194**, 1–42 (2012)
2. Andrews, K.T., Dumont, Y., M'Bengue, M.F., Purcell, J.M., Shillor, M.: Analysis and simulations of a nonlinear elastic dynamic beam. Z. Angew. Math. Phys. **63**, 1005–1019 (2012)
3. Anorld, V.I.: On teaching mathematics. Russ. Math. Surv. **53**(1), 229–236 (1998)
4. Aspnes, J., Goldberg, D., Yang, Y.R.: On the computational complexity of sensor network localization. Lecture Notes in Computer Science, vol. 3121, pp. 32–44. Springer (2004)
5. Atiyah, M.F.: Duality in Mathematics and Physics, lecture notes from the Institut de Matematica de la Universitat de Barcelona (IMUB) (2007)
6. Bader, B.W., Kolda, T.G.: Algorithm 862: MATLAB tensor classes for fast algorithm prototyping. ACM Trans. Math. Softw. **32**(4), 635–653 (2006)
7. Ball, J.M.: Some open problems in elasticity. Geometry, Mechanics, and Dynamics, pp. 3–59. Springer, New York (2002)
8. Bendsoe, M.P., Sigmund, O.: Topology Optimization: Theory, Methods and Applications. 370 pp. Springer, Berlin (2003)
9. Bengue, F.M., Shillor, M.: Regularity result for the problem of vibrations of a nonlinear beam. Electron. J. Differ. Equ. **27**, 1–12 (2008)
10. Burachik, R.S., Freire, W.P., Kaya, C.Y.: Interior Epigraph Directions method for nonsmooth and nonconvex optimization via generalized augmented Lagrangian duality. J. Glob. Optim. **60**, 501–529 (2014)
11. Burer, S., Letchford, A.N.: Non-convex mixed-integer nonlinear programming: a survey. Surv. Oper. Res. Manag. Sci. **17**, 97–106 (2012)
12. Cai, K., Gao, D.Y., Qin, Q.H.: Post-buckling solutions of hyper-elastic beam by canonical dual finite element method. Math. Mech. Solids **19**(6), 659–671 (2014)
13. Chen, Y.: Canonical Duality Theory for General Global Optimization Problems and Applications. PhD. Thesis, School of Science, Information Technology, and Engineering, Federation University Australia (2014)
14. Chen Y., Gao, D.Y.: Global solutions to spherical constrained quadratic minimization via canonical dual approach (2014). http://arXiv.org/abs/1308.4450
15. Chen, Y., Gao, D.Y.: Global solutions to nonconvex optimization of 4th-order polynomial and log-sum-exp functions. J. Glob. Optim. (2014). doi:10.1007/s10898-014-0244-5
16. Ciarlet, P.G.: Mathematical Elasticity, Volume 1: Three Dimensional Elasticity, 449 pp. North-Holland (1998)
17. Ciarlet, P.G.: Linear and Nonlinear Functional Analysis with Applications. SIAM, Philadelphia (2013)
18. Copeland, E.J., Kibble, T.W.B.: Cosmic strings and superstrings. Proc. R. Soc. A **466**(2115), 623–657 (2010)
19. Clark, D.M., Davey, B.A.: Natural Dualities for the Working Algebraist, 356 pp. Cambridge University Press, Cambridge (1998)
20. Davier, M.: The Higgs boson of the standard model. C. R. Phys. **8**, 986–998 (2007)
21. Ekeland, I.: Nonconvex duality. In: Gao, D.Y. (ed.) Proceedings of IUTAM Symposium on Duality, Complementarity and Symmetry in Nonlinear Mechanics, pp. 13–19. Kluwer Academic Publishers (2004)

22. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland, Amsterdam (1976)
23. Fang, S.C., Gao, D.Y., Sheu, R.L., Wu, S.Y.: Canonical dual approach to solving 0–1 quadratic programming problems. J. Ind. Manag. Optim. **4**(1), 125–142 (2008)
24. Fang, S.C., Gao, D.Y., Sheu, R.L., Xing, W.: Global optimization for a class of fractional programming problems. J. Glob. Optim. **45**(3), 337–353 (2009)
25. Feng, X., Glowinski, R., Neilan, M.: Recent developments in numerical methods for fully nonlinear second order partial differential equations. SIAM Rev. **55**(2), 205–267 (2013)
26. Floudas, C.A., Gounaris, C.E.: A review of recent advances in global optimization. J. Glob. Optim. **45**, 3–38 (2009)
27. Floudas, C.A., Pardalos, P.M.: Encyclopedia of Optimization, 2nd edn. Springer, New York (2009)
28. Gao, D.Y.: On the complementary bounding theorems for limit analysis. Int. J. Solids Struct. **24**, 545–556 (1988)
29. Gao, D.Y.: Panpenalty finite element programming for limit analysis. Comput. Struct. **28**, 749–755 (1988)
30. Gao, D.Y.: Dynamically loaded rigid-plastic analysis under large deformation. Q. Appl. Math. **48**, 731–739 (1990)
31. Gao, D.Y.: On the extremum potential variational principles for geometrical nonlinear thin elastic shell. Sci. China (Scientia Sinica) (A) **33**(1), 324–331 (1990)
32. Gao, D.Y.: On the extremum variational principles for nonlinear elastic plates. Q. Appl. Math. **48**, 361–370 (1990)
33. Gao, D.Y.: Complementary principles in nonlinear elasticity. Sci. China (Scientia Sinica) (A) (Chinese Ed.) **33**(4), 386–394 (1990)
34. Gao, D.Y.: Bounding theorem on finite dynamic deformations of plasticity. Mech. Res. Commun. **17**, 33–39 (1990)
35. Gao, D.Y.: Extended bounding theorems for nonlinear limit analysis. Int. J. Solids Struct. **27**, 523–531 (1991)
36. Gao, D.Y.: Global extremum criteria for nonlinear elasticity. J. Appl. Math. Phys. (ZAMP) **43**, 924–937 (1992)
37. Gao, D.Y.: Duality theory in nonlinear buckling analysis for von Karman equations. Stud. Appl. Math **94**, 423–444 (1995)
38. Gao, D.Y.: Limit analysis of plastic shells subjected to large deflections. Eur. J. Mech., A/Solids **14**(3), 459–472 (1995)
39. Gao, D.Y.: Nonlinear elastic beam theory with applications in contact problem and variational approaches. Mech. Res. Commun. **23**(1), 11–17 (1996)
40. Gao, D.Y.: Complementary finite element method for finite deformation nonsmooth mechanics. J. Eng. Math. **30**, 339–353 (1996)
41. Gao, D.Y.: Complementarity and duality in natural sciences. Philosophical Study in Modern Science and Technology (in Chinese), pp. 12–25. Tsinghua University Press, Beijing (1996)
42. Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory. Appl. Mech. Rev. **50**(11), S64–S71 (1997)
43. Gao, D.Y.: Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications. IMA J. Appl. Math. **61**, 199–235 (1998)
44. Gao, D.Y.: Bi-complementarity and duality: A framework in nonlinear equilibria with applications to the contact problems of elastoplastic beam theory. J. Math. Anal. Appl. **221**, 672–697 (1998)
45. Gao, D.Y.: Pure complementary energy principle and triality theory in finite elasticity. Mech. Res. Comm. **26**(1), 31–37 (1999)
46. Gao, D.Y.: Duality-Mathematics. Wiley Encyclopedia of Electronical and Electronical Engineering **6**, 68–77 (1999)
47. Gao, D.Y.: General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics. Meccanica **34**, 169–198 (1999)

48. Gao, D.Y.: Finite deformation beam models and triality theory in dynamical post-buckling analysis. Int. J. Non-Linear Mech. **35**, 103–131 (2000)
49. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications, xviii + 454 pp. Springer, New York (2000)
50. Gao, D.Y.: Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications. Nonlinear Anal. **42**(7), 1161–1193 (2000)
51. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
52. Gao, D.Y.: Finite deformation beam models and triality theory in dynamical post-buckling analysis. Int. J. Non-Linear Mech. **5**, 103–131 (2000)
53. Gao, D.Y.: Bi-duality in nonconvex optimization. In: Floudas, C.A., Pardalos, P.D. (eds.) Encyclopedia of Optimization, vol. 1, pp. 477–482. Springer, New York (2001)
54. Gao, D.Y.: Tri-duality in global optimization. In: Floudas, C.A., Pardalos, P.D. (eds.) Encyclopedia of Optimization, vol. 1, pp. 485–491. Springer, New York (2001)
55. Gao, D.Y.: Complementarity, polarity and triality in nonsmooth, nonconvex and nonconservative Hamilton systems. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. **359**, 2347–2367 (2001)
56. Gao, D.Y.: Duality and triality in non-smooth, nonconvex and nonconservative systems: a survey, new phenomena and new results. In: Baniotopoulos, C. (ed.) Nonsmooth/Nonconvex Mechanics with Applications in Engineering, pp. 1–14. Thessaloniki, Greece (2002)
57. Gao, D.Y.: Nonconvex semi-linear problems and canonical dual solutions. In: Gao, D.Y., Ogden, R.W. (eds.) Advances in Mechanics and Mathematics, vol. II, pp. 261–312. Springer (2003)
58. Gao, D.Y.: Perfect duality theory and complete solutions to a class of global optimization problems. Optimization **52**(4–5), 467–493 (2003)
59. Gao, D.Y.: Canonical duality theory and solutions to constrained nonconvex quadratic programming. J. Glob. Optim. **29**, 377–399 (2004)
60. Gao, D.Y.: Complementary variational principle, algorithm, and complete solutions to phase transitions in solids governed by Landau-Ginzburg equation. Math. Mech. Solid **9**, 285–305 (2004)
61. Gao, D.Y.: Sufficient conditions and perfect duality in nonconvex minimization with inequality constraints. J. Ind. Manag. Optim. **1**, 59–69 (2005)
62. Gao, D.Y.: Complete solutions and extremality criteria to polynomial optimization problems. J. Glob. Optim. **35**, 131–143 (2006)
63. Gao, D.Y.: Duality in distributed-parameter control of nonconvex and nonconservative dynamical systems with applications. Nonlinear Dyn. Syst. Theory **6**, 257–279 (2006)
64. Gao, D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. J. Indust. Manag. Optim. **3**(2), 293–304 (2007)
65. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solutions for global optimization. Comput. Chem. Eng. **33**, 1964–1972 (2009)
66. Gao, D.Y.: Analytical solutions to general anti-plane shear problems in finite elasticity. Contin. Mech. Thermodyn. **28**, 175–194 (2016)
67. Gao, D.Y.: Canonical duality-triality: unified understanding on bifurcation, chaos, and NP-hard problems in complex systems. To appear
68. Gao, D.Y., Hajilarov, E.: On analytic solutions to 3-d finite deformation problems governed by St Venant-Kirchhoff material. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Canonical Duality Theory. Springer (2016). arXiv:1504.02865
69. Gao, D.Y., Lu, X.J.: An optimal mass transfer problem in the limit case. Math. Method Appl, Sci (2014)
70. Gao, D.Y., Motreanu, D.: Handbook of Nonconvex Analysis and Applications, 680 pp. International Press, Boston (2010)
71. Gao, D.Y., Ogden, R.W.: Multi-solutions to non-convex variational problems with implications for phase transitions and numerical computation. Q. J. Mech. Appl. Math. **61**, 497–522 (2008)
72. Gao, D.Y., Ogden, R.W.: Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem. Zeits. Ang. Math. Physik **59**, 498–517 (2008)

73. Gao, D.Y., Onat, E.T.: Rate variational extremum principles for finite elastoplasticity. Appl. Math. Mech. **11**(7), 659–667 (1990)
74. Gao, D.Y., Ruan, N.: Solutions and optimality criteria for nonconvex quadratic-exponential minimization problem. Math. Method Oper. Res. **67**(3), 479–491 (2008)
75. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. **47**, 463–484 (2010)
76. 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 (2009)
77. Gao, D.Y., Strang, G.: Geometric nonlinearity: Potential energy, complementary energy, and the gap function. Q. Appl. Math. **47**(3), 487–504 (1989)
78. Gao, D.Y., Strang, G.: Dual extremum principles in finite deformation elastoplastic analysis. Acta Appl. Math. **17**, 257–267 (1989)
79. 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)
80. Gao, D.Y., Wu, C.: On the triality theory in global optimization. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Advances in Canonical Duality Theory. Springer (2016). arXiv:1104.2970
81. Gao, D.Y., Yang, W.H.: Multi-duality in minimal surface type problems. Stud. Appl. Math. **95**, 127–146. MIT (1995)
82. Gao, D.Y., Yang, W.C.: Complete solutions to minimal distance problem between two nonconvex surfaces. Optimization **57**(5), 705–714 (2008)
83. 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)
84. Gao, D.Y., Ogden, R.W., Stavroulakis, G.: Nonsmooth and Nonconvex Mechanics: Modelling. Analysis and Numerical Methods. Kluwer Academic Publishers, Dordrecht (2001)
85. Gao, D.Y., Ruan, N., Sherali, H.: Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality. J. Glob. Optim. **45**, 473–497 (2009)
86. Gao, D.Y., Ruan, N., Sherali, H.: Canonical dual solutions for fixed cost quadratic programs. In: Chinchuluun, A., et al. (eds.) Optimization and Optimal Control. Springer Optimization and Its Applications, vol. 39, pp. 139–156 (2010)
87. Gao, D.Y., Ruan, N., Pardalos, P.: Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization. In: Boginski, V., Commander, C.W., Pardalos, P., Ye, Y. (eds.) Sensors: Theory, Algorithms, and Applications, vol. 61, pp. 37–54. Springer, New York (2012)
88. Gao, D.Y., Watson, L.T., Easterling, D.R., Thacker, W.I., Billups, S.C.: Solving the canonical dual of box- and integer-constrained nonconvex quadratic programs via a deterministic direct search algorithm. Optim. Methods Softw. **28**(2), 313–326 (2013)
89. Gao, T.K.: Complete solutions to a class of eighth-order polynomial optimization problems. IMA J. Appl. Math. (2013). Published online July 21, 2013 doi:10.1093/imamat/hxt033
90. Grossmann, I.E.: Review of nonlinear mixed-integer and disjunctive programming techniques. Optim. Eng. **3**, 227–252 (2002)
91. Holzapfel, G.A.: Nonlinear Solid Mechanics: A Continuum Approach for Engineering. Wiley, Chichester (2000). ISBN 978-0471823193
92. Ioniță, A.: Finite Element Analysis of the Deformation of a Rubber Diaphragm, Ph.D. Thesis, Department of Engineering Science and Mechanics, Virginia Tech. 142 pp (2001)
93. Jaffe, A., Witten, E.: The millennium prize problems. In: Carlson, J., Jaffe, A., Wiles, A. (eds.) Quantum Yang-Mills Theory, pp. 129–152. Clay Mathematics Institute (2006)
94. Karp, R.: Reducibility among combinatorial problems. In: Miller, R.E., Thatcher, J.W. (eds.) Complexity of Computer Computations, pp. 85–103. Plenum Press, New York (1972)
95. Koiter, W.T.: On the complementary energy theorem in nonlinear elasticity theory. In: Fichera, G. (ed.) Trends in the Application of Pure Mathematics to Mechanics. Pitman (1976)
96. Kuttler, K.L., Purcell, J., Shillor, M.: Analysis and simulations of a contact problem for a nonlinear dynamic beam with a crack. Q. J. Mech. Appl. Math. **65**, 1-2-5 (2012)

97. Lagrange, J.L.: Méchanique analitique, La veuve Desaint, Paris. Mecanique analytique, vol. 2, 2nd edn, pp. 1811–1815. Courcier, Paris (1788)
98. Landau, L.D., Lifshitz, E.M.: Mechanics, 3rd edn. Butterworth-Heinemann (1976). ISBN 978-0-750-62896-9
99. Latorre, V.: A Potential Reduction Method for Canonical Duality, with an Application to the Sensor Network Localization Problem (2014). Preprint available at arXiv:1403.5991
100. Latorre, V., Gao, D.Y.: Canonical duality for RBF neural networks. Neuralcomputings **134**, 189–197 (2014)
101. Latorre, V., Gao, D.Y.: Canonical duality for solving general nonconvex constrained problems. Optim. Lett. (2015). http://link.springer.com/article/10.1007/s11590-015-0860-0/fulltext.html
102. Latorre, V., Gao, D.Y.: Global optimal trajectory in chaos and NP-hardness (2015). arXiv:1512.08343
103. Latorre, V., Sagratella, S.: A canonical duality approach for the solution of affine quasi-variational inequalities. J. Glob. Optim. **1**, 1 (2014). doi:10.1007/s10898-014-0236-5
104. Lee, S.J., Shield, R.T.: Variational principles in finite elastics. J. Appl. Math. Phys. (ZAMP) **31**, 437–453 (1980)
105. Levinson, M.: The complementary energy theorem in finite elasticity. Trans. ASME, Ser. E. J. Appl. Mech. **87**, 826–828 (1965)
106. Li, J., Kuttler, K.L., Shillor, M.: Existence for dynamic contact of a stochastic viscoelastic Gao beam. Nonlinear Anal. Ser, B Real World Appl (2014)
107. Li, S.F., Gupta, A.: On dual configuration forces. J. Elast. **84**, 13–31 (2006)
108. Li, C., Zhou, X., Gao, D.Y.: Stable trajectory of logistic map. Nonlinear Dyn. (2014). doi:10.1007/s11071-014-1433-y
109. Liu, I.S.: Further remarks on Euclidean objectivity and the principle of material frame-indifference. Contin. Mech. Thermodyn. **17**, 125–133 (2005)
110. Machalova, J., Netuka, H.: Bending of a nonlinear beam reposing on an unilateral foundation. Appl. Comput. Mech. **5**, 45–54 (2011)
111. Marsden, J.E., Hughes, T.J.R.: Mathematical Foundations of Elasticity. Prentice-Hall, Englewood Cliffs (1983)
112. Morales-Silva, D.M., Gao, D.Y.: Complete solutions and triality theory to a nonconvex optimization problem with double-well potential in R^n. Numer. Algebra Contr. Optim. **3**(2), 271–282 (2013)
113. Morales-Silva, D.M., Gao, D.Y.: Canonical duality theory and triality for solving general nonconstrained global optimization problems. Math. Mech. Complex Syst. **3**(2), 139–161 (2015)
114. Morales-Silva, D.M., Gao, D.Y.: On the minimal distance between two surfaces. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Canonical Duality Theory: Unified Methodology for Multidisciplinary Study. Springer (2017). Reprint online at arXiv:1210.1618
115. Morris, S.A.: Pontryagin Duality and Structure of Locally Compact Abelian Groups, 128 pp. Cambridge University Press, Cambridge (1977)
116. Murdoch, A.I.: Objectivity in classical continuum physics: a rationale for discarding the 'principle of invariance under superposed rigid body motions' in favour of purely objective considerations. Contin. Mech. Thermodyn. **15**, 309–320 (2003)
117. Murdoch, A.I.: On criticism of the nature of objectivity in classical continuum physics. Contin. Mech. Thermodyn. **17**(2), 135–148 (2005)
118. Oden, J.T., Reddy, J.N.: On dual complementary variational principles in mathematical physics. Int. J. Eng. Sci. **12**, 1–29 (1973)
119. Oden, J.T., Reddy, J.N.: Variational Methods in Theoretical Mechanics. Springer, Berlin (1983)
120. Ogden, R.W.: Non-Linear Elastic Deformations. Ellis Horwood/Dover, Chichester/Mineola (1984/1997)
121. Pardalos, P.M., Vavasis, S.A.: Quadratic programming with one negative eigenvalue is NP-hard. J. Glob. Optim. **1**, 15–22 (1991)

122. Panza, M.: The origins of analytic mechanics in the 18th century. In: Jahnke, H.N. (ed.) A History of Analysis, pp. 137–153 (2003)
123. Pardalos, P.M., Romeijn, H.E., Tuy, H.: Recent developments and trends in global optimization. J. Comput. Appl. Math. **124**(1–2), 209–228 (2000)
124. Raoult, A.: Quasiconvex envelopes in nonlinear elasticity. In: Schröder, J., Neff, P. (eds.) Rank-One Convexity in Applied Mechanics, pp. 17–52. Springer (2010)
125. Reissner, E.: On a variational theorem for finite elastic deformations. J. Math. Phys. **32**, 129–153 (1953)
126. Ruan, N., Gao, D.Y.: Canonical duality approach for nonlinear dynamical systems. IMA J. Appl. Math. **79**, 313–325 (2014)
127. Ruan, N., Gao, D.Y.: Global optimal solutions to a general sensor network localization problem. Perform. Eval. **75–76**, 1–16 (2014)
128. 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)
129. Santos, H.A.F.A., Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Nonlinear Mech. **7**, 240–247 (2011)
130. Saxe, J.: Embeddability of weighted graphs in k-space is strongly NP-hard. In: Proceedings of 17th Allerton Conference in Communications, Control, and Computing, pp. 480–489. Monticello, IL (1979)
131. Strang, G.: Introduction to Applied Mathematics. Wellesley-Cambridge Press (1986)
132. Stumpf, H.: The derivation of the dual extremum and complementary stationary principles in geometrical nonlinear shell theory. Ing. Arch. **48**, 221–237 (1979)
133. Strugariu, R., Voisei, M.D., Zălinescu, C.: Counter-examples in bi-duality, triality and tri-duality. Discret. Contin. Dyn. Syst. Ser. A **31**(4), 1453–1468 (2011)
134. Todd, M.: Semidefinit optimization. Acta Numerica **10**, 515–560 (2001)
135. Vavasis, S.: Quadratic programming is in NP. Info. Proc. Lett. **36**, 73–77 (1990)
136. Voisei, M.D., Zălinescu, C.: Some remarks concerning Gao-Strang's complementary gap function. Appl. Anal. **90**(6), 1111–1121 (2011)
137. Voisei, M.D., Zălinescu, C.: Counter examples to some triality and tri-duality results. J. Glob. Optim. **49**, 173–183 (2011)
138. Voisei, M.D., Zălinescu, C.: A counter-example to "Minimal distance between two nonconvex surfaces". Optimization **60**(5), 593–602 (2011)
139. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Global extremal conditions for multi-integer quadratic programming. J. Ind. Manag. Optim. **4**(2), 213–225 (2008)
140. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Canonical dual approach to solving the maximum cut problem. J. Glob. Optim. **54**, 341–351 (2012)
141. Wu, C., Gao, D.Y.: Canonical primal-dual method for solving non-convex minimization problems. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Canonical Duality Theory. Springer
142. Wu, C., Gao, D.Y., Teo, K.L.: A direct optimization method for low group delay FIR filter design. Sig. Process. **93**(7), 1764–1772 (2013)
143. Yau, S.T., Gao, D.Y.: Obstacle problems for von Kámán equations. Adv. Appl. Math. **13**, 123–141 (1992)
144. 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)
145. Zhou, X.J., Gao, D.Y., Yang, C.H.: Canonical primal-dual algorithm for solving fourth-order polynomial minimization problems. Appl. Math. Comput. **227**, 246–255 (2014)
146. Zhou, X.J., Gao, D.Y., Yang, C.H.: Global solutions to a class of CEC benchmark constrained optimization problems. Optim. Lett. vol. 10, pp. 457–472 (2016). doi:10.1007/s11590-014-0784-0

Analytic Solutions to Large Deformation Problems Governed by Generalized Neo-Hookean Model

David Yang Gao

Abstract This paper addresses some fundamental issues in nonconvex analysis. By using pure complementary energy principle proposed by the author, a class of fully nonlinear partial differential equations in nonlinear elasticity is able to convert a unified algebraic equation, a complete set of analytical solutions are obtained in dual space for 3-D finite deformation problems governed by generalized neo-Hookean model. Both global and local extremal solutions to the nonconvex variational problem are identified by a triality theory. Connection between challenges in nonlinear analysis and NP-hard problems in computational science is revealed. Results show that Legendre–Hadamard condition can only guarantee ellipticity for generalized convex problems. For nonconvex systems, the ellipticity depends not only on the stored energy, but also on the external force field. Uniqueness is proved based on a generalized quasiconvexity and a generalized ellipticity condition. Application is illustrated for nonconvex logarithm stored energy.

1 Nonconvex Variational Problem and Challenges

Minimum total potential energy principle in nonlinear elasticity has always presented fundamental challenging problems not only in continuum mechanics, but also in nonlinear analysis and computational sciences. This paper intends to solve, under certain conditions, the following minimum potential variational problem (($\mathscr{P}$) for short):

$$(\mathscr{P}): \quad \min \left\{ \Pi(\boldsymbol{\chi}) = \int_{\mathscr{B}} W(\nabla \boldsymbol{\chi}) \mathrm{d}\mathscr{B} - \int_{S_t} \boldsymbol{\chi} \cdot \mathbf{t} \mathrm{d}S | \;\; \boldsymbol{\chi} \in \mathscr{X}_c \right\}, \tag{1}$$

where the unknown deformation $\boldsymbol{\chi}(\mathbf{x}) = \{\chi_i(x_j)\} \in \mathscr{X}_a$ is a vector-valued mapping $\mathscr{B} \subset \mathbb{R}^3 \to \omega \subset \mathbb{R}^3$ from a given material particle $\mathbf{x} = \{x_i\} \in \mathscr{B}$ in the undeformed

D.Y. Gao (✉)
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_2

body to a position vector in the deformed configuration ω. The body is fixed on the boundary $S_x \subset \partial\mathscr{B}$, while on the remaining boundary $S_t = S_x \cap \partial\mathscr{B}$, the body is subjected to a given surface traction $\mathbf{t}(\mathbf{x})$. In this paper, we let $\mathscr{X}_a$ as a *geometrically admissible space* defined by

$$\mathscr{X}_a = \{\boldsymbol{\chi} \in \mathscr{W}^{1,1}(\mathscr{B}; \mathbb{R}^3) | \ \boldsymbol{\chi}(\mathbf{x}) = 0 \ \ \forall \mathbf{x} \in S_x\} \tag{2}$$

where $\mathscr{W}^{1,1}$ is the standard notation for Sobolev space, i.e., a function space in which both $\boldsymbol{\chi}$ and its weak derivative $\nabla\boldsymbol{\chi}$ have a finite $L^1(\mathscr{B})$ norm. For homogeneous hyperelastic body, the strain energy $W(\mathbf{F})$ is assumed to be C^1 on its domain $\mathscr{F}_c \subset \mathbb{R}^{3\times 3}$, in which certain necessary *constitutive constraints* are included, such as

$$\det \mathbf{F} > 0, \quad W(\mathbf{F}) \geq 0 \ \ \forall \mathbf{F} \in \mathscr{F}_c, \quad W(\mathbf{F}) \to \infty \text{ as } \|\mathbf{F}\| \to \infty. \tag{3}$$

Thus, the *kinetically admissible space* in $(\mathscr{P})$ is simply defined by

$$\mathscr{X}_c = \{\boldsymbol{\chi} \in \mathscr{X}_a | \ \nabla\boldsymbol{\chi} \in \mathscr{F}_c\} \tag{4}$$

which is essentially nonconvex due to nonlinear constraints such as $\det(\nabla\boldsymbol{\chi}) > 0$. Also, the stored energy $W(\mathbf{F})$ is in general nonconvex in order to model real-world problems such as post-buckling and phase transitions, etc. Therefore, the nonconvex variational problem $(\mathscr{P})$ has usually multiple local optimal solutions.

Let $\mathscr{X}_b \subset \mathscr{X}_c$ be a subspace with two additional conditions

$$\mathscr{X}_b = \{\boldsymbol{\chi} \in \mathscr{X}_c | \ \ \boldsymbol{\chi} \in C^2(\mathscr{B}; \mathbb{R}^3), \quad W(\mathbf{F}(\boldsymbol{\chi})) \in C^2(\mathscr{F}_c; \mathbb{R})\}. \tag{5}$$

If $\partial\mathscr{B}$ is sufficiently regular, the criticality condition $\delta\Pi(\boldsymbol{\chi}; \delta\boldsymbol{\chi}) = 0 \ \ \forall \delta\boldsymbol{\chi} \in \mathscr{X}_b$ leads to a nonlinear boundary value problem

$$(BVP): \quad \begin{cases} -\nabla \cdot \boldsymbol{\sigma}(\nabla\boldsymbol{\chi}) = 0 \quad \text{in } \mathscr{B}, \\ \mathbf{N} \cdot \boldsymbol{\sigma}(\nabla\boldsymbol{\chi}) = \mathbf{t} \quad \text{on } S_t, \quad \boldsymbol{\chi} = 0 \quad \text{on } S_x \end{cases} \tag{6}$$

where, $\mathbf{N} \in \mathbb{R}^3$ is a unit vector normal to $\partial\mathscr{B}$, and $\boldsymbol{\sigma}(\mathbf{F})$ is the first Piola–Kirchhoff stress (force per unit undeformed area), defined by

$$\boldsymbol{\sigma} = \nabla W(\mathbf{F}), \quad \text{or } \sigma_{ij} = \frac{\partial W(\mathbf{F})}{\partial F_{ij}}, \quad i, j = 1, 2, 3. \tag{7}$$

Remark 1 (Nonconvexity, Multi-Solutions, and NP-Hard Problems) The stored energy $W(\mathbf{F})$ in nonlinear elasticity is generally nonconvex. It turns out that the fully nonlinear (BVP) could have multiple solutions $\{\boldsymbol{\chi}_k(\mathbf{x})\} \in \mathscr{X}_c \subset \mathbb{R}^\infty$ at each material point $\mathbf{x} \in \mathscr{B}_s \subset \mathscr{B}$. As long as the continuous domain $\mathscr{B}_s \neq \emptyset$, this solution set $\{\boldsymbol{\chi}_k(\mathbf{x})\}(k = 1, \ldots, K)$ can form infinitely many (K^∞) solutions to (BVP) even $\mathscr{B} \subset \mathbb{R}$. It is impossible to use traditional convexity and ellipticity conditions to identify global minimizer among all these local solutions. Gao and

Ogden discovered in [10] that for certain given external force field, both global and local extremum solutions are nonsmooth and cannot be obtained by Newton-type numerical methods. Therefore, Problem ($\mathscr{P}$) is much more difficult than (BVP). In computational mechanics, any direct numerical method for solving ($\mathscr{P}$) will lead to a nonconvex minimization problem in $\mathbb{R}^n$, which could have K^n local solutions. Due to the lack of global optimality condition, it is fundamentally difficult to solve nonconvex minimization problems by traditional methods within polynomial time. Therefore, in computational sciences most nonconvex minimization problems are considered to be NP-hard (Nondeterministic Polynomial-time hard).

Direct methods for solving nonconvex variational problems in finite elasticity have been studies extensively during the last 50 years and many generalized convexities, such as poly-, quasi- and rank-one convexities, have been proposed. For a given function $W : \mathscr{F}_c \to \mathbb{R}$, the following statements are well-known (see [16])[1]:

$$\text{convex} \Rightarrow \text{ polyconvex } \Rightarrow \text{ quasiconvex } \Rightarrow \text{ rank-one convex.}$$

Although the generalized convexities have been well studied for general function $W(\mathbf{F})$ on matrix space $\mathbb{R}^{m\times n}$, these mathematical concepts provide only necessary conditions for local minimal solutions, and cannot be applied to general (nonconvex) finite deformation problems. In reality, the stored energy $W(\mathbf{F})$ must be nonconvex in order to model real-world phenomena. Strictly speaking, due to certain necessary constitutive constraints such as $\det \mathbf{F} > 0$ and objectivity condition, etc., even the domain $\mathscr{F}_c$ is not convex, therefore, it is not appropriate to discuss convexity of the stored energy $W(\mathbf{F})$ in general nonlinear elasticity. How to identify global optimal solution has been a fundamental challenging problem in nonconvex analysis and computational science. ■

Remark 2 (Canonical Duality, Gap Function, and Global Extremality)
The objectivity is a necessary constraint for any hyperelastic model. A real-valued function $W : \mathscr{F}_c \to \mathbb{R}$ is objective iff there exists a function $V(\mathbf{C})$ such that $W(\mathbf{F}) = V(\mathbf{F}^T\mathbf{F})\ \ \forall \mathbf{F} \in \mathscr{F}_c$ (see [1]). By the fact that the right Cauchy–Green tensor $\mathbf{C}$ is an objective measure on a convex domain $\mathscr{E}_a = \{\mathbf{C} \in \mathbb{R}^{3\times 3} |\ \mathbf{C} = \mathbf{C}^T,\ \ \mathbf{C} \succ 0\}$, it is possible and natural to discuss the convexity of $V(\mathbf{C})$. A real-valued function $V : \mathscr{E}_a \to \mathbb{R}$ is called *canonical* if the duality relation $\boldsymbol{\xi}^* = \nabla V(\boldsymbol{\xi}) : \mathscr{E}_a \to \mathscr{E}_a^*$ is one-to-one and onto [5]. The canonical duality is necessary for modeling natural phenomena, which lays a foundation for the canonical duality theory [5]. This theory was developed from Gao and Strang's original work in 1989 [11] for general non-convex/nonsmooth variational problems in finite deformation theory. The key idea of this theory is assuming the existence of a geometrically admissible (objective) measure $\boldsymbol{\xi} = \Lambda(\mathbf{F})$ and a canonical function $V(\boldsymbol{\xi})$ such that the following *canonical transformation* holds

$$\boldsymbol{\xi} = \Lambda(\mathbf{F}) : \mathscr{F}_a \to \mathscr{E}_a \quad \Rightarrow \quad W(\mathbf{F}) = V(\Lambda(\mathbf{F})). \tag{8}$$

[1] It was proved recently that rank-one convexity also implies polyconvexity for isotropic, objective, and isochoric elastic energies in the two-dimensional case [15].

Gao and Strang discovered that the directional derivative $\Lambda_t(\mathbf{F}) = \delta\Lambda(\mathbf{F})$ is adjoined with the equilibrium operator, while its complementary operator $\Lambda_c(\mathbf{F}) = \Lambda(\mathbf{F}) - \Lambda_t(\mathbf{F})\mathbf{F}$ leads to a so-called *complementary gap function*, which recovers duality gaps in traditional duality theories and provides a sufficient condition for identifying both global and local extremal solutions [2, 5, 12]. ■

The canonical duality theory has been applied for solving a large class of nonconvex, nonsmooth, discrete problems in multidisciplinary fields of nonlinear analysis, nonconvex mechanics, global optimization, computational sciences, etc. A comprehensive review is given recently in [12]. The main goal of this paper is to show author's recent analytical solutions [7] for general anti-plane shear problems can be easily generalized for solving finite deformation problems governed by generalized neo-Hookean materials. Some insightful results are obtained on generalized convexity and ellipticity in nonlinear analysis.

2 Complete Solutions to Generalized Neo-Hookean Material

By the fact that the right Cauchy–Green strain $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ is an objective tensor, its three principal invariants

$$I_1(\mathbf{C}) = \mathrm{tr}\,\mathbf{C}, \quad I_2(\mathbf{C}) = \frac{1}{2}[(\mathrm{tr}\,\mathbf{C})^2 - \mathrm{tr}\,(\mathbf{C}^2)], \quad I_3(\mathbf{C}) = \det\mathbf{C} \tag{9}$$

are also objective functions of $\mathbf{F}$. Clearly, for isochoric deformations we have $I_3(\mathbf{C}) = 1$. The elastic body is said to be *generalized neo-Hookean material* if the stored energy depends only on I_1, i.e., there exists a function $V(I_1)$ such that $W(\mathbf{F}) = V(I_1(\mathbf{C}(\mathbf{F})))$. Since $I_1 = \mathrm{tr}\,(\mathbf{F}^T\mathbf{F}) > 0 \;\; \forall\mathbf{F} \in \mathscr{F}_c$, the domain of $V(I_1)$ is a convex (positive) cone

$$\mathscr{E}_a = \{\xi \in L^p(\mathscr{B}) \mid \xi(\mathbf{x}) > 0 \;\; \forall\mathbf{x} \in \mathscr{B}\}, \tag{10}$$

it is possible to discuss the convexity of $V(I_1)$ on $\mathscr{E}_a$. Furthermore, we assume that $V(I_1)$ is a $C^2(\mathscr{E}_a)$ canonical function. Then the canonical transformation (8) for the generalized neo-Hookean model is

$$\xi = \Lambda(\mathbf{F}) = \mathrm{tr}\,(\mathbf{F}^T\mathbf{F}) : \mathscr{F}_c \to \mathscr{E}_a, \quad W(\mathbf{F}) = V(\xi(\mathbf{F})). \tag{11}$$

For a given external force $\mathbf{t}(\mathbf{x})$ on S_t, we introduce a *statically admissible space*

$$\mathscr{T}_a = \left\{\mathbf{T} \in \mathscr{W}^{1,1}(\mathscr{B}; \mathbb{R}^{3\times 3}) \mid \; \nabla\cdot\mathbf{T} = 0 \;\; \text{in}\,\mathscr{B}, \;\; \mathbf{N}\cdot\mathbf{T} = \mathbf{t} \;\; \text{on}\, S_t\right\}. \tag{12}$$

Thus for any given $\mathbf{T} \in \mathscr{T}_a$, the primal problem $(\mathscr{P})$ for the generalized neo-Hookean material can be written in following canonical form:

$$(\mathscr{P})_{\mathbf{T}}: \quad \min \left\{ \Pi_{\mathbf{T}}(\nabla \boldsymbol{\chi}) = \int_{\mathscr{B}} G(\nabla \boldsymbol{\chi}) \, \mathrm{d}\mathscr{B} \ | \ \forall \boldsymbol{\chi} \in \mathscr{X}_c \right\}, \tag{13}$$

where $\mathscr{X}_c = \{\boldsymbol{\chi} \in \mathscr{X}_a | \ \Lambda(\nabla \boldsymbol{\chi}) \in \mathscr{E}_a\}$ and the integrand $G : \mathscr{F}_a \to \mathbb{R}$ is defined by

$$G(\mathbf{F}) = V(\Lambda(\mathbf{F})) - \mathrm{tr}\,(\mathbf{F}^T \mathbf{T}). \tag{14}$$

By the fact that $\det \mathbf{F} > 0$ is not a variational constraint and the certain constitutive constraints, such as coercivity and objectivity, have been naturally relaxed by the canonical transformation, the domain of $G(\mathbf{F})$ is simply $\mathscr{F}_a = \mathbb{R}^{3\times 3}$.

Let $\mathrm{SO}(3) = \{\mathbf{R} \in \mathbb{R}^{3\times 3} | \ \ \mathbf{R}^T = \mathbf{R}^{-1}, \ \ \det \mathbf{R} = 1\}$ and

$$\mathscr{R} = \{\mathbf{R}(\mathbf{x}) \in L^1[\mathscr{B}, \mathbb{R}^{3\times 3}] | \ \ \mathbf{R}(\mathbf{x}) \in \mathrm{SO}(3) \ \ \forall \mathbf{x} \in \mathscr{B}\}. \tag{15}$$

Theorem 1 *For any given* $\mathbf{T} \in \mathscr{T}_a$*, if* $\bar{\boldsymbol{\chi}} \in \mathscr{X}_c$ *is a stationary solution to* $(\mathscr{P})_{\mathbf{T}}$*, then it is also a stationary solution to* $(\mathscr{P})$*.*

For any given rotation field $\mathbf{R}(\mathbf{x}) \in \mathscr{R}$ *such that* $\mathbf{R}^T\mathbf{T} \in \mathscr{T}_a$*, then* $\Pi_{\mathbf{T}}(\mathbf{F}) = \Pi_{\mathbf{T}}(\mathbf{RF})$*.*

For any uniform rotation $\mathbf{R} \in SO(3)$ *such that* $\mathbf{R}^T\mathbf{T} \in \mathscr{T}_a$*, if* $\bar{\boldsymbol{\chi}}$ *is a stationary solution to* $(\mathscr{P})$*, then* $\mathbf{R}\bar{\boldsymbol{\chi}}$ *is also a stationary solution to* $(\mathscr{P})$*.*

Proof. For any given $\mathbf{T} \in \mathscr{T}_a$, the stationary condition for the canonical variational problem $(\mathscr{P})_{\mathbf{T}}$ leads to the following canonical boundary value problem

$$(BVP)_{\mathbf{T}}: \quad \begin{cases} \nabla \cdot (2\zeta \nabla \boldsymbol{\chi}) = \nabla \cdot \mathbf{T} = 0 \quad \text{in } \mathscr{B}, \\ \mathbf{N} \cdot (2\zeta \nabla \boldsymbol{\chi}) = \mathbf{N} \cdot \mathbf{T} = \mathbf{t} \ \ \text{on } S_t, \quad \boldsymbol{\chi} = 0 \quad \text{on } S_x \end{cases} \tag{16}$$

which is identical to (BVP) since

$$\boldsymbol{\sigma} = \nabla W(\mathbf{F}) = \frac{\partial V(\xi)}{\partial \xi} \frac{\partial \xi}{\partial \mathbf{F}} = 2\zeta \mathbf{F}, \quad \zeta = \nabla V(\xi).$$

By the objectivity of $\xi = \Lambda(\mathbf{F}) = \Lambda(\mathbf{RF}) \ \ \forall \mathbf{R}(\mathbf{x}) \in \mathscr{R}$ and the fact that

$$\int_{\mathscr{B}} \mathrm{tr}\,[(\mathbf{R}\nabla \boldsymbol{\chi})^T \mathbf{T}] \mathrm{d}\mathscr{B} = \int_{\mathscr{B}} \mathrm{tr}\,[(\nabla \boldsymbol{\chi})^T (\mathbf{R}^T \mathbf{T})] \mathrm{d}\mathscr{B} = \int_{S_t} \boldsymbol{\chi} \cdot \mathbf{t} \mathrm{d}S \ \ \forall \mathbf{R}^T\mathbf{T} \in \mathscr{T}_a,$$

we have $\Pi_{\mathbf{T}}(\mathbf{F}) = \Pi_{\mathbf{T}}(\mathbf{RF}) \ \ \forall \mathbf{R}(\mathbf{x}) \in \mathscr{R}$. Particularly, for any uniform $\mathbf{R} \in \mathrm{SO}(3)$ such that $\mathbf{R}^T\mathbf{T} \in \mathscr{T}_a$, we have $\Pi(\boldsymbol{\chi}) = \Pi_{\mathbf{T}}(\mathbf{RF}(\boldsymbol{\chi}))$. □

Theorem 1 is important for understanding the canonical duality theory.

By the canonical assumption on $V(\xi)$, the duality relation $\zeta = \nabla V(\xi) : \mathscr{E}_a \to \mathscr{E}_a^*$ is invertible. The complementary energy can be defined uniquely by the Legendre transformation

$$V^*(\zeta) = \{\xi\zeta - V(\xi) | \ \zeta = \nabla V(\xi)\}. \tag{17}$$

Clearly, the function $V : \mathscr{E}_a \to \mathbb{R}$ is canonical if and only if the following *canonical duality relations* hold on $\mathscr{E}_a \times \mathscr{E}_a^*$

$$\zeta = \nabla V(\xi) \quad \Leftrightarrow \quad \xi = \nabla V^*(\zeta) \quad \Leftrightarrow \quad V(\xi) + V^*(\zeta) = \xi\zeta. \tag{18}$$

Using $V(\xi) = \xi\zeta - V^*(\zeta)$, the nonconvex function $G(\mathbf{F})$ can be written as the standard Gao and Strang total complementary function $\Xi : \mathscr{X}_a \times \mathscr{E}_a^* \to \mathbb{R}$

$$\Xi(\boldsymbol{\chi}, \zeta) = \int_{\mathscr{B}} \left[\Lambda(\nabla\boldsymbol{\chi})\zeta - V^*(\zeta) - \mathrm{tr}((\nabla\boldsymbol{\chi})^T\mathbf{T}) \right] \mathrm{d}\mathscr{B}. \tag{19}$$

The canonical dual function can be obtained by the *canonical dual transformation*:

$$\Pi^d(\zeta) = \mathrm{sta}\{\Xi(\boldsymbol{\chi}, \zeta) |\ \boldsymbol{\chi} \in \mathscr{X}_a\} = \int_{\mathscr{B}} G^d(\zeta)\mathrm{d}\mathscr{B}, \tag{20}$$

where the notation $\mathrm{sta}\{\Xi(\boldsymbol{\chi}, \zeta) |\ \boldsymbol{\chi} \in \mathscr{X}_a\}$ stands for finding (partial) stationary point $\boldsymbol{\chi} \in \mathscr{X}_a$ of $\Xi(\boldsymbol{\chi}, \zeta)$ for a given $\zeta \in \mathscr{S}_a$, and

$$G^d(\zeta) = -V^*(\zeta) - \frac{1}{4}\zeta^{-1}\tau^2, \qquad \tau^2 = \mathrm{tr}(\mathbf{T}^T\mathbf{T}). \tag{21}$$

Let $\mathscr{S}_a \subset \mathscr{E}_a^*$ be a canonical dual feasible space defined by

$$\mathscr{S}_a = \{\zeta \in \mathscr{E}_a^* |\ \zeta^{-1}\tau^2 \in L^1(\mathscr{B})\}. \tag{22}$$

Thus, the pure complementary energy principle, first proposed in 1998 [3], leads to the following canonical dual variational problem

$$(\mathscr{P}^d) : \quad \mathrm{sta}\left\{ \Pi^d(\zeta) = \int_{\mathscr{B}} G^d(\zeta)\mathrm{d}\mathscr{B} \ |\ \zeta \in \mathscr{S}_a \right\}. \tag{23}$$

Since the canonical dual variable ζ is a scalar-valued function, the criticality condition for this variational problem leads to a so-called *canonical dual algebraic equation* (see [5]):

$$4\zeta^2\nabla V^*(\zeta) = \tau^2(\mathbf{x}) \quad \forall \mathbf{x} \in \mathscr{B}. \tag{24}$$

Note that $\nabla V^*(\zeta) : \mathscr{E}_a^* \to \mathscr{E}_a$ is also one-to-one and onto, this equation has at least one solution for any given $\tau^2 = \mathrm{tr}(\mathbf{T}^T\mathbf{T}) \geq 0$ and $\zeta = 0$ only if $\tau = 0$. Therefore, although there is an inverse term ζ^{-1} in $G^d(\zeta)$, this canonical dual function is well-defined on $\mathscr{S}_a$. Due to the nonlinearity, the solution to (24) may not be unique [5, 7, 9, 10]. By the pure complementary energy principle proposed by Gao in 1999 (see [5]), we have

Theorem 2 (**Complementary Dual Principle**) *For any given* $\mathbf{T} \in \mathscr{T}_a$*, the following statements are equivalent:*

1) $(\bar{\boldsymbol{\chi}}, \bar{\varsigma})$ is a stationary point of $\Xi(\boldsymbol{\chi}, \zeta)$;
2) $\bar{\boldsymbol{\chi}}$ is a stationary solution to $(\mathscr{P})$;
3) $\bar{\varsigma}$ is a stationary solution to $(\mathscr{P}^d)$.
Moreover, we have

$$\Pi(\bar{\boldsymbol{\chi}}) = \Xi(\bar{\boldsymbol{\chi}}, \bar{\varsigma}) = \Pi^d(\bar{\varsigma}) \tag{25}$$

Proof. For any given $\mathbf{T} \in \mathscr{T}_a$, the stationary condition of $\Xi(\boldsymbol{\chi}, \zeta)$ leads to the canonical equilibrium equations

$$\Lambda(\mathbf{F}(\bar{\boldsymbol{\chi}})) = \nabla V^*(\bar{\varsigma}), \tag{26}$$

$$2\bar{\varsigma}\mathbf{F}(\bar{\boldsymbol{\chi}}) = \mathbf{T} \in \mathscr{T}_a \tag{27}$$

By the canonical duality, (26) is equivalent to $\bar{\varsigma} = \nabla V(\xi)$ with $\xi = \Lambda(\nabla\bar{\boldsymbol{\chi}})$. Thus, $\bar{\boldsymbol{\chi}}$ must be a stationary solution to $(\mathscr{P})_{\mathbf{T}}$ and also a stationary solution to $(\mathscr{P})$ due to Theorem 1.

By solving (27) we have $\mathbf{F}(\bar{\boldsymbol{\chi}}) = \frac{1}{2\bar{\varsigma}}\mathbf{T}$. Substituting this into (26) leads to the canonical dual equation (24). Thus, $\bar{\varsigma}$ is a stationary solution to $(\mathscr{P}^d)$.

The equivalence and the Eq. (25) can be proved by

$$\text{sta}\{\Pi_{\mathbf{T}}(\nabla\boldsymbol{\chi}) |\ \boldsymbol{\chi} \in \mathscr{X}_c\} = \text{sta}\{\Xi(\boldsymbol{\chi}, \zeta) |\ (\boldsymbol{\chi}, \zeta) \in \mathscr{X}_a \times \mathscr{E}_a^*\} = \text{sta}\{\Pi^d(\zeta) |\ \zeta \in \mathscr{S}_a\}$$

and Theorem 1. □

Theorem 3 (**Pure Complementary Energy Principle**) *For any given nontrivial $\mathbf{t} \neq 0$ and $\boldsymbol{\chi} \in \mathscr{X}_a$ such that $\mathbf{T} \in \mathscr{T}_a \neq \emptyset$, (24) has at least one solution $\zeta_k \neq 0$, the deformation gradient defined by $\mathbf{F}_k = \nabla\boldsymbol{\chi}_k = \zeta_k^{-1}\mathbf{T}$ is a critical point of $\Pi(\boldsymbol{\chi})$ and $\Pi(\boldsymbol{\chi}_k) = \Pi^d(\zeta_k)$.*

Moreover, if $\nabla \times (\zeta_k^{-1}\mathbf{T}) = 0$, then the deformation vector defined by

$$\boldsymbol{\chi}_k(\mathbf{x}) = \frac{1}{2}\int_{\mathbf{x}_0}^{\mathbf{x}} \zeta_k^{-1}\mathbf{T} \cdot d\mathbf{x} \tag{28}$$

along any path from $\mathbf{x}_0 \in S_x$ to $\mathbf{x} \in \mathscr{B}$ is a solution to $(BVP)_{\mathbf{T}}$ in the sense that it satisfies both equilibrium equation and boundary conditions in (16).

Proof. By the canonical duality relations in (18) we know that $\xi_k = \nabla V^*(\zeta_k) > 0$. Thus, for a given nontrivial $\mathbf{t}(\mathbf{x})$, there exists a nontrivial $\tau^2(\mathbf{x}) = \text{tr}(\mathbf{T}^T\mathbf{T})$ in $\mathscr{B}$ such that the canonical dual algebraic equation (24) have at least one nontrivial solution $\zeta_k(\mathbf{x})$ in $\mathscr{B}$.

Since the critical point ζ_k is a solution to (24), we have

$$\xi_k = \text{tr}(\mathbf{F}_k^T\mathbf{F}_k) = \frac{1}{4}\zeta_k^{-2}\text{tr}(\mathbf{T}^T\mathbf{T}) = \nabla V^*(\zeta_k) \quad \Rightarrow \quad \mathbf{F}_k = \frac{1}{2}\zeta_k^{-1}\mathbf{T} \tag{29}$$

subjected to any given rotation field $\mathbf{R}(\mathbf{x}) \in \mathscr{R}$. By the fact that the canonical dual solution ζ_k defined by (24) is independent of the rotation field, the canonical duality leads to

$$G^d(\zeta_z) = \Xi(\mathbf{F}_k, \zeta_k) = V(\Lambda(\mathbf{F}_k)) - \mathrm{tr}(\mathbf{F}_k^T \mathbf{T}) = G(\mathbf{F}_k).$$

This shows $\Pi(\boldsymbol{\chi}_k) = \Pi^d(\zeta_k)$.

To prove $\boldsymbol{\chi}_k$ defined by (28) is a solution to $(BVP)_{\mathbf{T}}$, we simply substitute $\nabla\boldsymbol{\chi}_k = \mathbf{F}_k = \frac{1}{2}\zeta_k^{-1}\mathbf{T}$ into $(BVP)_{\mathbf{T}}$ to have all necessary equilibrium conditions satisfied. Therefore, $\boldsymbol{\chi}_k$ defined by (28) is a solution to $(BVP)_{\mathbf{T}}$. □

This pure complementary energy principle shows that by the canonical dual transformation, the fully nonlinear partial differential equation in $(BVP)_{\mathbf{T}}$ can be converted to an algebraic equation (24), which can be solved to obtain a complete set of solutions (see [7, 8]). In literature, this pure complementary energy principle is known as the Gao principle [14].

Since $\mathscr{S}_a$ is nonconvex, in order to identify global and local optimal solutions, we need the following convex subsets

$$\mathscr{S}_a^+ = \{\zeta \in \mathscr{S}_a |\ \zeta > 0\}, \quad \mathscr{S}_a^- = \{\zeta \in \mathscr{S}_a |\ \zeta < 0\}. \tag{30}$$

Then by the canonical duality-triality theory developed in [5] we have the following theorem.

Theorem 4 *Suppose that $V : \mathscr{E}_a \to \mathbb{R}$ is convex and for a given $\mathbf{T} \in \mathscr{T}_a$ such that $\{\zeta_k\}$ is a solution set to (24), $\mathbf{F}_k = \frac{1}{2}\zeta_k^{-1}\mathbf{T}$, and $\boldsymbol{\chi}_k$ is defined by (28), we have the following statements.*

1. *If $\zeta_k \in \mathscr{S}_a^+$, then $\nabla^2 W(\mathbf{F}_k) \succ 0$ and $\boldsymbol{\chi}_k$ is a global minimal solution to $(\mathscr{P})$.*
2. *If $\zeta_k \in \mathscr{S}_a^-$ and $\nabla^2 W(\mathbf{F}_k) \succ 0$, then $\boldsymbol{\chi}_k$ is a local minimal solution to $(\mathscr{P})$.*
3. *If $\zeta_k \in \mathscr{S}_a^-$ and $\nabla^2 W(\mathbf{F}_k) \prec 0$, then $\boldsymbol{\chi}_k$ is a local maximal solution to $(\mathscr{P})$.*

If $\{\zeta_k\} \subset \mathscr{S}_a^+$, then $\{\boldsymbol{\chi}_k\}$ is a convex set. The solution of $(\mathscr{P})$ is unique if $\{\zeta_k\} \subset \mathscr{S}_a^+$.

Proof. By using chain rule for $W(\mathbf{F}) = V(\xi(\mathbf{F}))$ we have $\nabla W(\mathbf{F}) = 2\mathbf{F}[\nabla V(\xi)] = 2\zeta\mathbf{F}$, and

$$\nabla^2 W(\mathbf{F}) = 2\zeta \mathbf{I} \otimes \mathbf{I} + 4h(\xi)\mathbf{F} \otimes \mathbf{F}, \tag{31}$$

where $\mathbf{I}$ is an identity tensor in $\mathbb{R}^{3\times 3}$, $h(\xi) = \nabla^2 V(\xi) \geq 0$ due to the convexity of V on $\mathscr{E}_a$. Therefore, $\nabla^2 W(\mathbf{F}_k) \succ 0$ if $\zeta_k \in \mathscr{S}_a^+$.

To prove $\boldsymbol{\chi}_k$ is a global minimizer of $(\mathscr{P})$, we follow Gao and Strang's work in 1989 [11]. By the convexity of $V(\xi)$ on its convex domain $\mathscr{E}_a$, we have

$$V(\xi) - V(\xi_k) \geq (\xi - \xi_k)\zeta_k \quad \forall \xi,\ \xi_k \in \mathscr{E}_a, \quad \zeta_k = \nabla V(\xi_k). \tag{32}$$

For any given variation $\delta\boldsymbol{\chi}$, we let $\boldsymbol{\chi} = \boldsymbol{\chi}_k + \delta\boldsymbol{\chi}$. Then we have [11]

$$\Lambda(\nabla\boldsymbol{\chi}) = \mathrm{tr}[(\nabla\boldsymbol{\chi})^T(\nabla\boldsymbol{\chi})] = \Lambda(\nabla\boldsymbol{\chi}_k) + \Lambda_t(\nabla\boldsymbol{\chi}_k)(\nabla\delta\boldsymbol{\chi}) - \Lambda_c(\nabla\delta\boldsymbol{\chi}), \tag{33}$$

where $\Lambda_t(\mathbf{F})\delta\mathbf{F} = 2\mathrm{tr}[\mathbf{F}^T(\delta\mathbf{F})]$ and $\Lambda_c(\delta\boldsymbol{\chi}) = -\Lambda(\delta\boldsymbol{\chi})$. Clearly, $\Lambda(\mathbf{F}) = \Lambda_t(\mathbf{F})\mathbf{F} + \Lambda_c(\mathbf{F})$. Then combining the inequality (32) and (33), we have

$$\begin{aligned}\Pi(\boldsymbol{\chi}) - \Pi(\boldsymbol{\chi}_k) &\geq \int_{\mathscr{B}} 2\zeta_k \mathrm{tr}[(\nabla\boldsymbol{\chi}_k)^T(\nabla\delta\boldsymbol{\chi})]\mathrm{d}\mathscr{B} - \int_{S_t} \delta\boldsymbol{\chi}\cdot\mathbf{t}\mathrm{d}S + \int_{\mathscr{B}} \zeta_k \mathrm{tr}[(\nabla\boldsymbol{\chi})^T(\nabla\boldsymbol{\chi})]\mathrm{d}\mathscr{B} \\ &= \int_{\mathscr{B}} [2\zeta_k(\nabla\boldsymbol{\chi}_k) - \mathbf{T}] : (\nabla\delta\boldsymbol{\chi})\mathrm{d}\mathscr{B} + G_{ap}(\delta\boldsymbol{\chi}, \zeta_k) \ \ \forall\boldsymbol{\chi},\ \delta\boldsymbol{\chi} \in \mathscr{X}_c \end{aligned} \tag{34}$$

for any given $\mathbf{T} \in \mathscr{T}_a$, where

$$G_{ap}(\boldsymbol{\chi}, \zeta) = \int_{\mathscr{B}} -\Lambda_c(\nabla\boldsymbol{\chi})\zeta\,\mathrm{d}\mathscr{B} = \int_{\mathscr{B}} \zeta\,\mathrm{tr}[(\nabla\boldsymbol{\chi})^T(\nabla\boldsymbol{\chi})]\mathrm{d}\mathscr{B} \tag{35}$$

is the *complementary gap function* introduced by Gao and Strang in [11]. If $\boldsymbol{\chi}_k$ is a critical point of $\Pi(\boldsymbol{\chi})$, then we have

$$\int_{\mathscr{B}} [2(\nabla\boldsymbol{\chi}_k)\zeta_k - \mathbf{T}] : (\nabla\delta\boldsymbol{\chi})\mathrm{d}\mathscr{B} = 0 \ \ \forall\delta\boldsymbol{\chi} \in \mathscr{X}_c, \ \ \forall\mathbf{T} \in \mathscr{T}_a$$

Thus, we have $\Pi(\boldsymbol{\chi}) - \Pi(\boldsymbol{\chi}_k) \geq G_{ap}(\delta\boldsymbol{\chi}, \zeta_k) \geq 0 \ \ \forall\delta\boldsymbol{\chi} \in \mathscr{X}_c$ if $\zeta_k \in \mathscr{S}_a^+$. This shows that $\boldsymbol{\chi}_k$ is a global minimizer of $(\mathscr{P})$.

To prove the local extremality, we replace $\mathbf{F}_k$ in (31) by $\mathbf{F}_k = \frac{1}{2}\zeta_k^{-1}\mathbf{T}$ such that

$$\mathbf{G}(\zeta_k) = \nabla^2 W(\mathbf{F}_k) = 2\zeta_k \mathbf{I}\otimes\mathbf{I} + \zeta_k^{-2}h(\xi_k)\mathbf{T}\otimes\mathbf{T}, \tag{36}$$

where $\xi_k = \nabla V^*(\zeta_k)$. Clearly, for a given $\mathbf{T} \in \mathscr{T}_a$ such that $\zeta_k \in \mathscr{S}_a^-$, the Hessian $\nabla^2 W(\mathbf{F}_k)$ could be either positive or negative definite. The total potential $\Pi(\boldsymbol{\chi}_k)$ is locally convex if the *Legendre condition* $\nabla^2 W(\nabla\boldsymbol{\chi}_k) \succeq 0$ holds, locally concave if $\nabla^2 W(\nabla\boldsymbol{\chi}_k) \prec 0$. Since $\boldsymbol{\chi}_k$ is a global minimizer when $\zeta_k \in \mathscr{S}_a^+$, therefore, for $\zeta_k \in \mathscr{S}_a^-$, the stationary solution $\boldsymbol{\chi}_k$ is a local minimizer if $\nabla^2 W(\nabla\boldsymbol{\chi}_k) \succ 0$ and, by the triality theory [5, 12], $\boldsymbol{\chi}_k$ is the biggest local maximizer if $\nabla^2 W(\nabla\boldsymbol{\chi}_k) \prec 0$.

If $\{\zeta_k\} \subset \mathscr{S}_a^+$, then all the solutions $\{\boldsymbol{\chi}_k\}$ are global minimizers and form a convex set. Since $\Pi^d(\zeta)$ is strictly concave on the open convex set $\mathscr{S}_a^+$, the condition $\{\zeta_k\} \subset \mathscr{S}_a^+$ implies the unique solution of (24). In this case, Problems $(\mathscr{P})_{\mathbf{T}}$ has at most one solution. □

Theorem 5 (**Triality Theory**) *For any given* $\mathbf{T} \in \mathscr{T}_a \neq \emptyset$*, let* ζ_k *be a critical point of* $(\mathscr{P}^d)$*, the vector* $\boldsymbol{\chi}_k$ *be defined by (28), and* $\mathscr{X}_o \times \mathscr{S}_o \subset \mathscr{X}_c \times \mathscr{S}_a^-$ *a neighborhood*[2] *of* $(\boldsymbol{\chi}_k, \zeta_k)$*.*

If $\zeta_k \in \mathscr{S}_a^+$*, then*

[2]The neighborhood $\mathscr{X}_o$ of $\boldsymbol{\chi}_k$ in the canonical duality theory means that $\boldsymbol{\chi}_k$ is the only one critical point of $\Pi(\boldsymbol{\chi})$ on $\mathscr{X}_o$ (see [5]).

$$\Pi(\boldsymbol{\chi}_k) = \min_{\boldsymbol{\chi} \in \mathscr{X}_c} \Pi(\boldsymbol{\chi}) = \max_{\zeta \in \mathscr{S}_a^+} \Pi^d(\zeta) = \Pi^d(\zeta_k). \tag{37}$$

If $\zeta_k \in \mathscr{S}_a^-$ *and* $\mathbf{G}(\zeta_k) \succ 0$, *then*

$$\Pi(\boldsymbol{\chi}_k) = \min_{\boldsymbol{\chi} \in \mathscr{X}_o} \Pi(\boldsymbol{\chi}) = \min_{\zeta \in \mathscr{S}_o} \Pi^d(\zeta) = \Pi^d(\zeta_k). \tag{38}$$

If $\zeta_k \in \mathscr{S}_a^-$ *and* $\mathbf{G}(\zeta_k) \prec 0$, *then*

$$\Pi(\boldsymbol{\chi}_k) = \max_{\boldsymbol{\chi} \in \mathscr{X}_o} \Pi(\boldsymbol{\chi}) = \max_{\zeta \in \mathscr{S}_o} \Pi^d(\zeta) = \Pi^d(\zeta_k). \tag{39}$$

This theorem shows that for convex canonical function V, the triality theory can be used to identify both global and local extremum solutions to the variational problem $(\mathscr{P})$ and the nonconvex minimum variational problem $(\mathscr{P})_{\mathbf{T}}$ is canonically equivalent to the following concave maximization problem over an open convex set $\mathscr{S}_a^+$, i.e.,

$$(\mathscr{P}^\sharp)_{\mathbf{T}}: \quad \max \left\{ \Pi^d(\zeta) = \int_{\mathscr{B}} G^d(\zeta) \mathrm{d}\mathscr{B} \mid \zeta \in \mathscr{S}_a^+ \right\}, \tag{40}$$

which is much easier to solve than directly for obtaining global optimal solution of $(\mathscr{P})$.

3 Generalized Quasiconvexity, G-Ellipticity, and Uniqueness

Ellipticity is a classical concept originally from linear partial differential systems, where the deformation is a scalar-valued function $\chi : \mathscr{B} \to \mathbb{R}$ and stored energy is a quadratic function $W(\boldsymbol{\gamma}) = \frac{1}{2}\boldsymbol{\gamma}^T \mathbf{H} \boldsymbol{\gamma}$ of $\boldsymbol{\gamma} = \nabla \chi \in \mathbb{R}^3$. The linear operator

$$L[\chi] = -\nabla \cdot [\mathbf{H}(\nabla \chi)] = -[h_{ij}\chi_{,j}]_{,i}$$

is called elliptic if $\mathbf{H} = \{h_{ij}\}$ is positive definite. In this case, the function $G(\boldsymbol{\gamma}) = W(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T \boldsymbol{\tau}$ is convex and its level set is an ellipse for any given $\boldsymbol{\tau} \in \mathbb{R}^3$. This concept has been extended to nonlinear analysis. The fully nonlinear partial differential equation in (BVP) (6) is called elliptic if the following Legendre–Hadamard (LH) condition holds

$$(\mathbf{a} \otimes \mathbf{a}) : \nabla^2 W(\mathbf{F}) : (\boldsymbol{\eta} \otimes \boldsymbol{\eta}) \geq 0 \ \ \forall \mathbf{a}, \boldsymbol{\eta} \in \mathbb{R}^3, \quad \forall \mathbf{F} \in \mathscr{F}_a. \tag{41}$$

The (BVP) is called strong elliptic if the inequality holds strictly. In this case, (BVP) has at most one solution. In vector space, the LH condition is equivalent to Legendre condition $\nabla^2 W(\boldsymbol{\gamma}) \succeq 0 \ \ \forall \boldsymbol{\gamma} \in \mathbb{R}^n$.

Clearly, the LH condition is only a sufficient condition for local minimizer of the variational problem ($\mathscr{P}$). In order to identify ellipticity, one must to check LH condition for all local solutions, which is impossible for general fully nonlinear problems. Also, the traditional ellipticity definition depends only on the stored energy $W(\mathbf{F})$ regardless of the linear term in $G(\mathbf{F}) = W(\mathbf{F}) - \text{tr}\,(\mathbf{F}^T\mathbf{T})$. This definition works only for convex systems since the linear term tr $(\mathbf{F}^T\mathbf{T})$ can't change the convexity of $G(\mathbf{F})$. But this is not true for nonconvex systems. To see this, let us consider the St. Venant–Kirchhoff material

$$W(\mathbf{F}) = \frac{1}{2}\mathbf{E} : \mathbf{H} : \mathbf{E}, \quad \mathbf{E} = \frac{1}{2}[(\mathbf{F})^T(\mathbf{F}) - \mathbf{I}], \tag{42}$$

where $\mathbf{I}$ is a unit tensor in $\mathbb{R}^{3\times 3}$. Clearly, this function is not even rank-one convex. A special case of this model in $\mathbb{R}^n$ is the well-known double-well potential $W(\boldsymbol{\gamma}) = \frac{1}{2}(\frac{1}{2}|\boldsymbol{\gamma}|^2 - 1)^2$. If we let $\xi = \Lambda(\boldsymbol{\gamma}) = \frac{1}{2}|\boldsymbol{\gamma}|^2 - 1$ be an objective measure, we have the canonical function $V(\xi) = \frac{1}{2}\xi^2$. In this case, the canonical dual algebraic equation (24) is a cubic equation (see [5]) $2\zeta^2(\zeta + 1) = \tau^2$, which has at most three real solutions $\{\zeta_k(\mathbf{x})\}$ at each $\mathbf{x} \in \mathscr{B}$ satisfying $\zeta_1 \geq 0 \geq \zeta_2 \geq \zeta_3$. It was proved in [5] (Theorem 3.4.4, page 133) that for a given force $\mathbf{t}(\mathbf{x})$, if $\tau^2(\mathbf{x}) > 8/27 \;\; \forall \mathbf{x} \in \mathscr{B} \subset \mathbb{R}$, then $(BVP)_{\mathbf{T}}$ has only one solution on $\mathscr{B}$. If $\tau^2(\mathbf{x}) < 8/27 \;\; \forall \mathbf{x} \in \mathscr{B}_s \subset \mathscr{B}$, then $(BVP)_{\mathbf{T}}$ has three solutions $\{\chi_k(\mathbf{x})\}$ at each $\mathbf{x} \in \mathscr{B}_s$, i.e., $\Pi(\chi)$ is nonconvex on $\mathscr{B}_s$. It was shown by Gao and Ogden that these solutions are nonsmooth if $\tau(\mathbf{x})$ changes its sign on $\mathscr{B}_s$ [10].

Analytical solutions for general 3-D finite deformation problem ($\mathscr{P}$) were first proposed by Gao in 1998–1999 [3, 4]. It is proved recently [8] that for St Venant–Kirchhoff material, the problem ($\mathscr{P}$) could have 24 critical solutions at each material point $\mathbf{x} \in \mathscr{B}$, but only one global minimizer. The solution is unique if the external force is sufficiently large.

For a given function $G : \mathscr{F}_a \to \mathbb{R}$, its *level set* and *sub-level set* of height $\alpha \in \mathbb{R}$ are defined, respectively, as the following

$$\mathscr{L}_\alpha(G) = \{\mathbf{F} \in \mathscr{F}_a \mid \; G(\mathbf{F}) = \alpha\}, \quad \mathscr{L}_\alpha^\flat(G) = \{\mathbf{F} \in \mathscr{F}_a \mid \; G(\mathbf{F}) \leq \alpha\}, \quad \alpha \in \mathbb{R}. \tag{43}$$

The geometrical explanation for ellipticity and Theorem 4 is illustrated by Fig. 1, which shows that the nonconvex function $G(\boldsymbol{\gamma}) = \frac{1}{2}(\frac{1}{2}|\boldsymbol{\gamma}|^2 - 1)^2 - \boldsymbol{\gamma}^T\boldsymbol{\tau}$ depends sensitively on the external force $\boldsymbol{\tau} \in \mathbb{R}^2$. If $|\boldsymbol{\tau}|$ is big enough, $G(\boldsymbol{\gamma})$ has only one minimizer and its level set is an ellipse (Fig. 1b). Otherwise, $G(\boldsymbol{\gamma})$ has multiple local minimizers and its level set is not an ellipse. For $\boldsymbol{\tau} = 0$, it is well-known Mexican hat in theoretical physics (Fig. 1a). Figure 1 shows that although $G(\boldsymbol{\gamma})$ has only one global minimizer for certain given $\boldsymbol{\tau}$, the function is still nonconvex. Such a function is called quasiconvex in the context of global optimization. In order to distinguish this type of functions with Morry's quasiconvexity in nonconvex analysis, a generalized definition in a tensor space $\mathscr{F}_a \subset \mathbb{R}^{m\times n}$ could be convenient.

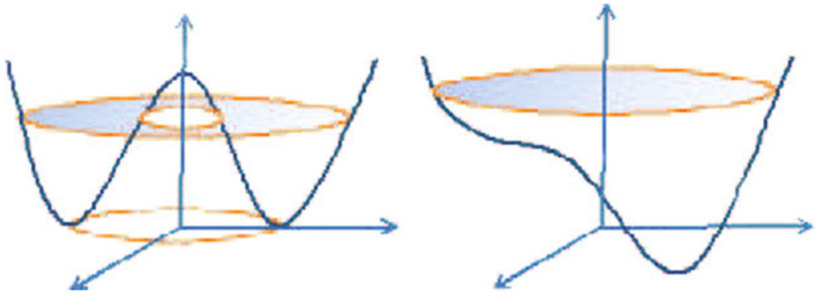


Fig. 1 Graphs and level sets of $G(\mathbf{x})$ for $\boldsymbol{\tau} = 0$ (*left*) and $\boldsymbol{\tau} \neq 0$ (*right*)

Definition 1 (**G-Quasiconvexity**). A function $G : \mathscr{F}_a \subset \mathbb{R}^{m\times n} \to \mathbb{R}$ is called G-quasiconvex if its domain $\mathscr{F}_a$ is convex and

$$G(\theta\mathbf{F} + (1-\theta)\mathbf{T}) \leq \max\{G(\mathbf{F}), G(\mathbf{T})\} \;\; \forall \mathbf{F},\; \mathbf{T} \in \mathscr{F}_a, \quad \forall\theta \in [0, 1]. \tag{44}$$

It is called strictly G-quasiconvex if the inequality holds strictly.

Moreover, we may need a definition of generalize ellipticity for nonconvex systems.

Definition 2 (**G-Ellipticity**). For a given function $G : \mathscr{F}_a \to \mathbb{R}$ and $\alpha \in \mathbb{R}$, its level set $\mathscr{L}_\alpha(G)$ is said to be a G-ellipse if it is a closed, simply connected set. For a given $\mathbf{t}$ such that $\mathbf{T} \in \mathscr{T}_a$, the (BVP) is said to be G-elliptic if the total potential function $G(\mathbf{F})$ is G-quasiconvex on $\mathscr{F}_a$. (BVP) is strongly G-elliptic if $G(\mathbf{F})$ is strictly G-quasiconvex.

Lemma 1 *For a given function* $G : \mathscr{F}_a \subset \mathbb{R}^{m\times n} \to \mathbb{R}$,

$$G(\mathbf{F}) \textit{ is G-quasiconvex} \Leftrightarrow \mathscr{L}^\flat_\alpha(G) \textit{ is convex} \Leftrightarrow \mathscr{L}_\alpha(G) \textit{ is a G-ellipse } \forall\alpha \in \mathbb{R}\,.$$

$$G(\mathbf{F}) \textit{ is convex} \Rightarrow \textit{is rank-one convex} \Rightarrow \textit{G-quasiconvex} \Rightarrow (BVP) \textit{ is G-elliptic}.$$

This statement shows an important fact in nonconvex systems, i.e., the total number of solutions to a nonlinear equation depends not only on the stored energy, but also (mainly) on the external force field. The nonlinear partial differential equation in (BVP) is elliptic only if it is G-elliptic. (BVP) has at most one solution if $G(\mathbf{F})$ is strictly G-quasiconvex on $\mathscr{F}_a$.

Remark 3 (**Existence and Uniqueness**) *Suppose that the canonical function* $V : \mathscr{E}_a \to \mathbb{R}$ *is convex, then* $\nabla V^*(\zeta) > 0$ *is a monotonic operator on* $\mathscr{E}^*_a$. *If for a given* $\mathbf{t} : S_t \to \mathbb{R}^3$ *such that* $\mathbf{T} \in \mathscr{T}_a \neq \emptyset$ *and* $\tau^2(\mathbf{x}) = tr(\mathbf{T}^T\mathbf{T}) \neq 0 \;\; \forall\mathbf{x} \in \mathscr{B}$, *then the nonconvex variational problem* $(\mathscr{P})$ *has at least one nontrivial solution a.e. in* $\mathscr{B}$. *It has a unique nontrivial solution if there exists a constant* τ_c *such that* $\tau^2(\mathbf{x}) = tr(\mathbf{T}^T\mathbf{T}) \geq \tau_c^2 \;\; \forall\mathbf{x} \in \mathscr{B}$.

In global optimization, the most simple quadratic integer programming problem

$$(\mathscr{P})_i: \quad \min \left\{ \Pi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{x}^T\mathbf{t} \;\mid\; \mathbf{x} = \{x_i\}^n \in \{0, 1\}^n \subset \mathbb{R}^n \right\}$$

could have up to 2^n local minimizers, which cannot be solved directly by traditional deterministic methods in polynomial time due to the indefinite matrix $\mathbf{Q}$ and the integer constraint. Such a nonconvex discrete optimization problem is considered as NP-hard in computer science. However, by using canonical transformation $\boldsymbol{\xi} = \Lambda(\mathbf{x}) = \{x_i(x_i - 1)\} \in \mathbb{R}^n$, the canonical dual of this discrete problem is a concave maximization over a convex set in continuous space [12]. It was proved in [6] that there exists a positive vector $\boldsymbol{\tau} = \{\tau_i\}^n > \mathbf{0} \in \mathbb{R}^n$, if $\{|t_i| \leq \tau_i\}^n$, then $\mathscr{S}_a^+ \neq \emptyset$ and $(\mathscr{P})_i$ is not NP-hard. The decision variable is simply $\{x_i\} = \{0 \text{ if } t_i < -\tau_i,\ 1 \text{ if } t_i > \tau_i\}$ (Theorem 8, [6]). Thus, the canonical duality theory can be used to identify NP-hard problems [12].

4 Applications in Anti-plane Shear Deformation

Now let us consider a special case that the homogeneous elastic body $\mathscr{B} \subset \mathbb{R}^3$ is a cylinder with generators parallel to the $\mathbf{e}_3$ axis and with cross section a sufficiently nice region $\Omega \subset \mathbb{R}^2$ in the $\mathbf{e}_1 \times \mathbf{e}_2$ plane. The so-called anti-plane shear deformation is defined by (see [13])

$$\boldsymbol{\chi}(\mathbf{x}) = \{x_1, \quad x_2, \quad x_3 + u(x_1, x_2)\} : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3, \tag{45}$$

where (x_1, x_2, x_3) are cylindrical coordinates in the reference configuration $\mathscr{B}$ relative to a cylindrical basis $\{\mathbf{e}_i\}$, $i = 1, 2, 3$. On $\Gamma_\chi \subset \partial\Omega$, the homogenous boundary condition is given $u(x_\alpha) = 0 \;\; \forall x_\alpha \in \Gamma_\chi, \;\; \alpha = 1, 2$. On the remaining boundary $\Gamma_t = \partial\Omega \cap \Gamma_\chi$, the cylinder is subjected to the shear force

$$\mathbf{t}(\mathbf{x}) = t(\mathbf{x})\mathbf{e}_3 \;\; \forall \mathbf{x} \in \Gamma_t,$$

where $t : \Gamma_t \to \mathbb{R}$ is a prescribed function. For this anti-plane shear deformation we have

$$\mathbf{F} = \nabla\boldsymbol{\chi} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ u_{,1} & u_{,2} & 1 \end{pmatrix}, \quad \mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = \begin{pmatrix} 1 + u_{,1}^2 & u_{,1}u_{,2} & u_{,1} \\ u_{,1}u_{,2} & 1 + u_{,2}^2 & u_{,2} \\ u_{,1} & u_{,2} & 1 \end{pmatrix}, \tag{46}$$

where $u_{,\alpha}$ represents $\partial u/\partial x_\alpha$ for $\alpha = 1, 2$. By the notation $|\nabla u|^2 = u_{,1}^2 + u_{,2}^2$, we have

$$I_1(\mathbf{C}) = I_2(\mathbf{C}) = 3 + |\nabla u|^2, \quad I_3(\mathbf{C}) \equiv 1, \tag{47}$$

Clearly, both $\mathbf{F}$ and $I_1(\mathbf{C})$ depend only on the shear strain $\boldsymbol{\gamma} = \nabla u = \{u_{,\alpha}\}$, therefore, the strain energy can be equivalently written in the forms of

$$W(\mathbf{F}(\boldsymbol{\gamma})) = V(\xi(\boldsymbol{\gamma})) = \hat{W}(\boldsymbol{\gamma}) \tag{48}$$

where $\hat{W}(\boldsymbol{\gamma})$ is a real-valued function.

The fact $\det \mathbf{F} \equiv 1$ shows that the anti-plane shear state (45) is an isochoric deformation. Therefore, the kinetically admissible displacement space $\mathscr{X}_c$ can be simply replaced by a convex set

$$\mathscr{U}_c = \{u(\mathbf{x}) \in \mathscr{W}^{1,1}(\Omega; \mathbb{R})| \;\; u(\mathbf{x}) = 0 \;\; \forall \mathbf{x} = \{x_\alpha\} \in \Gamma_\chi\}. \tag{49}$$

Thus, in terms of $\xi = \Lambda(\boldsymbol{\gamma}) = I_1 - 3 = |\boldsymbol{\gamma}|^2$ and $W(\mathbf{F}(\boldsymbol{\gamma})) = V(\Lambda(\boldsymbol{\gamma}))$, for any given

$$\boldsymbol{\tau} \in \mathscr{T}_a = \{\boldsymbol{\tau} \in C^1[\Omega; \mathbb{R}^2]| \;\; \nabla \cdot \boldsymbol{\tau} = 0 \quad \text{in } \Omega, \quad \mathbf{n} \cdot \boldsymbol{\tau} = t \quad \text{on } \Gamma_t\}$$

Problem $(\mathscr{P})_{\mathbf{T}}$ for the anti-plane shear deformation (45) has the following form

$$(\mathscr{P})_s: \quad \min \left\{ \Pi(u) = \int_\Omega G(\nabla u) \mathrm{d}\Omega \;\mid\; u \in \mathscr{U}_c \right\}, \quad G(\boldsymbol{\gamma}) = V(\Lambda(\boldsymbol{\gamma})) - \boldsymbol{\gamma}^T \boldsymbol{\tau} \tag{50}$$

Under certain regularity conditions, the associated mixed boundary value problem is

$$(BVP)_s: \quad \begin{cases} \nabla \cdot (2\zeta \nabla u) = 0 \quad \text{in } \Omega, \\ \mathbf{n} \cdot (2\zeta \nabla u) = t \quad \text{on } \Gamma_t, \quad u = 0 \quad \text{on } \Gamma_\chi \end{cases} \tag{51}$$

where $\mathbf{n} = \{n_\alpha\} \in \mathbb{R}^2$ is a unit vector norm to $\partial\Omega$, and $\zeta = \nabla V(\xi), \;\; \xi = |\nabla u|^2$.

If $\Gamma_\chi = \partial\Omega$, then $(BVP)_s$ is a Dirichlet boundary value problem, which has only trivial solution due to zero input. For Neumann boundary value problem $\Gamma_t = \partial\Omega$, the external force field must be such that

$$\int_{\Gamma_t} t(\mathbf{x}) \mathrm{d}\Gamma = 0$$

for overall force equilibrium. In this case, if $\bar{\boldsymbol{\chi}}$ is a solution to $(BVP)_s$, then $\boldsymbol{\chi} = \bar{\boldsymbol{\chi}} + \mathbf{c}$ is also a solution for any vector $\mathbf{c} \in \mathbb{R}^3$ since the cylinder is not fixed. Therefore, the mixed boundary value problem $(BVP)_s$ is necessary for anti-plane shear deformation to have a unique solution.

By the fact that the only unknown u is a scalar-valued function, anti-plane shear deformations are one of the simplest classes of deformations that solids can undergo [13]. Indeed, if $V(\xi)$ is a canonical function on $\mathscr{E}_a = \{\xi \in L^p(\Omega)| \;\; \xi(\mathbf{x}) \geq 0 \;\; \forall \mathbf{x} \in \Omega\}$ and for any given $\boldsymbol{\tau} \in \mathscr{T}_a$ such that $\tau = |\boldsymbol{\tau}|$, the canonical dual problem has a very simple form

$$(\mathscr{P}^d)_s: \quad \text{sta}\left\{\Pi^d(\zeta) = \int_\Omega \left[-V^*(\zeta) - \frac{1}{4}\zeta^{-1}\tau^2\right] \mathrm{d}\Omega \mid \zeta \in \mathscr{S}_a\right\}. \tag{52}$$

Since $\Lambda(u) = |\nabla u|^2$, the canonical dual algebraic equation (24) for this problem is

$$4\zeta^2 \nabla V^*(\zeta) = \tau^2(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega. \tag{53}$$

Corollary 1 *For any given nontrivial shear force $t(\mathbf{x}) \neq 0$ on Γ_t such that $\boldsymbol{\tau} \in \mathscr{T}_a \neq \emptyset$, the canonical dual problem $(\mathscr{P}^d)_s$ has at least one nontrivial solution ζ_k. If $\nabla \times (\zeta_k^{-1}\boldsymbol{\tau}) = 0$, the scale-valued function*

$$u_k(\mathbf{x}) = \frac{1}{2}\int_{\mathbf{x}_0}^{\mathbf{x}} \zeta_k^{-1}\boldsymbol{\tau} \cdot d\mathbf{x} \tag{54}$$

along any path from $\mathbf{x}_0 \in \Gamma_\chi$ to $\mathbf{x} \in \Omega$ is a critical point of $\Pi(u)$ and $\Pi(u_k) = \Pi^d(\zeta_k)$.

If $\zeta_k \in \mathscr{S}_a^+$, then u_k is a global minimizer of $(\mathscr{P})_s$.
If $\zeta_k \in \mathscr{S}_a^-$ and $\mathbf{G}(\zeta_k) \succ 0$, then u_k is a local minimizer of $(\mathscr{P})_s$.
If $\zeta_k \in \mathscr{S}_a^-$ and $\mathbf{G}(\zeta_k) \prec 0$, then u_k is a local maximizer of $(\mathscr{P})_s$.

Example. Applications of the canonical duality theory to general anti-plane shear problems have been demonstrated for solving convex exponential and nonconvex polynomial stored energies recently in [7]. In this paper, the following generalized neo-Hookean model is considered

$$V(\xi) = c_1(I_1 - 3) + c_2(I_1 - 3)\log(I_1 - 3) \tag{55}$$

where $c_1,\ c_2$ are positive material constants. Clearly, $V(\xi)$ is convex in $\xi = I_1 - 3$, but

$$\hat{W}(\boldsymbol{\gamma}) = V(I_1(\boldsymbol{\gamma})) = c_1|\boldsymbol{\gamma}|^2 + c_2|\boldsymbol{\gamma}|^2\log|\boldsymbol{\gamma}|^2$$

is a double-well function of the shear strain $\boldsymbol{\gamma} = \nabla u$ (see Fig. 2).

It is easy to check

$$\zeta = \nabla V(\xi) = c_1 + c_2(\log\xi + 1) : \mathscr{E}_a \to \mathscr{E}_a^* = L^q(\Omega)$$

is one-to-one and onto, where q is a dual number of $p \geq 1$, i.e., $1/p + 1/q = 1$. The complementary energy can be obtained easily

$$V^*(\zeta) = \text{sta}\{\xi\zeta - V(\xi)|\ \ \xi \in \mathscr{E}_a\} = c_2\exp[c_2^{-1}(\zeta - c_1) - 1]$$

In this case, the canonical dual algebraic equation is

$$\zeta^2\exp\left[\frac{\zeta - c_1}{c_2} - 1\right] = \tau^2(\mathbf{x}) \ \ \forall \mathbf{x} \in \Omega. \tag{56}$$

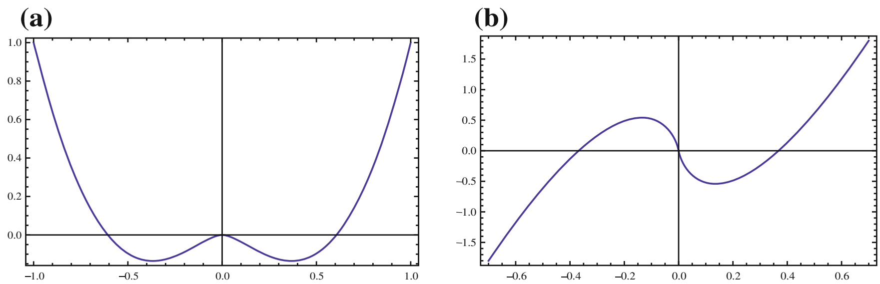


Fig. 2 Graphs of $\hat{W}(\boldsymbol{\gamma})$ (**a**) and its derivative (**b**) ($c_1 = c_2 = 1$)

Let $h^2(\zeta) = \zeta^2 \exp[(\zeta - c_1)/c_2 - 1]$ be the left hand side function in the canonical dual algebraic equation (56). By solving $h'(\zeta_c) = 0$ we known that at $\zeta_c = -2c_2$, $h(\zeta)$ has a local maximum

$$h_{\max}(\zeta_c) = \eta = 2c_2\sqrt{\exp[-3 - c_1/c_2]}.$$

From the graphs of the canonical dual algebraic curve $h(\zeta)$ given in Fig. 3 we can see that the canonical dual algebraic equation (56) may have at most three real solutions in the order of $\zeta_1 \geq 0 \geq \zeta_2 \geq \zeta_3$ depending on $\tau = |\boldsymbol{\tau}(\mathbf{x})|, \;\; \mathbf{x} \in \Omega$ (see Fig. 3b). The Eq. (56) has a unique solution if $\tau > \eta$. In this case, the total strain grand $G(\gamma)$ is strictly G-quasiconvex (see Fig. 4). Figure 5 shows the graphs of $G(\gamma)$ and its canonical dual $G^d(\zeta)$ for $\tau < \eta$. In this case, the function $G(\gamma)$ is nonconvex and has three critical points. The triality theory holds for $G(\gamma)$ and its canonical dual $G^d(\zeta)$

$$G(\gamma_1) = \min_{\gamma \geq 0} G(\gamma) = \max_{\zeta > 0} G^d(\zeta) = G^d(\zeta_1).$$

$$G(\gamma_2) = \min_{\gamma \in \mathscr{G}_o} G(\gamma) = \min_{\zeta > -2c_2} G^d(\zeta) = G^d(\zeta_2).$$

$$G(\gamma_3) = \max_{\gamma \in \mathscr{G}_o} G(\gamma) = \max_{\zeta < -2c_2} G^d(\zeta) = G^d(\zeta_3),$$

where $\mathscr{G}_o$ is a neighborhood of γ_i $(i = 1, 2)$.

5 Conclusions

In summary, the following conclusions can be obtained.

1. The pure complementary energy principle and canonical duality-triality theory developed in [5] are useful for solving general nonlinear boundary value problems in nonlinear elasticity.

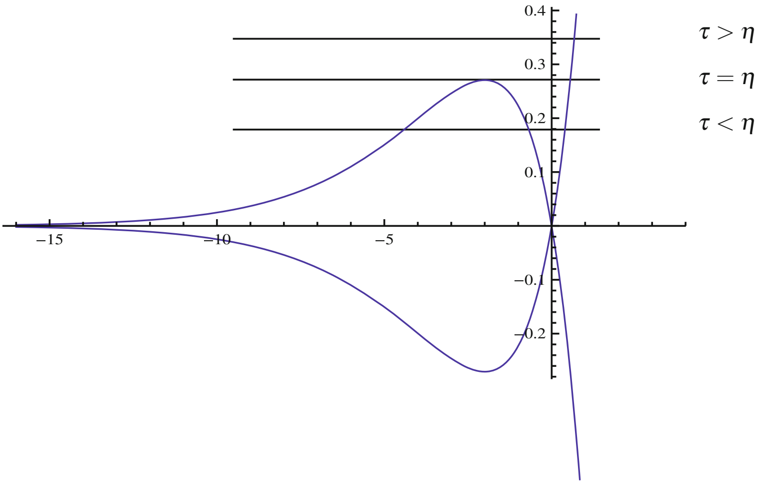
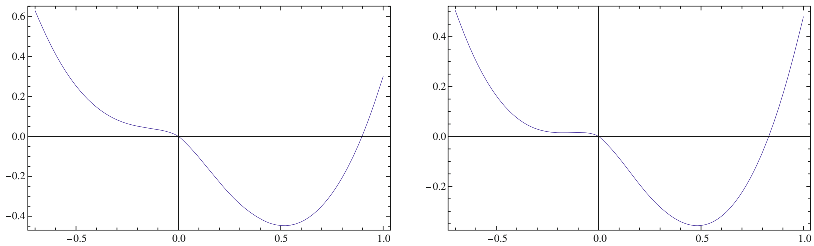
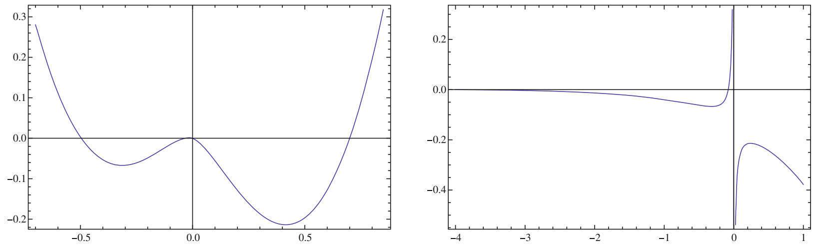


Fig. 3 Dual algebraic curve $h(\zeta)$ $(c_1 = c_2 = 1)$



(a) $\tau > \eta$ (b) $\tau = \eta$

Fig. 4 Graphs of G-quasiconvex $G(\gamma)$ $(c_1 = c_2 = 1)$



(a) Graph of $G(\gamma)$ (b) Graph of $G^d(\zeta)$

Fig. 5 Graphs of $G(\gamma)$ and $G^d(\zeta)$ for $\tau < \eta$ $(c_1 = c_2 = 1)$

2. Both convexity of the total potential and ellipticity condition of the associated fully nonlinear boundary value problem depend not only on the stored energy function, but also sensitively on the external force field.
3. The Legendre–Hadamard condition is only a necessary ellipticity condition for convex systems. The triality theory provides a sufficient condition to identify both global and local extremum solutions for nonconvex problems.

These conclusions are naturally included in the canonical duality-triality theory developed by the author and his coworkers during the last 25 years [5]. Extensive applications have been given in multidisciplinary fields of biology, chaotic dynamics, computational mechanics, information theory, phase transitions, post-buckling, operations research, industrial and systems engineering, etc. (see recent review article [12]).

Acknowledgements The research was supported by US Air Force Office of Scientific Research (AFOSR FA9550-10-1-0487).

References

1. Ciarlet, P.G.: Linear and Nonlinear Functional Analysis with Applications. SIAM, Philadelphia (2013)
2. Gao, D.Y.: Global extremum criteria for nonlinear elasticity. ZAMP **43**, 924–937 (1992)
3. Gao, D.Y.: Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications IMAJ. Appl. Math. **61**, 199–235 (1998)
4. Gao, D.Y.: General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics. Meccanica **34**, 169–198 (1999)
5. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications, xviii + 454pp. Springer, New York (2000)
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.: Analytical solutions to general anti-plane shear problem in finite elasticity. Contin. Mech. Theor. 2015 (2015)
8. Gao, D.Y., Hajilarov, E.: Analytic solutions to three-dimensional finite deformation problems governed by St Venant-Kirchhoff material. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Canonical Duality Theory. Springer (2017). arXiv:1504.02865v1
9. Gao, D.Y., Ogden, R.W.: Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem. ZAMP **59**, 498–517 (2008)
10. Gao, D.Y., Ogden, R.W.: Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation. Q. J. Mech. Appl. Math. **61**(4), 497–522 (2008)
11. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Q. Appl. Math. **47**, 487–504 (1989)
12. Gao, D.Y., Ruan, N., Latorre, V.: Canonical duality-triality: bridge between nonconvex analysis/mechanics and global optimization in complex systems. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Canonical Duality Theory. Springer (2017). arXiv:1410.2665v2
13. Horgan, C.O.: Anti-plane shear deformations in linear and nonlinear solid mechanics. SIAM Rev. **37**(1), 53–81 (1995)
14. Li, S.F., Gupta, A.: On dual configuration forces. J. Elast. **84**, 13–31 (2006)

15. Martin, R.J., Ghiba, I.-D., Neff, P.: Rank-one convexity implies polyconvexity for isotropic, objective and isochoric elastic energies in the two-dimensional case (2015). http://www.researchgate.net/publication/279632850
16. Schröder, J., Neff, P. (eds.): Poly-. Quasi- and Rank-One Convexity in Applied Mechanics. Springer, New York (2010)

Analytic Solutions to 3-D Finite Deformation Problems Governed by St Venant–Kirchhoff Material

David Yang Gao and Eldar Hajilarov

Abstract This paper presents a detailed study on analytical solutions to a general nonlinear boundary-value problem in finite deformation theory. Based on canonical duality theory and the associated pure complementary energy principle in nonlinear elasticity proposed by Gao in (Mech Res Commun 26:31–37, 1999, [6], Wiley Encyclopedia of Electrical and Electronics Engineering, 1999, [7], Meccanica 34:169–198, 1999, [8]), we show that the general nonlinear partial differential equation for deformation is actually equivalent to an algebraic (tensor) equation in stress space. For St Venant–Kirchhoff materials, this coupled cubic algebraic equation can be solved principally to obtain all possible solutions. Our results show that for any given external source field such that the statically admissible first Piola–Kirchhoff stress field has nonzero eigenvalues, the problem has a unique global minimal solution, which is corresponding to a positive-definite second Piola–Kirchhoff stress **T**, and at most eight local solutions corresponding to negative-definite **T**. Additionally, the problem could have 15 unstable solutions corresponding to indefinite **T**. This paper demonstrates that the canonical duality theory and the pure complementary energy principle play fundamental roles in nonconvex analysis and finite deformation theory.

1 Nonconvex Variational Problem and Motivation

A large class of finite deformation problems in nonlinear elasticity can be formulated on the basis of a variational principle ($\mathscr{P}$) in which it is required to minimize certain nonconvex potential energy. Typically, this takes the form

D.Y. Gao (✉) · E. Hajilarov
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia
e-mail: d.gao@federation.edu.au

E. Hajilarov
e-mail: e.hajilarov@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics
and Mathematics 37, DOI 10.1007/978-3-319-58017-3_3

$$(\mathscr{P}): \quad \min_{\boldsymbol{\chi} \in \mathscr{X}_a} \left\{ \Pi(\boldsymbol{\chi}) = \int_{\Omega} W(\nabla \boldsymbol{\chi}) \mathrm{d}\Omega + \int_{\Omega} \phi(\boldsymbol{\chi}) \rho \mathrm{d}\Omega - \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} \mathrm{d}\Gamma \right\}, \quad (1)$$

where $\boldsymbol{\chi}$ represents the deformation field (a bijection), $W(\mathbf{F})$ is the strain energy per unit reference volume, which is a nonlinear differentiable function of the deformation gradient $\mathbf{F} = \nabla \boldsymbol{\chi}$, and ∇ is the gradient operator in a simply connected domain (the reference configuration of the body) $\Omega \subset \mathbb{R}^3$ with boundary $\partial\Omega = \Gamma = \Gamma_t \cup \Gamma_\chi$ such that $\Gamma_t \cap \Gamma_\chi = \emptyset$. Each material point in Ω is labeled by its position vector $\mathbf{X}$ and the corresponding point in the deformed configuration is denoted by $\mathbf{x}\ (= \boldsymbol{\chi}(\mathbf{X}))$. The body force $\mathbf{f}$ (per unit mass) is taken to be conservative with potential $\phi(\mathbf{x})$ so that $\mathbf{f} = -\mathrm{grad}\phi$, and ρ is the reference mass density. On the part Γ_t of the boundary the surface traction $\mathbf{t}$ is prescribed to be of dead-load type, while on Γ_χ the deformation $\boldsymbol{\chi}$ is given. The notation $\mathscr{X}_a$ identifies a *kinematically admissible space* of deformations $\boldsymbol{\chi}$, defined by

$$\mathscr{X}_a = \{\boldsymbol{\chi} \in \mathscr{W}^{1,p}(\Omega; \mathbb{R}^3) \mid \nabla \boldsymbol{\chi} \in \mathscr{F}_a, \quad \boldsymbol{\chi} = \boldsymbol{\chi}_0 \quad \text{on } \Gamma_\chi\}, \quad (2)$$

where $\mathscr{W}^{1,p}$ is the Sobolev space, i.e. a function space in which both $\boldsymbol{\chi}$ and its weak derivative $\nabla \boldsymbol{\chi}$ have a finite $L^p(\Omega)$ norm. $\mathscr{F}_a = \{\mathbf{F} \in \mathscr{L}^p(\Omega; \mathbb{R}^{3\times 3}) |\ \det \mathbf{F} > 0\}$ denotes the admissible deformation gradient space with $p > 1$. Clearly, solutions $\boldsymbol{\chi} \in \mathscr{X}_a$ of the problem $(\mathscr{P})$ are not necessarily to be smooth.

The criticality condition $\delta \Pi(\boldsymbol{\chi}) = 0$ leads to a mixed boundary-value problem (BVP), namely

$$(BVP): \quad \begin{cases} \nabla \cdot [\nabla_{\mathbf{F}} W(\nabla \boldsymbol{\chi})] + \rho \mathbf{f} = \mathbf{0} & \text{in } \Omega, \\ \mathbf{n} \cdot [\nabla_{\mathbf{F}} W(\nabla \boldsymbol{\chi})] = \mathbf{t} & \text{on } \Gamma_t, \end{cases} \quad (3)$$

where $\nabla_{\mathbf{F}} W(\nabla \boldsymbol{\chi}) = \partial W(\mathbf{F})/\partial \mathbf{F}$ (in components $\partial W/\partial F_{i\alpha}$), $\mathbf{n}$ is the unit outward normal to Γ_t and, in component form, we adopt the conventions $\nabla \cdot \boldsymbol{\tau} = \{\partial \tau_{i\alpha}/\partial X_\alpha\}$ and $\boldsymbol{\tau} \cdot \mathbf{n} = \{\tau_{i\alpha} n_\alpha\}$. Note that $\nabla \cdot \boldsymbol{\tau}$ is defined in the weak sense where $\nabla \boldsymbol{\chi}$ is discontinuous. In general, it is rarely possible to solve this nonlinear boundary-value problem by use of direct methods. Indeed, the strain energy $W(\mathbf{F})$ is a nonconvex function of $\mathbf{F}$, the problems $(\mathscr{P})$ and (BVP) are not equivalent, and *(BVP)* may possess multiple solutions. Identification of the global minimizer of the variational problem $(\mathscr{P})$ is a fundamentally difficult task in nonconvex analysis. From the point of view of numerical analysis, any numerical discretization of the problem $(\mathscr{P})$ leads to a nonconvex minimization problem, and it is well known in global optimization theory that most nonconvex minimization problems are NP-hard [11–13].

Duality principles play fundamental roles in sciences and engineering, especially in continuum mechanics and variational analysis. For linear elasticity, since the stored strain energy W is a convex function of the (infinitesimal) strain tensor, it is well known that each potential variational (primal) problem is linked a unique equivalent (dual) complementary variational problem via the conventional Legendre transformation. This one-to-one duality relation is also known as the complementary variational principle, which has been well studied with extensive applications in both mathematical physics and engineering mechanics (see Arthurs, Nobel-Sewell, Oden-Reddy, Tabarrok-Rimrott, etc.).

In finite deformation theory, if the stored-energy density $W(\mathbf{F})$ is a strictly convex function of the deformation gradient tensor $\mathbf{F}$ over the field Ω, then the first Piola–Kirchhoff stress tensor can be uniquely determined by $\boldsymbol{\tau} = \nabla W(\mathbf{F})$ and the complementary energy density W^* can be obtained explicitly by the Legendre transformation:

$$W^*(\boldsymbol{\tau}) = \left\{ \mathbf{F} : \boldsymbol{\tau} - W(\mathbf{F}) \;\middle|\; \boldsymbol{\tau} = \nabla W(\mathbf{F}) \right\}, \tag{4}$$

where $\mathbf{F} : \boldsymbol{\tau}$ is defined as $\mathrm{tr}(\mathbf{F} \cdot \boldsymbol{\tau}^{\mathrm{T}})$ and $^{\mathrm{T}}$ signifies the transpose. In this case, the complementary variational problem can be defined as

$$\min_{\boldsymbol{\tau} \in \mathscr{T}_a} \left\{ \Pi^c(\boldsymbol{\tau}) = \int_\Omega W^*(\boldsymbol{\tau}) \mathrm{d}\Omega - \int_{\Gamma_\chi} \boldsymbol{\chi}_0 \cdot \boldsymbol{\tau} \cdot \mathbf{n} \mathrm{d}\Gamma \right\}, \tag{5}$$

where $\mathscr{T}_a$ is the *statically admissible space* defined by

$$\mathscr{T}_a = \left\{ \boldsymbol{\tau} \in \mathscr{L}^q(\Omega) \;\middle|\; \nabla \cdot \boldsymbol{\tau} + \rho \mathbf{f} = \mathbf{0} \text{ in } \Omega, \; \boldsymbol{\tau} \cdot \mathbf{n} = \mathbf{t} \text{ on } \Gamma_t \right\}, \tag{6}$$

where q is the conjugate number of p, i.e. it is given by $1/p + 1/q = 1$. This complementary variational problem was first studied by Levinson [25]. The well-known Levinson principle states that if $\bar{\boldsymbol{\tau}}$ is a solution of the complementary variational problem (5), then the deformation field $\bar{\boldsymbol{\chi}}$ defined through the inverse constitutive law $\mathbf{F}(\bar{\boldsymbol{\chi}}) = \nabla W^*(\bar{\boldsymbol{\tau}})$ is a solution of the potential variational problem (1) and the complementarity condition

$$\Pi(\bar{\boldsymbol{\chi}}) + \Pi^c(\bar{\boldsymbol{\tau}}) = 0$$

holds. This principle can be proved easily by using the traditional Lagrangian duality theory (see [9–11]).

The Levinson principle is simply the counterpart in finite deformation theory of the complementary variational principle in linear elasticity. In finite deformation theory, the stored strain energy $W(\mathbf{F})$ is in general nonconvex such that the stress-deformation relation $\boldsymbol{\tau} = \nabla W(\mathbf{F})$ is not uniquely invertible [29] and the complementary energy function W^* cannot be defined explicitly via the Legendre transformation. Although by the Fenchel transformation

$$W^\sharp(\boldsymbol{\tau}) = \max_{\mathbf{F}} \{ \mathbf{F} : \boldsymbol{\tau} - W(\mathbf{F}) \},$$

the Fenchel–Moreau type dual problem can be formulated in the form of

$$\max_{\boldsymbol{\tau} \in \mathscr{T}_a} \left\{ \Pi^\sharp(\tau) = \int_{\Gamma_\chi} \boldsymbol{\chi}_0 \cdot \boldsymbol{\tau} \cdot \mathbf{n} \mathrm{d}\Gamma - \int_\Omega W^\sharp(\boldsymbol{\tau}) \mathrm{d}\Omega \right\}, \tag{7}$$

the nonconvexity of W leads only to the so-called *weak duality theorem*

$$\min_{\boldsymbol{\chi} \in \mathscr{X}_a} \Pi(\boldsymbol{\chi}) \geq \max_{\boldsymbol{\tau} \in \mathscr{T}_a} \Pi^\sharp(\boldsymbol{\tau})$$

due to the Fenchel–Young inequality $W(\mathbf{F}) \geq \mathbf{F} : \boldsymbol{\tau} - W^\sharp(\boldsymbol{\tau})$. In nonconvex analysis, the nonzero $\theta = \min_{\boldsymbol{\chi} \in \mathscr{X}_a} \Pi(\boldsymbol{\chi}) - \max_{\boldsymbol{\tau} \in \mathscr{T}_a} \Pi^\sharp(\boldsymbol{\tau}) > 0$ is called the *duality gap*. This duality gap shows that the well-developed Fenchel–Moreau duality theory can be used to solve mainly convex problems.

In finite deformation theory, the well-known Hellinger–Reissner principle [21, 30] and the Fraeijs de Veubeke principle [36] hold for both convex and nonconvex problems. However, these principles are not considered as *pure complementary variational principles* since the Hellinger–Reissner principle involves both the displacement field and the second Piola–Kirchhoff stress tensor; while the Fraeijs de Veubeke principle has both the rotation tensor and the first Piola–Kirchhoff stress as its variational arguments. The existence of a pure complementary variational principle in general finite deformation theory has been discussed by many researchers over several decades (see, for example, [22–24, 26, 28, 29]). Moreover, since the extremality condition in nonconvex variational analysis and global optimization is fundamentally difficult to resolve, none of the classical complementary-dual variational principles in finite deformation theory can be used for reliable numerical computations.

Canonical duality theory provides a potentially useful methodology for solving a large class of nonconvex problems in complex systems. This theory consists mainly of (1) a *canonical dual transformation*, which can be used to formulate perfect dual problems in nonconvex systems; (2) a *complementary-dual variational principle,* which allows a unified analytical solution form in terms of the canonical dual solutions; (3) a *triality theory*, which provides sufficient criteria for identifying both global and local extrema. The original idea of the canonical dual transformation was introduced by Gao and Strang [19] in finite deformation systems. In order to recover the duality gap in nonconvex variational problems, they discovered a so-called *complementary gap function*, which leads to a complementary-dual variational principle in finite deformation mechanics. They proved that if this gap function is positive on a dual feasible space, the generalized Hellinger–Reissner energy is a saddle-functional. It turns out that this gap function provides a sufficient condition for global optimal solution in nonconvex variational problems. Seven years later, it was realized that the negative gap function could be used to identify local extrema. Therefore, a triality theory was first proposed in post-buckling problems of a large deformation beam model [4], and a pure complementary energy principle was eventually obtained in [6]. This principle can be used to obtain a general analytical solution for 3D large deformation elasto-plasticity [8]. It was shown by Gao and Ogden (see [16, 17]) that for one-dimensional nonlinear elasticity problems, both global and local minimal solutions are usually nonsmooth and can't be obtained by any Newton type of numerical methods. For finite dimensional systems, the canonical duality theory has been successfully applied for solving a large class of challenging problems in nonlinear elasticity [3], computational mechanics [1, 20, 34] and global optimization with extensive applications in computational biology [7, 37], chaotic dynamical systems [27, 32], discrete and network optimization [15, 18, 31, 33].

The purpose of this paper is to illustrate the application of the pure complementary variational principle in combination with triality theory by solving a general nonconvex variational problem governed by St Venant–Kirchhoff material. The paper is organized as follows. Section 2 presents a brief review on the canonical duality theory in nonlinear elasticity. Some fundamental issues in nonlinear elasticity are addressed, including the reasons why the Legendre–Hadamard condition provides only necessary condition for local minima, how the Gao–Strang gap function and the triality theory can be used to identify both global and local extremal solutions. In Sect. 3 we show that for the St Venant–Kirchhoff materials, the pure complementary variational problem can be solved principally to obtain all possible solutions. Some concluding remarks are contained in Sect. 4.

2 Canonical Duality Theory and Complementary Variational Principle

It is known that the stored-energy function $W : \mathcal{F}_a \to \mathbb{R}$ must obey certain physical laws and requirements in continuum mechanics, such as the principle of material frame-indifference [35], which lay a mathematical foundation for the canonical duality theory. Let

$$\mathrm{SO}(3) = \{\mathbf{Q} \in \mathbb{R}^{3\times 3} |\ \mathbf{Q}^T = \mathbf{Q}^{-1}, \quad \det \mathbf{Q} = 1\} \tag{8}$$

be the special orthogonal group.

Definition 1 (Objectivity and Isotropy [9]).

(D1) Objective Set and Objective Function*: A subset $\mathcal{F}_a \subset \mathbb{R}^{3\times 3}$ is said to be* objective *if for every $\mathbf{F} \in \mathcal{F}_a$ and every $\mathbf{Q} \in SO(3)$, $\mathbf{QF} \in \mathcal{F}_a$. A scalar-valued function $W : \mathcal{F}_a \to \mathbb{R}$ is said to be* objective *if its domain is objective and*

$$W(\mathbf{QF}) = W(\mathbf{F}) \ \ \forall \mathbf{F} \in \mathcal{F}_a, \ \forall \mathbf{Q} \in SO(3). \tag{9}$$

(D2) Isotropic Set and Isotropic Function*: A subset $\mathcal{F}_a \subset \mathbb{R}^{3\times 3}$ is said to be* isotropic *if for every $\mathbf{F} \in \mathcal{F}_a$ and every $\mathbf{Q} \in SO(3)$, $\mathbf{FQ} \in \mathcal{F}_a$. A scalar-valued function $W : \mathcal{F}_a \to \mathbb{R}$ is said to be* isotropic *if its domain is isotropic and*

$$W(\mathbf{FQ}) = W(\mathbf{F}) \ \ \forall \mathbf{F} \in \mathcal{F}_a, \ \forall \mathbf{Q} \in SO(3). \tag{10}$$

The objectivity implies that the constitutive law of material is independent with the observer (coordinate free). While the isotropy means that the material possesses certain symmetry. Generally speaking, the deformation gradient $\mathbf{F}$ is a two-point tensor, which is not considered as a strain measure. The *right Cauchy–Green tensor* $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ is a (Lagrange type) strain measure which is objective (rotation free), i.e.,

$$\mathbf{C}(\mathbf{QF}) = (\mathbf{QF})^T(\mathbf{QF}) = \mathbf{F}^T\mathbf{Q}^T\mathbf{QF} = \mathbf{C}(\mathbf{F}) \;\; \forall \mathbf{Q} \in \mathrm{SO}(3).$$

Dually, the *left Cauchy–Green tensor* $\mathbf{B} = \mathbf{F}^T\mathbf{F}$ is an isotropic function of $\mathbf{F}$.

In continuum mechanics, the objectivity is also known as *the principle of frame-indifference*. According to P.G. Ciarlet, the stored-energy function of a hyper-elastic material is objective if and only if there exists a function $U(\mathbf{C})$ such that $W(\mathbf{F}) = U(\mathbf{C}(\mathbf{F}))$ (see Theorem 4.2-1 in [2]). This principle lays a foundation for the canonical duality theory.

Indeed, the canonical dual transformation was developed from the concept of the objectivity. The key step of this transformation is the introduction of a geometrically admissible strain measure $\xi = \boldsymbol{\Lambda}(\boldsymbol{\chi}) : \mathscr{X}_a \to \mathscr{E}_a \subset \mathbb{R}^{3\times 3}$ and the canonical function $U(\xi) : \mathscr{E}_a \to \mathbb{R}$ such that the nonconvex stored-energy $W(\mathbf{F})$ can be written in the canonical form $W(\nabla\boldsymbol{\chi}) = U(\boldsymbol{\Lambda}(\boldsymbol{\chi}))$. According to [9], a convex differentiable real-valued function $U(\boldsymbol{\xi})$ is said to be canonical on its domain $\mathscr{E}_a$ if the duality relation $\boldsymbol{\xi}^* = \nabla U(\boldsymbol{\xi}) : \mathscr{E}_a \to \mathscr{E}_a^*$ is invertible such that the conjugate function $U^*(\boldsymbol{\xi}^*)$ of $U(\xi)$ can be defined uniquely by the Legendre transformation

$$U^*(\boldsymbol{\xi}^*) = \{\boldsymbol{\xi} : \boldsymbol{\xi}^* - U(\boldsymbol{\xi}) |\; \boldsymbol{\xi}^* = \nabla U(\boldsymbol{\xi}) \;\; \forall \boldsymbol{\xi} \in \mathscr{E}_a\}. \tag{11}$$

By the theory of convex analysis, it is easy to prove that the following canonical duality relations hold on $\mathscr{E}_a \times \mathscr{E}_a^*$

$$\boldsymbol{\xi}^* = \nabla U(\boldsymbol{\xi}) \;\Leftrightarrow\; \boldsymbol{\xi} = \nabla U^*(\boldsymbol{\xi}^*) \;\Leftrightarrow\; U(\boldsymbol{\xi}) + U^*(\boldsymbol{\xi}^*) = \boldsymbol{\xi} : \boldsymbol{\xi}^* \tag{12}$$

and the pair $(\boldsymbol{\xi}, \boldsymbol{\xi}^*)$ is called the *canonical dual pair* on $\mathscr{E}_a \times \mathscr{E}_a^*$.

Thus, on replacing $W(\nabla\boldsymbol{\chi})$ in the total potential energy $\Pi(\boldsymbol{\chi})$ by its canonical form $W(\nabla\boldsymbol{\chi}) = U(\boldsymbol{\Lambda}(\boldsymbol{\chi}))$, and we take the body force to be a constant, so that $\phi(\boldsymbol{\chi}) = -\mathbf{f}\cdot\boldsymbol{\chi}$, the minimal potential energy variational problem (1) can be written in the following canonical form

$$(\mathscr{P}) : \min_{\boldsymbol{\chi}\in\mathscr{X}_a} \left\{ \Pi(\boldsymbol{\chi}) = \int_\Omega [U(\boldsymbol{\Lambda}(\boldsymbol{\chi})) - \rho\boldsymbol{\chi}\cdot\mathbf{f}]\mathrm{d}\Omega - \int_{\Gamma_t} \boldsymbol{\chi}\cdot\mathbf{t}\mathrm{d}\Gamma \right\}. \tag{13}$$

Furthermore, in terms of $\boldsymbol{\zeta} = \boldsymbol{\xi}^*$ and by the Fenchel–Young equality

$$U(\boldsymbol{\Lambda}(\boldsymbol{\chi})) = \boldsymbol{\Lambda}(\boldsymbol{\chi}) : \boldsymbol{\zeta} - U^*(\boldsymbol{\zeta}),$$

the so-called *total complementary energy functional* [19] $\Xi : \mathscr{X}_a \times \mathscr{E}_a^* \to \mathbb{R}$ can be written, in the present context, as

$$\Xi(\boldsymbol{\chi}, \boldsymbol{\zeta}) = \int_\Omega \left[\boldsymbol{\Lambda}(\boldsymbol{\chi}) : \boldsymbol{\zeta} - U^*(\boldsymbol{\zeta}) - \rho\boldsymbol{\chi}\cdot\mathbf{f}\right]\mathrm{d}\Omega - \int_{\Gamma_t} \boldsymbol{\chi}\cdot\mathbf{t}\mathrm{d}\Gamma. \tag{14}$$

For a given statically admissible field $\boldsymbol{\tau} \in \mathscr{T}_a$, this total complementary functional can be written in the following form

$$\Xi_{\boldsymbol{\tau}}(\boldsymbol{\chi}, \boldsymbol{\zeta}) = \int_{\Gamma_\chi} \boldsymbol{\chi}_0 \cdot \boldsymbol{\tau} \cdot \mathbf{n} \mathrm{d}\Gamma + \int_{\Omega} \big[\boldsymbol{\Lambda}(\boldsymbol{\chi}) : \boldsymbol{\zeta} - U^*(\boldsymbol{\zeta}) - (\nabla \boldsymbol{\chi}) : \boldsymbol{\tau}\big] \mathrm{d}\Omega. \quad (15)$$

For a given $\boldsymbol{\zeta} \in \mathscr{E}_a^*$, the *canonical dual functional* $\Pi^d(\boldsymbol{\zeta})$ is then defined by

$$\Pi^d(\boldsymbol{\zeta}) = \big\{\Xi(\boldsymbol{\chi}, \boldsymbol{\zeta}) \mid \delta_{\boldsymbol{\chi}} \Xi(\boldsymbol{\chi}, \boldsymbol{\zeta}) = 0\big\} = F^{\boldsymbol{\Lambda}}(\boldsymbol{\zeta}) - \int_{\Omega} U^*(\boldsymbol{\zeta}) \mathrm{d}\Omega, \quad (16)$$

where $F^{\boldsymbol{\Lambda}}(\boldsymbol{\zeta})$ is defined by the so-called $\boldsymbol{\Lambda}$-conjugate transformation [9, 13]

$$F^{\boldsymbol{\Lambda}}(\boldsymbol{\zeta}) = \text{sta}\left\{\int_{\Omega} [\boldsymbol{\Lambda}(\boldsymbol{\chi}) : \boldsymbol{\zeta} - \rho \boldsymbol{\chi} \cdot \mathbf{f}] \mathrm{d}\Omega - \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} \mathrm{d}\Gamma \;\Big|\; \boldsymbol{\chi} \in \mathscr{X}_a\right\}, \quad (17)$$

with sta indicating the stationary value at fixed $\boldsymbol{\zeta} \in \mathscr{E}_a^*$. In terms of $\boldsymbol{\tau} \in \mathscr{T}_a$, we have the following form

$$F_{\boldsymbol{\tau}}^{\boldsymbol{\Lambda}}(\boldsymbol{\zeta}) = \int_{\Gamma_\chi} \boldsymbol{\chi}_0 \cdot \boldsymbol{\tau} \cdot \mathbf{n} \mathrm{d}\Gamma + \text{sta}\left\{\int_{\Omega} [\boldsymbol{\Lambda}(\boldsymbol{\chi}) : \boldsymbol{\zeta} - (\nabla \boldsymbol{\chi}) : \boldsymbol{\tau}] \mathrm{d}\Omega \;\Big|\; \boldsymbol{\chi} \in \mathscr{X}_a\right\}. \quad (18)$$

In finite deformation theory,

$$\Pi_{\boldsymbol{\tau}}^d(\boldsymbol{\zeta}) = F_{\boldsymbol{\tau}}^{\boldsymbol{\Lambda}}(\boldsymbol{\zeta}) - \int_{\Omega} U^*(\boldsymbol{\zeta}) \mathrm{d}\Omega \quad (19)$$

is also called the *pure complementary energy functional*, which was first proposed in [6].

Theorem 1 (Complementary-Dual Variational Principle [8]**)** *For a given statically admissible field $\boldsymbol{\tau} \in \mathscr{T}_a$, the following statements are equivalent:*

1. $(\bar{\boldsymbol{\chi}}, \bar{\boldsymbol{\zeta}})$ *is a critical point of* $\Xi_{\boldsymbol{\tau}}(\boldsymbol{\chi}, \boldsymbol{\zeta})$*;*
2. $\bar{\boldsymbol{\chi}}$ *is a critical point of* $\Pi(\boldsymbol{\chi})$*;*
3. $\bar{\boldsymbol{\zeta}}$ *is a critical point of* $\Pi_{\boldsymbol{\tau}}^d(\boldsymbol{\zeta})$.

Moreover, we have

$$\Pi(\bar{\boldsymbol{\chi}}) = \Xi(\bar{\boldsymbol{\chi}}, \bar{\boldsymbol{\zeta}}) = \Xi_{\boldsymbol{\tau}}(\bar{\boldsymbol{\chi}}, \bar{\boldsymbol{\zeta}}) = \Pi_{\boldsymbol{\tau}}^d(\bar{\boldsymbol{\zeta}}). \quad (20)$$

This theorem shows that to find a critical solution to the nonconvex total potential $\Pi(\boldsymbol{\chi})$ is equivalent to find a critical point of its canonical dual function $\Pi_{\boldsymbol{\tau}}^d(\boldsymbol{\zeta})$. For a given $\boldsymbol{\tau} \in \mathscr{T}_a$, different choice of the geometrical measure $\boldsymbol{\Lambda}(\boldsymbol{\chi})$ will leads to different, but equivalent, $\Pi_{\boldsymbol{\tau}}^d(\boldsymbol{\zeta})$ on a subset $\mathscr{S}_a \subset \mathscr{E}_a^*$.

In finite deformation theory, the canonical duality relation is also known as the Hill *work conjugate* and the canonical function $U(\boldsymbol{\xi})$ is called strain energy-density.

According to Hill, for a given hyper-elastic material, there exist a class of strain measures $\boldsymbol{\xi}$ and the associated canonical functions $U(\boldsymbol{\xi})$ such that the associated stress can by defined uniquely by the canonical duality relation $\boldsymbol{\xi}^* = \nabla U(\boldsymbol{\xi})$. There are many canonical strain measures in finite elasticity and many of these strain measures belong to the well-known Hill–Seth strain family

$$\mathbf{E}^{(\eta)} = \frac{1}{2\eta}[\mathbf{C}^\eta - \mathbf{I}],$$

where $\mathbf{I}$ is an identity tensor in $\mathbb{R}^{3\times 3}$ and η is a real number.

Canonical duality theory and pure complementary energy principle for general strain measures have been studied in [9]. In this paper, we consider only the Green–St Venant strain tensor $\mathbf{E}^{(1)}$, simply denoted as $\mathbf{E}$. In this case, the geometrical operator

$$\mathbf{E} = \Lambda(\boldsymbol{\chi}) = \frac{1}{2}[(\nabla\boldsymbol{\chi})^T(\nabla\boldsymbol{\chi}) - \mathbf{I}] : \mathscr{X}_a \to \mathscr{E}_a \tag{21}$$

is a quadratic operator and its domain can be defined by

$$\mathscr{E}_a = \{\mathbf{E} \in \mathscr{L}^{p/2}(\Omega; \mathbb{R}^{3\times 3}) |\ \mathbf{E} = \mathbf{E}^T,\ (2\mathbf{E} + \mathbf{I}) \succ 0\}. \tag{22}$$

We assume that the associated strain energy density $U(\mathbf{E}) : \mathscr{E}_a \to \mathbb{R}$ is convex such that the conjugate stress $\boldsymbol{\zeta}$ of $\mathbf{E}$, denoted by $\mathbf{T}$, can be defined uniquely by the constitutive law

$$\mathbf{T} = \nabla U(\mathbf{E}) : \mathscr{E}_a \to \mathscr{E}_a^*. \tag{23}$$

This associated stress $\mathbf{T}$ is the well-known second Piola–Kirchhoff stress, which is well-defined on $\mathscr{E}_a^* = \{\mathbf{T} \in \mathscr{L}^{p/(p-2)}(\Omega; \mathbb{R}^{3\times 3}) |\ \ \mathbf{T} = \mathbf{T}^T\}$. In this case, the pure complementary energy $\Pi^d_{\boldsymbol{\tau}}$ has the form of

$$\Pi^d_{\boldsymbol{\tau}}(\mathbf{T}) = \int_{\Gamma_\chi} \boldsymbol{\chi}_0 \cdot \boldsymbol{\tau} \cdot \mathbf{n} \mathrm{d}\Gamma - \int_\Omega \left[\frac{1}{2}\mathrm{tr}(\boldsymbol{\tau} \cdot \mathbf{T}^{-1} \cdot \boldsymbol{\tau} + \mathbf{T}) + U^*(\mathbf{T})\right] \mathrm{d}\Omega, \tag{24}$$

which is well-defined on the canonical dual space

$$\mathscr{S}_a = \{\mathbf{T} \in \mathscr{E}_a^* |\ \ \mathrm{tr}(\boldsymbol{\tau} \cdot \mathbf{T}^{-1} \cdot \boldsymbol{\tau}) \in \mathscr{L}^1(\Omega; \mathbb{R})\ \ \forall \boldsymbol{\tau} \in \mathscr{T}_a\}. \tag{25}$$

Therefore, the canonical dual problem is to find the critical point $\bar{\mathbf{T}} \in \mathscr{S}_a$ such that

$$(\mathscr{P}^d) : \quad \Pi^d_{\boldsymbol{\tau}}(\bar{\mathbf{T}}) = \mathrm{sta}\{\Pi^d_{\boldsymbol{\tau}}(\mathbf{T}) |\ \mathbf{T} \in \mathscr{S}_a\}. \tag{26}$$

Theorem 2 (Analytical Solution Form [9]**)** *For a given $\boldsymbol{\tau} \in \mathscr{T}_a$, if $\bar{\mathbf{T}}$ is a critical point of $\Pi^d_{\boldsymbol{\tau}}(\mathbf{T})$, then along any path from $\mathbf{X}_0 \in \Gamma_\chi$ to $\mathbf{X} \in \Omega$, the deformation defined by*

$$\bar{\boldsymbol{\chi}} = \int_{\mathbf{X}_0}^{\mathbf{X}} \boldsymbol{\tau} \cdot \bar{\mathbf{T}}^{-1} \cdot \mathrm{d}\mathbf{X} + \boldsymbol{\chi}_0(\mathbf{X}_0) \tag{27}$$

is a critical solution to ($\mathscr{P}$). *Moreover, if*

$$\nabla \times (\boldsymbol{\tau} \cdot \bar{\mathbf{T}}^{-1}) = \mathbf{0}, \tag{28}$$

then $\bar{\boldsymbol{\chi}}$ *is a closed form solution to the boundary-value problem (BVP) (3).*

The proof of this theorem can be found in [5, 6, 8]. In fact, the criticality condition $\delta \Pi^d_{\boldsymbol{\tau}}(\mathbf{T}) = 0$ leads to the following dual tensor equation:

$$\mathbf{T} \cdot \left[\mathbf{I} + 2(\nabla U^*(\mathbf{T}))\right] \cdot \mathbf{T} = \boldsymbol{\tau}^{\mathrm{T}} \cdot \boldsymbol{\tau}, \tag{29}$$

which is equivalent to

$$\nabla U^*(\bar{\mathbf{T}}) = \frac{1}{2}\left((\boldsymbol{\tau} \cdot \bar{\mathbf{T}}^{-1})^T \boldsymbol{\tau} \cdot \bar{\mathbf{T}}^{-1} - \mathbf{I}\right).$$

This is actually the constitutive law $\mathbf{E} = \boldsymbol{\Lambda}(\bar{\boldsymbol{\chi}}) = \frac{1}{2}[\mathbf{F}^T\mathbf{F} - \mathbf{I}] = \nabla U^*(\bar{\mathbf{T}})$ subjected to $\mathbf{F} = \boldsymbol{\tau} \cdot \bar{\mathbf{T}}^{-1}$. Therefore, if the compatibility condition $\nabla \times \mathbf{F} = \mathbf{0}$, in index notation

$$\frac{\partial F_{i\alpha}}{\partial X_\beta} = \frac{\partial F_{i\beta}}{\partial X_\alpha},$$

holds, then $\mathbf{F}$ is the deformation gradient and $\bar{\boldsymbol{\chi}}$ is a solution to *(BVP)*.

Remark 1 (PDE ⇔ Algebraic Equation) *Theorem 2 shows that by the pure complementary energy principle, the nonlinear partial differential equation (BVP) is equivalently converted to a canonical dual tensor equation (29), which can be solved to obtain the stress field* $\bar{\mathbf{T}}$ *for certain materials. From the Eq. (29) we know that* $\mathbf{T} = \mathbf{0}$ *if* $\boldsymbol{\tau} = \mathbf{0}$. *Therefore, although* $\mathbf{T}^{-1}$ *appears in* $\Pi^d_{\boldsymbol{\tau}}(\mathbf{T})$, *this pure complementary energy is well-defined on* $\mathscr{S}_a$. *The Eq. (27) presents an analytical solution form to the boundary-value problem in terms of the canonical dual stress field* $\bar{\mathbf{T}}$ *and the statically admissible* $\boldsymbol{\tau} \in \mathscr{T}_a$. *Of course, this is purely formal and in general it is not easy to obtain the solution for general practices unless the deformation compatibility condition (28) holds.*

It has been assumed here that the relation between $\mathbf{T}$ *and* $\mathbf{E}$ *is invertible. This certainly holds in a neighborhood of the (stress-free) reference configuration since the canonical strain energy* $U(\mathbf{E})$ *is convex in such a neighborhood. It is a reasonable assumption to extend this to a sufficiently large domain that includes deformations of practical interest. Finite element implementations of nonlinear elasticity are usually based on the variables* $\mathbf{T}$ *and* $\mathbf{E}$ *and the associated tangent tensor* $\partial\mathbf{T}/\partial\mathbf{E} = \nabla^2 U(\mathbf{E})$, *which is assumed to be positive definite. It is always possible to select forms of the strain-energy function* W *such that this is the case, although the possibility of its failure for particular materials is not in general ruled out.*

In terms of the deformation $\chi \in \mathscr{X}_a$ and the second Piola–Kirchhoff stress $\mathbf{T} \in \mathscr{E}_a^*$, the total complementary functional $\Xi(\chi, \mathbf{T})$ can be written as

$$\Xi_{\tau}(\chi, \mathbf{T}) = \int_{\Omega} \left[\mathbf{E}(\chi) : \mathbf{T} - U^*(\mathbf{T}) - (\nabla \chi) : \tau\right] \mathrm{d}\Omega + \int_{\Gamma_\chi} \chi_0 \cdot \tau \cdot \mathbf{n} \mathrm{d}\Gamma \tag{30}$$

which is actually the well-known Hellinger–Reissner energy if the first Piola–Kirchhoff stress is replaced by external force field. From the nonlinear canonical dual tensor equation (29) we know that for a given $\tau \in \mathscr{T}_a$, the pure complementary energy $\Pi_{\tau}^d(\mathbf{T})$ may have multiple critical points. In order to identify the global extremum, we need to introduce the following subspaces:

$$\mathscr{S}_a^+ = \{\mathbf{T} \in \mathscr{S}_a |\ \mathbf{T} \succ 0\}, \quad \mathscr{S}_a^- = \{\mathbf{T} \in \mathscr{S}_a |\ \mathbf{T} \prec 0\}. \tag{31}$$

Theorem 3 *Suppose for a given $\tau \in \mathscr{T}_a$, the pair $(\bar{\chi}, \bar{\mathbf{T}})$ is an isolated critical point of $\Xi_{\tau}(\chi, \mathbf{T})$. If $\bar{\mathbf{T}} \in \mathscr{S}_a^+$, then $\bar{\chi}$ is a global minimizer of $\Pi(\chi)$ on $\mathscr{X}_a$ if and only if $\bar{\mathbf{T}}$ is a global maximizer of $\Pi_{\tau}^d(\mathbf{T})$ on $\mathscr{S}_a^+$, i.e.,*

$$\Pi(\bar{\chi}) = \min_{\chi \in \mathscr{X}_a} \Pi(\chi) \quad \Leftrightarrow \quad \max_{\mathbf{T} \in \mathscr{S}_a^+} \Pi_{\tau}^d(\mathbf{T}) = \Pi_{\tau}^d(\bar{\mathbf{T}}). \tag{32}$$

If $\bar{\mathbf{T}} \in \mathscr{S}_a^-$, then $\bar{\chi}$ is a local maximizer of $\Pi(\chi)$ if and only if $\bar{\mathbf{T}}$ is a local maximizer of $\Pi_{\tau}^d(\mathbf{T})$, i.e., on a neighborhood $\mathscr{X}_o \times \mathscr{S}_o \subset \mathscr{X}_a \times \mathscr{S}_a^-$,

$$\Pi(\bar{\chi}) = \max_{\chi \in \mathscr{X}_o} \Pi(\chi) \quad \Leftrightarrow \quad \max_{\mathbf{T} \in \mathscr{S}_o} \Pi_{\tau}^d(\mathbf{T}) = \Pi_{\tau}^d(\bar{\mathbf{T}}). \tag{33}$$

If $\bar{\mathbf{T}} \in \mathscr{S}_a^-$ and $\nabla_{\mathbf{F}}^2 W(\nabla \bar{\chi}) \succ 0$, then $\bar{\chi}$ is a local minimizer of $\Pi(\chi)$.

Remark 2 (The Complementary Gap Function and Triality Theory)
Theorem 3 shows that the extremality of the primal solution χ depends on its canonical dual solution $\mathbf{S}$. This result was first discovered by Gao and Strang in [19], i.e. they proved that $\bar{\chi}(\bar{\mathbf{S}})$ is a global minimizer of $\Pi(\chi)$ if the complementary gap function satisfies

$$G_{ap}(\chi, \bar{\mathbf{S}}) = \int_{\Omega} \frac{1}{2}[(\nabla \chi)^T (\nabla \chi) + \mathbf{I}] : \bar{\mathbf{S}} \mathrm{d}\Omega \geq 0 \ \ \forall \chi \in \mathscr{X}_a \tag{34}$$

Since $G_{ap}(\chi, \bar{\mathbf{S}})$ is quadratic in χ, this gap function is positive for any given $\chi \in \mathscr{X}_a$ if $\bar{\mathbf{S}} \succeq 0$. Replacing $\mathbf{F} = \nabla \chi$ by $\mathbf{F} = \tau \cdot \mathbf{S}^{-1}$, this gap function can be written as the so-called pure gap function

$$G_{ap}(\chi(\mathbf{S}), \mathbf{S}) = \int_{\Omega} \frac{1}{2} tr(\tau \cdot \mathbf{S}^{-1} \cdot \tau + \mathbf{S}) \mathrm{d}\Omega, \tag{35}$$

which is a main term in the pure complementary energy $\Pi^d_{\boldsymbol{\tau}}(\mathbf{S})$ *in addition to* $U^*(\mathbf{S})$. *Comparing* $\Pi^d_{\boldsymbol{\tau}}(\mathbf{S})$ *with* $\Pi^\sharp(\boldsymbol{\tau})$ *given by (7), we can understand that this gap function not only recovers the duality gap in the Fenchel–Moreau duality theory, but also provides a global extremality condition for nonconvex variational problem* $(\mathscr{P})$.

To see this in detail, let us consider the canonical transformation $W(\mathbf{F}) = U(\mathbf{E}(\mathbf{F}))$. *By chain rule we have*

$$\frac{\partial^2 W(\mathbf{F})}{\partial F^i_\alpha \partial F^j_\beta} = \delta^{ij} S_{\alpha\beta} + \sum_{\theta,\nu=1}^{3} F^i_\theta H_{\theta\alpha\beta\nu} F^j_\nu, \tag{36}$$

where $\mathbf{H} = \{H_{\theta\alpha\beta\nu}\} = \nabla^2 U(\mathbf{E})$. *By the convexity of the canonical function* $U(\mathbf{E})$, *we have* $\mathbf{H} \succ 0$. *Therefore, if* $\mathbf{S} = \{S_{\alpha\beta}\} \in \mathscr{S}^+_a$, *the Hessian* $\nabla^2 W(\mathbf{F}) \succ 0$ *and, by Gao and Strang [19], the associated deformation field* $\boldsymbol{\chi}$ *is a global minimizer of* $\Pi(\boldsymbol{\chi})$. *The statement (32) shows that the nonconvex minimization problem* $(\mathscr{P})$ *is equivalent to a concave maximization dual problem over a convex space* $\mathscr{S}^+_a$, *i.e.,*

$$\max\{\Pi^d_{\boldsymbol{\tau}}(\mathbf{T}) | \ \mathbf{T} \in \mathscr{S}^+_a\}, \tag{37}$$

which is much easier than the nonconvex primal problem $(\mathscr{P})$. *The global optimality condition* $\mathbf{S} \in \mathscr{S}^+_a$ *is a strong case of Gao and Strang's positive gap function (34).*

Subsequently, in a study of post-buckling analysis for a nonlinear beam theory, it was found that if the dual solution $\bar{\mathbf{T}}$ *is negative definite in the domain* Ω, *the solution* $\bar{\boldsymbol{\chi}}$ *could be either a local minimizer or a local maximizer of the total potential energy. To see this, we substitutive* $\mathbf{F} = \boldsymbol{\tau} \cdot \mathbf{T}^{-1}$ *into (36) to obtain*

$$\frac{\partial^2 W(\mathbf{F})}{\partial F^i_\alpha \partial F^j_\beta} = \delta^{ij} S_{\alpha\beta} + \sum_{\theta,\nu,\delta,\lambda=1}^{3} \tau^i_\theta S^{-1}_{\theta\delta} H_{\delta\alpha\beta\nu} S^{-1}_{\nu\lambda} \tau^j_\lambda \tag{38}$$

which shows that even if $\mathbf{T} \prec 0$, *the Hessian matrix* $\nabla^2 W(\mathbf{F})$ *could be either positive or negative definite, depending on the eigenvalues of* $\mathbf{T} \in \mathscr{S}^-_a$. *Thus, in addition to the double-max duality (33), we have the so-called double-min duality*

$$\Pi(\bar{\boldsymbol{\chi}}) = \min_{\boldsymbol{\chi} \in \mathscr{X}_o} \Pi(\boldsymbol{\chi}) \quad \Leftrightarrow \quad \min_{\mathbf{T} \in \mathscr{S}_o} \Pi^d_{\boldsymbol{\tau}}(\mathbf{T}) = \Pi^d_{\boldsymbol{\tau}}(\bar{\mathbf{T}}), \tag{39}$$

which holds under certain condition (see [12]). For this reason, a so-called triality theory was proposed first in post-buckling analysis of a large deformed beam model [4], and then in general nonconvex mechanics [8, 9]. This triality theory reveals an important fact in nonconvex analysis, i.e. for a given statically admissible field $\boldsymbol{\tau} \in \mathscr{T}_a$, *if the canonical dual equation (29) has multiple solutions* $\{\mathbf{S}_k\}$ *in a subset* $\Omega_o \subset \Omega$, *then the boundary-value problem (BVP) could have an infinite number of solutions* $\{\boldsymbol{\chi}_k(\mathbf{x})\}$ *in* Ω. *The well-known Legendre–Hadamard (L-H) condition is only a necessary condition for a local minimal solution, while the triality theory can identify not only the global minimizers, but also both local minimizers and local* *maximizers. It is known that an elliptic equation is corresponding to a convex variational problem. If the boundary-value problem (3) has multiple solutions* $\{\boldsymbol{\chi}_k(\mathbf{x})\}$ *at one material point* $\mathbf{x} \in \Omega$*, the total potential* $\Pi(\boldsymbol{\chi})$ *is not convex and the operator* $A(\boldsymbol{\chi}) = \nabla \cdot [\nabla_{\mathbf{F}} W(\nabla \boldsymbol{\chi})]$ *may not be elliptic at* $\mathbf{x} \in \Omega$ *even if the L-H condition holds at certain* $\boldsymbol{\chi}_k(\mathbf{x})$*.*

The pure complementary energy principle and triality theory play a fundamental role not only in nonconvex analysis, but also in computational science and global optimization (see [12, 14, 15]).

3 Application to St Venant–Kirchhoff Material

For St. Venant–Kirchhoff material, the canonical energy function $U(\mathbf{E})$ has the most simple form:

$$U(\mathbf{E}) = \mu \mathrm{tr}(\mathbf{E}^2) + \frac{1}{2}\lambda(\mathrm{tr}\mathbf{E})^2. \tag{40}$$

The second Piola–Kirchhoff stress depends linearly on the Green–St Venant strain via the Hooke's law:

$$\mathbf{S} = \nabla U(\mathbf{E}) = 2\mu \mathbf{E} + \lambda(\mathrm{tr}\mathbf{E})\mathbf{I} = \mathbf{H} : \mathbf{E}, \tag{41}$$

where $\mathbf{H}$ is the Hooke tensor for St Venant–Kirchhoff material. The complementary energy is

$$U^*(\mathbf{S}) = \frac{1}{4\mu}\mathrm{tr}(\mathbf{S}^2) - \frac{\lambda}{4\mu(3\lambda + 2\mu)}(\mathrm{tr}\mathbf{S})^2, \tag{42}$$

and hence

$$\mathbf{E} = \nabla U^*(\mathbf{T}) = \frac{1}{2\mu}\mathbf{T} - \frac{\lambda}{2\mu(3\lambda + 2\mu)}(\mathrm{tr}\mathbf{T})\mathbf{I} \equiv \mathbf{H}^{-1} : \mathbf{S}. \tag{43}$$

By the canonical dual tensor equation (29), we have

$$\mathbf{S}^2 + 2\mathbf{T}(\mathbf{H}^{-1} : \mathbf{T})\mathbf{T} = \mathbf{T}^2 + \frac{1}{\mu}\mathbf{T}^3 - \frac{\lambda}{\mu(3\lambda + 2\mu)}(\mathrm{tr}\mathbf{T})\mathbf{T}^2 = \boldsymbol{\tau}^T \boldsymbol{\tau}. \tag{44}$$

The diagonalization of this tensor equation leads to the following coupled cubic algebraic system:

$$S_i^2 + \frac{1}{\mu}S_i^3 - \frac{\lambda}{\mu(3\lambda + 2\mu)}(S_1 + S_2 + S_3)S_i^2 = \tau_i^2 \quad i = 1, 2, 3. \tag{45}$$

For convenience, we make the following substitutions in (45):

$$S_i = \mu\zeta_i, \quad \tau_i^2 = \mu^2\sigma_i, \quad i = 1, 2, 3,$$

and $k = \frac{\lambda}{3\lambda+2\mu} < 1/3$ (due to $\mu > 0$). So, the system (45) can be written as follows

$$\zeta_i^3 + \zeta_i^2 - k(\zeta_1 + \zeta_2 + \zeta_3)\zeta_i^2 = \sigma_i, \quad i = 1, 2, 3. \tag{46}$$

3.1 Auxiliary Equation

In this section we will study solutions of the following equation:

$$G(\zeta, q, \sigma) = \zeta^3 + (1 - kq)\zeta^2 - \sigma = 0, \tag{47}$$

where $\sigma > 0$, $0 < k < \frac{1}{3}$, and q is an arbitrary real number. Also, since $\sigma > 0$, we can assume that $\zeta \neq 0$.

Since the parameter q in this section is assumed to be independent on ζ, the following results are similar to one-dimensional nonlinear elasticity problems studied by Gao [9, 10], Gao and Ogden [16].

Lemma 3.1 *If* $\zeta_1, \zeta_2, \zeta_3$ *are solutions of the equations* $G(\zeta, q, \sigma_1) = 0$, $G(\zeta, q, \sigma_2) = 0$, $G(\zeta, q, \sigma_3) = 0$ *correspondingly, and* $\zeta_1 + \zeta_2 + \zeta_3 = q$, *then* $\zeta_1, \zeta_2, \zeta_3$ *satisfy (46).*

Proof. Obvious. □

Lemma 3.2 *Equation (47) has exactly one positive solution. It has negative solutions iff*

$$q \leq \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}})$$

There is only one negative solution if and only if $q = \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}})$.

Proof. To check that there is exactly one positive root one can apply the Descartes' rule of signs. To prove the rest, let's fix q, σ and notice that $G(\zeta, q, \sigma) = 0$ has negative solutions iff it has at least two different solutions. This will happen iff the values of the function at local minimum and maximum have different signs. The extremums of G are at $\zeta_0 = -\frac{2}{3}(1 - kq)$ and 0. Since the value of G at 0 is $-\sigma < 0$, we find when $G(\zeta_0, q, \sigma) \geq 0$. Solving this inequality we get $q \leq \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}})$.

Corollary 3.1 *The equation* $G(\zeta, 0, \sigma) = 0$ *has negative solution(s) iff*

$$\sigma \leq \frac{4}{27}.$$

Proof. Apply Lemma 3.2 to $q = 0$. □

Lemma 3.3 *Let's fix $\sigma > 0$ and assume that ζ_0, q_0 satisfy (47), and $\zeta_0 \neq 0$, $\zeta_0 \neq -\sqrt[3]{2\sigma}$. Then there exists a unique continuously differentiable function $\zeta(q)$, such that $\zeta(q_0) = \zeta_0$, $\zeta(q)$ and q both satisfy (47) and*

$$\frac{d\zeta}{dq} = \frac{k\zeta^3}{\zeta^3 + 2\sigma}.$$

Moreover, there are three possibilities ("branches") for $\zeta(q)$:

(a) If $\zeta_0 \in (-\infty, -\sqrt[3]{2\sigma})$, then the range of $\zeta(q)$ is $(-\infty, -\sqrt[3]{2\sigma})$, the domain is $(-\infty, \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}}))$, and $\zeta(q)$ is monotonically increasing.
(b) If $\zeta_0 \in (-\sqrt[3]{2\sigma}, 0)$, then the range of $\zeta(q)$ is $(-\sqrt[3]{2\sigma}, 0)$, the domain is $(-\infty, \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}})))$, and $\zeta(q)$ is monotonically decreasing.
(c) If $\zeta_0 \in (0, +\infty)$, then the range of $\zeta(q)$ is $(0, +\infty)$, the domain is $(-\infty, +\infty)$, and $\zeta(q)$ is monotonically increasing.

Proof. Let's fix σ and find q from (47)

$$q(\zeta) = \frac{\zeta^3 + \zeta^2 - \sigma}{k\zeta^2}.$$

Since, $\frac{dq}{d\zeta} = \frac{\zeta^3 + 2\sigma}{k\zeta^3}$ and $\sigma > 0$ it is obvious that $q(\zeta)$ is monotonically increasing in the intervals $\zeta \in (-\infty, -\sqrt[3]{2t})$ and $\zeta \in (0, +\infty)$ and is monotonically decreasing in the interval $\zeta \in (-\sqrt[3]{2t}, 0)$. The corresponding intervals for q are $(-\infty, \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}}))$, $(-\infty, +\infty)$, and $(-\infty, \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}}))$. Also, one can easily check that

$$\frac{d\zeta}{dq} = \frac{k\zeta^3}{\zeta^3 + 2\sigma}.$$

Thus, the lemma is proved. □

Definition 3.1 *The three branches of $\zeta(q, \sigma)$ (σ is fixed) described in Lemma 3.3 will be denoted as follows:*

(a) $\zeta^1(q, \sigma)$ is a positive branch with the domain $(-\infty, +\infty)$ and range $(0, +\infty)$;
(b) $\zeta^3(q, \sigma) < \zeta^2(q, \sigma)$ are two negative branches with the domain $(-\infty, \frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}}))$ and ranges $(-\infty, -\sqrt[3]{2\sigma})$ and $(-\sqrt[3]{2\sigma}, 0)$ correspondingly.(Note that Corollary 3.1 implies that $\frac{1}{k}(1 - 3\sqrt[3]{\frac{\sigma}{4}}) \leq 0$)

Definition 3.2 *Let's introduce the following notations:*

(a) $\bar{\zeta}^i(q, \sigma) = \zeta^i(q, \sigma) - \frac{q}{3}$, $i = 1, 2, 3$;
(b) $F^{i,j,k}(q, \sigma_1, \sigma_2, \sigma_3) = \bar{\zeta}^i(q, \sigma_1) + \bar{\zeta}^j(q, \sigma_2) + \bar{\zeta}^k(q, \sigma_3)$, $i, j, k = 1, 2, 3$.

Lemma 3.4 *The following statements are true:*

(a) *For* $i = 1, 2, 3$

$$\bar{\zeta}^i(q,\sigma) = -\frac{(1-3k)\zeta^i(q,\sigma)^3 + \zeta^i(q,\sigma)^2 - \sigma}{3k\zeta^i(q,\sigma)^2}$$

and

$$\frac{d\bar{\zeta}^i}{dq} = -\frac{(1-3k)\zeta^i(q,\sigma)^3 + 2\sigma}{3(\zeta^i(q,\sigma)^3 + 2\sigma)}.$$

(b) $\zeta^1(0,\sigma) = \bar{\zeta}^1(0,\sigma) > 0$, $\zeta^2(0,\sigma) = \bar{\zeta}^2(0,\sigma) < 0$, *and* $\zeta^3(0,\sigma) = \bar{\zeta}^3(0,\sigma) < 0$.

(c) *For a fixed* σ, $\bar{\zeta}^1(q,\sigma)$ *is monotonically decreasing in* q *and* $\lim_{q\to+\infty}\bar{\zeta}^1(q,\sigma) = -\infty$.

(d) *For a fixed* σ,

$$\lim_{q\to-\infty}\bar{\zeta}^2(q,\sigma) = +\infty \quad and \quad \lim_{q\to-\infty}\bar{\zeta}^3(q,t) = +\infty.$$

(e) *For fixed* $\sigma_1, \sigma_2, \sigma_3$, *each of* $F^{i,j,k}(q,\sigma_1,\sigma_2,\sigma_3)$, $i, j, k = 1, 2, 3$, *is continuous. Moreover,* $F^{1,1,1}(q,\sigma_1,\sigma_2,\sigma_3)$ *is monotonically decreasing in* q.

Proof. To check (a), first substitute $q(\zeta) = \frac{\zeta^3+\zeta^2-\sigma}{k\zeta^2}$ into $\bar{\zeta}^i(q,\sigma) = \zeta^i(q,\sigma) - \frac{q}{3}$, $i = 1, 2, 3$. Expression for $\frac{d\bar{\zeta}^i}{dq}$ can be obtained either by direct differentiation of the previously obtained expression for $\zeta^i(q,\sigma)$ or subtracting $\frac{1}{3}$ from $\frac{d\zeta}{dq} = \frac{k\zeta^3}{\zeta^3+2\sigma}$.
(b) is obvious.
To prove (c), recall, that $k < \frac{1}{3}$, and use formulas from (a). To prove (d), recall, that $k < \frac{1}{3}$, and use the first formula from (a). (e) immediately follows from (a) and (c). □

Lemma 3.5 *Solutions,* $\zeta^1(0,\sigma), \zeta^2(0,\sigma), \zeta^3(0,\sigma)$, *of the equation* $G(\zeta,0,\sigma) = \zeta^3 + \zeta^2 - \sigma = 0$, $0 < \sigma \le \frac{4}{27}$, *enjoy the following properties:*

(a) *If* $\sigma = \frac{4}{27}$ *the solutions are* $\zeta^1(0,\frac{4}{27}) = \frac{1}{3}$, $\zeta^2(0,\frac{4}{27}) = \zeta^3(0,\frac{4}{27}) = -\frac{2}{3}$

(b) *If* $0 < \sigma_1 < \sigma_2 \le 0$, *then*

$$0 < \zeta^1(0,\sigma_1) < \zeta^1(0,\sigma_2) \le \frac{1}{3}$$

and

$$-1 < \zeta^3(0,\sigma_1) < \zeta^3(0,\sigma_2) \le -\frac{2}{3} \le \zeta^2(0,\sigma_2) < \zeta^2(0,\sigma_1) < 0$$

(c) $\zeta^1(0,\sigma) + \zeta^2(0,\sigma) < 0$

Proof. (a) can be checked directly.
To prove (b), one can either apply the implicit function theorem to $H(\zeta,\sigma) = G(\zeta,0,\sigma) = 0$. Or, less formally, draw the graph of $y = \zeta^3 + \zeta^2 - \frac{4}{27}$ and observe what happens to its roots when the graph is shifted upward until it becomes $y = \zeta^3 + \zeta^2$.
(c) Obviously, $\zeta^1(0,\sigma) + \zeta^2(0,\sigma) + \zeta^3(0,\sigma) = -1$. So, $\zeta^1(0,\sigma) + \zeta^2(0,\sigma) = -1 - \zeta^3(0,\sigma) < 0$, since $\zeta^3(0,\sigma) > -1$.

3.2 Solutions of the St. Venant–Kirchhoff Material

We are now ready to present our main results.

Theorem 4 *For any given force field* $\mathbf{f} : \Omega \to \mathbb{R}^d$ *and the surface traction* $\mathbf{t} : \Gamma_t \to \mathbb{R}^d$ *such that the statically admissible stress* $\boldsymbol{\tau} \in \mathscr{T}_a$ *has nonzero eigenvalues almost ever where in* Ω*, the canonical dual problem* $(\mathscr{P}^d)$ *has a unique positive critical solution* $\mathbf{T} \in \mathscr{S}_a^+$.

Proof. We need to prove that for arbitrarily given $\sigma_1, \sigma_2, \sigma_3 > 0$, the system of equations (46) has a unique positive solution $(\zeta_1, \zeta_2, \zeta_3)$, such that all $\zeta_i > 0$, $i = 1, 2, 3$. From Lemma 3.4(b), it follows that

$$F^{1,1,1}(0, \sigma_1, \sigma_2, \sigma_3) > 0.$$

From Lemma 3.4(c), it follows that for some $q_1 > 0$, large enough,

$$F^{1,1,1}(q_1, \sigma_1, \sigma_2, \sigma_3) < 0.$$

Therefore, since $F^{1,1,1}$ is continuous and monotonically decreasing in q (Lemma 3.4(e)), there exists a unique q_0, $0 < q_0 < q_1$, such that

$$F^{1,1,1}(q_0, \sigma_1, \sigma_2, \sigma_3) = 0.$$

i.e.

$$\zeta^1(q_0, \sigma_1) + \zeta^1(q_0, \sigma_2) + \zeta^1(q_0, \sigma_3) = q_0.$$

So, from Lemma 3.1 it follows that $\zeta^1(q_0, \sigma_1)$, $\zeta^1(q_0, \sigma_2)$, $\zeta^1(q_0, \sigma_3)$ form a positive solution of (46), which are eigenvalues of the second Piola–Kirchhoff stress $\mathbf{T}$. Therefore, Problem $(\mathscr{P}^d)$ has a unique global maximizer $\mathbf{T} \in \mathscr{S}_a^+$. □

Theorem 5 *For any given force field* $\mathbf{f} : \Omega \to \mathbb{R}^d$ *and the surface traction* $\mathbf{t} : \Gamma_t \to \mathbb{R}^d$ *such that the eigenvalues of the statically admissible stress tensor function* $\boldsymbol{\tau} \in \mathscr{T}_a$ *satisfy* $0 < \sigma_1, \sigma_2, \sigma_3 < \frac{4}{27}$, *the total complementary energy* $\Pi^d_{\boldsymbol{\tau}}(\mathbf{T})$ *has eight negative solutions* $\mathbf{T}_k \in \mathscr{S}_a^-$, $k = 1, \ldots, 8$.

Proof. We need to prove that for arbitrarily given $0 < \sigma_1, \sigma_2, \sigma_3 < \frac{4}{27}$, the system of equations (46) has 8 solutions $(\zeta_1, \zeta_2, \zeta_3)$, such that all $\zeta_i < 0$, $i = 1, 2, 3$. From Corollary 3.1 it follows that each of the equations $G(\zeta, 0, \sigma_i)$, has two negative solutions, $\zeta^2(0, \sigma_i) > \zeta^3(0, \sigma_i)$, $i = 1, 2, 3$. From Lemma 3.4(b), it follows that for $i, j, k = 2, 3$

$$F^{i,j,k}(0, \sigma_1, \sigma_2, \sigma_3) < 0.$$

From Lemma 3.4(d) it follows that there exists $q_1 < 0$ such that

$$F^{i,j,k}(q_1, \sigma_1, \sigma_2, \sigma_3) > 0.$$

Therefore, since $F^{i,j,k}$ is continuous in q (Lemma 3.4(e)), there exists q_0, $0 > q_0 > q_1$, such that

$$F^{i,j,k}(q_0, \sigma_1, \sigma_2, \sigma_3) = 0,$$

i.e.

$$\zeta^i(q_0, \sigma_1) + \zeta^j(q_0, \sigma_2) + \zeta^k(q_0, \sigma_3) = q_0.$$

So, from Lemma 3.1 it follows that $\zeta^i(q_0, \sigma_1)$, $\zeta^j(q_0, \sigma_2)$, $\zeta^k(q_0, \sigma_3)$ form a negative solution of (46). Since, each of i, j, k can be chosen independently from the set $\{2, 3\}$, we have total 8 different negative solutions. □

Theorem 6 *For any given force field* $\mathbf{f} : \Omega \to \mathbb{R}^d$ *and the surface traction* $\mathbf{t} : \Gamma_t \to \mathbb{R}^d$ *such that the eigenvalues of the statically admissible stress tensor function* $\boldsymbol{\tau} \in \mathcal{T}_a$ *satisfy* $0 < \sigma_1, \sigma_2, \sigma_3 < \frac{4}{27}$*, the total complementary energy* $\Pi^d_{\boldsymbol{\tau}}(\mathbf{T})$ *has at least 15 mixed stationary points, i.e., some eigenvalues* ζ_i*,* $i = 1, 2, 3$*, of* $\mathbf{T}$ *are positive, some are negative.*

Proof. Each of the equations $G(\zeta, 0, \sigma_i)$, has one positive and two negative solutions: $\zeta^1, \zeta^2, \zeta^3$. Applying Lemma 3.5 it is easy to check that

(1) for $i, j = 2, 3$,

$$F^{1,i,j}(0, \sigma_1, \sigma_2, \sigma_3) < 0,\ F^{i,1,j}(0, \sigma_1, \sigma_2, \sigma_3) < 0$$

(2) $F^{2,3,1}(0, \sigma_1, \sigma_2, \sigma_3) < 0$, $F^{3,2,1}(0, \sigma_1, \sigma_2, \sigma_3) < 0$, $F^{3,3,1}(0, \sigma_1, \sigma_2, \sigma_3) < 0$.
(3) $F^{1,1,2}(0, \sigma_1, \sigma_2, \sigma_3) < 0$, $F^{1,1,3}(0, \sigma_1, \sigma_2, \sigma_3) < 0$, $F^{1,3,1}(0, \sigma_1, \sigma_2, \sigma_3) < 0$, $F^{3,1,1}(0, \sigma_1, \sigma_2, \sigma_3) < 0$.

For each of these 15 combinations, $F^{a,b,c}$, there exists $q_1 < 0$ such that $F^{a,b,c}$ $(q_1, \sigma_1, \sigma_2, \sigma_3) > 0$.
Therefore, since $F^{a,b,c}$ is continuous in q (Lemma 3.4(e)), there exists q_0, $0 > q_0 > q_1$, such that

$$F^{a,b,c}(q_0, l_1, l_2, l_3) = 0$$

i.e.,

$$\zeta^a(q_0, \sigma_1) + \zeta^b(q_0, \sigma_2) + \zeta^c(q_0, \sigma_3) = q_0$$

So, from Lemma 3.1 it follows that $\zeta^a(q_0, \sigma_1)$, $\zeta^b(q_0, \sigma_2)$, $\zeta^c(q_0, \sigma_3)$ form a mixed solution of (46).
Obviously, these 15 combinations result in different mixed stationary points of $\Pi^d_{\boldsymbol{\tau}}(\mathbf{T})$. □

4 Conclusions

We have illustrated that by using the canonical duality theory, the nonconvex minimal potential problem ($\mathscr{P}$) is canonically dual to a concave maximization problem in a convex stress space $\mathscr{S}_a^+$, which can be solved by well-developed numerical methods. By the pure complementary energy principle, the general nonlinear partial differential equation in nonlinear elasticity is actually equivalent to an algebraic (tensor) equation, which can be solved for certain materials to obtain all possible stress solutions. Both global and local extremal solutions can be identified by the triality theory, while the Legendre–Hadamard condition is only necessary for local minimizers. Our results shows that for St. Venant–Kirchhoff material, the nonlinear boundary-value problem could have 24 solutions at each material point, but only one global minimizer if the statically admissible stress $\boldsymbol{\tau} \neq 0$. It is important to have a detailed study on these solutions in the future.

Acknowledgements This research was supported by the US Air Force Office of Scientific Research under the grant AFOSR FA9550-17-1-0151. Results presented in Sect. 3 were discussed with Professor Ray Ogden from University of Glasgow.

References

1. Cai, K., Gao, D.Y., Qin, Q.H.: Post-buckling solutions of hyper-elastic beam by canonical dual finite element method. Math. Mech. Solids **19**(6), 659–671 (2014)
2. Ciarlet, P.G.: Mathematical Elasticity. Volume I: Three-Dimensional Elasticity. Elsevier Science Publishers B.V., North-Holland (1988)
3. Gao, D.Y.: Global extremum criteria for nonlinear elasticity. J. Appl. Math. Phys. (ZAMP) **43**, 924–937 (1992)
4. Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory. Appl. Mech. Rev. **50**, S64–S71 (1997)
5. Gao, D.Y.: Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications. IMA J. Appl. Math. **61**, 199–235 (1998)
6. Gao, D.Y.: Pure complementary energy principle and triality theory in finite elasticity. Mech. Res. Commun. **26**, 31–37 (1999)
7. Gao, D.Y.: Duality-Mathematics. Wiley Encyclopedia of Electrical and Electronics Engineering, vol. 6, pp. 68–77. Wiley, New York (1999)

8. Gao, D.Y.: General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics. Meccanica **34**, 169–198 (1999)
9. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications, xviii + 454pp. Kluwer Academic Publishers, Boston (2000)
10. Gao, D.Y.: Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications. Nonlinear Anal. **42**, 1161–1193 (2000)
11. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Global Optim. **17**, 127–160 (2000)
12. Gao, D.Y.: Nonconvex semi-linear problems and canonical dual solutions. Gao, D.Y., Ogden, R.W. (eds.) Advances in Mechanics and Mathematics, vol. II, pp. 261–312. Kluwer Academic Publishers, Boston (2003)
13. Gao, D.Y.: Perfect duality theory and complete set of solutions to a class of global optimization. Optimization **52**, 467–493 (2003)
14. Gao, D.Y.: Complementary variational principle, algorithm, and complete solutions to phase transitions in solids governed by Landau-Ginzburg equation. Math. Mech. Solids **9**, 285–305 (2004)
15. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solutions for global optimization. Comput. Chem. Eng. **33**, 1964–1972 (2009)
16. Gao, D.Y., Ogden, R.W.: Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation. Q. J. Mech. Appl. Math. **61**, 497–522 (2008)
17. Gao, D.Y., Ogden, R.W.: Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem. ZAMP **59**, 498–517 (2008)
18. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Global Optim. **47**, 463–484 (2010)
19. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Q. Appl. Math. **47**, 487–504 (1989)
20. 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)
21. Hellinger, E.: Die allgemeine Ansätze der Mechanik der Kontinua. Encyklopädie der Mathematischen Wissenschaften IV **4**, 602–694 (1914)
22. Koiter, W.T.: On the complementary energy theorem in nonlinear elasticity theory. In: Fichera, G. (ed.) Trends in Applications of Pure Mathematics to Mechanics. Pitman, London (1976)
23. Lee, S.J., Shield, R.T.: Variational principles in finite elastics. J. Appl. Math. Phys. (ZAMP) **31**, 437–453 (1980)
24. Lee, S.J., Shield, R.T.: Applications of variational principles in finite elasticity. J. Appl. Math. Phys. (ZAMP) **31**, 454–472 (1980)
25. Levinson, M.: The complementary energy theorem in finite elasticity. J. Appl. Mech. **87**, 826–828 (1965)
26. Li, S.F., Gupta, A.: On dual configuration forces. J. Elast. **84**, 13–31 (2006)
27. Li, C., Zhou, X., Gao, D.Y.: Stable trajectory of logistic map. Nonlinear Dyn. (2014). doi:10.1007/s11071-014-1433-y
28. Oden, J.T., Reddy, J.N.: Variational Methods in Theoretical Mechanics. Springer, Berlin (1983)
29. Ogden, R.W.: A note on variational theorems in non-linear elastostatics. Math. Proc. Camb. Philos. Soc. **77**, 609–615 (1975)
30. Reissner, E.: On a variational theorem for finite elastic deformations. J. Math. Phys. **32**, 129–135 (1953)
31. Ruan, N., Gao, D.Y.: Global optimal solutions to a general sensor network localization problem. Perform. Eval. **75–76**, 1–16 (2014)
32. Ruan, N., Gao, D.Y.: Canonical duality approach for nonlinear dynamical systems. IMA J. Appl. Math. **79**, 313–325 (2014)
33. Ruan, N., Gao, D.Y., Jiao, Y.: Canonical dual least square method for solving general nonlinear systems of equations. Comput. Optim. Appl. **47**, 335–347 (2010)

34. Santos, H.A.F.A., Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Nonlinear Mech. **47**, 240–247 (2011). doi:10.1016/j.ijnonlinmec.2011.05.012
35. Truesdell, C., Noll, W.: The Non-Linear Field Theories of Mechanics, 3rd edn. Springer, Berlin (2004)
36. Veubeke, B.F.: A new variational principle for finite elastic displacements. Int. J. Eng. Sci. **10**, 745–763 (1972)
37. 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

Remarks on Analytic Solutions and Ellipticity in Anti-plane Shear Problems of Nonlinear Elasticity

David Yang Gao

Abstract This paper revisits a well-studied anti-plane shear deformation problem formulated by Knowles in 1976. It shows that a homogenous hyper-elasticity for general anti-plane shear deformation must be governed by a generalized neo-Hookean model. Based on minimum total potential principle, a well-determined fully nonlinear system is obtained for isochoric deformation, which admits nontrivial states of finite anti-plane shear without ellipticity constraint. By a pure complementary energy principle, a complete set of analytical solutions is obtained, both global and local extremal solutions are identified by a triality theory. It is proved that the Legendre condition (i.e., the strong ellipticity) does not necessary to guarantee a unique solutions. The uniqueness depends not only on the stored energy, but also on the external force. Knowles' over-determined system is simply due to a pseudo-Lagrange multiplier $p(x_1, x_2)$ and two self-balanced equilibrium equations in the plane. The constitutive condition in his theorems is naturally satisfied with $b = \lambda/2$.

1 Remarks on Nonconvex Variational Problem and Challenges

Minimum total potential energy principle plays a fundamental role in continuum mechanics, especially for hyper-elasticity. One important feature is that the equilibrium equations obtained (under certain regularity conditions) by this principle are naturally compatible. Therefore, instead of the local method adopted by Knowles [6, 7], the discussion of this paper begins from the minimum potential variational problem (($\mathscr{P}$) for short):

$$(\mathscr{P}): \quad \min \left\{ \Pi(\boldsymbol{\chi}) = \int_{\mathscr{B}} W(\nabla \boldsymbol{\chi}) d\mathscr{B} - \int_{S_t} \boldsymbol{\chi} \cdot \mathbf{t} dS | \;\; \boldsymbol{\chi} \in \mathscr{X}_c \right\}, \tag{1}$$

D.Y. Gao (✉)
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_4

where the unknown deformation $\boldsymbol{\chi}(\mathbf{x}) = \{\chi_i(x_j)\} \in \mathscr{X}_a$ is a vector-valued mapping from a given material particle $\mathbf{x} = \{x_i\} \in \mathscr{B}$ in the undeformed body to a position vector $\boldsymbol{\chi} \in \omega$ in the deformed configuration in $\mathbb{R}^3$. The body is fixed on the boundary $S_x \subset \partial\mathscr{B}$, while on the remaining boundary $S_t = S_x \cap \partial\mathscr{B}$, the body is subjected to a given surface traction $\mathbf{t}(\mathbf{x})$. The admissible deformation space $\mathscr{X}_a$ in this paper is assumed to be

$$\mathscr{X}_a = \{\boldsymbol{\chi} \in \mathscr{W}^{1,1}(\mathscr{B}; \mathbb{R}^3) | \;\; \boldsymbol{\chi}(\mathbf{x}) = 0 \;\; \forall \mathbf{x} \in S_x\}, \tag{2}$$

where $\mathscr{W}^{1,1}$ is the standard notation for Sobolev space, i.e., a function space in which both $\boldsymbol{\chi}$ and its weak derivative $\nabla\boldsymbol{\chi}$ have a finite $L^1(\mathscr{B})$ norm. Clearly, a function in $\mathscr{W}^{1,1}$ is not necessarily to be smooth, or even continuous. For homogeneous hyperelastic body, the strain energy $W(\mathbf{F})$ is assumed to be C^1 on its domain $\mathscr{F}_a \subset \mathbb{R}^{3\times 3}$, in which certain necessary *constitutive constraints* are included, such as

$$\det \mathbf{F} > 0, \quad W(\mathbf{F}) \geq 0 \;\; \forall \mathbf{F} \in \mathscr{F}_a, \quad W(\mathbf{F}) \to \infty \text{ as } \|\mathbf{F}\| \to \infty. \tag{3}$$

For incompressible materials, the condition $\det \mathbf{F} = 1$ should be included. Finally, $\mathscr{X}_c = \{\boldsymbol{\chi} \in \mathscr{X}_a | \; \nabla\boldsymbol{\chi} \in \mathscr{F}_a\}$ is the *kinetically admissible space*, which is nonconvex due to nonlinear constraints such as $\det(\nabla\boldsymbol{\chi}) > 0$. Also, the stored energy $W(\mathbf{F})$ is in general nonconvex in order to model real-world problems. Thus, the nonconvex problem $(\mathscr{P})$ has usually multiple local optimal solutions. Let $\mathscr{X}_b \subset \mathscr{X}_c$ be a subspace with two additional conditions

$$\mathscr{X}_b = \{\boldsymbol{\chi} \in \mathscr{X}_c | \;\; \boldsymbol{\chi} \in C^2(\mathscr{B}; \mathbb{R}^3), \quad W(\mathbf{F}(\boldsymbol{\chi})) \in C^2(\mathscr{F}_a; \mathbb{R})\}, \tag{4}$$

the criticality condition $\delta\Pi(\boldsymbol{\chi}; \delta\boldsymbol{\chi}) = 0 \;\; \forall \delta\boldsymbol{\chi} \in \mathscr{X}_b$ leads to a nonlinear boundary value problem

$$(BVP): \quad \begin{cases} -\nabla \cdot \boldsymbol{\sigma}(\nabla\boldsymbol{\chi}) = 0 & \text{in } \mathscr{B}, \\ \mathbf{N} \cdot \boldsymbol{\sigma}(\nabla\boldsymbol{\chi}) = \mathbf{t} \quad \text{on } S_t, & \boldsymbol{\chi} = 0 \quad \text{on } S_x \end{cases} \tag{5}$$

where $\mathbf{N} \in \mathbb{R}^3$ is a unit vector normal to $\partial\mathscr{B}$, and $\boldsymbol{\sigma}(\mathbf{F})$ is the first Piola–Kirchhoff stress (force per unit undeformed area), defined by

$$\boldsymbol{\sigma} = \nabla W(\mathbf{F}), \quad \text{or } \; \sigma_{ij} = \frac{\partial W(\mathbf{F})}{\partial F_{ij}}, \quad i, j = 1, 2, 3. \tag{6}$$

Remark 1 (Isochoric Deformation, pseudo-Lagrange Multiplier) Strictly speaking, there is an inequality constraint in $\mathscr{X}_c$, i.e., the admissible deformation condition $\det(\nabla\boldsymbol{\chi}) > 0$. According to the mathematical theory of variational inequality, the problem *(BVP)* should be subjected to the following *KKT conditions*

$$p \leq 0, \quad \det(\nabla\boldsymbol{\chi}) > 0, \quad p \det(\nabla\boldsymbol{\chi}) = 0, \tag{7}$$

where p is a Lagrange multiplier and $p \leq 0$ is called the condition of constraint qualification. The equality $p \det(\nabla \boldsymbol{\chi}) = 0$ is the well-known *complementarity condition* in variational inequality theory, by which we must have $p = 0$ in order to guarantee the inequality constraint $\det(\nabla \boldsymbol{\chi}) > 0$. Therefore, this constraint is actually not active to the problem $(\mathscr{P})$. Such an inactive constraint is not a variational constraint.

For incompressible deformation, the inequality condition $\det(\nabla \boldsymbol{\chi}) > 0$ in $\mathscr{X}_c$ should be replaced by an equality constraint $\det(\nabla \boldsymbol{\chi}) = 1$. In this case, $(\mathscr{P})$ is a constrained variational problem. The KKT conditions (7) should be replaced by (see [8])

$$p \neq 0, \quad \det \mathbf{F}(\boldsymbol{\chi}) = 1, \quad p(\det \mathbf{F}(\boldsymbol{\chi}) - 1) = 0 \tag{8}$$

and we must have $p(\mathbf{x}) \neq 0$ for a.e. $\mathbf{x} \in \mathscr{B}$ in order to ensure $\det \mathbf{F}(\boldsymbol{\chi}) - 1 = 0$. The associated *(BVP)* should be

$$(BVP)_p : \quad \begin{cases} -\nabla \cdot \boldsymbol{\sigma}(\nabla \boldsymbol{\chi}, p) = 0, \quad \det(\nabla \boldsymbol{\chi}) = 1 \quad \text{in } \mathscr{B}, \\ \mathbf{N} \cdot \boldsymbol{\sigma}(\nabla \boldsymbol{\chi}, p) = \mathbf{t} \quad \text{on } S_t, \quad \boldsymbol{\chi} = 0 \quad \text{on } S_x, \end{cases} \tag{9}$$

in which, $\boldsymbol{\sigma}(\mathbf{F}, p) = \nabla W(\mathbf{F}) - p\mathbf{F}^{-T}$, where $\mathbf{F}^{-T} = (\mathbf{F}^T)^{-1}$. In this case, we have two variables $(\boldsymbol{\chi}, p)$ and two equations in $\mathscr{B}$, and thus, the problem $(BVP)_p$ is a well-defined system.

For isochoric (i.e., volume preserving) deformation, say the anti-plane shear problems, the condition $\det \mathbf{F} = 1$ is naturally satisfied all most every where (*a.e.*) in $\mathscr{B}$. In this case, the trivial condition $\det \mathbf{F} = 1$ is not a variational constraint of $(\mathscr{P})$ for incompressible material and the complementarity condition $p(\det \mathbf{F} - 1) = p(1 - 1) \equiv 0$ cannot lead to $p(\mathbf{x}) \neq 0$ $a.e.$ in $\mathscr{B}$. The arbitrary parameter $p(\mathbf{x})$ can be called *pseudo-Lagrange multiplier*, which is not an unknown variable for $(BVP)_p$. Otherwise, the $(BVP)_p$ is an over-determined system. This fact in KKT theory is important for understanding Knowles' anti-plane shear problem. ■

Physically speaking, the hydrostatic pressure p is not necessarily to be zero for isochoric deformations of incompressible materials. There are many examples in the literature, see the book by Ogden [9] as well as many papers by Rivlin on volume-preserving deformations of isotropic materials (simple shear, torsion, flexure, etc.).[1]

2 Anti-plane Shear Deformation Problems

The so-called anti-plane shear deformation studied by Knowles' [6] is defined by

$$\boldsymbol{\chi}(\mathbf{x}) = \left\{\lambda^{-\frac{1}{2}} x_1, \quad \lambda^{-\frac{1}{2}} x_2, \quad \lambda x_3 + u(x_1, x_2)\right\}, \tag{10}$$

[1]Personal communications with David Steigmann, Ray Ogden, and C. Horgan.

where (x_1, x_2, x_3) are cylindrical coordinates in the reference configuration $\mathscr{B}$ relative to a cylindrical basis $\{\mathbf{e}_i\}$, $i = 1, 2, 3$, the constant $\lambda > 0$ is a given pre-stretch, and $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ is the amount of shear in the planes normal to $\mathbf{e}_3$. On $\Gamma_\chi \subset \partial\Omega$, the homogenous boundary condition is given $u(x_\alpha) = 0 \;\; \forall x_\alpha \in \Gamma_\chi, \;\; \alpha = 1, 2$. On the remaining boundary $\Gamma_t = \partial\Omega \cap \Gamma_\chi$, the cylinder is subjected to the shear force $\mathbf{t}(\mathbf{x}) = t(\mathbf{x})\mathbf{e}_3 \;\; \forall \mathbf{x} \in \Gamma_t$, where $t : \Gamma_t \to \mathbb{R}$ is a prescribed function. For this anti-plane shear deformation we have

$$\mathbf{F} = \nabla\boldsymbol{\chi} = \begin{pmatrix} \lambda^{-\frac{1}{2}} & 0 & 0 \\ 0 & \lambda^{-\frac{1}{2}} & 0 \\ u_{,1} & u_{,2} & \lambda \end{pmatrix}, \quad \mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = \begin{pmatrix} \lambda^{-1} + u_{,1}^2 & u_{,1}u_{,2} & \lambda u_{,1} \\ u_{,1}u_{,2} & \lambda^{-1} + u_{,2}^2 & \lambda u_{,2} \\ \lambda u_{,1} & \lambda u_{,2} & \lambda^2 \end{pmatrix}, \tag{11}$$

where $u_{,\alpha}$ represents $\partial u/\partial x_\alpha$ for $\alpha = 1, 2$. By the notations $\boldsymbol{\gamma} = \nabla u = \{u_{,1}, u_{,2}\} \in \mathbb{R}^2$ and $|\boldsymbol{\gamma}|^2 = \gamma_1^2 + \gamma_2^2$, the principal invariants of $\mathbf{C}$ are

$$I_1(\mathbf{C}) = \lambda_1 + |\boldsymbol{\gamma}|^2, \quad I_2(\mathbf{C}) = \lambda_2 + \lambda^{-1}|\boldsymbol{\gamma}|^2, \quad I_3(\mathbf{C}) \equiv 1, \tag{12}$$

where $\lambda_1 = \lambda^2 + 2\lambda^{-1}, \;\; \lambda_2 = \lambda^{-2} + 2\lambda$. Particularly,

$$I_1(\mathbf{C}) = I_2(\mathbf{C}) = 3 + |\boldsymbol{\gamma}|^2 \quad \text{if } \lambda = 1. \tag{13}$$

Lemma 1 *For any given $\lambda > 0$, the homogenous hyper-elasticity for general anti-plane shear deformation must be governed by a generalized neo-Hookean model, i.e., there exists a real-valued function $V(I_1)$ such that $W(\mathbf{F}(\boldsymbol{\gamma})) = V(I_1(\boldsymbol{\gamma}))$.*

The proof is elementary, i.e., by the fact that $I_2 = \lambda^{-1}I_1 + a$, $a = \lambda_2 - \lambda^{-1}\lambda_1$, and $I_3 = 1$, there must exist real-valued functions $\bar{W}(I_1, I_2)$ and $V(I_1)$ such that

$$W(\mathbf{F}) = \bar{W}(I_1, I_2) = \bar{W}(I_1, \lambda^{-1}I_1 + a) = V(I_1), \quad \forall I_1 \geq \lambda_1, \quad \lambda > 0. \tag{14}$$

The fact $\det \mathbf{F} \equiv 1$ shows that the anti-plane shear state (10) is an isochoric deformation. The kinetically admissible space $\mathscr{X}_c$ in problem $(\mathscr{P})$ can be simply replaced by a convex set

$$\mathscr{U}_c = \{u(x_1, x_2) \in \mathscr{W}^{1,1}(\Omega; \mathbb{R}) | \;\; u(\mathbf{x}) = 0 \;\; \forall \mathbf{x} \in \Gamma_\chi\}. \tag{15}$$

Thus, in term of d $W(\mathbf{F}(\boldsymbol{\gamma})) = V(I_1(\boldsymbol{\gamma}))$, Problem $(\mathscr{P})$ for the general anti-plane shear deformation problem has the following form:

$$(\mathscr{P})_s: \;\; \min\left\{\Pi(u) = \int_\Omega V(I_1(\nabla u))\mathrm{d}\Omega - \int_{\Gamma_t} ut\mathrm{d}\Gamma \;\;|\;\; u \in \mathscr{U}_c\right\}. \tag{16}$$

By the fact that the stored energy $V(I_1)$ is coercive, i.e. $V(I_1(\boldsymbol{\gamma})) \to \infty$ as $|\boldsymbol{\gamma}| \to \infty$, this minimum variational problem is bounded below. Let $\mathscr{U}_b$ be a smooth subset of $\mathscr{U}_c$ defined by

$$\mathscr{U}_b = \{u \in \mathscr{U}_c |\ u \in C^2(\Omega; \mathbb{R}),\quad V(\xi) \in C^2(\mathscr{E}_a; \mathbb{R})\}. \tag{17}$$

Under certain regularity conditions for both Ω and $\partial\Omega$, the associated *(BVP)* is

$$(BVP)_s : \quad \begin{cases} \nabla \cdot (2\zeta \nabla u) = 0 & \text{in } \Omega, \\ \mathbf{n} \cdot (2\zeta \nabla u) = t & \text{on } \Gamma_t, \quad u = 0 \quad \text{on } \Gamma_\chi \end{cases} \tag{18}$$

where $\mathbf{n} = \{n_\alpha\} \in \mathbb{R}^2$ is a unit vector norm to $\partial\Omega$, and $\zeta = \nabla V(\xi),\quad \xi = I_1 = \lambda_1 + |\nabla u|^2$.

If $\Gamma_\chi = \partial\Omega$, then $(BVP)_s$ is a Dirichlet boundary value problem, which has only trivial solution due to zero input. For Neumann boundary value problem $\Gamma_t = \partial\Omega$, the external force field must be such that

$$\int_{\Gamma_t} t(\mathbf{x}) \mathrm{d}\Gamma = 0 \tag{19}$$

for overall force equilibrium. In this case, if $\bar{u}$ is a solution to $(BVP)_s$, then $u = \bar{u} + u_c$ is also a solution for any constant u_c since the cylinder is not fixed in x_3 direction. By the fact that the only unknown u is a scalar-valued function, the $(BVP)_s$ is well-defined and the condition $\Gamma_t \neq \emptyset$ is necessary for anti-plane shear deformation to have nontrivial solutions.

Although the anti-plane shear deformations are one of the simplest classes of deformations that solids can undergo [5], to solve either the nonconvex variational problem $(\mathscr{P})_s$ or the fully nonlinear partial differential equation in $(BVP)_s$ is still difficult by direct methods. However, it shows in [3] that these two problems can be solved easily using canonical duality theory. Note that $\mathbf{F}$ is an affine function of $\boldsymbol{\gamma} = \nabla u \in \mathscr{G}_a \subset \mathbb{R}^2$, it is also mathematically equivalent to assume the existence of a real-valued function $\hat{W} : \mathscr{G}_a \to \mathbb{R}$ such that

$$W(\mathbf{F}(\boldsymbol{\gamma})) = V(I_1(\boldsymbol{\gamma})) = \hat{W}(\boldsymbol{\gamma}) \ \ \forall \boldsymbol{\gamma} \in \mathscr{G}_a \tag{20}$$

holds for general anti-plane shear deformation problems without any additional constitutive constraints added in $\mathscr{G}_a$. Instead of $|\boldsymbol{\gamma}|^2$ used in [3], this paper adopts I_1 as the canonical strain measure:

$$\xi = \Lambda(\nabla u) = |\nabla u|^2 + \lambda_1,\quad \Lambda : \mathscr{U}_c \to \mathscr{E}_a = \{\xi \in L^q(\Omega)|\ \ \xi(\mathbf{x}) \geq \lambda_1 \quad \text{in } \Omega\}, \tag{21}$$

where L^q is the standard notation of Lebesgue integrable space with $q \geq 1$. Clearly, $\mathscr{E}_a$ is a convex set. Therefore, it is able to discuss the convexity of the stored energy $W(\mathbf{F}(\boldsymbol{\gamma})) = V(\xi(\boldsymbol{\gamma}))$ on $\mathscr{E}_a$. First, we let $V : \mathscr{E}_a \to \mathbb{R}$ be a canonical function, i.e.

the duality relation

$$\zeta = \nabla V(\xi) : \mathscr{E}_a \to \mathscr{E}_a^* = \{\zeta \in L^{q'}(\Omega) |\ \zeta(\mathbf{x}) \geq -\lambda_1 \quad \forall \mathbf{x} \in \Omega\}$$

is one-to-one and onto, where $q' = -q/(1-q)$ is a dual number of q. Thus, the complementary energy $V^* : \mathscr{E}_a^* \to \mathbb{R}$ can be uniquely defined by the Legendre transformation

$$V^*(\zeta) = \text{sta}\{\xi\zeta - V(\xi) |\ \ \xi \in \mathscr{E}_a\}, \tag{22}$$

where sta$\{g(\xi)|\xi \in \mathscr{E}_a\}$ stands for finding stationary value of g on $\mathscr{E}_a$. Let

$$\mathscr{T}_a = \{\boldsymbol{\tau} \in C^1(\Omega; \mathbb{R}^2) |\ \ \nabla \cdot \boldsymbol{\tau} = 0 \quad \text{in } \Omega, \quad \mathbf{n} \cdot \boldsymbol{\tau} = t \quad \text{on } \Gamma_t\}. \tag{23}$$

Thus, for a given $t(\mathbf{x})$ on Γ_t such that $\boldsymbol{\tau} \in \mathscr{T}_a$ and $\tau^2 = |\boldsymbol{\tau}|^2$, the canonical dual problem can be obtained easily on the canonical dual space $\mathscr{S}_a = \{\zeta \in \mathscr{E}_a^* |\ \ \zeta^{-1}\tau^2 \in L^1(\Omega)\}$ [1]

$$(\mathscr{P}^d)_s : \quad \text{sta}\left\{\Pi^d(\zeta) = \int_\Omega \left[\lambda_1\zeta - V^*(\zeta) - \frac{1}{4}\zeta^{-1}\tau^2\right] \mathrm{d}\Omega \ |\ \ \zeta \in \mathscr{S}_a\right\}. \tag{24}$$

Theorem 1 (Pure Complementary Energy Principle) *For a given pre-stretch $\lambda > 0$ and a nontrivial shear force $t(\mathbf{x})$ on Γ_t such that $\boldsymbol{\tau} \in \mathscr{T}_a \neq \emptyset$, the pure complementary energy $\Pi^d(\zeta)$ has at least one nontrivial critical solution ζ_k defined by the following algebraic equation*

$$4\zeta^2[\nabla V^*(\zeta) - \lambda_1] = \tau^2. \tag{25}$$

The function defined by

$$u_k(\mathbf{x}) = \frac{1}{2}\int_{\mathbf{x}_0}^{\mathbf{x}} \zeta_k^{-1}\boldsymbol{\tau} \cdot d\mathbf{x} \tag{26}$$

along any path from $\mathbf{x}_0 \in \Gamma_\chi$ to $\mathbf{x} \in \Omega$ is a critical point of $\Pi(u)$ and $\Pi(u_k) = \Pi^d(\zeta_k)$.

The proof of this theorem is a special application of the general pure complementary energy principle proposed in [1]. This theorem shows that the fully nonlinear PDE in $(BVP)_s$ is variationally equivalent to a *canonical dual algebraic equation* (25), which can be solved completely to obtain all possible solutions $\{\zeta_k\}$. Clearly, $\zeta_k = 0$ only if $\tau = 0$. It is easy to verify that that each u_k satisfies both equilibrium equation and boundary conditions in $(BVP)_s$.

In order to identify global and local optimal solutions, we need the following convex subsets

$$\mathscr{S}_a^+ = \{\zeta \in \mathscr{S}_a |\ \zeta(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in \Omega\}, \qquad \mathscr{S}_a^- = \{\zeta \in \mathscr{S}_a |\ \zeta(\mathbf{x}) < 0 \quad \forall \mathbf{x} \in \Omega\}. \tag{27}$$

Theorem 2 *Suppose that $V(\xi)$ is convex on $\mathscr{E}_a$ and for a given $\boldsymbol{\tau} \in \mathscr{T}_a$ such that $\{\zeta_k\}$ is a solution set to (25), $\boldsymbol{\gamma}_k = \frac{1}{2}\zeta_k^{-1}\boldsymbol{\tau}$, and u_k is defined by (26), we have the following statements:*

1. *If $\zeta_k \in \mathscr{S}_a^+$, then $\nabla^2 W(\boldsymbol{\gamma}_k) \succ 0$ and u_k is a global minimal solution to $(\mathscr{P})_s$.*
2. *If $\zeta_k \in \mathscr{S}_a^-$ and $\nabla^2 W(\boldsymbol{\gamma}_k) \succ 0$, then u_k is a local minimal solution to $(\mathscr{P})_s$.*
3. *If $\zeta_k \in \mathscr{S}_a^-$ and $\nabla^2 W(\boldsymbol{\gamma}_k) \prec 0$, then u_k is a local maximal solution to $(\mathscr{P})_s$.*

If $\{\zeta_k\} \subset \mathscr{S}_a^+$, then $(\mathscr{P})_s$ has a unique solution on $\mathscr{U}_c$.

Proof. To prove u_k is a global minimizer of $(\mathscr{P})_s$, we follow the standard canonical duality theory. By the convexity of $V(\xi)$ on its convex domain $\mathscr{E}_a$, we have

$$V(\xi) - V(\xi_k) \geq (\xi - \xi_k)\zeta_k \quad \forall \xi,\ \xi_k \in \mathscr{E}_a, \quad \zeta_k = \nabla V(\xi_k). \tag{28}$$

For any given variation δu, let $u = u_k + \delta u$, $\xi = \Lambda(\nabla u)$, $\quad \xi_k = \Lambda(\nabla u_k)$. Then

$$\Pi(u) - \Pi(u_k) \geq \int_\Omega [2\zeta_k(\nabla u_k) - \boldsymbol{\tau}]^T(\nabla \delta u)\mathrm{d}\Omega + G_{ap}(\delta u, \zeta_k) \ \ \forall u,\ \delta u \in \mathscr{U}_c \tag{29}$$

for any given $\boldsymbol{\tau} \in \mathscr{T}_a$, where

$$G_{ap}(u, \zeta) = \int_\Omega |\nabla u|^2 \zeta \mathrm{d}\Omega \tag{30}$$

is the Gao–Strang *complementary gap function* [1]. If u_k is a critical point of $\Pi(u)$, then we have $2\zeta_k(\nabla u_k) = \boldsymbol{\tau}$ Thus, we have $\Pi(u) - \Pi(u_k) \geq G_{ap}(\delta u, \zeta_k) \geq 0 \ \ \forall \delta u \in \mathscr{U}_c$ if $\zeta_k \in \mathscr{S}_a^+$. This shows that u_k is a global minimizer of $(\mathscr{P})_s$.

Using chain rule for $\hat{W}(\boldsymbol{\gamma}) = V(\xi(\boldsymbol{\gamma}))$ we have $\nabla \hat{W}(\boldsymbol{\gamma}) = 2\boldsymbol{\gamma}[\nabla V(\xi)] = 2\zeta\boldsymbol{\gamma}$, and

$$\nabla^2 \hat{W}(\boldsymbol{\gamma}) = 2\zeta \mathbf{I} + 4h(\xi)\boldsymbol{\gamma} \otimes \boldsymbol{\gamma}, \tag{31}$$

where $\mathbf{I}$ is an identity tensor in $\mathbb{R}^{2\times 2}$, $h(\xi) = \nabla^2 V(\xi) \geq 0$ due to the convexity of V on $\mathscr{E}_a$. Therefore, $\nabla^2 W(\boldsymbol{\gamma}_k) \succ 0$ if $\zeta_k \in \mathscr{S}_a^+$. To prove the local extremality, we replace $\boldsymbol{\gamma}_k$ in (31) by $\boldsymbol{\gamma}_k = \frac{1}{2}\zeta_k^{-1}\boldsymbol{\tau}$ such that

$$\mathbf{G}(\zeta_k) = \nabla^2 \hat{W}(\boldsymbol{\tau}_k) = 2\zeta_k \mathbf{I} + \zeta_k^{-2} h(\xi_k)\boldsymbol{\tau} \otimes \boldsymbol{\tau}, \tag{32}$$

where $\xi_k = \nabla V^*(\zeta_k)$. Clearly, for a given $\boldsymbol{\tau} \in \mathscr{T}_a$ such that $\zeta_k \in \mathscr{S}_a^-$, the Hessian $\nabla^2 \hat{W}(\boldsymbol{\gamma}_k)$ could be either positive or negative definite. The total potential $\Pi(u_k)$ is locally convex if the *Legendre condition* $\nabla^2 \hat{W}(\nabla u_k) \succ 0$ holds, locally concave if $\nabla^2 \hat{W}(\nabla u_k) \prec 0$. Since u_k is a global minimizer when $\zeta_k \in \mathscr{S}_a^+$, therefore, for $\zeta_k \in \mathscr{S}_a^-$, the stationary solution u_k is a local minimizer if $\nabla^2 \hat{W}(\nabla u_k) \succ 0$ and, by the triality theory [1], u_k is a biggest local maximizer if $\nabla^2 \hat{W}(\nabla u_k) \prec 0$.

If $\{\zeta_k\} \subset \mathscr{S}_a^+$, then all the solutions $\{u_k\}$ are global minimizers and form a convex set. Since $\Pi^d(\zeta)$ is strictly concave on the open convex set $\mathscr{S}_a^+$, the condition

$\{\zeta_k\} \subset \mathscr{S}_a^+$ implies the unique solution of (25). In this case, Problem $(BVP)_s$ has at most one solution. □

Remark 2 (Legendre Condition, Ellipticity, and Global Optimality) In terms of $\hat{W}(\boldsymbol{\gamma})$, the equilibrium equation in $(BVP)_s$ can be written as

$$L[u] = -\nabla \cdot (\nabla_{\boldsymbol{\gamma}} \hat{W}(\nabla u)) = -(h_{\alpha\beta} u_\alpha)_{,\beta} = 0 \quad \text{in } \Omega$$

where $\mathbf{H}(\boldsymbol{\gamma}) = \{h_{\alpha\beta}\} = \nabla^2 \hat{W}(\boldsymbol{\gamma})$. In is known that the operator $L[u]$ is elliptic if

$$\mathbf{H}(\boldsymbol{\gamma}) = \nabla^2 \hat{W}(\boldsymbol{\gamma}) \succeq 0 \;\; \forall \boldsymbol{\gamma}(\mathbf{x}) \;\; a.e. \text{ in } \Omega. \tag{33}$$

This is the well-known *Legendre condition*. For nonlinear elasticity, the problem $(\mathscr{P})_s$ could have multiple critical solutions $\{u_k(\mathbf{x})\}$ at each $\mathbf{x} \in \Omega_s \subseteq \Omega$. As long as $\Omega_s \neq \varnothing$, the boundary value problem $(BVP)_s$ should have infinitely many solutions (see [4]). Thus, it is impossible to use Legendre condition to identify global minimal solution among all these infinitely many local solutions. Also, the traditional ellipticity definition depends only on the stored energy $\hat{W}(\mathbf{F})$ regardless of the external energy (i.e., the linear term) in $\Pi(u)$. This definition works only for convex systems since the linear term cannot change the convexity of the integrand $G(\boldsymbol{\gamma}) = \hat{W}(\boldsymbol{\gamma}) - \boldsymbol{\gamma}^T \boldsymbol{\tau}$ in the total potential $\Pi(u)$. But this is not true for nonconvex systems.

To see this, let us consider the most simple *double-well potential* $\hat{W}(\boldsymbol{\gamma}) = \frac{1}{2}(|\boldsymbol{\gamma}|^2 - 1)^2$. If we let $\xi = \Lambda(\boldsymbol{\gamma}) = |\boldsymbol{\gamma}|^2 - 1$ be a canonical measure (correspondingly $\lambda_1 = -1$), we have the canonical function $V(\xi) = \frac{1}{2}\xi^2$. In this case, the canonical dual algebraic equation (25) is a cubic equation (see [1]) $4\zeta^2(\zeta + 1) = \tau^2$, which has at most three real solutions $\{\zeta_k(\mathbf{x})\}$ at each $\mathbf{x} \in \Omega$ satisfying $\zeta_1 \geq 0 \geq \zeta_2 \geq \zeta_3$. It was proved in [1] (Theorem 3.4.4, page 133) that for a given force $\mathbf{t}(\mathbf{x})$, if $\tau^2(\mathbf{x}) > 2/27 \;\; \forall \mathbf{x} \in \Omega \subset \mathbb{R}$, then $(BVP)_s$ has only one solution on Ω. If $\tau^2(\mathbf{x}) < 2/27 \;\; \forall \mathbf{x} \in \Omega_s \subset \Omega$, then $(BVP)_s$ has three solutions $\{u_k(\mathbf{x})\}$ at each $\mathbf{x} \in \Omega_s$, i.e., $\Pi(u)$ is nonconvex on Ω_s. It was shown by Gao and Ogden that these solutions are nonsmooth if $\tau(\mathbf{x})$ changes its sign on Ω_s [4].

Theorem 2 shows that the Legendre condition is only necessary but not sufficient condition for global optimality.

The sufficient condition is simply

$$\zeta_k \in \mathscr{S}_a^+ \quad \Leftrightarrow \quad G_{ap} = \int_\Omega \zeta_k |\nabla u|^2 \mathrm{d}\Omega \geq 0 \;\; \forall u \in \mathscr{U}_c, \tag{34}$$

which was first proposed by Gao and Strang in 1989 (see [1]). Under this condition, the nonconvex minimum variational problem $(\mathscr{P})_s$ is equivalent to a concave maximum dual problem over the convex set $\mathscr{S}_a^+$, which can be solved easily to obtain all possible global optimal solutions (see [3]). ■

For a given function $G(\boldsymbol{\gamma}) : \mathscr{G}_a \to \mathbb{R}$, its *level set* and *sub-level set* of height $\alpha \in \mathbb{R}$ are defined respectively by

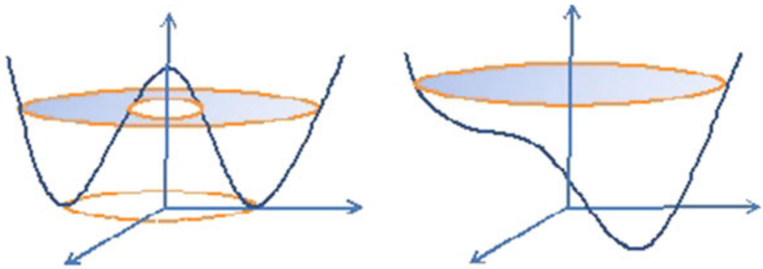


Fig. 1 Graphs and level sets of $G(\boldsymbol{\gamma})$ for $\boldsymbol{\tau} = 0$ (*left*) and $\boldsymbol{\tau} \neq 0$ (*right*)

$$\mathscr{L}_\alpha(G) = \{\boldsymbol{\gamma} \in \mathscr{G}_a \mid G(\boldsymbol{\gamma}) = \alpha\}, \quad \mathscr{L}_\alpha^\flat(G) = \{\boldsymbol{\gamma} \in \mathscr{G}_a \mid G(\boldsymbol{\gamma}) \leq \alpha\}, \quad \alpha \in \mathbb{R}. \tag{35}$$

The geometrical explanation for ellipticity and Theorem 2 is illustrated in Fig. 1, which shows that the nonconvex function $G(\boldsymbol{\gamma}) = \frac{1}{2}(|\boldsymbol{\gamma}|^2 - 1)^2 - \boldsymbol{\gamma}^T\boldsymbol{\tau}$ depends sensitively on the external force $\boldsymbol{\tau} \in \mathbb{R}^2$. If $|\boldsymbol{\tau}|$ is bigger enough, $G(\boldsymbol{\gamma})$ has only one minimizer and its level set is an ellipse (Fig. 1b). Otherwise, $G(\boldsymbol{\gamma})$ has multiple local minimizers and its level set is not an ellipse. For $\boldsymbol{\tau} = 0$, it is well-known Mexican hat in theoretical physics (Fig. 1a).

Figure 1 shows that although $G(\boldsymbol{\gamma})$ has only one global minimizer for certain given $\boldsymbol{\tau}$, the function is still nonconvex. Such a function is called quasiconvex in the context of global optimization. In order to distinguish this type of functions with Morry's quasiconvexity in nonconvex analysis, we need generalized definitions of quasiconvexity and ellipticity.

Definition 1 (G-Quasiconvexity, G-Ellipse, and G-Ellipticity) A function $G : \mathscr{G}_a \subset \mathbb{R}^n \to \mathbb{R}$ is called G-quasiconvex if its domain $\mathscr{G}_a$ is convex and

$$G(\theta\boldsymbol{\gamma} + (1-\theta)\boldsymbol{v}) \leq \max\{G(\boldsymbol{\gamma}), G(\boldsymbol{v})\} \;\; \forall \boldsymbol{\gamma}, \;\; \boldsymbol{v} \in \mathscr{G}_a, \quad \forall \theta \in [0, 1]. \tag{36}$$

It is called strictly G-quasiconvex if the inequality holds strictly.

A level set $\mathscr{L}_\alpha(G)$ is said to be a G-ellipse if it is a closed, simply connected set $\forall \alpha \in \mathbb{R}$.

For a given $t(\mathbf{x})$ on Γ_t such that $\boldsymbol{\tau} \in \mathscr{T}_a \neq \emptyset$, the $(BVP)_s$ is said to be G-elliptic if the total potential function $G(\boldsymbol{\gamma})$ is G-quasiconvex on $\mathscr{G}_a$. $(BVP)_s$ is strongly G-elliptic if $G(\boldsymbol{\gamma})$ is strictly G-quasiconvex.

Lemma 2 *For a given function* $G : \mathscr{G}_a \subset \mathbb{R}^n \to \mathbb{R}$,

$$G \text{ is G-quasiconvex} \Leftrightarrow \mathscr{L}_\alpha^\flat(G) \text{ is convex} \Leftrightarrow \mathscr{L}_\alpha(G) \text{ is a G-ellipse} \, \forall \alpha \in \mathbb{R}. \tag{37}$$

$$G \text{ is convex} \Rightarrow \text{G-quasiconvex} \Rightarrow (BVP)_s \text{ is G-elliptic}. \tag{38}$$

This statement shows an important fact in nonconvex systems, i.e. the total number of solutions to a nonlinear equation depends not only on the stored energy, but also (mainly) on the external force field. The nonlinear partial differential equation in

$(BVP)_s$ is elliptic only if it is G-elliptic. $(BVP)_s$ has at most one solution if $G(\boldsymbol{\gamma})$ is strictly G-quasiconvex on $\mathscr{G}_a$.

Theorem 3 (Uniqueness) *Suppose that the canonical function $V : \mathscr{E}_a \to \mathbb{R}$ is convex and for a given $t(\mathbf{x})$ on Γ_t such that $\boldsymbol{\tau} \in \mathscr{T}_a \neq \emptyset$, then there exists a constant τ_c such that if $\tau^2(\mathbf{x}) = |\boldsymbol{\tau}|^2 \geq \tau_c^2 \;\; \forall \mathbf{x} \in \Omega$, the total potential function $G(\boldsymbol{\gamma})$ is strictly G-quasiconvex and $(\mathscr{P})_s$ has at most one solution.*

This theorem can be proved easily by the fact that $\nabla V^*(\zeta) \geq \lambda_1$ and is monotone on $\mathscr{E}_a$.

3 Remarks on Knowles' Over-Determined Problem

Now let us revisit Knowles' work in 1976 [6]. Instead of the minimal potential variational problem $(\mathscr{P})$, Knowles started from the strong form of $(\mathscr{P})$, i.e., div $\boldsymbol{\sigma} = 0$ in the boundary value problem $(BVP)_p$ given in (9) with general constitutive law for incompressible materials

$$\boldsymbol{\sigma} = \nabla_{\mathbf{F}} \bar{W}(I_1(\mathbf{F}), I_2(\mathbf{F})) - p\mathbf{F}^{-T}, \tag{39}$$

where $\nabla_{\mathbf{F}} \bar{W} = \partial \bar{W}/\partial \mathbf{F}$. For the same anti-plane shear deformation problem (10), he ended up with three equilibrium equations (i.e., Eqs. (2.19) and (2.20) in [6])[2]:

$$q_{,\alpha} + \left(2\bar{W}_2 u_{,\alpha} u_{,\beta}\right)_{,\beta} - p_{,3} u_{,\alpha} = 0, \tag{40}$$

$$\left[2(\bar{W}_1 + \lambda^{-1}\bar{W}_2) u_{,\beta}\right]_{,\beta} - \lambda^{-1} p_{,3} = 0, \tag{41}$$

where $\bar{W}_\alpha = \partial \bar{W}/\partial I_\alpha = \zeta_\alpha, \;\; \alpha, \beta = 1, 2$ and $q = \lambda p - 2\bar{W}_1 - 2(\lambda^2 + \lambda^{-1} + |\nabla u|^2)\bar{W}_2$.

The first two equations in (40) are corresponding to the general equilibrium equation $\sigma_{ij,j} = 0$ in $\mathbf{e}_1$ and $\mathbf{e}_2$ directions; while the third one (41) is in $\mathbf{e}_3$ direction. Knowles indicated (Eq. (2.22) in [6]) that the hydrostatic pressure p is linear in x_3, i.e.

$$p = c x_3 + \bar{p}(x_1, x_2), \tag{42}$$

where c is a constant. Saccomandi emphasized recently that p is the Lagrange multiple associated with the incompressibility constraint, which must be in the form of (42) and $c \neq 0$ for general incompressible material [13].

Clearly, for a given strain energy $W(\mathbf{F}) = \bar{W}(I_1, I_2)$, the governing equations obtained by Knowles constitute an over-determined system in general, i.e. two unknowns (u, p) but three equations. In order to solve this over-determined problem,

[2] There is a mistake in [6], i.e., $\bar{W}_1$ in Knowles' equation (2.19) should be $\bar{W}_2$.

Knowles believed that the stored energy $\bar{W}(I_1, I_2)$ should have some restrictions and he proved the following theorem.

Theorem 4 (**Knowles, 1976** [6]) If the stored energy $\bar{W}(I_1, I_2)$ is such that the ellipticity condition[3]

$$\frac{d[2R(\bar{W}_1 + \lambda^{-1}\bar{W}_2)]}{dR} > 0 \quad \forall R \geq 0, \quad \lambda > 0 \tag{43}$$

holds, then the associated incompressible elastic material admits nontrivial states of anti-plane shear for a given pre-stretch λ if and only if $\bar{W}(I_1, I_2)$ also satisfies the following constitutive constraint (i.e., Eq. (3.22) in [6])

$$b\bar{W}_1 + (b\lambda^{-1} - 1)\bar{W}_2 = 0, \tag{44}$$

for some constant b, for all values of I_1, I_2 such that $I_1 = \lambda_1 + R^2, \quad I_2 = \lambda_2 + \lambda^{-1}R^2, \quad R = |\boldsymbol{\gamma}|$.

First, by Lemma 1 we know that $W(\mathbf{F}) = \bar{W}(I_1, I_2) = V(I_1)$ hold for any given anti-plane shear deformation. There is no need to have both I_1, I_2 as variables. Therefore, the following trivial result shows immediately that Knowles' condition (44) is not a constitutive constraint.

Lemma 3 *For any given stored energy $W(\mathbf{F}) = \bar{W}(I_1, I_2)$ such that $I_1 = \lambda_1 + |\nabla u|^2, \quad I_2 = \lambda_2 + \lambda^{-1}|\nabla u|^2$, Knowles' constitutive condition (44) is automatically satisfied for $b = \frac{1}{2}\lambda$.*

The proof of this statement is elementary: by chain rule and $I_1 = \lambda I_2 + \lambda_1 - \lambda\lambda_2$, then

$$\bar{W}_2 = \frac{\partial \bar{W}}{\partial I_1}\frac{\partial I_1}{\partial I_2} = \lambda \bar{W}_1 \Rightarrow \quad b\bar{W}_1 + (b\lambda^{-1} - 1)\bar{W}_2 = (2b - \lambda)\bar{W}_1 = 0 \;\; \forall b = \frac{1}{2}\lambda.$$

To check if the Lagrange multiplier $p = p(x_1, x_2, x_3)$ must be in Knowles' formula (42), we use mathematical theory of Lagrange duality. For any given real-valued function $\phi(\mathbf{x}) \in L^q(\mathscr{B})$, the Lagrange multiplier p for the equality constraint $\phi(\mathbf{x}) = 0$ must be in the dual space $L^{q'}(\mathscr{B})$ such that $1/q + 1/q' = 1$. Since $\mathbf{F}$ depends only on $(x_1, x_2) \in \Omega$, the constraint $\phi(\mathbf{x}) = \det \mathbf{F}(u) - 1$ is defined on $\Omega \subset \mathscr{B}$, its Lagrange multiplier $p(\mathbf{x})$ must be defined on Ω. Indeed, by simple calculation for the form (42)

$$\int_{\mathscr{B}} \phi(x_1, x_2)p(x_1, x_2, x_3)\mathrm{d}\mathscr{B} = \int_{\Omega} \phi(x_1, x_2)[\int p(x_1, x_2, x_3)\mathrm{d}x_3]\mathrm{d}\Omega$$

one can easily find that the Lagrange multiplier is independent of x_3. Thus, we must have $c \equiv 0$ and $p = p(x_1, x_2)$ for any anti-plane shear deformation. For this reason

[3]i.e., Eq. (3.5) in [6].

and $\bar{W}_2 = \lambda \bar{W}_1$, $\zeta = \nabla V(\xi) = \bar{W}_1$, the Eq. (41) (i.e., (2.20) in [6]) is identical to the equation in $(BVP)_s$:

$$\left[2(\bar{W}_1 + \lambda^{-1}\bar{W}_2)u_{,\beta}\right]_{,\beta} = 0 \quad \Leftrightarrow \quad \nabla \cdot [2\zeta \nabla u] = 0 \quad \text{in } \Omega. \tag{45}$$

Now we need to check the other two equilibrium equations in Knowles' over-determined system. Instead of the local analysis, we use the well-known *virtual work principle*

$$\int_{\mathscr{B}} \operatorname{tr}(\boldsymbol{\sigma} \cdot \delta \mathbf{F}(\boldsymbol{\chi})) d\mathscr{B} = \int_{S_t} \mathbf{t} \cdot \delta \boldsymbol{\chi}, \quad \forall \boldsymbol{\chi} \in \mathscr{X}_c, \tag{46}$$

which holds for any given deformation problem regardless of constitutive laws. On $\mathscr{X}_b \subset \mathscr{X}_c$, we have the following strong complementarity conditions :

$$(\delta \boldsymbol{\chi}) \cdot (\operatorname{div} \boldsymbol{\sigma}) = 0 \quad \text{in } \mathscr{B}, \quad (\delta \boldsymbol{\chi}) \cdot \boldsymbol{\sigma} \cdot \mathbf{N} = (\delta \boldsymbol{\chi}) \cdot \mathbf{t} \quad \text{on } S_t. \tag{47}$$

The fact that the anti-plane shear deformation (10) has no displacements in $\mathbf{e}_1$ and $\mathbf{e}_2$ directions, i.e., $\delta \chi_\alpha \equiv 0 \;\; \forall \alpha = 1, 2 \;\; a.e.$ in Ω, the vector div $\boldsymbol{\sigma}$ is not necessarily to be zero in these directions. This shows that the additional two equilibrium equations (40), i.e. (2.19) in the paper [6], cannot be obtained from the virtual work principle. By the fact that the boundary value problem $(BVP)_s$ is well-determined by Eq. (45), these two extra equations are useless for the problem considered.

To understand the "function" of the hydrostatic pressure $p(\mathbf{x})$ in Knowles' over-determined problem, we use the KKT complementarity condition in (8), i.e. $p(\det \mathbf{F} - 1) = 0$. Since the anti-plane shear state is an isochoric deformation, the equality $\det \mathbf{F}(u) \equiv 1$ is trivially satisfied all most every where in Ω for any materials. For incompressible material, $p(\mathbf{x})$ is a pseudo-Lagrange multiplier and can be an arbitrarily given parameter. For compressible material, $p(\mathbf{x})$ could be zero, but in either case, the only function of this parameter is to balance the extra two Eq. (40), which cannot be obtained by the virtual work principle. This shows that the governing equations obtained by the minimum total potential principle are always compatible.

Finally, let us exam the ellipticity condition in Knowles's theorem. The following theorem is important in nonlinear analysis.

Theorem 5 *The ellipticity condition (43) is neither necessary nor sufficient for the nonlinear partial differential equation (45) to admit nontrivial states of anti-plane shear.* $(BVP)_s$ *has at least one solution only if* $t(\mathbf{x}) \neq 0$ *on* Γ_t *such that* $\mathscr{T}_a \neq \emptyset$.

For any given convex function $\bar{W}(I_1, I_2)$ *and the external force* $t(\mathbf{x}) \neq 0$ *on* Γ_t, *Eq. (45) is strongly G-elliptic if for every solution* ζ_1 *of (25)*

$$\zeta_1 = \bar{W}_1(I_1, I_2) > 0. \tag{48}$$

Proof. Let $\boldsymbol{\xi} = \{I_1, I_2\}$. Using chain rule for $\hat{W}(\boldsymbol{\gamma}) = \bar{W}(\boldsymbol{\xi}(\boldsymbol{\gamma}))$

$$\nabla \hat{W}(\boldsymbol{\gamma}) = \nabla_{\boldsymbol{\gamma}} \bar{W}(\boldsymbol{\xi}(\boldsymbol{\gamma})) = 2\boldsymbol{\gamma}(\bar{W}_1 + \lambda^{-1}\bar{W}_2),$$

thus, Knowles' ellipticity condition (43) is actually a special case of the strong Legendre condition $\nabla^2 \hat{W}(\boldsymbol{\gamma}) \succ 0$, which can only guarantee the convexity of $W(\mathbf{F}) = \hat{W}(\boldsymbol{\gamma})$, i.e., under this condition, Problem $(BVP)_s$ has at most one solution. Clearly, $(BVP)_s$ has a trivial solution if $t(\mathbf{x}) = 0$ on Γ_t. Therefore, Knowles' ellipticity condition (43) is not sufficient to admit a nontrivial solution.

By the canonical duality theory we know that for nonconvex stored energy $W(\mathbf{F}) = \hat{W}(\boldsymbol{\gamma})$, Problem $(\mathscr{P})_s$ can have multiple nontrivial solutions if $t(\mathbf{x}) \neq 0$ on Γ_t such that $\mathscr{T}_a \neq \emptyset$. Therefore, the condition (43) is also not necessary to admit a nontrivial solution.

By simple calculation for (43), we have

$$2(\bar{W}_1 + \lambda^{-1}\bar{W}_2) + 4R^2(\bar{W}_{11} + 2\lambda^{-1}\bar{W}_{12} + \lambda^{-2}\bar{W}_{22}) > 0, \tag{49}$$

which is a strong case for (31), where $\bar{W}_{\alpha\beta} = \partial^2 \bar{W}/\partial I_\alpha \partial I_\beta$. If the canonical function $\bar{W}(I_1, I_2)$ is convex in $\boldsymbol{\xi} = \{I_1, I_2\}$, we have

$$\bar{W}_{11} + 2\lambda^{-1}\bar{W}_{12} + \lambda^{-2}\bar{W}_{22} \geq 0 \;\; \forall \{I_1, I_2\} \in \mathbb{R}^2, \quad \lambda > 0. \tag{50}$$

By the facts that $\zeta_2 = \bar{W}_2 = \lambda \bar{W}_1 = \lambda\zeta_1$ and $\zeta_1 = \nabla V(I_1) = \bar{W}_1$, we know that the condition (49) holds as long as

$$2(\bar{W}_1 + \lambda^{-1}\bar{W}_2) = 4\zeta_1 > 0.$$

Thus, by Theorems 2 and 3 we know that the function $G(\boldsymbol{\gamma})$ is strictly G-quasiconvex and (45) is strongly G-elliptic. In this case, $(BVP)_s$ has at most one solution. □

Combining Theorems 1, 2, 5 and Lemma 3 we know that Knowles' constitutive constraints (43) and (44) are neither necessary nor sufficient for the existence of nontrivial states of anti-plane shear. Actually, this ellipticity condition even disallows many possible nontrivial local solutions in nonconvex problems. Indeed, it was shown in [3, 4] that for any given nonconvex stored energy $W(\mathbf{F}(\boldsymbol{\gamma})) = \bar{W}(I_1(\boldsymbol{\gamma}), I_2(\boldsymbol{\gamma})) = \hat{W}(\boldsymbol{\gamma})$ and nontrivial external force $t(\mathbf{x}) \neq 0$ such that $|t(\mathbf{x})| \leq \tau_c$ holds at certain $\mathbf{x} \in \Omega$, the minimum potential variational problem $(\mathscr{P})_s$ can have multiple solutions $\{u_k\}$ in Banach space $\mathscr{U}_c$ and can be obtained analytically by the canonical duality theory. Both global and local minimum solutions could be nonsmooth if $\boldsymbol{\tau}(\mathbf{x})$ changes its sign in Ω. While Knowles' over-determined system admits only a unique smooth solution in C^2 due to the additional ellipticity restriction on $\bar{W}$. Therefore, Knowles' over-determined system is a very special case of the variational problem $(\mathscr{P})_s$.

4 Conclusions

In summary, the following conclusions can be obtained.

1. The pure complementary energy principle and canonical duality–triality theory developed in [1] are powerful for solving general nonlinear boundary value problems in nonlinear elasticity.
2. Both convexity of the total potential and ellipticity condition of the associated fully nonlinear boundary value problem depend not only on the stored energy function, but also sensitively on the external force field.
3. The triality theory provides a sufficient condition to identify both global and local extremum solutions for nonconvex problems.
4. General anti-plane shear deformation problems must be governed by the generalized neo-Hookean model.
5. Unless the KKT theory is wrong, the incompressibility is not a variational constraint for any anti-plane shear deformation problem, the pseudo-Lagrange multiplier p depends only on (x_1, x_2), which is not a variable for the problem.
6. Unless the virtual work principle is wrong, there is only one equilibrium equation for general anti-plane shear deformation problems. The two extra equations in Knowles' over-determined system are not required.
7. Unless the minimum potential variational principle is wrong, the constitutive conditions required by Theorems in [6, 7] are neither necessary nor sufficient for general homogeneous materials to admit nontrivial states of anti-plane shear.

The first three conclusions are naturally included in the canonical duality–triality theory developed by the author and his coworkers during the last 25 years [1]. Extensive applications have been given in multidisciplinary fields of biology, chaotic dynamics, computational mechanics, information theory, phase transitions, post-buckling, operations research, industrial and systems engineering, etc.

The last four conclusions are obtained recently when the author was involved in discussions with colleagues on anti-plane shear deformation problems. As highly cited papers [6, 7], Knowles' over-determined system has been extensively applied to many anti-plane shear deformation problems in literature, see recent papers [10–12]. The main motivation for this paper and [2] is due to the recent challenges (see [13]).

Acknowledgements Insightful discussions with Professor C. Horgan from University of Virginia and Professor Martin Ostoja-Starzewski from University of Illinois are sincerely acknowledged. Reviewer's important comments and constructive suggestions are sincerely acknowledged. The research was supported by US Air Force Office of Scientific Research (AFOSR FA9550-10-1-0487).

References

1. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications, xviii +, Kluwer Academic Publishers, Dordrecht (2000) p. 454
2. Gao, D.Y.: Duality in G. Saccomandi's Challenge on Analytical Solutions to Anti-plane Shear Problem in Finite Elasticity (2015). https://arxiv.org/abs/1511.03374
3. Gao, D.Y.: Analytical solutions to general anti-plane shear problem in finite elasticity. Continuum Mech. Theorm. **28**(1), 175–194 (2016)
4. Gao, D.Y., Ogden, R.W.: Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation. Quarterly J. Mech. Appl. Math. **61**(4), 497–522 (2008)
5. Horgan, C.O.: Anti-plane shear deformations in linear and nonlinear solid mechanics. SIAM Rev. **37**(1), 53–81 (1995)
6. Knowles, J.K.: On finite anti-plane shear for imcompressible elastic materials. J. Aust. Math. Soc. **19**, 400–415 (1976)
7. Knowles, J.K.: A note on anti-plane shear for compressible materials in finite elastostatics. J. Aust. Math. Soc. B **20**, 17 (1977)
8. Latorre, V., Gao, D.Y.: Canonical duality for solving general nonconvex constrained problems. Optim. Lett. (2015). doi:10.1007/s11590-015-0860-0
9. Ogden, R.W.: Non-Linear Elastic Deformations. Ellis Horwood, Dover (1984/97)
10. Pucci, E., Rajagopal, K.R., Saccomandi, G.: On the determination of semi-inverse solutions of nonlinear Cauchy elasticity: The not so simple case of anti-plane shear. J. Eng. Sci. Int. (2014). doi:10.1016/j.ijengsci.2014.02.033
11. Pucci, E., Saccomandi, G.: The anti-plane shear problem in non-linear elasticity revisited. J. Elast. **113**, 167–177 (2013)
12. Pucci, E., Saccomandi, G.: Secondary motions associated with anti-plane shear in nonlinear isotropic elasticity. Q. Jl Mech. Appl. Math **66**(2), 221–239 (2013)
13. Saccomandi, G., Gao, D.Y.: Analytical solutions to general anti-plane shear problems in finite elasticity. Continuum Mech, Thermodyn. (2015)

Canonical Duality Method for Solving Kantorovich Mass Transfer Problem

Xiaojun Lu and David Yang Gao

Abstract This paper addresses analytical solution to the Kantorovich mass transfer problem. Through an ingenious approximation mechanism, the Kantorovich problem is first reformulated as a variational form, which is equivalent to a nonlinear differential equation with Dirichlet boundary. The existence and uniqueness of the solution can be demonstrated by applying the canonical duality theory. Then, using the canonical dual transformation, a perfect dual maximization problem is obtained, which leads to an analytical solution to the primal problem. Its global extremality for both primal and dual problems can be identified by a triality theory. In addition, numerical maximizers for the Kantorovich problem are provided under different circumstances. Finally, the theoretical results are verified by applications to Monge's problem. Although the problem is addressed in one-dimensional space, the theory and method can be generalized to solve high-dimensional problems.

1 Introduction

The Monge–Kantorovich mass transfer model is widely used in modern economic activities, medical science, and mechanical processes. In these respects, some typical examples include the logistics of transport for industrial products, purification of blood in the kidneys and livers, shape optimization, etc. Interesting readers can refer to [1, 2, 9, 23, 24, 28, 29] for more details.

The original transfer problem, which was proposed by Monge [28], investigated how to move one mass distribution to another one with the least amount of work. In this paper, we consider the Monge–Kantorovich problem in the 1-D case. Let

X. Lu (✉)
School of Mathematics and Jiangsu Key Laboratory of Engineering Mechanics,
Southeast University, Nanjing 211189, China
e-mail: lvxiaojun1119@hotmail.de

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

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics
and Mathematics 37, DOI 10.1007/978-3-319-58017-3_5

$\Omega = [a, b]$ and $\Omega^* = [c, d]$, $a, b, c, d \in \mathbb{R}$ and denote $U := \Omega \cup \Omega^* = [a, b] \cup [c, d]$. Here we focus on the closed case, and other bounded cases can be discussed similarly. Moreover, f^+ and f^- are two nonnegative density functions in Ω and Ω^*, respectively, and satisfy the normalized balance condition

$$\int_\Omega f^+ dx = \int_{\Omega^*} f^- dx = 1.$$

Let $c : \Omega \times \Omega^* \to [0, +\infty)$ be a cost function, which indicates the work required to move a unit mass from the position x to a new position y. There are many types of cost functions while dealing with different problems [2, 5, 9, 27]. In Monge's problem, the cost function is proportional to the distance $|x - y|$,

$$c(x, y) = |x - y|.$$

The Monge's problem consists in finding an optimal mass transfer mapping $\mathbf{s}^* : \Omega \to \Omega^*$ to minimize the cost functional $I(\mathbf{s})$:

$$I(\mathbf{s}^*) = \min_{\mathbf{s}\in\mathscr{N}} \left\{ I[\mathbf{s}] := \int_\Omega |x - \mathbf{s}(x)| f^+(x) dx \right\}, \tag{1}$$

where $\mathbf{s} : \Omega \to \Omega^*$ belongs to the class $\mathscr{N}$ of measurable mappings driving $f^+(x)$ to $f^-(y)$,

$$\mathbf{s}_\# f^+ = f^-,$$

which means, for $\forall x \in \Omega$,

$$f^+(x) = f^-(\mathbf{s}(x))|\det(J(\mathbf{s}(x)))|,$$

where $J(\mathbf{s}(x))$ is the Jacobian matrix of the mapping $\mathbf{s}$.

In the 1940s, Kantorovich [23, 24] relaxed Monge's transfer problem (1) and proposed the task of finding a Kantorovich potential $u^* \in \mathscr{L}$ solving

$$K[u^*] = \max_{w\in\mathscr{L}} \left\{ K[w] := \int_U w f dz = \int_U w(f^+ - f^-) dz \right\}, \tag{2}$$

where $\mathscr{L}$ is the class of functionals $w : U \to \mathbb{R}$ satisfying

$$\mathrm{Lip}[w] := \sup_{x\neq y} \frac{|w(x) - w(y)|}{|x - y|} \leq 1.$$

As a matter of fact, the Kantorovich's problem (2) is not a perfect maximization dual of Monge's minimization problem (1). Following the procedure of [5, 9], one can prove the *dual criteria for optimality* in the bounded case.

Lemma 1.1. *Let* $\mathbf{s}^* \in \mathscr{N}$ *and* $u^* \in \mathscr{L}$. *If the following identity holds,*

$$u^*(x) - u^*(\mathbf{s}^*(x)) = |x - \mathbf{s}^*(x)|,$$

then

- $\mathbf{s}^*$ *is an optimal mass transfer mapping in Monge's problem (1);*
- u^* *is a Kantorovich potential maximizing Kantorovich's problem (2);*
- *The minimum* $I[\mathbf{s}^*]$ *in (1) is equal to the maximum* $K[u^*]$ *in (2);*
- *Every optimal mass transfer mapping* $\mathbf{s}^*$ *and Kantorovich potential* u^* *satisfy the above identity.*

Due to the implicitness of u^*, L.C. Evans, W. Gangbo, and J. Moser [6, 8, 9] provided an ODE recipe to build $\mathbf{s}^*$ by solving a flow problem involving Du. This method is indeed useful but very complicated. In 2001, L.A. Caffarelli, M. Feldman, R.J. McCann, N.S. Trudinger, and X.J. Wang showed a much simpler approach to construct optimal mappings by decomposition of transfer sets and measure theory [5, 30]. Once an analytical Kantorovich potential u^* is found, by checking the above identity, one can immediately judge whether it is possible to construct a suitable optimal mapping $\mathbf{s}^*$ by virtue of u^*. However, due to the nonuniform convexity of the cost function $c(x, y)$, it is difficult to find optimal mass allocation. In order to gain some insight into this problem, many approximating mechanisms were introduced. For example, L.A. Caffarelli, W. Gangbo, R.J. McCann and X.J. Wang [4, 13, 14, 30], etc. utilized an approximation of strictly convex cost functions

$$c_\varepsilon(x, y) = |x - y|^{1+\varepsilon} \quad \varepsilon > 0.$$

The existence and uniqueness of the optimal mapping $\mathbf{s}_\varepsilon^*$ can be proved by convex analysis. Then let ε tends to 0, and one can construct an optimal mapping $\mathbf{s}^*$ using transfer rays and transfer sets invoked by L.C. Evans and W. Gangbo [8]. In addition, N.S. Trudinger and X.J. Wang used the approximation

$$c_\varepsilon(x, y) = \sqrt{\varepsilon^2 + |x - y|^2}$$

in the discussion of regularity [27, 30]. All the above-mentioned approximations concentrate upon the cost function $c(x, y)$. In this paper, we are eager to explore whether the approximation of Kantorovich's problem can bring more useful information.

Let $\mathscr{L}_0$ be a subset of $\mathscr{L}$,

$$\mathscr{L}_0 := \left\{\phi \in W_0^{2,\infty}(U) \cap C(U) \,\middle|\, |\phi_x| \le 1, \phi = 0 \text{ on } \Omega \cap \Omega^*\right\},$$

where $W_0^{2,\infty}(U)$ is a Sobolev space. Here, when $\Omega \cap \Omega^* = \emptyset$, $C(U)$ represents $C(\Omega)$ and $C(\Omega^*)$. We restrict our discussion of Kantorovich's problem (2) in $\mathscr{L}_0$, namely,

$$K[u] = \max_{w \in \mathscr{L}_0} \left\{ K[w] := \int_U w f dz = \int_U w(f^+ - f^-) dz \right\}. \tag{3}$$

In the survey paper [10], L.C. Evans proposed a sequence of approximated dual problems of (3). Now we explain the mechanism. We consider a sequence of approximated primal problems

$$(\mathscr{P}^{(k)}) : \min_{w_k \in \mathscr{L}_0} \left\{ J^{(k)}[w_k] := \int_U \left(H^{(k)}(w_{k,x}) - w_k f \right) dx \right\}, \tag{4}$$

where $w_{k,x}$ is the derivative of w_k with respect to x, $H^{(k)} : \mathbb{R} \to \mathbb{R}^+$ is defined as

$$H^{(k)}(\gamma) := \frac{1}{k} \mathrm{e}^{\frac{k}{2}(\gamma^2 - 1)},$$

and $J^{(k)}$ is called the potential energy functional. Notice that when $|\gamma| \leq 1$, then $\lim_{k \to \infty} H^{(k)}(\gamma) = 0$ uniformly. From [10], it is clear that

$$- \lim_{k \to \infty} \min_{w_k \in \mathscr{L}_0} J^{(k)}[w_k] = \max_{w \in \mathscr{L}_0} K[w].$$

Consequently, once a sequence of functions $\{u_k^*\}_k$ satisfying $J^{(k)}[u_k^*] = \min_{w_k \in \mathscr{L}_0} J^{(k)}[w_k]$ globally is obtained, then it will help us find a Kantorovich potential $u = \lim_{k \to \infty} u_k^*$ which solves (3).

In this paper, we investigate analytic solutions to the Kantorovich potential u^* of problem (3) using *canonical duality theory*. This theory was developed from Gao and Strang's original work on nonconvex/nonsmooth variational problems [21]. During the last few years, considerable effort has been taken to illustrate these nonconvex problems from the theoretical point of view [16, 17]. Interesting readers can refer to [18–20, 22].

Before we state the main results, we introduce some useful notations.

- θ_k is the corresponding Gâteaux derivative of $H^{(k)}$ with respect to $w_{k,x}$ given by

$$\theta_k(x) = \mathrm{e}^{\frac{k}{2}(w_{k,x}^2 - 1)} w_{k,x}.$$

- $\Phi^{(k)} : \mathscr{L}_0 \to L^\infty(U)$ is a nonlinear geometric mapping given by

$$\Phi^{(k)}(w_k) := \frac{k}{2}(w_{k,x}^2 - 1).$$

For convenience's sake, denote

$$\xi_k := \Phi^{(k)}(w_k).$$

It is evident that ξ_k belongs to the function space $\mathscr{U}$ given by

$$\mathscr{U} := \left\{\phi \in L^\infty(U) \middle| \phi \leq 0\right\}.$$

- $\Psi^{(k)} : \mathscr{U} \to L^\infty(U)$ is a canonical energy defined as

$$\Psi^{(k)}(\xi_k) := \frac{1}{k}\mathrm{e}^{\xi_k},$$

which is a convex function with respect to ξ_k.
- ζ_k is the corresponding Gâteaux derivative of $\Psi^{(k)}$ with respect to ξ_k given by

$$\zeta_k = \frac{1}{k}\mathrm{e}^{\xi_k},$$

which is invertible with respect to ξ_k and belongs to the function space $\mathscr{V}^{(k)}$,

$$\mathscr{V}^{(k)} := \left\{\phi \in L^\infty(U) \middle| 0 \leq \phi \leq \frac{1}{k}\right\}.$$

- λ_k is defined as

$$\lambda_k := k\zeta_k,$$

and belongs to the function space $\mathscr{V}$,

$$\mathscr{V} := \left\{\phi \in L^\infty(U) \middle| 0 \leq \phi \leq 1\right\}.$$

Now we are ready to introduce the main theorems.

Theorem 1.2. *For positive density functions $f^+ \in C(\Omega)$, $f^- \in C(\Omega^*)$, we can always find a sequence of analytical functions $\{u_k^* \in \mathscr{L}_0\}_k$ minimizing the approximated problems (4) globally.*

By canonical duality method, we are able to find an analytical Kantorovich potential for (3).

Theorem 1.3. *For positive density functions $f^+ \in C(\Omega)$, $f^- \in C(\Omega^*)$, we can always find an analytical global maximizer $u \in \mathscr{L}_0$ for Kantorovich's mass transfer problem (3).*

Remark 1.4. *Generally speaking, there are plenty of approximating schemes, for example, one can also let*

$$H^{(k)}(\gamma) := \frac{1}{k}(\gamma^2 - 1)^2.$$

Then by following the procedure in dealing with double-well potentials in [19, 21], we could definitely find an analytical Kantorovich potential.

Remark 1.5. *Through applying the canonical duality method, we have devised a systematic procedure in finding an analytical minimizer. In fact, for other types of cost functions, for instance,* $c(x, y) = |x - y|^p, \quad p \in [1, \infty)$*, we can also use this method to construct an analytical Kantorovich potential* u^**. Compared with former results [8, 9], we obtain an explicit representation of Kantorovich potential, which helps us construct an optimal mapping* $\mathbf{s}^*$ *according to Lemma 1.1. This question will be discussed in detail in the application part.*

The rest of the paper is organized as follows. In Sect. 2, first we apply the canonical dual transformation to establish a sequence of perfect dual problems and a pure complementary energy principle. Next we explain the canonical duality theory and triality theory. In particular, the triality theory provides global extremum conditions for the problem (4). Afterward, we construct a sequence of analytical functions minimizing $J^{(k)}$ globally and a Kantorovich potential maximizing $K[w]$ of (3). In the final analysis, we use a product allocation model in 1-D to illustrate our theoretical results.

2 Proof of the Main Results: Technique of Canonical Duality Method

2.1 Proof of Lemma 1.1 in the Bounded Case:

Proof. Similar as [5], for any $\mathbf{s} \in \mathcal{N}$ and $w \in \mathcal{L}_0$, we compute

$$
\begin{aligned}
I[\mathbf{s}] &= \int_\Omega |x - \mathbf{s}(x)| f^+(x) dx \\
&= \int_U |x - \mathbf{s}(x)| f^+(x) dx \\
&\geq \int_U (w(x) - w(\mathbf{s}(x))) f^+(x) dx \\
&= \int_U w(x) f^+(x) dx - \int_U w(y) f^-(y) dy \\
&= K[w].
\end{aligned}
$$

Taking into account the given identity, we complete the proof.

2.2 Proof of Theorem 1.2:

Here we apply the variational method to discuss problem (4). Now we show an important lemma in this respect.

Lemma 2.1. *The Euler–Lagrange equation for* $(\mathscr{P}^{(k)})$ *takes the following form,*

$$\theta_{k,x} + f = (\mathrm{e}^{\frac{k}{2}(u_{k,x}^2-1)} u_{k,x})_x + f = 0, \quad \textit{in } U. \tag{5}$$

Remark 2.2. *The term* $\mathrm{e}^{\frac{k}{2}(u_{k,x}^2-1)}$ *is called the transport density. Clearly, like* $p-$*Laplace operator,* $\mathrm{e}^{\frac{k}{2}(u_{k,x}^2-1)}$ *is a highly nonlinear and nonlocal function of* $u_k \in \mathscr{L}_0$. *With the hidden boundary value* $u_k = 0$ *on* ∂U, *we are able to prove the existence and uniqueness of the solution of (5). This important fact will be explained later.*

Proof. Indeed, the Gâteaux derivative of $J^{(k)}$ with respect to u_k belongs to $L^1(U)$. For any given $\mu > 0$ and any test function $\phi \in \mathscr{L}_0$, by integrating by parts, we have

$$\lim_{\mu\to 0^+} \frac{J^{(k)}[u_k + \mu\phi] - J^{(k)}[u_k]}{\mu}$$

$$= \lim_{\mu\to 0^+} \int_U \left\{ \frac{\frac{1}{k}\mathrm{e}^{\frac{k}{2}((u_k+\mu\phi)_x^2-1)} - \frac{1}{k}\mathrm{e}^{\frac{k}{2}(u_{k,x}^2-1)}}{\mu} - \phi f \right\} dx$$

$$= \int_U \left\{ \frac{1}{k}\mathrm{e}^{\frac{k}{2}(u_{k,x}^2-1)} \lim_{\mu\to 0^+} \frac{\mathrm{e}^{\frac{k}{2}(\mu^2\phi_x^2+2\mu u_{k,x}\phi_x)} - 1}{\frac{k}{2}(\mu^2\phi_x^2 + 2\mu u_{k,x}\phi_x)} \cdot \frac{\frac{k}{2}(\mu^2\phi_x^2 + 2\mu u_{k,x}\phi_x)}{\mu} - \phi f \right\} dx$$

$$= \int_U \left\{ \frac{1}{k}\mathrm{e}^{\frac{k}{2}(u_{k,x}^2-1)} k u_{k,x}\phi_x - \phi f \right\} dx$$

$$= -\int_U \left\{ (\mathrm{e}^{\frac{k}{2}(u_{k,x}^2-1)} u_{k,x})_x \phi + \phi f \right\} dx.$$

Actually, since u_k and ϕ are both in $\mathscr{L}_0$, then for any given $\mu < 0$, when $\mu \to 0^-$, the above calculation still holds.

Now we are going to apply the canonical duality method invoked by David Y. Gao [17]. By Legendre transformation, we define a *Gao–Strang total complementary energy functional*.

Definition 2.3. With the notations in Sect. 1, we define a Gao–Strang total complementary energy $\Xi^{(k)}$ in the form

$$\Xi^{(k)}(u_k, \zeta_k) := \int_U \left\{ \Phi^{(k)}(u_k)\zeta_k - \Psi_*^{(k)}(\zeta_k) - f u_k \right\} dx, \tag{6}$$

where the function $\Psi_*^{(k)} : \mathscr{V}^{(k)} \to L^\infty(U)$ is defined as

$$\Psi_*^{(k)}(\zeta_k) := \xi_k\zeta_k - \Psi^{(k)}(\xi_k) = \zeta_k(\ln(k\zeta_k) - 1). \tag{7}$$

Next we introduce an important *criticality criterium* for the Gao–Strang total complementary energy functional.

Definition 2.4. $(\bar{u}_k, \bar{\zeta}_k) \in \mathscr{L}_0 \times \mathscr{V}^{(k)}$ is called a critical pair of $\Xi^{(k)}$ if and only if

$$D_{u_k} \Xi^{(k)}(\bar{u}_k, \bar{\zeta}_k) = 0, \tag{8}$$

and

$$D_{\zeta_k} \Xi^{(k)}(\bar{u}_k, \bar{\zeta}_k) = 0, \tag{9}$$

where D_{u_k}, D_{ζ_k} denote the partial Gâteaux derivatives of $\Xi^{(k)}$, respectively.

By variational method, we explore the criticality criterium (8) and (9). Indeed, on the one hand, we have the following observation from (8).

Lemma 2.5. *For a fixed $\zeta_k \in \mathscr{V}^{(k)}$, (8) leads to the equilibrium equation*

$$(\lambda_k \bar{u}_{k,x})_x + f = 0, \quad \textit{in } U. \tag{10}$$

Remark 2.6. *It is easy to check the equilibrium equation (10) is consistent with (5) except that the transport density is replaced by $\lambda_k = k\zeta_k$. We will use this fact to construct a sequence of analytical solutions later.*

Proof. Indeed, the partial Gâteaux derivative of $\Xi^{(k)}$ with respect to u_k belongs to $L^1(U)$. For $\forall \mu > 0$ and any test function $\phi \in \mathscr{L}_0$, by integrating by parts, we have

$$\begin{aligned}
&\lim_{\mu \to 0^+} \frac{\Xi^{(k)}(\bar{u}_k + \mu\phi, \zeta_k) - \Xi^{(k)}(\bar{u}_k, \zeta_k)}{\mu} \\
&= \lim_{\mu \to 0^+} \int_U \frac{\Lambda^{(k)}(\bar{u}_k + \mu\phi) - \Lambda^{(k)}(\bar{u}_k)}{\mu} \zeta_k dx - \int_U f\phi dx \\
&= \lim_{\mu \to 0^+} \int_U \frac{k(\bar{u}_{k,x} + \mu\phi_x)^2 - \bar{u}_{k,x}^2)}{2\mu} \zeta_k dx - \int_U f\phi dx \\
&= \int_U k\bar{u}_{k,x}\phi_x \zeta_k dx - \int_U f\phi dx \\
&= -\int_U \Big\{ (k\zeta_k \bar{u}_{k,x})_x \phi dx + f\phi \Big\} dx.
\end{aligned}$$

Since u_k and ϕ are both in $\mathscr{L}_0$, then for $\forall \mu < 0$, when $\mu \to 0^-$, the above calculation still holds.

On the other hand, from (9), we have the following observation.

Lemma 2.7. *For a fixed $u_k \in \mathscr{L}_0$, (9) is in fact the constructive law*

$$\Phi^{(k)}(u_k) = D_{\zeta_k}\Psi_*^{(k)}(\bar{\zeta}_k). \tag{11}$$

Remark 2.8. *It is worth noticing that (11) is consistent with the notations in Sect. 1.*

Proof. Indeed, the partial Gâteaux derivative of $\Xi^{(k)}$ with respect to ζ_k belongs to $L^1(U)$. For $\forall\mu > 0$ and any test function $\phi \in \mathscr{L}_0$, by integrating by parts, we have

$$\begin{aligned}
&\lim_{\mu\to 0^+} \frac{\Xi^{(k)}(u_k, \bar{\zeta}_k + \mu\phi) - \Xi^{(k)}(u_k, \bar{\zeta}_k)}{\mu} \\
&= \lim_{\mu\to 0^+} \int_U \left\{\Phi^{(k)}(u_k)\phi - \frac{\Psi_*^{(k)}(\bar{\zeta}_k + \mu\phi) - \Psi_*^{(k)}(\bar{\zeta}_k)}{\mu}\right\} dx \\
&= \int_U \left\{\Phi^{(k)}(u_k) - D_{\zeta_k}\Psi_*^{(k)}(\bar{\zeta}_k)\right\}\phi dx.
\end{aligned}$$

Since u_k and ϕ are both in $\mathscr{L}_0$, then for $\forall\mu < 0$, when $\mu \to 0^-$, the above calculation still holds.

Lemmas 2.5 and 2.7 indicate that $\bar{u}_k$ from the critical pair $(\bar{u}_k, \bar{\zeta}_k)$ solves (5). Now we introduce the canonical duality theory. For our purpose, we define the following *Gao–Strang pure complementary energy functional*.

Definition 2.9. From Definition 2.3, we define a Gao–Strang pure complementary energy $J_d^{(k)}$ in the form

$$J_d^{(k)}[\zeta_k] := \Xi^{(k)}(\bar{u}_k, \zeta_k), \tag{12}$$

where $\bar{u}_k$ solves (5).

For the sake of convenience, we give another representation of $J_d^{(k)}$ by the following lemma.

Lemma 2.10. *The pure complementary energy functional $J_d^{(k)}$ can be rewritten as*

$$J_d^{(k)}[\zeta_k] = -\frac{1}{2}\int_U \left\{\frac{\theta_k^2}{k\zeta_k} + k\zeta_k + 2\zeta_k(\ln(k\zeta_k) - 1)\right\} dx. \tag{13}$$

Remark 2.11. *$\bar{u}_k$ is included in this representation in an implicit manner, which will simplify our further discussion considerably.*

Proof. With Definition 2.3, by integrating by parts, we have

$$
\begin{aligned}
\Xi^{(k)}(\bar{u}_k, \zeta_k) &= \int_U \left\{\left(\frac{k}{2}(\bar{u}_{k,x}^2 - 1\right)\zeta_k - \Psi_*^{(k)}(\zeta_k) - f\bar{u}_k\right\}dx \\
&= \int_U \left\{k\zeta_k \bar{u}_{k,x}^2 - f\bar{u}_k\right\}dx \\
&\quad - \textstyle\int_U \left\{\frac{k}{2}(\bar{u}_{k,x}^2 - 1)\zeta_k + k\zeta_k + \zeta_k(\ln(k\zeta_k) - 1)\right\}dx \\
&= \underbrace{-\int_U \left\{(k\zeta_k\bar{u}_{k,x})_x + f\right\}\bar{u}_k dx}_{(I)} \\
&\quad \underbrace{-\frac{1}{2}\int_U \left\{k\zeta_k\bar{u}_{k,x}^2 + k\zeta_k + 2\zeta_k(\ln(k\zeta_k) - 1)\right\}dx.}_{(II)}
\end{aligned}
\tag{14}
$$

Since $\bar{u}_k$ solves (5), then the first part (I) disappears. Keeping in mind the definition of θ_k, we reach the conclusion immediately.

With the above discussion, next we establish the dual variational problem of (4).

$$
(\mathscr{P}_d^{(k)}) : \max_{\zeta_k \in \mathscr{V}^{(k)}} \left\{J_d^{(k)}[\zeta_k] = -\frac{1}{2}\int_U \left\{\frac{\theta_k^2}{k\zeta_k} + k\zeta_k + 2\zeta_k(\ln(k\zeta_k) - 1)\right\}dx\right\}. \tag{15}
$$

By variational calculus, we have the following lemma.

Lemma 2.12. *The variation of $J_d^{(k)}$ with respect to ζ_k leads to the dual algebraic equation(DAE), namely,*

$$
\theta_k^2 = k\bar{\zeta}_k^2(2\ln(k\bar{\zeta}_k) + k), \tag{16}
$$

where $\bar{\zeta}_k$ is from the critical pair $(\bar{u}_k, \bar{\zeta}_k)$.

Proof. Indeed, by calculating the Gâteaux derivative of $J_d^{(k)}$ with respect to ζ_k, we can prove the lemma immediately.

Remark 2.13. *Taking into account the notation of λ_k, we can rewrite (16) as*

$$
\theta_k^2 = \lambda_k^2 \ln(\mathrm{e}\lambda_k^{\frac{2}{k}}). \tag{17}
$$

From (17), we know that θ_k^2 is monotonously increasing with respect to $\lambda_k > \mathrm{e}^{-\frac{k}{2}}$.

As a matter of fact, we have the following asymptotic expansion of θ_k^2.

Lemma 2.14. *When $k \geq 3$, θ_k^2 has the expansion of the form*

$$\theta_k^2 = (1 - \frac{2}{k})\lambda_k^2 + \frac{2}{k}\lambda_k^3 + R_k(\lambda_k),$$

where the remainder term $|R_k(\lambda_k)| \leq \frac{1}{k}$ uniformly for any $\lambda_k \in [e^{-\frac{k}{2}}, 1]$. In particular, for a fixed $x \in U$, if $\lim_{k\to\infty} \lambda_k = \lambda$ and $\lim_{k\to\infty} \theta_k = \theta$ in L^∞, then we have the limit version of (17),

$$\theta^2 = \lambda^2. \tag{18}$$

Remark 2.15. *In the 1-D case, later on we will demonstrate how to find the limit θ of the sequence $\{\theta_k\}_k$ as k tends to infinity.*

Proof. Since $\lambda_k \in [e^{-\frac{k}{2}}, 1]$, we can rewrite (17) using Taylor's expansion formula for $\ln \lambda_k$ at the point 1,

$$\theta_k^2 = \lambda_k^2\Big(1 + \frac{2}{k}(\lambda_k - 1) - \frac{1}{k\eta_k^2}(\lambda_k - 1)^2\Big) = (1 - \frac{2}{k})\lambda_k^2 + \frac{2}{k}\lambda_k^3 - \frac{1}{k}\frac{\lambda_k^2}{\eta_k^2}(\lambda_k - 1)^2,$$

where $\eta_k \in (\lambda_k, 1)$. It is evident that

$$|\frac{1}{k}\frac{\lambda_k^2}{\eta_k^2}(\lambda_k - 1)^2| \leq |\frac{1}{k}\frac{\lambda_k^2}{\lambda_k^2}(\lambda_k - 1)^2| \leq \frac{1}{k}.$$

This concludes our proof.

By comparing (5) with (10), we deduce that an analytical solution of (5) can be given as

$$\bar{u}_k(x) = \int_{x_0}^{x} \frac{\theta_k(t)}{\lambda_k(t)}dt + C, \tag{19}$$

where $x, x_0 \in U$. Together with (18), we see that

$$\lim_{k\to\infty} |\bar{u}_{k,x}| = 1,$$

which is consistent with the conclusion in [9]. Summarizing the above discussion, we have the following duality theorem.

Theorem 2.16 (Canonical Duality Theory)**.** *For positive density functions $f^+ \in C(\Omega)$, $f^- \in C(\Omega^*)$, if θ_k is a solution of the Euler–Lagrange equation (5), which is not identically equal to 0, then (17) has a unique positive root $\bar{\lambda}_k$ due to the monotonicity property. Furthermore, an analytical function given by*

$$\bar{u}_k(x) = \int_{x_0}^{x} \frac{\theta_k(t)}{\bar{\lambda}_k(t)}dt + C \tag{20}$$

is a local minimizer of (4) and satisfies the following duality identity locally,

$$J^{(k)}[\bar{u}_k] = J_d^{(k)}[\bar{\zeta}_k], \tag{21}$$

where $(\bar{u}_k, \bar{\zeta}_k)$ *is a critical pair for* $\Xi^{(k)}$.

Proof. It suffices to prove the identity (21). Indeed, this identity is obtained by direct variational calculus of $J^{(k)}[u_k]$ and $J_d^{(k)}[\zeta_k]$ in (4) and (15), respectively.

$$J^{(k)}[\bar{u}_k] = \Xi^{(k)}(\bar{u}_k, \bar{\zeta}_k) = J_d^{(k)}[\bar{\zeta}_k]. \tag{22}$$

Remark 2.17. *Theorem 2.16 demonstrates that the maximization of the pure complementary energy functional* $J_d^{(k)}$ *is perfectly dual to the minimization of the potential energy functional* $J^{(k)}$. *In effect, the identity (22) indicates there is no duality gap between them.*

Up to now, we have constructed a critical pair $(\bar{u}_k, \bar{\zeta}_k)$ satisfying (22) locally. Next we verify that $\bar{u}_k$ and $\bar{\zeta}_k$ are exactly a global minimizer for $J^{(k)}$ and a global maximizer for $J_d^{(k)}$, respectively. In the following theorem, we apply the *triality theory* to obtain the extremum conditions for the critical pair.

Theorem 2.18 (Triality Theory)**.** *For positive density functions* $f^+ \in C(\Omega)$, $f^- \in C(\Omega^*)$, *we have,* θ_k *is the unique solution of the Euler–Lagrange equation (5) with hidden Dirichlet boundary. Moreover,* $\bar{\zeta}_k$ *is a global maximizer of* $J_d^{(k)}$ *over* $\mathscr{V}^{(k)}$, *and the corresponding* $\bar{u}_k$ *in the form of (20) is a global minimizer of* $J^{(k)}$ *over* $\mathscr{L}_0$, *namely,*

$$J^{(k)}(u_k^*) = J^{(k)}(\bar{u}_k) = \min_{u_k \in \mathscr{L}_0} J^{(k)}(u_k) = \max_{\zeta_k \in \mathscr{V}^{(k)}} J_d^{(k)}(\zeta_k) = J_d^{(k)}(\bar{\zeta}_k). \tag{23}$$

Proof. We divide our proof into three parts. In the first and second parts, we discuss the uniqueness of θ_k. Extremum conditions will be illustrated in the third part.

First Part:

Without loss of generality, we consider the disjoint case $\Omega = [a, b]$ and $\Omega^* = [c, d]$, $b < c$. In Ω, we have a general solution for the nonlinear differential equation (5) in the form of

$$\theta_k(x) = -F(x) + C_k, \quad F(x) := \int_a^x f^+(x)dx, \quad x \in [a, b].$$

Since $f^+ > 0$, then $F \in C[a, b]$ is a strictly increasing function with respect to $x \in [a, b]$ and consequently is invertible. Let F^{-1} be its inverse function, which is also a strictly increasing function, then

$$F^{-1} : [0, 1] \to [a, b].$$

From Remark 2.13, we see that there exists a unique piecewise continuous function $\lambda_k(x) > \mathrm{e}^{-\frac{k}{2}}$ except for the point $x = F^{-1}(C_k)$. By paying attention to the fact that $\bar{u}_k(a) = 0$, we represent the analytical solution $\bar{u}_k$ in the following form:

$$\bar{u}_k(x) = \int_a^x \frac{-F(x) + C_k}{\lambda_k(x)} dx, \quad x \in [a, b].$$

Since

$$\lim_{x \to F^{-1}(C_k)^-} \frac{-F(x) + C_k}{\lambda_k(x)} = 0, \quad \lim_{x \to F^{-1}(C_k)^+} \frac{-F(x) + C_k}{\lambda_k(x)} = 0,$$

thus $\bar{u}_k$ is continuous at the point $x = F^{-1}(C_k)$. As a result, $\bar{u}_k \in C[a, b]$. Recall that

$$\bar{u}_k(b) = \int_a^{F^{-1}(C_k)} \frac{-F(x) + C_k}{\lambda_k(x)} dx + \int_{F^{-1}(C_k)}^b \frac{-F(x) + C_k}{\lambda_k(x)} dx = 0,$$

and we can determine the constant $C_k \in (0, 1)$ uniquely. Indeed, let

$$M(t) := \int_a^{F^{-1}(t)} \frac{-F(x) + t}{\lambda_k(x, t)} dx + \int_{F^{-1}(t)}^b \frac{-F(x) + t}{\lambda_k(x, t)} dx,$$

where $\lambda_k(x, t)$ is from (17). It is evident that λ_k depends on C_k. As a matter of fact, M is strictly increasing with respect to $t \in (0, 1)$, which leads to

$$C_k = M^{-1}(0).$$

Indeed, for $t_1 < t_2$, $t_1, t_2 \in (0, 1)$, by keeping in mind the identity (17), we have

$$\begin{aligned}
M(t_1) &= \int_a^{F^{-1}(t_1)} \frac{-F(x) + t_1}{\lambda_k(x, t_1)} dx + \int_{F^{-1}(t_1)}^b \frac{-F(x) + t_1}{\lambda_k(x, t_1)} dx \\
&= \int_a^{F^{-1}(t_1)} \frac{-F(x) + t_1}{\lambda_k(x, t_1)} dx + \int_{F^{-1}(t_1)}^{F^{-1}(t_2)} \frac{-F(x) + t_1}{\lambda_k(x, t_1)} dx + \int_{F^{-1}(t_2)}^b \frac{-F(x) + t_1}{\lambda_k(x, t_1)} dx \\
&< \int_a^{F^{-1}(t_1)} \frac{-F(x) + t_2}{\lambda_k(x, t_2)} dx + \int_{F^{-1}(t_1)}^{F^{-1}(t_2)} \frac{-F(x) + t_2}{\lambda_k(x, t_2)} dx + \int_{F^{-1}(t_2)}^b \frac{-F(x) + t_2}{\lambda_k(x, t_2)} dx \\
&= M(t_2).
\end{aligned}$$

More information concerned with C_k will be explained in the proof of Theorem 1.3.

Second Part:

Applying the similar procedure, we see that

$$\theta_k(x) = G(x) - D_k, \quad G(x) := \int_c^x f^-(x)dx, \quad x \in [c, d],$$

where the constant $D_k \in (0, 1)$. Since $f^- > 0$, then $G \in C[c, d]$ is a strictly increasing function with respect to $x \in [c, d]$ and consequently is invertible. Let G^{-1} be its inverse function, which is also a strictly increasing function, then

$$G^{-1} : [0, 1] \to [c, d].$$

We can represent the analytical solution $\bar{u}_k$ in the following form:

$$\bar{u}_k(x) = \int_c^x \frac{G(x) - D_k}{\lambda_k(x)} dx, \quad x \in [c, d].$$

Since

$$\lim_{x \to G^{-1}(D_k)^-} \frac{G(x) - D_k}{\lambda_k(x)} = 0, \quad \lim_{x \to G^{-1}(D_k)^+} \frac{G(x) - D_k}{\lambda_k(x)} = 0,$$

thus $\bar{u}_k$ is continuous at the point $x = G^{-1}(D_k)$. As a result, $\bar{u}_k \in C[c, d]$. Recall that

$$\bar{u}_k(d) = \int_c^{G^{-1}(D_k)} \frac{G(x) - D_k}{\lambda_k(x)} dx + \int_{G^{-1}(D_k)}^d \frac{G(x) - D_k}{\lambda_k(x)} dx = 0,$$

and we can determine the constant $D_k \in (0, 1)$ uniquely. Indeed, let

$$N(t) := \int_c^{G^{-1}(t)} \frac{G(x) - t}{\lambda_k(x, t)} dx + \int_{G^{-1}(t)}^d \frac{G(x) - t}{\lambda_k(x, t)} dx,$$

where $\lambda_k(x, t)$ is from (17). As a matter of fact, N is strictly decreasing with respect to $t \in (0, 1)$, which leads to

$$D_k = N^{-1}(0).$$

Indeed, for $t_1 < t_2, t_1, t_2 \in (0, 1)$, by keeping in mind the identity (17), we have

$$
\begin{aligned}
N(t_1) &= \int_c^{G^{-1}(t_1)} \frac{G(x)-t_1}{\lambda_k(x,t_1)}dx + \int_{G^{-1}(t_1)}^{d} \frac{G(x)-t_1}{\lambda_k(x,t_1)}dx \\
&= \int_c^{G^{-1}(t_1)} \frac{G(x)-t_1}{\lambda_k(x,t_1)}dx + \int_{G^{-1}(t_1)}^{G^{-1}(t_2)} \frac{G(x)-t_1}{\lambda_k(x,t_1)}dx + \int_{G^{-1}(t_2)}^{d} \frac{G(x)-t_1}{\lambda_k(x,t_1)}dx \\
&> \int_c^{G^{-1}(t_1)} \frac{G(x)-t_2}{\lambda_k(x,t_2)}dx + \int_{G^{-1}(t_1)}^{G^{-1}(t_2)} \frac{G(x)-t_2}{\lambda_k(x,t_2)}dx + \int_{G^{-1}(t_2)}^{d} \frac{G(x)-t_2}{\lambda_k(x,t_2)}dx \\
&= N(t_2).
\end{aligned}
$$

Furthermore, the other cases, such as $b = c$ and $b > c$, can also be discussed similarly due to the fact that $\bar{u}_k = 0$ on $\Omega \cap \Omega^*$. Therefore, θ_k is uniquely determined in U and the analytic solution $\bar{u}_k \in C(U)$.

Third Part:

In order to prove the extremum of the critical pair, we recall the second variational formula for both $J^{(k)}$ and $J_d^{(k)}$.

On the one hand, for any test function $\phi \in \mathscr{L}_0$ satisfying $\phi_x \neq 0$ a.e. in U, the second variational form $\delta_\phi^2 J^{(k)}$ with respect to ϕ is equal to

$$
\int_U \frac{d^2}{dt^2}\Big\{H^{(k)}((\bar{u}_k + t\phi)_x)\Big\}\Big|_{t=0} dx = \int_U \mathrm{e}^{\frac{k}{2}(\bar{u}_{k,x}^2 - 1)}\Big\{k(\bar{u}_{k,x}\phi_x)^2 + \phi_x^2\Big\}dx. \tag{24}
$$

On the other hand, for any test function $\psi \in \mathscr{V}^{(k)}$ satisfying $\psi \neq 0$ a.e. in U, the second variational form $\delta_\psi^2 J_d^{(k)}$ with respect to ψ is equal to

$$
\begin{aligned}
&-\frac{1}{2}\int_U \frac{d^2}{dt^2}\Big\{\frac{\theta_k^2}{k(\zeta_k + t\psi)} + 2(\zeta_k + t\psi)\Big(\ln(k(\zeta_k + t\psi)) - 1\Big)\Big\}\Big|_{t=0} dx \\
= &-\int_U \Big\{\frac{\theta_k^2\psi^2}{k\zeta_k^3} + \frac{\psi^2}{\zeta_k}\Big\}dx.
\end{aligned}
$$

From (24) and (25), we know immediately that

$$
\delta_\phi^2 J^{(k)}(\bar{u}_k) > 0, \quad \delta_\psi^2 J_d^{(k)}(\bar{\zeta}_k) < 0. \tag{25}
$$

Then Theorem 2.16 and the uniqueness of θ_k discussed in the first and second parts complete our proof.

Consequently, we reach the conclusion of Theorem 1.2.

2.3 Proof of Theorem 1.3:

Proof. Without loss of generality, we still consider the disjoint case, $b < c$. First we show an important lemma which describes the asymptotic behavior of C_k and D_k when k tends to infinity.

Lemma 2.19. *When $b < c$, the sequences of $\{C_k\}_k$ and $\{D_k\}_k$ are given in the proof of Theorem 2.18, then we have*

$$\lim_{k\to\infty} C_k = F(\frac{a+b}{2}), \tag{26}$$

$$\lim_{k\to\infty} D_k = G(\frac{c+d}{2}). \tag{27}$$

Proof. Recall the identity

$$\bar{u}_k(b) = \int_a^{F^{-1}(C_k)} \frac{-F(x)+C_k}{\lambda_k(x)}dx + \int_{F^{-1}(C_k)}^b \frac{-F(x)+C_k}{\lambda_k(x)}dx = 0. \tag{28}$$

Since

$$\lim_{k\to\infty} \frac{-F(x)+C_k}{\lambda_k(x)} = 1, \quad x \in [a, F^{-1}(C_k)),$$

$$\lim_{k\to\infty} \frac{-F(x)+C_k}{\lambda_k(x)} = -1, \quad x \in (F^{-1}(C_k), b],$$

then for $\forall \varepsilon > 0$, there exists an $N \in \mathbb{N}^+$, such that for $\forall k > N$, the following inequalities hold:

$$(1-\varepsilon)(F^{-1}(C_k)-a) \le \int_a^{F^{-1}(C_k)} \frac{-F(x)+C_k}{\lambda_k(x)}dx \le (1+\varepsilon)(F^{-1}(C_k)-a), \tag{29}$$

$$(-1-\varepsilon)(b-F^{-1}(C_k)) \le \int_{F^{-1}(C_k)}^b \frac{-F(x)+C_k}{\lambda_k(x)}dx \le (-1+\varepsilon)(b-F^{-1}(C_k)). \tag{30}$$

Combining (29)–(31) together, we have

$$\frac{a+b}{2} - \frac{b-a}{2}\varepsilon \le F^{-1}(C_k) \le \frac{a+b}{2} + \frac{b-a}{2}\varepsilon. \tag{31}$$

Then (27) follows immediately. It is obvious that we can prove (28) in a similar manner.

As a result, we define the limit of θ_k in L^∞ as

$$\theta(x) := \begin{cases} \lim\limits_{k\to\infty}(-F(x)+C_k) = F_{ab}(x), \quad F_{ab}(x) = -F(x) + F(\dfrac{a+b}{2}),\ x\in[a,b], \\ \\ \lim\limits_{k\to\infty}(G(x)-D_k) = G_{cd}(x), \quad G_{cd}(x) = G(x) - G(\dfrac{c+d}{2}), \quad\ x\in[c,d]. \end{cases}$$

Next, according to (18), we define the limit of λ_k in L^∞ as

$$\lambda(x) := \begin{cases} |F_{ab}(x)|,\ x\in[a,b], \\ \\ |G_{cd}(x)|,\ x\in[c,d]. \end{cases}$$

Finally, we calculate the limit of $\bar{u}_k$ in $\mathscr{L}_0$ as follows:

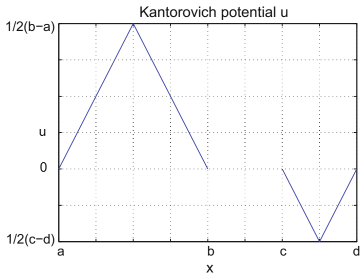
$$u(x) := \begin{cases} \displaystyle\int_a^x \frac{F_{ab}(x)}{|F_{ab}(x)|}dx = x - a, & x\in[a, \frac{a+b}{2}], \\ \displaystyle\int_a^{\frac{a+b}{2}} \frac{F_{ab}(x)}{|F_{ab}(x))|}dx + \int_{\frac{a+b}{2}}^x \frac{F_{ab}(x)}{|F_{ab}(x)|}dx = -x + b, & x\in(\frac{a+b}{2}, b], \\ \displaystyle\int_c^x \frac{G_{cd}(x)}{|G_{cd}(x)|}dx = -x + c, & x\in[c, \frac{c+d}{2}], \\ \displaystyle\int_c^{\frac{c+d}{2}} \frac{G_{cd}(x)}{|G_{cd}(x)|}dx + \int_{\frac{c+d}{2}}^x \frac{G_{cd}(x)}{|G_{cd}(x)|}dx = x - d, & x\in(\frac{c+d}{2}, d]. \end{cases} \tag{32}$$

This solution is illustrated in Fig. 1. Several other cases can be proved similarly, and the corresponding Kantorovich potentials are depicted in Figs. 1, 2, 3 and 4. As a result, we have constructed a global maximizer for Kantorovich's mass transfer problem (3) in 1-D.

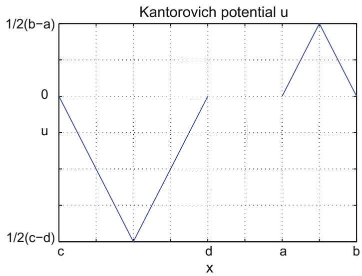
2.4 Application to Monge's Problem

During the past few decades, Monge's and Kantorovich's problems have been the subject of active inquiry, since it covers the domains of optimization, probability theory, partial differential equations, allocation mechanism in economics and membrane filtration in biology, etc. In this application part, we apply the main theorems to solve a product allocation model in 1-D.

We want to transport some products from $[a, b]$ to $[c, d]$. Assume that the products are distributed uniformly in $[a, b]$, that means, the density function f^+ satisfies



(a) $a < b < c < d,\ \ b-a \geq d-c$



(b) $c < d < a < b,\ \ b-a \leq d-c$

Fig. 1 The unique continuous Kantorovich potential of Problem (3) while Ω and Ω^* are disjoint in 1-D

$$f^+(x) = \frac{1}{b-a}, \quad x \in [a,b].$$

Figure 1: When $f^- > 0$ in $[c,d]$, according to Theorem 1.3, one can check that the unique Kantorovich potential does not satisfy the dual criteria for optimality. Therefore, in this case, the Kantorovich problem (3) is not a perfect dual problem of Monge's problem (1). We know, the optimal mapping should be $\mathbf{s}^*(x) = c$, in which case, the density f^- is a δ-function in the form of

$$f^- = \begin{cases} \infty & \text{if } x = c, \\ 0 & \text{if } x \in (c,d], \end{cases}$$

satisfying

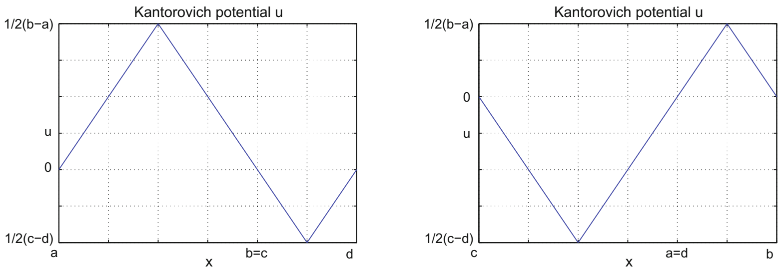
$$\int_c^d f^-(x)dx = 1.$$

Figure 2: When $f^- > 0$ in $[c,d]$, according to Theorem 1.3, one can check that the unique Kantorovich potential satisfies the dual criteria for optimality only for $x \in [\frac{a+b}{2}, b],\ \ y \in [c, \frac{c+d}{2}]$. Therefore, in this case, the Kantorovich problem (3) is not a perfect dual problem of Monge's problem (1). We know, the optimal mapping should be $\mathbf{s}^*(x) = c$.

Figure 3: When $f^- > 0$ in $[c,d]$, according to Theorem 1.3, one can check that the unique Kantorovich potential satisfies the dual criteria for optimality only when we choose

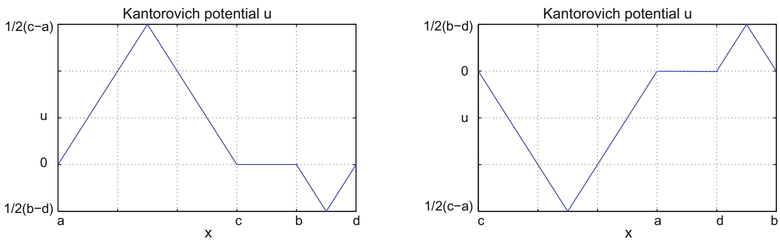
$$\mathbf{s}(x) = \begin{cases} c & \text{if } x \in [\frac{a+c}{2}, c] \ \text{ and } y = c, \\ x & \text{if } x, y \in (c,b), \\ y & \text{if } x = b \ \text{ and } y \in [b, \frac{b+d}{2}]. \end{cases}$$

Therefore, in this case, the Kantorovich problem (3) is not a perfect dual problem of Monge's problem (1). We know, the optimal mapping should be



(a) $a < b = c < d,\ \ b - a \geq d - c$ (b) $c < d = a < b,\ \ b - a \leq d - c$

Fig. 2 The unique continuous Kantorovich potential of Problem (3) while Ω and Ω^* have a unique common point in 1-D



(a) $a < c < b < d,\ \ b - a \geq d - c$ (b) $c < a < d < b,\ \ b - a \leq d - c$

Fig. 3 The unique continuous Kantorovich potential of Problem (3) while $\Omega \cap \Omega^*$ have more than one common point and $\Omega \not\subseteq \Omega^*$ or $\Omega^* \not\subseteq \Omega$ in 1-D

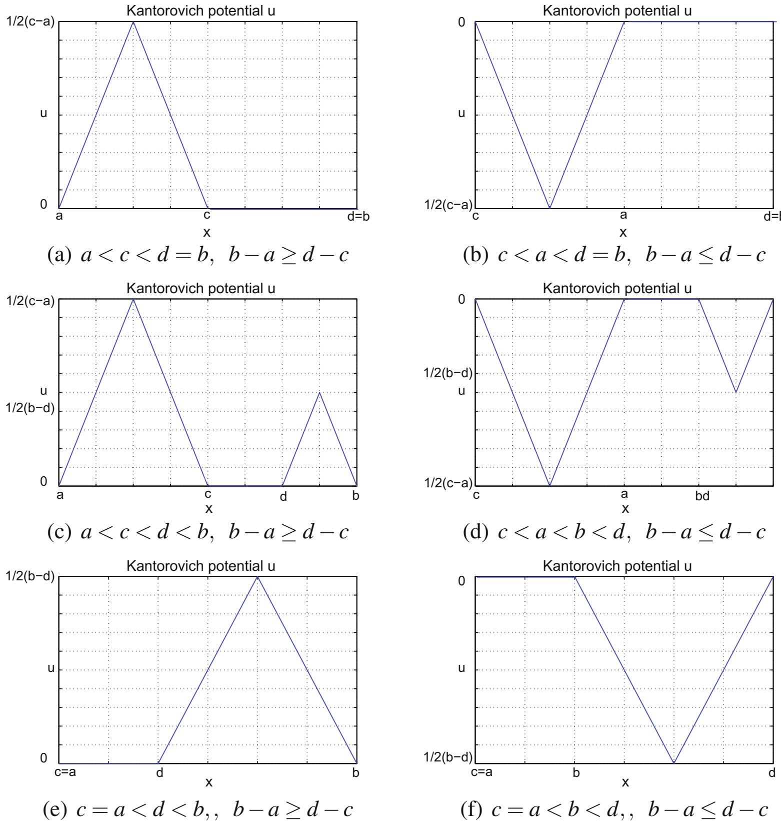
$$\mathbf{s}^*(x) = \begin{cases} c & \text{if } x \in [a, c], \\ x & \text{if } x \in (c, b]. \end{cases}$$

Figure 4a, c, e: When $f^- > 0$ in $[c, d]$, according to Theorem 1.3, one can check that the unique Kantorovich potential satisfies the dual criteria for optimality only when we choose

$$\mathbf{s}(x) = \begin{cases} c & \text{if } x \in [\frac{a+c}{2}, c] \ \text{ and } y = c, \\ x & \text{if } x, y \in (c, d], \\ d & \text{if } x \in (d, \frac{b+d}{2}] \ \text{ and } y = d. \end{cases}$$

Therefore, in this case, the Kantorovich problem (3) is not a perfect dual problem of Monge's problem (1). We know, the optimal mapping should be

$$\mathbf{s}^*(x) = \begin{cases} c & \text{if } x \in [a, c], \\ x & \text{if } x \in (c, d], \\ d & \text{if } x \in (d, b]. \end{cases}$$



(a) $a<c<d=b,\ \ b-a\geq d-c$
(b) $c<a<d=b,\ \ b-a\leq d-c$
(c) $a<c<d<b,\ \ b-a\geq d-c$
(d) $c<a<b<d,\ \ b-a\leq d-c$
(e) $c=a<d<b,,\ \ b-a\geq d-c$
(f) $c=a<b<d,,\ \ b-a\leq d-c$

Fig. 4 The unique continuous Kantorovich potential of Problem (3) while $\Omega\subseteq\Omega^*$ or $\Omega^*\subseteq\Omega$ in 1-D

Figure 4b, d, f: When $f^->0$ in $[a,b]$, according to Theorem 1.3, one can check that the unique Kantorovich potential does satisfy the dual criteria for optimality when we choose $\mathbf{s}^*(x)=x,\quad x\in[a,b]$. Therefore, the Kantorovich problem (3) in this case is a perfect dual problem of Monge's problem (1).

Acknowledgements The main results in this paper were obtained during a research collaboration in the Federation University Australia in August, 2015. The first author wishes to thank Professor David Gao for his hospitality and financial support. This project is partially supported by US Air Force Office of Scientific Research (AFOSR FA9550-10-1-0487 and FA9550-17-1-0151). This project is also supported by Jiangsu Planned Projects for Postdoctoral Research Funds (1601157B), Shanghai University Start-up Grant for Shanghai 1000-Talent Program Scholars, National Natural Science Foundation of China (NSFC 61673104, 71673043, 71273048, 71473036, 11471072), the Scientific Research Foundation for the Returned Overseas Chinese Scholars, Fundamental Research Funds for the Central Universities (2014B15214, 2242017K40086), Open Research Fund Program of Jiangsu

Key Laboratory of Engineering Mechanics, Southeast University (LEM16B06). In particular, the authors also express their deep gratitude to the referees for their careful reading and useful remarks.

References

1. Ambrosio, L.: Lecture Notes on Optimal Transfer Problems, preprint
2. Ambrosio, L.: Optimal transport maps in Monge-Kantorovich problem. ICM **3**, 1–3 (2002)
3. Bourgain, J., Brezis, H.: Sur l'équation div $u = f$. C. R. Acad. Sci. Paris, Ser. **I334**, 973–976 (2002)
4. Caffarelli, L.A.: Allocation maps with general cost functions, in partial differential equations and applications. Lect. Notes Pure Appl. Math. **177**, 29–35 (1996)
5. Caffarelli, L.A., Feldman, M., MCcann, R.J.: Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. J. AMS **15**, 1–26 (2001)
6. Dacorogna, B., Moser, J.: On a partial differential equation involving the Jacobian determinant. Ann. Inst. H. Poincaré **7**, 1–26 (1990)
7. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. Dunod, Paris (1976)
8. Evans, L.C., Gangbo, W.: Differential equations methods for the Monge-Kantorovich mass transfer problem. Mem. Am. Math. Soc. **653** (1999)
9. Evans, L.C.: Partial differential equations and Monge-Kantorovich mass transfer(survey paper)
10. Evans, L.C.: Three singular variational problems, preprint (2002)
11. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence (2002)
12. Gangbo, W., McCann, R.J.: Optimal maps in Monge's transport problem, preprint (1995)
13. Gangbo, W., McCann, R.J.: Optimal maps in Monge's mass transport problem. C. R. Acad. Sci. Paris Sér. I Math. **321**, 1653–1658 (1995)
14. Gangbo, W., McCann, R.J.: The geometry of optimal transportation. Acta Math. **177**, 113–161 (1996)
15. Gangbo, W., Święch, A.: Optimal maps for the multidimensional Monge-Kantorovich problem, preprint (1996)
16. Gao, D.Y.: Duality, triality and complementary extremum principles in non-convex parametric variational problems with applications. IMA J. Appl. Math. **61**, 199–235 (1998)
17. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory. Methods and Applications. Kluwer Academic Publishers, Boston (2000)
18. Gao, D.Y.: Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications. Nonlinear Anal. **42**(7), 1161–1193 (2000)
19. Gao, D.Y., Lu, X.: Multiple solutions for non-convex variational boundary value problems in higher dimensions, preprint (2013)
20. Gao, D.Y., Ogden, R.W.: Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation. Q. Jl Mech. Appl. Math. **61**(4), 497–522 (2008)
21. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Q. Appl. Math. **47**(3), 487–504 (1989)
22. Gao, D.Y., Ogden, R.W., Stravroulakis, G.: Nonsmooth and Nonconvex Mechanics: Modelling. Analysis and Numerical Methods. Kluwer Academic Publishers, Boston (2001)
23. Kantorovich, L.V.: On the transfer of masses. Dokl. Akad. Nauk. SSSR **37**, 227–229 (1942). (Russian)
24. Kantorovich, L.V.: On a problem of Monge. Uspekhi Mat. Nauk. **3**, 225–226 (1948)
25. Li, Q.R., Santambrogio, F., Wang, X.J.: Regularity in Monge's mass transfer problem. J. Math. Pures Appl. **102**, 1015–1040 (2014)
26. Lions, J.L., Magenes, E.: Problèmes aux limites non homogènes et applications I-III, Dunod, Paris, 1968-1970

27. Ma, X.N., Trudinger, N.S., Wang, X.J.: Regularity of potential functions of the optimal transportation problem. Arch. Ration. Mech. Anal. **177**, 151–183 (2005)
28. Monge, G.: Mémoire sur la théorie des déblais et de remblais, Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année, pp. 666–704 (1781)
29. Sudakov, V.N.: Geometric problems in the theory of infinite-dimensional probability distributions. Proc. Steklov Inst. **141**, 1–178 (1979)
30. Trudinger, N.S., Wang, X.J.: On the Monge mass transfer problem. Calc. Var. **13**, 19–31 (2001)
31. Vershik, A.M.: Some remarks on the infinite-dimensional problems of linear programming. Russian Math. Surv. **25**, 117–124 (1970)

Triality Theory for General Unconstrained Global Optimization Problems

David Yang Gao and Changzhi Wu

Abstract Triality theory is proved for a general unconstrained global optimization problem. The method adopted is simple but mathematically rigorous. Results show that if the primal problem and its canonical dual have the same dimension, the triality theory holds strongly in the tri-duality form as it was originally proposed. Otherwise, both the canonical min-max duality and the double-max duality still hold strongly, but the double-min duality holds weakly in a super-symmetrical form as it was expected. Additionally, a complementary weak saddle min-max duality theorem is discovered. Therefore, an open problem on this statement left in 2003 is solved completely. This theory can be used to identify not only the global minimum, but also the largest local minimum, maximum, and saddle points. Application is illustrated. Some fundamental concepts in optimization and remaining challenging problems in canonical duality theory are discussed.

1 Introduction

The general global optimization problem to be solved is proposed in the following form:

$$(\mathscr{P}): \quad \text{ext}\left\{ \Pi(\boldsymbol{x}) = W(\boldsymbol{x}) + \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{f}\rangle \mid \boldsymbol{x} \in \mathbb{R}^n \right\}, \tag{1}$$

where $W(\boldsymbol{x})$ is a nonconvex function, $A \in \mathbb{R}^{n\times n}$ is a given symmetric matrix, $\boldsymbol{f} \in \mathbb{R}^n$ is a given (source) vector, $\langle *, * \rangle$ is an inner product in $\mathbb{R}^n$, and the notation ext$\{*\}$ stands for finding global extrema of the function given in $\{*\}$, including both global minimum and the largest local minimum and maximum. In order to have this general problem making sense in reality, the nonconvex function $W(\boldsymbol{x})$ should obey certain fundamental rules in systems theory. In this paper, we shall need the following assumptions for the nonconvex function $W(\boldsymbol{x})$.

D.Y. Gao (✉) · C. Wu
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_6

(A1) The nonconvex function $W(\boldsymbol{x})$ is twice continuously differentiable.

(A2) There exits a *geometrical operator*

$$\Lambda(\boldsymbol{x}) = \left\{ \frac{1}{2}\boldsymbol{x}^T B^k \boldsymbol{x} + \boldsymbol{b}_k^T \boldsymbol{x} \right\} : \ \mathbb{R}^n \to \mathbb{R}^m \tag{2}$$

and a strictly convex function $V : \mathbb{R}^m \to \mathbb{R}$ such that

$$W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x})), \tag{3}$$

where $B^k \in \mathbb{R}^{n\times n}$ and $\boldsymbol{b}_k \in \mathbb{R}^n$, $k = 1, \cdots, m$.

(A3). The critical points of problem $(\mathscr{P})$ are non-singular, i.e., if $\nabla \Pi(\bar{\boldsymbol{x}}) = 0$, then $\det(\nabla^2 \Pi(\bar{\boldsymbol{x}})) \neq 0$.

Based on Assumption (A2), the general problem (1) can be reformulated in the following canonical form:

$$(\mathscr{P}) : \ \ \text{ext} \left\{ \Pi(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x})) + \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{f}\rangle \mid \boldsymbol{x} \in \mathbb{R}^n \right\}. \tag{4}$$

This problem arises extensively in many fields of engineering and sciences, including Euclidean distance geometry [5, 20], computational biology [6, 29, 45], numerical methods for solving a large class of nonconvex variational problems in mathematical physics [13, 27, 35], and much more.

Actually, the assumption (A2) is the so-called *canonical transformation* introduced in [10]. The idea of this transformation was from Gao and Strang's original work [22] on nonconvex variational problems in large deformation theory, where the geometrical operator $\Lambda(\chi) = \frac{1}{2}(\nabla\chi)^T(\nabla\chi)$ is a (pure) quadratic measure of the deformation gradient $\varepsilon = \nabla\chi$, which is the so-called *right Cauchy–Green deformation tensor*, and $W(\nabla\chi)$ is a stored strain energy. In continuum physics, a real-valued function is called stored energy which must obey certain physical (constitutive) laws.

Objectivity is a basic concept in science, which is often attributed with the property of scientific measurements that can be measured independently of the observer. General description of the objectivity can be easily found on internet and in many mathematical physics textbooks (see [28, 33]). Mathematical definitions of the objective set and objective function are given in the book [10] (Chap. 6, p. xxx).

Let

$$\mathscr{Q} = \{Q \in \mathbb{R}^{m\times m} |\ Q^T = Q^{-1}, \ \ \det Q = 1\}$$

be a proper orthogonal rotation group.

Definition 1 (Objectivity and Isotropy). A subset $\mathscr{Y}_a \subset \mathbb{R}^m$ is said to be *objective* if $Q\boldsymbol{y} \in \mathscr{Y}_a \ \ \forall \boldsymbol{y} \in \mathscr{Y}_a$ and $\forall Q \in \mathscr{Q}$. A real-valued function $T : \mathscr{Y}_a \to \mathbb{R}$ is said to be *objective* if its domain is objective and

$$T(Q\boldsymbol{y}) = T(\boldsymbol{y}) \;\; \forall \boldsymbol{y} \in \mathscr{Y}_a \;\; \text{and} \;\; \forall Q \in \mathscr{Q}. \tag{5}$$

A subset $\mathscr{Y}_a \subset \mathbb{R}^m$ is said to be *isotropic* if $\boldsymbol{y}Q^T \in \mathscr{Y}_a \;\; \forall \boldsymbol{y} \in \mathscr{Y}_a$ and $\forall Q \in \mathscr{Q}$. A real-valued function $T : \mathscr{Y}_a \to \mathbb{R}$ is said to be *isotropic* if its domain is isotropic and

$$T(\boldsymbol{y}Q^T) = T(\boldsymbol{y}) \;\; \forall \boldsymbol{y} \in \mathscr{Y}_a \;\; \text{and} \;\; \forall Q \in \mathscr{Q}. \tag{6}$$

Geometrically speaking, the objectivity means that the function $T(\boldsymbol{y})$ does not depend on rotation, but on certain measure (norm) of its variable $\boldsymbol{y}$. Therefore, the most simple objective function is the l_2-norm $T(\boldsymbol{y}) = \|\boldsymbol{y}\|$ since $\|Q\boldsymbol{y}\|^2 = \boldsymbol{y}^T Q^T Q \boldsymbol{y} = \boldsymbol{y}^T \boldsymbol{y} = \|\boldsymbol{y}\|^2 \;\; \forall Q \in \mathscr{Q}$. While the isotropy implies that the function $T(\boldsymbol{y})$ possesses a certain symmetry. By the fact that $(\boldsymbol{x}Q^T)(\boldsymbol{x}Q^T)^T = \boldsymbol{x}\boldsymbol{x}^T \succeq 0 \;\; \forall Q \in \mathscr{Q}$, the concept of isotropy plays important role in Semi-Definite Programming (SDP) and integer programming [14, 19].

Using finite element discretization for the deformation field $u(\boldsymbol{x})$, the nonconvex variational problems in infinite-dimensional space can be reduced to the canonical global optimization problem ($\mathscr{P}$) (see [27, 35]). It is known in continuum physics that the stored energy W is usually a nonconvex function of the linear measure ∇u (which is not a strain measure), but $V(e)$ is convex in term of the objective measure $e = \Lambda(u)$. Therefore, by this quadratic objective operator $\Lambda(u)$, a *complementary gap function* was discovered by Gao and Strang in nonconvex variational analysis and the complementary variational principle was recovered in fully nonlinear equilibrium problems of mathematical physics.[1] They also proved that the nonnegative gap function can be used to identify global minimizer of the nonconvex problem. Seven years later, it was discovered that the negative gap function can be used to identify the largest local minimum and maximum. Therefore, the *triality theory* was first proposed in nonconvex mechanics [7, 8], and then generalized to global optimization [11]. This triality theory is composed of a canonical min-max duality and two pairs of double-min, double-max dualities, which reveals an intrinsic duality pattern in complex systems and has been used successfully for solving a wide class of challenging problems in nonconvex analysis and global optimization [10]. However, it was realized in 2003 [12, 13] that the double-min duality holds conditionally under "certain additional conditions". Recently, this problem is partly solved for a class of fourth-order polynomial optimization problems [24, 37].

The objectivity lays a foundation for mathematical physics and systems theory. Generally speaking, if the function $W(\boldsymbol{x})$ is convex, there exists a (geometrically) linear operator $D : \mathbb{R}^n \to \mathscr{Y}_a \subset \mathbb{R}^m$ and a convex objective function $T : \mathscr{Y}_a \to \mathbb{R}$ such that $W(\boldsymbol{x}) = T(D\boldsymbol{y})$. Then the problem ($\mathscr{P}$) is called *geometrically linear* (see

[1] In continuum physics, complementary variational principle means perfect duality since any duality gap will violate certain physical laws. The existence of a complementary variational principle was a well-known debate existing for several decades in large deformation theory (see [31]). This problem was partially solved by Gao and Strang's work, and solved completely in 1999 [9].

definition given in [10, 11]). Actually, the most popular Lagrangian in its original form is defined by[2] (see [30])

$$\Pi(\boldsymbol{x}) = T(D\boldsymbol{x}) - U(\boldsymbol{x}), \tag{7}$$

where the convex objective function $T(\boldsymbol{y})$ is a kinetic energy, $U(\boldsymbol{x})$ is a potential energy of the system, which could be either linear or convex on a subset $\mathscr{X}_a \subset \mathbb{R}^n$ such that $\Pi(\boldsymbol{x})$ is well-defined on the so-called *kinetically admissible space* $\mathscr{X}_k = \{\boldsymbol{x} \in \mathscr{X}_a |\ D\boldsymbol{x} \in \mathscr{Y}_a\}$ [10]. For Newtonian mechanics, $T(\boldsymbol{y})$ is a quadratic function, the objectivity is implied. While for Einsteins special relativity theory, the objective function $T(\boldsymbol{y})$ is strictly convex (see Chap. 2, [10]). In either case, the duality relation (constitutive law) $\boldsymbol{y}^* = \nabla T(\boldsymbol{y})$ is an one-to-one (canonical) mapping from $\mathscr{Y}_a$ to its dual space $\mathscr{Y}_a^*$ induced by a bilinear form $\langle *; *\rangle$. Therefore, the so-called *complementary energy* $T^*(\boldsymbol{y}^*)$ can be uniquely defined on $\mathscr{Y}_a^*$ by the classical Legendre transformation

$$T^*(\boldsymbol{y}^*) = \text{sta}\{\langle \boldsymbol{y}; \boldsymbol{y}^*\rangle - T(\boldsymbol{y}) |\ \boldsymbol{y} \in \mathscr{Y}_a\}$$

such that the original Lagrangian $\Pi(\boldsymbol{x})$ is equivalent to its mixed form

$$L(\boldsymbol{x}, \boldsymbol{y}^*) = \langle D\boldsymbol{x}; \boldsymbol{y}^*\rangle - T^*(\boldsymbol{y}^*) - U(\boldsymbol{x}) : \mathscr{X}_a \times \mathscr{Y}_a^* \to \mathbb{R}, \tag{8}$$

which is the standard form in mathematical optimization.

If $U(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f}\rangle$ is linear, the Lagrangian form $\Pi(\boldsymbol{x})$ is convex on $\mathscr{X}_k$ and its mixed form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ is a saddle function on $\mathscr{X}_a \times \mathscr{Y}_a^*$. Therefore, the traditional saddle Lagrangian duality theory links the convex primal problem $\min\{\Pi(\boldsymbol{x}) |\ \boldsymbol{x} \in \mathscr{X}_k\}$ to a unique dual problem

$$\max\left\{\Pi^*(\boldsymbol{y}^*) = -T^*(\boldsymbol{y}^*) |\ \ \boldsymbol{y}^* \in \mathscr{Y}_s^*\right\}, \tag{9}$$

where $\mathscr{Y}_s^* = \{\boldsymbol{y}^* \in \mathscr{Y}_a^* |\ D^*\boldsymbol{y}^* = \boldsymbol{f} \in \mathscr{X}_a^* \subset \mathbb{R}^n\}$ is the so-called *statically admissible space*, $D^* : \mathscr{Y}_a^* \to \mathscr{X}^*$ is an adjoint operator of D defined by $\langle D\boldsymbol{x}; \boldsymbol{y}^*\rangle = \langle \boldsymbol{x}, D^*\boldsymbol{y}^*\rangle$. The objectivity of this dual problem (both the target function T^* and the feasible set $\mathscr{Y}_a^*$)[3] is guaranteed by the objectivity of $T(\boldsymbol{y})$. By introducing a Lagrange multiplier $\boldsymbol{x}$, which must be a solution to the primal problem (see Lagrange multiplier's law in Sect. 1.5 [10]), to relax the equilibrium constraint $D^*\boldsymbol{y}^* = \boldsymbol{f}$ in $\mathscr{Y}_s^*$, the Lagrangian is exactly the mixed form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ and the one-to-one Lagrangian saddle min-max duality

---

[2] The Lagrangian form was first introduced by W. Hamilton in classical mechanics and denoted by $L = T - U$, which is the standard notation extensively used from dynamical systems to quantum field theory (see [30]).

[3] The equilibrium equation $D^*\boldsymbol{y}^* = \boldsymbol{f}$ in Newtonian systems is an invariant under the Galilean transformation, which is the combination of Newton's three laws, see Chap. 2, [10]); while for Einsteins special relativity theory, this abstract equation is an invariant under the Lorentz transformation.

$$\min_{\boldsymbol{x}\in\mathcal{X}_k} \Pi(\boldsymbol{x}) = \min_{\boldsymbol{x}\in\mathcal{X}_a} \max_{\boldsymbol{y}^*\in\mathcal{Y}_a^*} L(\boldsymbol{x}, \boldsymbol{y}^*) = \max_{\boldsymbol{y}^*\in\mathcal{Y}_a^*} \min_{\boldsymbol{x}\in\mathcal{X}_a} L(\boldsymbol{x}, \boldsymbol{y}^*) = \max_{\boldsymbol{y}^*\in\mathcal{Y}_s^*} \Pi^*(\boldsymbol{y}^*)$$

is called the *mono-duality* in canonical systems theory [10], which has been studied extensively in mathematical physics and convex analysis [2, 38].

If the function $U : \mathcal{X}_a \to \mathbb{R}$ is convex, the Lagrangian form $\Pi(\boldsymbol{x})$ (7) is the so-called *d.c. (difference of convex) function*. Since the mixed Lagrangian form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ is no longer a saddle function, the well-known Hamiltonian

$$H(\boldsymbol{x}, \boldsymbol{y}^*) = \langle D\boldsymbol{x}; \boldsymbol{y}^*\rangle - L(\boldsymbol{x}, \boldsymbol{y}^*) = T^*(\boldsymbol{y}^*) + U(\boldsymbol{x})$$

was introduced, which is convex and has been extensively used in dynamical systems. Actually, although the Lagrangian is not a saddle function in convex Hamiltonian systems, it is a so-called *super-critical function* [10], and if the total potential $U(\boldsymbol{x})$ is strictly convex on $\mathcal{X}_a \subset \mathbb{R}^n$ such that its Legendre conjugate $U^*(\boldsymbol{x}^*)$ can be uniquely defined on $\mathcal{X}_a^*$, then the canonical dual action of $\Pi(\boldsymbol{x})$ can still be defined by

$$\Pi^*(\boldsymbol{y}^*) = \max\{L(\boldsymbol{x}, \boldsymbol{y}^*) |\ \boldsymbol{x} \in \mathcal{X}_a\} = U^*(D^*\boldsymbol{y}^*) - T^*(\boldsymbol{y}^*)$$

on $\mathcal{Y}_s^* = \{\boldsymbol{y}^* \in \mathcal{Y}_a^* |\ D^*\boldsymbol{y}^* \in \mathcal{X}_a^*\}$, which is also a d.c. function. Therefore, instead of the mono-duality in static systems, convex Hamiltonian systems are controlled by the so-called *bi-duality theory*.

Bi-Duality Theorem [10]: *If* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}^*)$ *is a critical point of the Lagrangian* $L(\boldsymbol{x}, \boldsymbol{y}^*)$, *then* $\bar{\boldsymbol{x}}$ *is a critical point of* $\Pi(\boldsymbol{x})$, $\bar{\boldsymbol{y}}^*$ *is a critical point of* $\Pi^*(\boldsymbol{y}^*)$ *and* $\Pi(\bar{\boldsymbol{x}}) = L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}^*) = \Pi^*(\bar{\boldsymbol{y}}^*)$. *Moreover, if* $n = m$, *we have*

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x}\in\mathcal{X}_k} \Pi(\boldsymbol{x}) \quad \Leftrightarrow \quad \max_{\boldsymbol{y}^*\in\mathcal{Y}_s^*} \Pi^*(\boldsymbol{y}^*) = \Pi^*(\bar{\boldsymbol{y}}^*) \tag{10}$$

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathcal{X}_k} \Pi(\boldsymbol{x}) \quad \Leftrightarrow \quad \min_{\boldsymbol{y}^*\in\mathcal{Y}_s^*} \Pi^*(\boldsymbol{y}^*) = \Pi^*(\bar{\boldsymbol{y}}^*). \tag{11}$$

This bi-duality is actually a special case of the triality theory in geometrically linear systems, which was originally presented in Chap. 2 [10] for one-dimensional dynamical systems with a simple proof. This bi-duality reveals a stable periodical property in convex Hamiltonian systems.

The concepts of triality and tri-duality were originally proposed in nonconvex mechanics [8], where $W(\boldsymbol{x})$ is a nonconvex function (strain energy). By the fact that the linear operator $D : \mathbb{R}^n \to \mathbb{R}^m$ cannot change the convexity of the objective function, if the function $W(\boldsymbol{x})$ is nonconvex and $W(\boldsymbol{x}) = T(D\boldsymbol{x})$, the function $T(\boldsymbol{y})$ is still nonconvex and its Legendre conjugate $T^*(\boldsymbol{y}^*)$ can not be uniquely defined [36]. Although the Fenchel conjugate

$$T^\sharp(\boldsymbol{y}^*) = \sup\{\langle \boldsymbol{y}; \boldsymbol{y}^*\rangle - T(\boldsymbol{y}) |\ \boldsymbol{y} \in \mathcal{Y}_a\}$$

can be uniquely defined, the function

$$Ł(\boldsymbol{x}, \boldsymbol{y}^*) = \langle D\boldsymbol{x}; \boldsymbol{y}^* \rangle - T^\sharp(\boldsymbol{y}^*) - U(\boldsymbol{x}) \tag{12}$$

is not the traditional Lagrangian form and the associate saddle min-max duality theory will produce the so-called duality gap in nonconvex optimization.

Actually, in terms of $U(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f} \rangle - \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x} \rangle$, the total complementary function $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ defined by (20) can be written as

$$\Xi(\boldsymbol{x}, \boldsymbol{\varsigma}) = \langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma} \rangle - V^*(\boldsymbol{\varsigma}) - U(\boldsymbol{x}). \tag{13}$$

Comparing this $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ with either $Ł(\boldsymbol{x}, \boldsymbol{y}^*)$ or the mixed Lagrangian form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ we can see that the fundamental difference between the canonical duality theory and other methods is the canonical transformation $W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x}))$ instead of the linear transformation $W(\boldsymbol{x}) = T(D\boldsymbol{x})$ used in many other duality theories, including the Fenchel–Moreau–Rockafellar duality. In real applications, if the quadratic function $U(\boldsymbol{x})$ is nonconvex, the mixed Lagrangian form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ is nonconvex in $\boldsymbol{x}$ since D is linear. However, the total complementary function $\Xi(*, \boldsymbol{\varsigma}) : \mathscr{X} \subset \mathbb{R}^n \to \mathbb{R}$ is always convex for $\boldsymbol{\varsigma} \in \mathscr{S}_a^+$ and concave for $\boldsymbol{\varsigma} \in \mathscr{S}_a^-$ due to the geometrically nonlinear operator $\Lambda(\boldsymbol{x})$ and its canonical dual variable $\boldsymbol{\varsigma}$. Therefore, $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ was also called the *nonlinear Lagrangian* in [10] and the *extended Lagrangian* in [12]. If the geometrical operator $\Lambda(\boldsymbol{x})$ is quadratic and objective, the so-called Λ-transformation [12]

$$U^\Lambda(\boldsymbol{\varsigma}) = \text{sta}\{\langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma} \rangle - U(\boldsymbol{x}) |\ \ \boldsymbol{x} \in \mathscr{X}\} \tag{14}$$

is actually the *pure complementary gap function* which is obtained from the complementary gap function $G_{ap}(\boldsymbol{x}, \boldsymbol{\varsigma}) = \frac{1}{2}\langle \boldsymbol{x}, G(\boldsymbol{\varsigma})\boldsymbol{x} \rangle$ using the analytical solution form $\boldsymbol{x} = [G(\boldsymbol{\varsigma})]^{-1}F(\boldsymbol{\varsigma})$.

The canonical duality theory was originally developed from this concept [10], which is the reason why this theory can be applied not only for modeling and analysis of complex systems, but also for solving a large class of nonconvex/nonsmooth/ discrete problems in both mathematical physics and global optimization. In this paper, we shall need only the following weak assumptions for the nonconvex function $W(\boldsymbol{x})$.

(A1) The nonconvex function $W(\boldsymbol{x})$ is twice continuously differentiable.

(A2) There exits a *geometrical operator*

$$\Lambda(\boldsymbol{x}) = \left\{ \frac{1}{2}\boldsymbol{x}^T B^k \boldsymbol{x} + \boldsymbol{b}_k^T \boldsymbol{x} \right\} :\ \mathbb{R}^n \to \mathbb{R}^m \tag{15}$$

and a strictly convex function $V : \mathbb{R}^m \to \mathbb{R}$ such that

$$W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x})), \tag{16}$$

where $B^k \in \mathbb{R}^{n\times n}$ and $\boldsymbol{b}_k \in \mathbb{R}^n$, $k = 1, \cdots, m$.

(A3) The critical points of problem $(\mathscr{P})$ are non-singular, i.e., if $\nabla \Pi(\bar{\boldsymbol{x}}) = 0$, then $\det(\nabla^2 \Pi(\bar{\boldsymbol{x}})) \neq 0$.

Actually, the assumption (A2) is the so-called *canonical transformation* introduced in [10]. The idea of this transformation was from Gao and Strang's original work [22] on nonconvex variational problems in large deformation theory, where the geometrical operator $\Lambda(u) = \frac{1}{2}(\nabla u)^T(\nabla u)$ is a Cauchy–Riemann metric tensor field, which is an objective measure of the deformation gradient $\varepsilon = \nabla u$, and $W(\nabla u) = V(\Lambda(u))$ is a stored strain energy. Using finite element discretization for the deformation field $u(\boldsymbol{x})$, the nonconvex variational problems in infinite-dimensional space can be reduced to the canonical global optimization problem $(\mathscr{P})$ (see [27, 35]). It is known in continuum physics that the stored energy W is usually a nonconvex function of the linear measure ∇u (which is not a strain measure), but $V(e)$ is convex in term of the objective measure $e = \Lambda(u)$. Therefore, by this quadratic objective operator $\Lambda(u)$, a *complementary gap function* was discovered by Gao and Strang in nonconvex variational analysis and the complementary variational principle was recovered in fully nonlinear equilibrium problems of mathematical physics.[4] They also proved that the nonnegative gap function can be used to identify global minimizer of the nonconvex problem. Seven years later, it was discovered that the negative gap function can be used to identify the largest local minimum and maximum. Therefore, the *triality theory* was first proposed in nonconvex mechanics [7, 8], and then generalized to global optimization [11]. This triality theory is composed of a canonical min-max duality and two pairs of double-min, double-max dualities, which reveals an intrinsic duality pattern in complex systems and has been used successfully for solving a wide class of challenging problems in nonconvex analysis and global optimization [10]. However, it was realized in 2003 [12, 13] that the double-min duality holds conditionally under "certain additional conditions". Recently, this problem is partly solved for a class of fourth-order polynomial optimization problems [24, 37].

The aim of this paper is to prove the triality theory for the general nonconvex global optimization problem $(\mathscr{P})$. In the following sections, we first provide a brief review on the canonical duality theory and the associated triality theory. We will show that by the canonical transformation, the nonconvex primal problem $(\mathscr{P})$ can be reformulated as a canonical dual problem without duality gap. Section 3 presents a strong triality theory for the case that the primal problem and its canonical dual have the same dimension, i.e., $n = m$. We then show in Sect. 4 that this theory holds weakly for the case $n \neq m$. The "certain additional conditions" for the double-min duality are

---

[4] In continuum physics, complementary variational principle means perfect duality since any duality gap will violate certain physical laws. The existence of a complementary variational principle was a well-known debate existing for several decades in large deformation theory (see [31]). This problem was partially solved by Gao and Strang's work, and solved completely in 1999 [9].

provided. Application is illustrated in Sect. 5. The original definition of Lagrangian, Lagrangian duality and its difference with the canonical duality are discussed in Sect. 6. The paper is ended with some conclusion remarks and challenging problems.

2 Canonical Duality, Triality, and Open Problem

Based on the canonical transformation, i.e., Assumption (A2), the general problem (1) can be reformulated in the following canonical form:

$$(\mathscr{P}): \quad \text{ext}\left\{ \Pi(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x})) + \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{f}\rangle \mid \boldsymbol{x} \in \mathbb{R}^n \right\}. \tag{17}$$

This problem arises extensively in many fields of engineering and sciences, including Euclidean distance geometry [5, 20], computational biology [6, 29, 45], numerical methods for solving a large class of nonconvex variational problems in mathematical physics [13, 27, 35], and much more. Let

$$\mathscr{V}_a = \{\boldsymbol{\xi} \in \mathbb{R}^m \mid \ \boldsymbol{\xi} = \Lambda(\boldsymbol{x}) \ \ \forall \boldsymbol{x} \in \mathbb{R}^n\},$$

$$\mathscr{V}_a^* = \{\boldsymbol{\varsigma} \in \mathbb{R}^m \mid \ \boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) \ \ \forall \boldsymbol{\xi} \in \mathscr{V}_a\}.$$

By (A1) and (A2) we know that $V : \mathscr{V}_a \to \mathbb{R}$ is also a twice continuously differentiable. Therefore, its Legendre conjugate $V^* : \mathscr{V}_a^* \to \mathbb{R}$ can be uniquely defined as

$$V^*(\boldsymbol{\varsigma}) = \text{sta}\,\{\langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle - V(\boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \mathscr{V}_a\}, \tag{18}$$

where $\langle *; *\rangle$ is an inner product in $\mathbb{R}^m$ and sta{ } stands for finding stationary value of the expression given in { }. It is easy to verify that the canonical duality relations

$$\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) \quad \Leftrightarrow \quad \boldsymbol{\xi} = \nabla V^*(\boldsymbol{\varsigma}) \quad \Leftrightarrow \quad V(\boldsymbol{\xi}) + V^*(\boldsymbol{\varsigma}) = \langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle \tag{19}$$

hold on $\mathscr{V}_a \times \mathscr{V}_a^*$.

Substituting $V(\Lambda(\boldsymbol{x})) = \langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma}\rangle - V^*(\boldsymbol{\varsigma})$, the primal function $\Pi(\boldsymbol{x})$ can be reformulated as the total complementary function [10]

$$\Xi(\boldsymbol{x}, \boldsymbol{\varsigma}) = \frac{1}{2}\langle \boldsymbol{x}, G(\boldsymbol{\varsigma})\boldsymbol{x}\rangle - V^*(\boldsymbol{\varsigma}) - \langle \boldsymbol{x}, F(\boldsymbol{\varsigma})\rangle, \tag{20}$$

where

$$G(\boldsymbol{\varsigma}) = A + \sum_{k=1}^{m} \varsigma_k B^k, \quad F(\boldsymbol{\varsigma}) = \boldsymbol{f} - \sum_{k=1}^{m} \varsigma_k \boldsymbol{b}_k.$$

For a fixed ς, the criticality condition $\nabla_{\boldsymbol{x}} \Xi(\boldsymbol{x}, \varsigma) = 0$ leads to the following canonical equilibrium equation

$$G(\varsigma)\boldsymbol{x} = F(\varsigma), \tag{21}$$

which can be solved analytically to obtain[5] $\boldsymbol{x} = [G(\varsigma)]^{-1} F(\varsigma)$ for all ς in the canonical dual feasible space $\mathscr{S}_a$ defined by

$$\mathscr{S}_a = \{\varsigma \in \mathscr{V}_a^* |\ F(\varsigma) \in \mathscr{C}_{ol}\,(G\,(\varsigma))\},$$

where $\mathscr{C}_{ol}\,(G\,(\varsigma))$ is a space spanned by the columns of $G\,(\varsigma)$. Therefore, substituting this solution into the total complementary function Ξ, the canonical dual problem can be formulated as

$$(\mathscr{P}^d): \quad \text{ext}\left\{\Pi^d(\varsigma) = -\frac{1}{2}\langle [G(\varsigma)]^{-1} F(\varsigma), F(\varsigma)\rangle - V^*(\varsigma) \mid \varsigma \in \mathscr{S}_a\right\}. \tag{22}$$

The following theorem was originally presented in general nonconvex systems [10].

Theorem 1 (Analytical Solution and Complementary-dual principle). *Problem ($\mathscr{P}^d$) is canonically dual to ($\mathscr{P}$) in the sense that if $\bar{\varsigma}$ is a critical point of ($\mathscr{P}^d$), then*

$$\boldsymbol{x} = [G(\varsigma)]^{-1} F(\varsigma) \tag{23}$$

is a critical point of ($\mathscr{P}$), the pair $(\bar{\boldsymbol{x}}, \bar{\varsigma})$ is a critical point of $\Xi(\boldsymbol{x}, \varsigma)$, and

$$\Pi(\bar{\boldsymbol{x}}) = \Xi(\bar{\boldsymbol{x}}, \bar{\varsigma}) = \Pi^d(\bar{\varsigma}). \tag{24}$$

This theorem shows that there is no duality gap between the primal problem ($\mathscr{P}$) and its canonical dual ($\mathscr{P}^d$). Actually, in $\Xi(\boldsymbol{x}, \varsigma)$ the first term

$$G_{ap}(\boldsymbol{x}, \varsigma) = \frac{1}{2}\langle \boldsymbol{x}, G(\varsigma)\boldsymbol{x}\rangle \tag{25}$$

is the complementary gap function, first introduced by Gao and Strang in 1989 [22]. They proved that if this gap function is positive, the critical point $\bar{\varsigma}$ is a global maximizer of Π^d and the associated $\bar{\boldsymbol{x}}(\bar{\varsigma})$ is a global minimizer of the primal problem ($\mathscr{P}$). By introducing the following notations

$$\mathscr{S}_a^+ = \{\varsigma \in \mathscr{S}_a \mid G\,(\varsigma) \succeq 0\}, \tag{26}$$

$$\mathscr{S}_a^- = \{\varsigma \in \mathscr{S}_a \mid G\,(\varsigma) \prec 0\}, \tag{27}$$

[5] In this paper G^{-1} should be understood as a generalized inverse if $\det G = 0$ [11].

where $\mathbf{G}(\varsigma) \succeq 0$ means that $\mathbf{G}(\varsigma)$ is positive semi-definite and $\mathbf{G}(\varsigma) \prec 0$ means that $\mathbf{G}(\varsigma)$ is negative definite, the Gao and Strang canonical min-max duality theory can be stated as

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathbb{R}^n} \Pi(\boldsymbol{x}) = \max_{\varsigma\in\mathscr{S}_a^+} \Pi^d(\varsigma) = \Pi^d(\bar{\varsigma}). \tag{28}$$

This general result has been used extensively in nonconvex analysis and mechanics [10, 44]. In 1996, it was discovered by Gao that if the gap function is negative in a neighborhood $\mathscr{X}_o \times \mathscr{S}_o \subset \mathbb{R}^n \times \mathscr{S}_a^-$ of $(\bar{\boldsymbol{x}}, \bar{\varsigma})$, then either the double-max duality relation

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x}\in\mathscr{X}_o} \Pi(\boldsymbol{x}) = \max_{\varsigma\in\mathscr{S}_o} \Pi^d(\varsigma) = \Pi^d(\bar{\varsigma}) \tag{29}$$

holds or the double-min duality relation

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathscr{X}_o} \Pi(\boldsymbol{x}) = \min_{\varsigma\in\mathscr{S}_o} \Pi^d(\varsigma) = \Pi^d(\bar{\varsigma}). \tag{30}$$

Therefore, the triality theorem was formed by these three pairs of dualities and has been used extensively in nonconvex mechanics [10, 17] and global optimization [3, 21, 34]. However, it was realized in 2003 [12, 13] that if the dimensions of the primal problem and its canonical dual are different, the double-min duality (30) needs "certain additional conditions". For the sake of mathematical rigor, the double-min duality was not included in the triality theory and these additional constraints were left as an open problem (see Remark 1 in [12], also Theorem 3 and its Remark in a review article by Gao [13]). By the facts that the double-max duality (29) is always true and the double-min duality plays a key role in real-life applications, it was still included in the triality theory in the either-or form in many applications for the purposes of perfection in esthesis and some other reasons in reality. In the following sections, we will show that the triality theorem holds strongly for the problems it was originally proposed. Also we will explain the reasons why the "certain additional conditions" in the double-min duality were ignored.

3 Strong Triality Theory

In the case $n = m$, the triality theorem holds strongly in the following form.

Theorem 2 (Tri-duality Theorem). *Suppose that $\bar{\varsigma}$ is a critical point of the canonical problem $(\mathscr{P}^d)$ and $\bar{\boldsymbol{x}} = [G(\bar{\varsigma})]^{-1} F(\bar{\varsigma})$.*

If $\bar{\varsigma} \in \mathscr{S}_a^+$, then $\bar{\varsigma}$ is a global maximizer of Problem $(\mathscr{P}^d)$ in $\mathscr{S}_a^+$ if and only if $\bar{\boldsymbol{x}}$ is a global minimizer of Problem $(\mathscr{P})$, i.e., the following canonical min-max duality statement holds:

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathbb{R}^n} \Pi(\boldsymbol{x}) \Longleftrightarrow \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{31}$$

If $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^-$, then there exists a neighborhood $\mathscr{X}_o \times \mathscr{S}_o \subset \mathbb{R}^n \times \mathscr{S}_a^-$ of $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$ such that we have either the double-min duality statement

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathscr{X}_o} \Pi(\boldsymbol{x}) \Longleftrightarrow \min_{\boldsymbol{\varsigma}\in\mathscr{S}_o} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}), \tag{32}$$

or the double-max duality statement

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x}\in\mathscr{X}_o} \Pi(\boldsymbol{x}) \Longleftrightarrow \max_{\boldsymbol{\varsigma}\in\mathscr{S}_o} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{33}$$

Proof. If $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$ is a critical point of the total complementary function $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$, then by Theorem 1, we have $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$, and

$$\nabla^2\Pi^d(\bar{\boldsymbol{\varsigma}}) = -(\nabla\Lambda(\bar{\boldsymbol{x}}))^T[G(\bar{\boldsymbol{\varsigma}})]^{-1}\nabla\Lambda(\bar{\boldsymbol{x}}) - \nabla^2V^*(\bar{\boldsymbol{\varsigma}}), \tag{34}$$

$$\nabla^2\Pi(\bar{\boldsymbol{x}}) = G(\bar{\boldsymbol{\varsigma}}) + \nabla\Lambda(\bar{\boldsymbol{x}})\nabla^2V(\Lambda(\bar{\boldsymbol{x}}))(\nabla\Lambda(\bar{\boldsymbol{x}}))^T. \tag{35}$$

By the assumption (A2) we know that $V(\boldsymbol{\xi})$ is strictly convex, then,

$$\nabla^2(V(\Lambda(\bar{\boldsymbol{x}}))) = (\nabla^2V^*(\bar{\boldsymbol{\varsigma}}))^{-1} \succ 0, \tag{36}$$

where $\bar{\boldsymbol{\xi}} = \Lambda(\bar{\boldsymbol{x}})$. Substituting (36) into (35), we obtain

$$\nabla^2\Pi(\bar{\boldsymbol{x}}) = G(\bar{\boldsymbol{\varsigma}}) + \nabla\Lambda(\bar{\boldsymbol{x}})(\nabla^2V^*(\bar{\boldsymbol{\varsigma}}))^{-1}(\nabla\Lambda(\bar{\boldsymbol{x}}))^T. \tag{37}$$

- Proof of the canonical min-max duality statement (31) (this proof is a finite-dimensional version of Gao and Strangs original proof of Theorem 2 in nonconvex analysis [22]).

Suppose that $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^+$ is a critical point. Since $\Pi^d(\boldsymbol{\varsigma})$ is concave on $\mathscr{S}_a^+$, the critical point $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^+$ must be a global maximizer of $\Pi^d(\boldsymbol{\varsigma})$ on $\mathscr{S}_a^+$.

On the other hand, if $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^+$, the gap function $G_{ap}(\boldsymbol{x}, \bar{\boldsymbol{\varsigma}}) = \frac{1}{2}\langle\boldsymbol{x}, G(\bar{\boldsymbol{\varsigma}})\boldsymbol{x}\rangle$ is convex in $\boldsymbol{x} \in \mathbb{R}^n$. By the convexity of $V : \mathscr{V}_a \to \mathbb{R}$, we have [22]

$$\begin{aligned}\Pi(\boldsymbol{x}) - \Pi(\bar{\boldsymbol{x}}) &\geq \langle\nabla V(\Lambda(\bar{\boldsymbol{x}})); \Lambda(\boldsymbol{x}) - \Lambda(\bar{\boldsymbol{x}})\rangle + \frac{1}{2}\langle\boldsymbol{x}, A\boldsymbol{x}\rangle - \frac{1}{2}\langle\bar{\boldsymbol{x}}, A\bar{\boldsymbol{x}}\rangle - \langle\boldsymbol{x} - \bar{\boldsymbol{x}}, \boldsymbol{f}\rangle \\ &= G_{ap}(\boldsymbol{x}, \bar{\boldsymbol{\varsigma}}) - G_{ap}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}}) - \langle\boldsymbol{x} - \bar{\boldsymbol{x}}, F(\bar{\boldsymbol{\varsigma}})\rangle \\ &\geq \langle\boldsymbol{x} - \bar{\boldsymbol{x}}, G(\bar{\boldsymbol{\varsigma}})\bar{\boldsymbol{x}} - F(\bar{\boldsymbol{\varsigma}})\rangle = 0 \;\; \forall\boldsymbol{x} \in \mathbb{R}^n.\end{aligned}$$

Thus, $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1}F(\bar{\boldsymbol{\varsigma}})$ is a global minimizer of problem $(\mathscr{P})$. Furthermore, $\bar{\boldsymbol{\varsigma}}$ is also a global maximizer of Problem $(\mathscr{P}^d)$ in $\mathscr{S}_a^+$ and the statement (31) holds by Theorem 1.

- Proof of the double-min duality statement (32).

Suppose that $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ and $\bar{\boldsymbol{\varsigma}}$ is a local minimizer of problem $(\mathcal{P}^d)$. Then, we have $\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}) \succeq 0$ and

$$-(\nabla \Lambda(\bar{\boldsymbol{x}}))^T [G(\bar{\boldsymbol{\varsigma}})]^{-1} \nabla \Lambda(\bar{\boldsymbol{x}}) \succeq \nabla^2 V^*(\bar{\boldsymbol{\varsigma}}) \succ 0.$$

Thus, $\nabla \Lambda(\bar{\boldsymbol{x}})$ is invertible, which leads to

$$-G(\bar{\boldsymbol{\varsigma}}) \preceq \nabla \Lambda(\bar{\boldsymbol{x}})(\nabla^2 V^*(\bar{\boldsymbol{\varsigma}}))^{-1}(\nabla \Lambda(\bar{\boldsymbol{x}}))^T. \quad (38)$$

Therefore, we have

$$\nabla^2 \Pi(\bar{\boldsymbol{x}}) = G(\bar{\boldsymbol{\varsigma}}) + \nabla \Lambda(\bar{\boldsymbol{x}})(\nabla^2 V^*(\bar{\boldsymbol{\varsigma}}))^{-1}(\nabla \Lambda(\bar{\boldsymbol{x}}))^T \succ 0. \quad (39)$$

By the assumption (A3), $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$ is also a local minimizer of problem $(\mathcal{P})$. The reversed statement can be proved in the similar way. Thus, (32) holds.

- Proof of the double-max duality statement (33).

Suppose that $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ and $\bar{\boldsymbol{\varsigma}}$ is a local maximizer of problem $(\mathcal{P}^d)$. Then, $\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}) \preceq 0$. By Theorem 1, $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$ is a critical point of problem $(\mathcal{P})$. Due to the assumption (A3), $\nabla^2 \Pi(\bar{\boldsymbol{x}})$ is invertible. By the well-known Sherman–Morrison–Woodbury identity [4], $\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}})$ is also invertible. Furthermore,

$$(\nabla^2 \Pi(\bar{\boldsymbol{x}}))^{-1} = G(\bar{\boldsymbol{\varsigma}})^{-1} + G(\bar{\boldsymbol{\varsigma}})^{-1} \nabla \Lambda(\bar{\boldsymbol{x}})(\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}))^{-1}(\nabla \Lambda(\bar{\boldsymbol{x}}))^T G(\bar{\boldsymbol{\varsigma}})^{-1} \prec 0.$$

Thus, $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$ is also a local maximizer of problem $(\mathcal{P})$. Similarly, we can prove the reversed statement. Therefore, the triality theorem holds strongly for the case $n = m$. □

Remark 1. The tri-duality theorem provides global extremum criteria for three types solutions of the nonconvex problem $(\mathcal{P})$: a global minimizer $\bar{\boldsymbol{x}}(\bar{\boldsymbol{\varsigma}})$ if $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^+$ and a pair of the largest-valued local extrema, i.e., $\bar{\boldsymbol{x}}(\bar{\boldsymbol{\varsigma}})$ is a global maximizer (resp. minimizer) if $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ is a local maximizer (reps. minimizer). This pair of largest local extrema plays a critical role in nonconvex mechanics and phase transitions.

Remark 2. The tri-duality theorem can also be used to identify saddle points of the primal problem, i.e., $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ is a saddle point of $\Pi(\boldsymbol{\varsigma})$ if and only if $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$ is a saddle point of $\Pi(\boldsymbol{x})$. By the facts that the saddle points are not stable and do not exist physically, these points are excluded from the triality theory.

The triality theory was first discovered in post-buckling analysis of a large deformed elastic beam model proposed by Gao in 1996 [7, 8], where the primal functional is a double-well potential of a two-dimensional displacement field, and its canonical dual is the so-called *pure complementary energy* defined on a two-dimensional stress field. Therefore, the triality theory was first proposed in its strong form, i.e., the tri-duality theorem.

4 Triality Theory for General Case

We now consider the general case $m \neq n$. Suppose that $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{\varsigma}}$ are the critical points of problem $(\mathscr{P})$ and $(\mathscr{P}^d)$, respectively, where $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1}F(\bar{\boldsymbol{\varsigma}})$ and $G(\bar{\boldsymbol{\varsigma}})$ is invertible. In this case, we also can show that

$$\nabla^2 \Pi(\bar{\boldsymbol{x}}) = G(\bar{\boldsymbol{\varsigma}}) + \nabla \Lambda(\bar{\boldsymbol{x}})(\nabla^2 V^*(\bar{\boldsymbol{\varsigma}}))^{-1}(\nabla \Lambda(\bar{\boldsymbol{x}}))^T, \tag{40}$$

and

$$\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}) = -(\nabla \Lambda(\bar{\boldsymbol{x}}))^T [G(\bar{\boldsymbol{\varsigma}})]^{-1} \nabla \Lambda(\bar{\boldsymbol{x}}) - \nabla^2 V^*(\bar{\boldsymbol{\varsigma}}). \tag{41}$$

Suppose that $m < n$. By the Sherman–Morison–Woodbury Theorem in [4] and the assumption (A3), we have

$$[\nabla^2 \Pi(\bar{\boldsymbol{x}})]^{-1} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} + [G(\bar{\boldsymbol{\varsigma}})]^{-1} \nabla \Lambda(\bar{\boldsymbol{x}})(\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}))^{-1}(\nabla \Lambda(\bar{\boldsymbol{x}}))^T [G(\bar{\boldsymbol{\varsigma}})]^{-1}. \tag{42}$$

This shows that $\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}})$ is invertible. Similarly, we can show that $\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}})$ is also invertible if $m > n$.

Lemma 1. *Suppose that $m < n$, the critical point $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^-$ is a local minimizer of problem $(\mathscr{P}^d)$. Then, $\nabla^2 \Pi(\bar{\boldsymbol{x}})$ has m positive eigenvalues and $n - m$ negative eigenvalues, i.e., there exists two matrices $P_\flat \in \mathbb{R}^{n \times m}$ and $P_\sharp \in \mathbb{R}^{n \times (n-m)}$ such that*

$$P_\flat^T \nabla^2 \Pi(\bar{\boldsymbol{x}}) P_\flat \succ 0 \text{ and } P_\sharp^T \nabla^2 \Pi(\bar{\boldsymbol{x}}) P_\sharp \prec 0. \tag{43}$$

Proof. Since $\nabla^2 V^*(\bar{\boldsymbol{\varsigma}}) \succ 0$, there exists a invertible matrix $R \in \mathbb{R}^{m \times m}$ such that $\nabla^2 V^*(\bar{\boldsymbol{\varsigma}}) = R^T R$. Thus, we have

$$-(\nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1})^T [G(\bar{\boldsymbol{\varsigma}})]^{-1} \nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1} - I_{m \times m} \succ 0. \tag{44}$$

Note that $G(\bar{\boldsymbol{\varsigma}}) \prec 0$ and $\nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1} (\nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1}))^T \succeq 0$. There exists a matrix T such that

$$T^T G(\bar{\boldsymbol{\varsigma}}) T = \text{Diag}(-\lambda_1, \cdots, -\lambda_n), \tag{45}$$

and

$$T^T \nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1} (\nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1}))^T T = \text{Diag}(a_1, \cdots, a_m, 0, \cdots, 0), \tag{46}$$

where $\lambda_k > 0,\ k = 1, \cdots, n$, and $a_k > 0, k = 1, \cdots, m$. According to the decomposition theory of singular matrices, we know that there exist orthogonal matrices $U \in \mathbb{R}^{n \times n}$ and $E \in \mathbb{R}^{m \times m}$ such that

$$T^T \nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1} = U \begin{pmatrix} \sqrt{a_1} & & \\ & \ddots & \\ & & \sqrt{a_m} \\ 0 & \cdots & 0 \\ & \cdots & \\ 0 & \cdots & 0 \end{pmatrix} E. \tag{47}$$

In light of (46), we know that $U = I_{n\times n}$. Then, we have

$$\begin{aligned}(R^{-1})^T \nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}) R^{-1} &= -(\nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1})^T [G(\bar{\boldsymbol{\varsigma}})]^{-1} \nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1} - I_{m\times m} \\ &= -(T^T \nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1})^T [T^T G(\bar{\boldsymbol{\varsigma}}) T]^{-1} T \nabla \Lambda(\bar{\boldsymbol{x}}) R^{-1} - I_{m\times m} \\ &= E^T \mathrm{Diag}\left(\frac{a_1}{\lambda_1} - 1, \cdots, \frac{a_m}{\lambda_m} - 1\right) E \succ 0.\end{aligned}$$

Thus, $a_k > \lambda_k,\ k = 1, \cdots, m$. It is easy to verify that

$$T^T \nabla^2 \Pi(\bar{\boldsymbol{x}}) T = \mathrm{Diag}(a_1 - \lambda_1, \cdots, a_m - \lambda_m, -\lambda_{m+1}, \cdots, -\lambda_n). \tag{48}$$

This shows that $\nabla^2 \Pi(\bar{\boldsymbol{x}})$ has m positive eigenvalues and $n - m$ negative eigenvalues. Therefore, the matrix $P_\flat$ can be obtained by collecting all the eigenvectors corresponding to the positive eigenvalues and $P_\sharp$ can be obtained by collecting all the eigenvectors corresponding to the negative eigenvalues. □

In a similar way, we can prove the following lemma.

Lemma 2. *Suppose that $m > n$ and the critical point $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$ is a local minimizer of Problem ($\mathscr{P}$), where $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^-$. Then, $\nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}})$ has n positive eigenvalues and $m - n$ negative eigenvalues, i.e., there exists two matrices $Q_\flat \in \mathbb{R}^{m\times n}$ and $Q_\sharp \in \mathbb{R}^{m\times(m-n)}$ such that*

$$Q_\flat^T \nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}) Q_\flat \succ 0 \text{ and } Q_\sharp^T \nabla^2 \Pi^d(\bar{\boldsymbol{\varsigma}}) Q_\sharp \prec 0. \tag{49}$$

Let the m column vectors of $P_\flat$ be $\boldsymbol{p}_1^\flat, \cdots, \boldsymbol{p}_m^\flat$ and the n column vectors of $Q_\flat$ be $\boldsymbol{q}_1^\flat, \cdots, \boldsymbol{q}_n^\flat$, respectively. Clearly, $\boldsymbol{p}_1^\flat, \cdots, \boldsymbol{p}_m^\flat$ and $\boldsymbol{q}_1^\flat, \cdots, \boldsymbol{q}_n^\flat$ are two sets of linearly independent vectors, respectively. By introducing two subspaces

$$\mathscr{X}_\flat = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \bar{\boldsymbol{x}} + \theta_1 \boldsymbol{p}_1^\flat + \cdots + \theta_m \boldsymbol{p}_m^\flat,\ \theta_i \in \mathbb{R},\ i = 1, \cdots, m\}, \tag{50}$$

$$\mathscr{S}_\flat = \{\boldsymbol{\varsigma} \in \mathbb{R}^m \mid \boldsymbol{\varsigma} = \bar{\boldsymbol{\varsigma}} + \vartheta_1 \boldsymbol{q}_1^\flat + \cdots + \vartheta_n \boldsymbol{q}_n^\flat,\ \vartheta_i \in \mathbb{R},\ i = 1, \cdots, n\}, \tag{51}$$

the triality theory holds for general case in the following refined form.

Theorem 3. (Triality Theorem).

Suppose that $\bar{\boldsymbol{\varsigma}}$ is a critical point of problem ($\mathscr{P}^d$) and $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$.

If $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^+$, then the canonical min-max duality holds in the strong form of

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathbb{R}^n} \Pi(\boldsymbol{x}) \Leftrightarrow \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{52}$$

If $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^-$, then there exists a neighborhood $\mathscr{X}_o \times \mathscr{S}_o \subset \mathbb{R}^n \times \mathscr{S}_a^-$ of $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$ such that the double-max duality holds in the strong form of

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x}\in\mathscr{X}_o} \Pi(\boldsymbol{x}) \Leftrightarrow \max_{\boldsymbol{\varsigma}\in\mathscr{S}_o} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{53}$$

However, the double-min duality statement holds conditionally in the following symmetrical forms.

1. *If $m < n$ and $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^-$ is a local minimizer of $\Pi^d(\boldsymbol{\varsigma})$, then $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1}F(\bar{\boldsymbol{\varsigma}})$ is a saddle point of $\Pi(\boldsymbol{x})$ and the double-min duality holds weakly on $\mathscr{X}_o \cap \mathscr{X}_\flat \times \mathscr{S}_o$, i.e.,*

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathscr{X}_o\cap\mathscr{X}_\flat} \Pi(\boldsymbol{x}) = \min_{\boldsymbol{\varsigma}\in\mathscr{S}_o} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{54}$$

2. *If $m > n$ and $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1}F(\bar{\boldsymbol{\varsigma}})$ is a local minimizer of $\Pi(\boldsymbol{x})$, then $\bar{\boldsymbol{\varsigma}}$ is a saddle point of $\Pi^d(\boldsymbol{\varsigma})$ and the double-min duality holds weakly on $\mathscr{X}_o \times \mathscr{S}_o \cap \mathscr{S}_\flat$, i.e.,*

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathscr{X}_o} \Pi(\boldsymbol{x}) = \min_{\boldsymbol{\varsigma}\in\mathscr{S}_o\cap\mathscr{S}_\flat} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{55}$$

Proof. The proof of min-max duality statement (52) and the double-max duality statement (53) are the same to the proof of (31) and (33). We only need to prove (54) and (55).

Suppose that $m < n$ and $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^-$ is not only a local minimizer, but also a critical point of problem $(\mathscr{P}^d)$. By Lemma 1 we know that $\nabla^2\Pi(\bar{\boldsymbol{x}})$ has both positive and negative eigenvalues. Thus, $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1}F(\bar{\boldsymbol{\varsigma}})$ is a saddle point of problem $(\mathscr{P})$. We let

$$\varphi(t_1, \cdots, t_m) = \Pi(\bar{\boldsymbol{x}} + t_1\boldsymbol{p}_1^\flat + \cdots + t_m\boldsymbol{p}_m^\flat), \tag{56}$$

where $\boldsymbol{p}_1^\flat, \cdots, \boldsymbol{p}_m^\flat$ are the column vectors of $P_\flat$ defined in Lemma 1. By direct verification, we have

$$\nabla\varphi(0, \cdots, 0) = (\nabla\Pi(\bar{\boldsymbol{x}}))^T P_\flat = 0 \tag{57}$$

and

$$\nabla^2\varphi(0, \cdots, 0) = P_\flat^T \nabla^2\Pi(\bar{\boldsymbol{x}}) P_\flat \succ 0. \tag{58}$$

Thus, $(0, \cdots, 0)$ is a local minimizer of $\varphi(t_1, \cdots, t_m)$. Hence, the Eq. (54) holds. The statement (55) can be proved in the similar way. □

Remark 3. (NP-hard Problems and Perturbation). The canonical min-max duality (52) shows that the nonconvex minimization problem is equivalent to a concave maximization dual problem over a closed convex set $\mathcal{S}_a^+$. If $\Pi^d(\boldsymbol{\varsigma})$ has at least one critical point in $\mathcal{S}_a^+$, the global minimizer of $\Pi(\boldsymbol{x})$ can be easily obtained by the canonical duality theory. However, if $\Pi^d(\boldsymbol{\varsigma})$ has no critical points in $\mathcal{S}_a^+$, to find global minimizer for nonconvex function $\Pi(\boldsymbol{x})$ could be very difficult. If the vector $\boldsymbol{f} = \boldsymbol{0} \in \mathbb{R}^n$, the problem $(\mathcal{P})$ is homogenous. Moreover, if $\boldsymbol{b}_k = \boldsymbol{0} \in \mathbb{R}^n \ \forall k = 1, \cdots, m$, then the geometrical operator $\Lambda(\boldsymbol{x})$ is a pure quadratic measure (i.e., an objective measure in certain space). In this case, the vector $F(\boldsymbol{\varsigma}) = \boldsymbol{0}$, the set $\mathcal{S}_a^+$ is empty, and the canonical dual function $\Pi^d(\boldsymbol{\varsigma}) = -V^*(\boldsymbol{\varsigma})$ is concave, which has only a unique maximizer $\bar{\boldsymbol{\varsigma}}$. By the double-max duality we know that the corresponding primal solution $\bar{\boldsymbol{x}} = \boldsymbol{0}$ is a local maximizer if $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$. From the point view of systems theory, the pure quadratic operator $\Lambda(\boldsymbol{x})$ means that the system possesses certain symmetry. If there is no input ($\boldsymbol{f} = \boldsymbol{0}$), the primal function $\Pi(\boldsymbol{x})$ could have multiple global minimizers. It was indicated in [14] that a nonconvex minimization problem could be NP-hard if its canonical dual has no KKT (or critical) point in $\mathcal{S}_a^+$. In order to solve this type problems, several perturbation methods have been suggested in [19, 34, 43]. It is shown very recently that by the canonical duality theory, a class of NP-hard box/integer constrained programming problems are equivalent to unconstrained canonical dual problems in continuous space, which can be solved via deterministic methods [23].

Dual to $\mathcal{X}_\flat$ and $\mathcal{S}_\flat$, we can let

$$\mathcal{X}_\sharp = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \bar{\boldsymbol{x}} + \theta_1 \boldsymbol{p}_1^\sharp + \cdots + \theta_{n-m} \boldsymbol{p}_{n-m}^\sharp, \ \theta_i \in \mathbb{R}, \ i = 1, \cdots, n-m\},$$
$$\mathcal{S}_\sharp = \{\boldsymbol{\varsigma} \in \mathbb{R}^m \mid \boldsymbol{\varsigma} = \bar{\boldsymbol{\varsigma}} + \vartheta_1 \boldsymbol{q}_1^\sharp + \cdots + \vartheta_{m-n} \boldsymbol{q}_{m-n}^\sharp, \ \vartheta_i \in \mathbb{R}, \ i = 1, \cdots, m-n\},$$

where $\{\boldsymbol{p}_i^\sharp\}$ and $\{\boldsymbol{q}_i^\sharp\}$ are column vectors of $P_\sharp$ and $Q_\sharp$, respectively. Then, complementary to the weak double-min duality statements (54) and (55), we have the following weak saddle min-max duality theorem.

Theorem 4. (Weak Saddle Duality Theorem).
Suppose that $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ is a critical point of problem $(\mathcal{P}^d)$, the vector $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$, and $\mathcal{X}_o \times \mathcal{S}_o \subset \mathbb{R}^n \times \mathcal{S}_a^-$ is a neighborhood of $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$.

1. *If $m < n$ and $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ is a local minimizer of $\Pi^d(\boldsymbol{\varsigma})$, then $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$ is a saddle point of $\Pi(\boldsymbol{x})$ and the saddle max-min duality holds weakly on $\mathcal{X}_o \cap \mathcal{X}_\sharp \times \mathcal{S}_o$, i.e.,*

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x} \in \mathcal{X}_o \cap \mathcal{X}_\sharp} \Pi(\boldsymbol{x}) = \min_{\boldsymbol{\varsigma} \in \mathcal{S}_o} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{59}$$

2. *If $m > n$ and $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1} F(\bar{\boldsymbol{\varsigma}})$ is a local minimizer of $\Pi(\boldsymbol{x})$, then $\bar{\boldsymbol{\varsigma}}$ is a saddle point of $\Pi^d(\boldsymbol{\varsigma})$ and the saddle min-max duality holds weakly on $\mathcal{X}_o \times \mathcal{S}_o \cap \mathcal{S}_\sharp$, i.e.,*

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x} \in \mathcal{X}_o} \Pi(\boldsymbol{x}) = \max_{\boldsymbol{\varsigma} \in \mathcal{S}_o \cap \mathcal{S}_\sharp} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{60}$$

Remark 4. Theorem 3 shows that both the canonical min-max and double-max duality statements hold strongly for general cases; the double-min duality holds strongly for n = m but weakly for $n \neq m$ in a symmetrical form. The "certain additional conditions" are simply the intersection $\mathscr{X}_o \bigcap \mathscr{X}_\flat$ for $n > m$ and $\mathscr{S}_o \bigcap \mathscr{S}_\flat$ for $n < m$. Therefore, the open problem on the double-min duality left in 2003 [12, 13] is now solved! While from Theorem 4 we know that if $G(\bar{\boldsymbol{\varsigma}}) \prec 0$ and $n \geq m$, the solution $\bar{\boldsymbol{x}}(\bar{\boldsymbol{\varsigma}})$ could be a saddle point. Mathematically speaking, nonstable critical points do not produce any computational difficulties in numerical optimization. Also, in real-life problems the saddle point is not considered as a phase state and does not physically exist. These are the part of reasons why the saddle point in $\mathscr{S}_a^-$ is ignored by the triality theory.

The triality theory has been challenged recently by a large number of counterexamples in a series of more than seven papers (see [39, 41] and references cited therein). It was written in [39] that "Because our counterexamples are very simple, using quadratic functions defined on whole Hilbert (even finite dimensional) spaces, it is difficult to reinforce the hypotheses of the above mentioned results in order to keep the same conclusions and not obtain trivialities." It turns out that in addition to many conceptual mistakes (see Sect. 6), most of these counterexamples simply discuss the saddle points in $\mathscr{S}_a^-$ for the case $n \neq m$. In fact, these counterexamples address the same type of open problem for the double-min duality left unaddressed in [12, 13].[6] Indeed, by Theorem 3 we know that the double-min duality holds conditionally when $n \neq m$. Based on Theorems 2 and 3, we know that the saddle points could exist in $\mathscr{S}_a^-$ even if $n = m$; While by Theorem 4 one can easily construct many other *V-Z type counterexamples* which are physically useless.

5 Application

Let us consider the following quadratic-log optimization problem:

$$(\mathscr{P}): \ \text{ext} \left\{ \Pi(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T A \boldsymbol{x} - \sum_{k=1}^{m} \log(\frac{1}{2}\boldsymbol{x}^T B^k \boldsymbol{x} + d^k) - \boldsymbol{x}^T \boldsymbol{f} \ | \ \boldsymbol{x} \in \mathbb{R}^n \right\}, \tag{61}$$

where A is a positive definite matrix, $B^k, k = 1, \cdots, m$ are positive semi-definite matrices and $d^k > 0, k = 1, \cdots, m$. In this case, its canonical dual problem can be expressed as

$$(\mathscr{P}^d): \ \text{ext} \left\{ \Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\boldsymbol{f}^T G(\boldsymbol{\varsigma})^{-1} \boldsymbol{f} + \sum_{k=1}^{m} (d^k \varsigma_k + 1 + \log(-\varsigma_k)) \mid \boldsymbol{\varsigma} \in \mathscr{S}_a \right\}, \tag{62}$$

where $\mathscr{S}_a = \left\{ \boldsymbol{\varsigma} = \{\varsigma_i\} \in \mathbb{R}^m \mid -\frac{1}{d^k} \leq \varsigma_k < 0, k = 1, \cdots, m \right\}$.

[6] It is interesting to note that the references [12, 13] never been cited in any one of this set of papers.

Let $n = 2, \quad m = 1$, and

$$A = \begin{pmatrix} 1 & \\ & 2 \end{pmatrix}, \; B^1 = \begin{pmatrix} 5 & \\ & 4 \end{pmatrix}, \; \boldsymbol{f} = \begin{pmatrix} 0.5 \\ 0.1 \end{pmatrix} \text{ and } d^1 = 1.$$

In this case, we have $\mathscr{S}_a = \{\varsigma \in \mathbb{R} \mid -1 \le \varsigma < 0\}$ and

$$\Pi^d(\varsigma) = -\frac{1}{2}\left(\frac{0.5^2}{1+5\varsigma} + \frac{0.1^2}{2+4\varsigma} + \varsigma + 1 + \log(-\varsigma)\right).$$

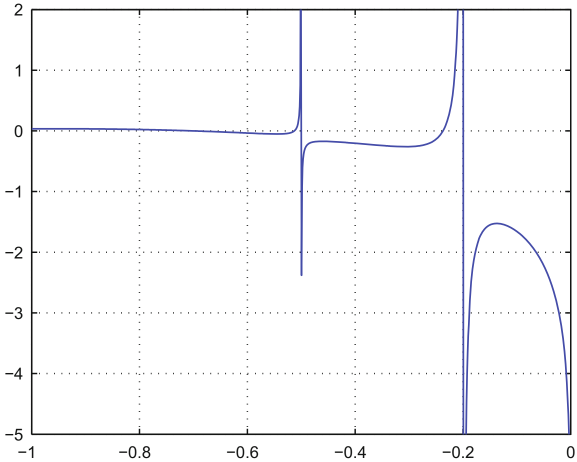


Fig. 1 Grapy of $\Pi^d(\varsigma)$

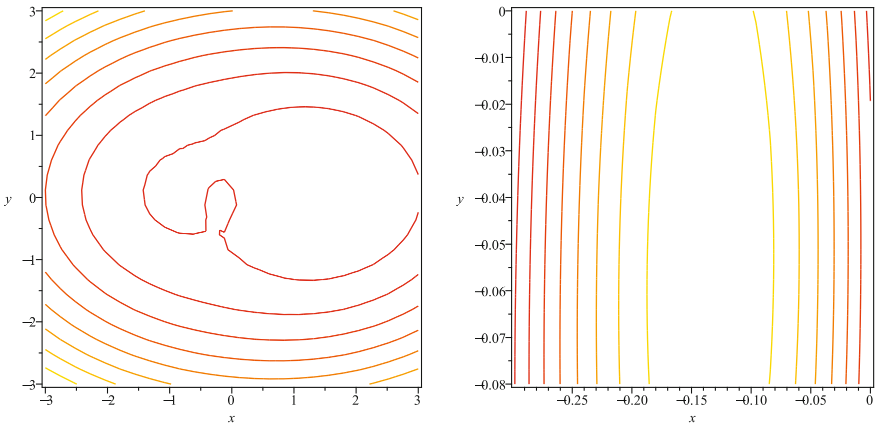


Fig. 2 Contours of $\Pi(\boldsymbol{x})$: **a** global minimizer of $\Pi(x)$ and **b** local maximizer of $\Pi(x)$

From Fig. 1 we can see that $\Pi^d(\varsigma)$ has one critical point $\bar{\varsigma}_1 = -0.13696432$ in $\mathscr{S}_a^+ = (-0.2, 0)$ and two critical points $\bar{\varsigma}_2 = -0.54470504$ and $\bar{\varsigma}_3 = -0.95209751$ in $\mathscr{S}_a^- = (-1, -0.5)$. Thus, by the triality theorem we know that $\bar{\boldsymbol{x}}_1 = (A + \bar{\varsigma}_1 B^1)^{-1}\boldsymbol{f} = [1.58640312, 0.06886375]^T$ is the global minimizer to the problem $(\mathscr{P})$. Since $\bar{\varsigma}_2 \in \mathscr{S}_a^-$ is a local minimizer, $\bar{\varsigma}_3 \in \mathscr{S}_a^-$ is a local maximizer to the problem $(\mathscr{P}^d)$, and $m = 1 < n = 2$, by Theorem 3 we know that $\bar{\boldsymbol{x}}_2 = (A + \bar{\varsigma}_2 B^1)^{-1}$ $\boldsymbol{f} = [-0.2901031, -0.5592211]^T$ is a saddle point of $\Pi(\boldsymbol{x})$, while $\bar{\boldsymbol{x}}_3 = (A + \bar{\varsigma}_3 B^1)^{-1}\boldsymbol{f} = [-0.13296148, -0.0552978]^T$ is a local maximizer to the problem $(\mathscr{P})$ (see Fig. 2). More applications can be found in [25].

6 Some Fundamental Concepts in Canonical Systems

Global optimization problem in mathematics is usually formulated in the following general form:

$$\min\{f(\boldsymbol{x}) |\ \boldsymbol{x} \in \mathscr{X} \subset \mathbb{R}^n\},$$

where the real-valued function $f(\boldsymbol{x})$ is simply assumed to be nonconvex (or Lipschitz, differentiable, etc.) on its feasible space $\mathscr{X} \in \mathbb{R}^n$, in which certain constraints are given. It is known that this problem could have a large number of local extrema and to identify global optima is a main challenging task in global optimization. If there is no detailed information available for the given function $f(\boldsymbol{x})$, it is difficult (may be impossible) to have a general theory and method for solving this general problem effectively. Also, to find the largest local extrema is fundamentally important in many real-life applications.[7]

Mathematics and physics (mechanics) have been complementary partners since Newton's time. It is known that the calculus of variation and mathematical optimization were originally developed from Euler–Lagrange mechanics. Also, the modern mathematical theory of convex analysis was started from J.J. Moreau's pioneering work in contact mechanics [32]. However, as V.I. Arnold pointed out [1]: "In the middle of the twentieth century it was attempted to divide physics and mathematics. The consequences turned out to be catastrophic." For example, in mathematical physics, the objectivity is directly related to some fundamental concepts and principles, such as geometrical nonlinearity, constitutive laws, and work-conjugate principle, etc. A function(al) can be called objective or free energy only if certain intrinsic constraints (physical laws) are satisfied (see [17]). Unfortunately, the objective function in mathematical optimization has been misused with other concepts such as cost function, energy function, and energy functional,[8] which leads to some conceptual mistakes. This section will discuss some important issues in classical Lagrangian mechanics/duality, mathematical optimization, and general systems theory.

---

[7] It should be emphasized here that to find the largest local maximum of $f(\boldsymbol{x})$ is not simply equivalent to solve the problem $\min\{-f(\boldsymbol{x}) |\ \boldsymbol{x} \in \mathscr{X}\}$.

[8] See the web page at http://en.wikipedia.org/wiki/Mathematical_optimization.

6.1 Canonical Systems

According to E. Tonti [40], in virtually every physical system there exists at least three types of variables:

(1) the configuration variable $\boldsymbol{x} \in \mathscr{X}$, which describes the state or output of the system, such as the Lagrangian generalized coordinates (or displacements) in analytical mechanics [30], decision variable in game theory, etc.

(2) the source variable $\boldsymbol{x}^* = \boldsymbol{f} \in \mathscr{X}^*$, which represents the input of the system, such as the external force in mechanics and charge density in theory of electrical field, etc.

(3) a pair of internal (or intermediate) variables $(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^*) \in \mathscr{E} \times \mathscr{E}^*$, which describes certain interior (constitutive) properties of the system, such as strain and stress in elasticity, velocity and momentum in dynamics, etc.

By the facts that the constitutive laws should be objective (coordinates-free) and physical variables appear always in one-to-one pairs (i.e., the Hill work-conjugacy principle in continuum mechanics [10]), it is reasonable to assume that for a given natural system, there exists a certain objective measure $\boldsymbol{\varepsilon} = \bar{\Lambda}(\boldsymbol{x}) : \mathscr{X}_a \subset \mathscr{X} \to \mathscr{E}_a \subset \mathscr{E}$ and a *stored energy* $\bar{W} : \mathscr{E}_a \to \mathbb{R}$ such that the constitutive duality relation $\boldsymbol{\varepsilon}^* = \nabla \bar{W}(\boldsymbol{\varepsilon}) : \mathscr{E}_a \to \mathscr{E}_a^* \subset \mathscr{E}^*$ is canonical (i.e., one-to-one on $\mathscr{E}_a \times \mathscr{E}_a^*$). Such a system is the so-called *canonical system* and is denoted as $\mathbb{S}_a = \{\langle \mathscr{X}_a, \mathscr{X}_a^* \rangle, \langle \mathscr{E}_a; \mathscr{E}_a^* \rangle; \bar{\Lambda}, C\}$ (see Chap. 4, [10]), where $C = \nabla \bar{W} : \mathscr{E}_a \to \mathscr{E}_a^*$ represents the constitutive mapping, $\langle *, * \rangle$ and $\langle *; * \rangle$ denote the bilinear forms on $\mathscr{X} \times \mathscr{X}^*$ and $\mathscr{E} \times \mathscr{E}^*$, respectively. The system is called *geometrically nonlinear* (resp. linear) if the geometrical operator $\bar{\Lambda}$ is nonlinear (resp. linear); the system is called *physically (or constitutively) nonlinear* (resp. linear) if the constitutive operator C is nonlinear (resp. linear); the system is called *fully nonlinear* (resp. linear) if it is both geometrically and physically nonlinear (resp. linear).

The most simple geometrically linear system is controlled by the quadratic function $\Pi(\boldsymbol{x}) = \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x} \rangle - \langle \boldsymbol{x}, \boldsymbol{f} \rangle$, where $A \in \mathbb{R}^{n \times n}$ is a symmetrical matrix.[9] If A is positive (semi) definite, by Cholesky decomposition we know that there exists a matrix $D : \mathbb{R}^n \to \mathbb{R}^m$ such that $A = D^T D$. Therefore, we have $\frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x} \rangle = \frac{1}{2}\langle D\boldsymbol{x}; D\boldsymbol{x} \rangle = T(D\boldsymbol{x})$ and $T(\boldsymbol{y})$ is an objective function of $\boldsymbol{y} = D\boldsymbol{x} \in \mathbb{R}^m$. By the fact that any symmetrical matrix can be written in difference of two positive definite matrices, it turns out that any given quadratic function can be written in the so-called d.c. (difference of convex functions) form.

6.2 Geometrically Linear Systems and Lagrangian Duality

In fact, for geometrically linear static systems, both the input and the configuration variables are time independent. In this case, the convex objective function $T(\boldsymbol{y})$ is

[9] The skew symmetric matrix $A_s = \frac{1}{2}(A - A^T)$ does not store energy since $\boldsymbol{x}^T A_s \boldsymbol{x} \equiv 0$.

the so-called internal (or stored) energy and $U(\boldsymbol{x})$ is the external potential, which should be linear $U(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f} \rangle$ such that its derivative $\nabla U(\boldsymbol{x}) = \boldsymbol{f}$ is a given source of the system. Therefore, the Lagrangian form $\Pi(\boldsymbol{x})$ represents the *total potential* of the system, which is convex on $\mathscr{X}_k$ and its mixed form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ is a saddle function on $\mathscr{X}_a \times \mathscr{Y}_a^*$. Therefore, the traditional saddle Lagrangian duality theory links the convex primal problem $\min\{\Pi(\boldsymbol{x}) | \ \boldsymbol{x} \in \mathscr{X}_k\}$ to a unique dual problem

$$\max \left\{ \Pi^*(\boldsymbol{y}^*) = -T^*(\boldsymbol{y}^*) | \ \ \boldsymbol{y}^* \in \mathscr{Y}_s^* \right\}, \tag{63}$$

where $\mathscr{Y}_s^* = \{\boldsymbol{y}^* \in \mathscr{Y}_a^* | \ D^*\boldsymbol{y}^* = \boldsymbol{f} \in \mathscr{X}_a^* \subset \mathbb{R}^n\}$ is the so-called *statically admissible space*. The objectivity of this dual problem is guaranteed by the objectivity of $T(\boldsymbol{y})$. By introducing a Lagrange multiplier $\boldsymbol{x}$, which must be a solution to the primal problem (see Lagrange multiplier's law in Sect. 1.5 [10]), to relax the equilibrium constraint $D^*\boldsymbol{y}^* = \boldsymbol{f}$ in $\mathscr{Y}_s^*$, the Lagrangian is exactly the mixed form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ and the one-to-one Lagrangian saddle min-max duality

$$\min_{\boldsymbol{x} \in \mathscr{X}_k} \Pi(\boldsymbol{x}) = \min_{\boldsymbol{x} \in \mathscr{X}_a} \max_{\boldsymbol{y}^* \in \mathscr{Y}_a^*} L(\boldsymbol{x}, \boldsymbol{y}^*) = \max_{\boldsymbol{y}^* \in \mathscr{Y}_a^*} \min_{\boldsymbol{x} \in \mathscr{X}_a} L(\boldsymbol{x}, \boldsymbol{y}^*) = \max_{\boldsymbol{y}^* \in \mathscr{Y}_s^*} \Pi^*(\boldsymbol{y}^*)$$

is called the *mono-duality* in canonical systems theory [10]. In mathematical economics, where the objective function $T(D\boldsymbol{x})$ is corresponding to the revenue, denoted by $R(\boldsymbol{x})$, and the potential $U(\boldsymbol{x})$ is the *cost function,* denoted by $C(\boldsymbol{x})$, then $\Pi(\boldsymbol{x}) = R(\boldsymbol{x}) - C(\boldsymbol{x})$ is the so-called total profit. For geometrically linear static problems, the cost function $C(\boldsymbol{x})$ is usually linear, while the revenue $R(\boldsymbol{x})$ is a concave objective function of certain measure (norm) of $\boldsymbol{x}$ in order to have maximum total profit $\Pi(\boldsymbol{x})$.

In geometrically linear dynamical systems, the convex function $T(\boldsymbol{y})$ is the kinetic energy and $U(\boldsymbol{x})$ represents the total potential of the system. In this case, the Lagrangian form $\Pi(\boldsymbol{x}) = T(D\boldsymbol{x}) - U(\boldsymbol{x})$ is the so-called *total action*, which is a d.c. (difference of convex) function. Since the mixed Lagrangian form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ is no longer a saddle function, the well-known Hamiltonian

$$H(\boldsymbol{x}, \boldsymbol{y}^*) = \langle D\boldsymbol{x}; \boldsymbol{y}^* \rangle - L(\boldsymbol{x}, \boldsymbol{y}^*) = T^*(\boldsymbol{y}^*) + U(\boldsymbol{x})$$

was introduced, which is convex and has been extensively used in dynamical systems. The criticality condition leads to the well-known *canonical Hamiltonian equations*

$$D\boldsymbol{x} = \nabla_{\boldsymbol{y}^*} H(\boldsymbol{x}, \boldsymbol{y}^*), \quad D^*\boldsymbol{y}^* = \nabla_{\boldsymbol{x}} H(\boldsymbol{x}, \boldsymbol{y}^*). \tag{64}$$

Actually, although the Lagrangian is not a saddle function in convex Hamiltonian systems, it is a so-called *super-critical function* [10], and if the total potential $U(\boldsymbol{x})$ is strictly convex on $\mathscr{X}_a \subset \mathbb{R}^n$ such that its Legendre conjugate $U^*(\boldsymbol{x}^*)$ can be uniquely defined on $\mathscr{X}_a^*$, then the canonical dual action of $\Pi(\boldsymbol{x})$ can still be defined by

$$\Pi^*(\boldsymbol{y}^*) = \max\{L(\boldsymbol{x}, \boldsymbol{y}^*) | \ \ \boldsymbol{x} \in \mathscr{X}_a\} = U^*(D^*\boldsymbol{y}^*) - T^*(\boldsymbol{y}^*)$$

on $\mathscr{Y}_s^* = \{\boldsymbol{y}^* \in \mathscr{Y}_a^* |\ D^*\boldsymbol{y}^* \in \mathscr{X}_a^*\}$, which is also a d.c. function. Therefore, instead of mono-duality in static systems, the convex Hamiltonian system is controlled by the so-called *bi-duality theory*.

Theorem 5. (Bi-Duality Theorem). *If $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}^*)$ is a critical point of the Lagrangian $L(\boldsymbol{x}, \boldsymbol{y}^*)$, then $\bar{\boldsymbol{x}}$ is a critical point of $\Pi(\boldsymbol{x})$, $\bar{\boldsymbol{y}}^*$ is a critical point of $\Pi^*(\boldsymbol{y}^*)$ and $\Pi(\bar{\boldsymbol{x}}) = L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}^*) = \Pi^*(\bar{\boldsymbol{y}}^*)$. Moreover, if $n = m$, we have either*

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x}\in\mathscr{X}_k} \Pi(\boldsymbol{x}) \quad\Leftrightarrow\quad \max_{\boldsymbol{y}^*\in\mathscr{Y}_s^*} \Pi^*(\boldsymbol{y}^*) = \Pi^*(\bar{\boldsymbol{y}}^*) \tag{65}$$

or

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathscr{X}_k} \Pi(\boldsymbol{x}) \quad\Leftrightarrow\quad \min_{\boldsymbol{y}^*\in\mathscr{Y}_s^*} \Pi^*(\boldsymbol{y}^*) = \Pi^*(\bar{\boldsymbol{y}}^*). \tag{66}$$

This bi-duality is actually a special case of the triality theory in geometrically linear systems, which was originally presented in Chap. 2 [10] for one-dimensional dynamical systems with a simple proof. This bi-duality reveals a stable periodical property in convex Hamiltonian systems.

6.3 Geometrically Nonlinear Systems and Canonical Duality

Problems in geometrically nonlinear systems are usually nonconvex. Due to the fact that the geometrically linear operator $D : \mathbb{R}^n \to \mathbb{R}^m$ cannot change the convexity of the objective function, if $W(\boldsymbol{x})$ is nonconvex and $W(\boldsymbol{x}) = T(D\boldsymbol{x})$, the function $T(\boldsymbol{y})$ is still nonconvex and its Legendre conjugate $T^*(\boldsymbol{y}^*)$ cannot be uniquely defined [36]. It turns out that traditional Lagrangian duality theory cannot be applied directly in this case. Although the Fenchel conjugate $T^\sharp(\boldsymbol{y}^*) = \sup\{\langle \boldsymbol{y}; \boldsymbol{y}^*\rangle - T(\boldsymbol{y}) |\ \boldsymbol{y} \in \mathscr{Y}_a\}$ can be uniquely defined, the function

$$\text{Ł}(\boldsymbol{x}, \boldsymbol{y}^*) = \langle D\boldsymbol{x}; \boldsymbol{y}^*\rangle - T^\sharp(\boldsymbol{y}^*) - U(\boldsymbol{x}) \tag{67}$$

is not the traditional Lagrangian form and the associate saddle min-max duality theory will produce the so-called duality gap in nonconvex optimization.

Actually, in terms of $U(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f}\rangle - \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x}\rangle$, the total complementary function $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ defined by (20) can be written as

$$\Xi(\boldsymbol{x}, \boldsymbol{\varsigma}) = \langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma}\rangle - V^*(\boldsymbol{\varsigma}) - U(\boldsymbol{x}). \tag{68}$$

Comparing this $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ with either Ł$(\boldsymbol{x}, \boldsymbol{y}^*)$ or the mixed Lagrangian form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ we can see that the fundamental difference between the canonical duality theory and other methods is the canonical transformation $W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x}))$ instead of the linear transformation $W(\boldsymbol{x}) = T(D\boldsymbol{x})$ used in many other duality theories, including the Fenchel–Moreau–Rockafellar duality. In real applications, if the quadratic function

$U(\boldsymbol{x})$ is nonconvex, the mixed Lagrangian form $L(\boldsymbol{x}, \boldsymbol{y}^*)$ is nonconvex in $\boldsymbol{x}$ since D is linear. However, the total complementary function $\Xi(*, \boldsymbol{\varsigma}) : \mathscr{X} \subset \mathbb{R}^n \to \mathbb{R}$ is always convex for $\boldsymbol{\varsigma} \in \mathscr{S}_a^+$ and concave for $\boldsymbol{\varsigma} \in \mathscr{S}_a^-$ due to the geometrically nonlinear operator $\Lambda(\boldsymbol{x})$ and its canonical dual variable $\boldsymbol{\varsigma}$. Therefore, $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ was also called the *nonlinear Lagrangian* in [10] and the *extended Lagrangian* in [12]. If the geometrical operator $\Lambda(\boldsymbol{x})$ is quadratic and objective, the so-called Λ-transformation [12]

$$U^{\Lambda}(\boldsymbol{\varsigma}) = \text{sta}\{\langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma}\rangle - U(\boldsymbol{x}) |\ \boldsymbol{x} \in \mathscr{X}\} \tag{69}$$

is actually the *pure complementary gap function* which is obtained from the complementary gap function $G_{ap}(\boldsymbol{x}, \boldsymbol{\varsigma}) = \frac{1}{2}\langle \boldsymbol{x}, G(\boldsymbol{\varsigma})\boldsymbol{x}\rangle$ using the analytical solution form $\boldsymbol{x} = [G(\boldsymbol{\varsigma})]^{-1} F(\boldsymbol{\varsigma})$.

The geometrical nonlinearity in continuum physics means large deformation (far from equilibrium states), which usually leads to bifurcation in static systems [35] and chaos in dynamical systems [13]. Therefore, geometrically nonlinear systems are usually nonconvex. This is the reason why the geometrical nonlinearity was emphasized in the title of Gao and Strang's original work [22], although the system they studied is fully nonlinear and governed by a nonconvex/nonsmooth total (super) potential functional

$$\Pi(u) = \bar{W}(\bar{\Lambda}(u)) + \bar{F}(u), \tag{70}$$

where $\bar{W}(e)$ is called the stored energy, which is a canonical function(al) such that the constitutive law $e^* = \partial \bar{W}(e)$ is invertible on its effective domain; while $\bar{F}(u)$ is an external energy, which must be linear on the statically admissible space such that its Gâteaux derivative $\partial \bar{F}(u) = -\bar{u}^*$ leads to the external force (source) field (under the sign convention). The geometrically nonlinear operator $e = \bar{\Lambda}(u)$ in Gao and Strang's work should be an objective measure in order to satisfy certain well-known deformation laws (see Chap. 6, [10]). Therefore, the complementary gap function $G_{ap}(u, e^*)$ was naturally introduced. This objective function lays a foundation for the triality theory.

Oppositely, in a recent paper entitled "Some remarks concerning Gao–Strang's complementary gap function" by Voisei and Zalinescu [41], they choose quadratic functions as the external energy $\bar{F}(u)$ (see Examples 2, 4 and 5 in [41]), and piecewise linear function (see Example 1 in [41]) as the stored energy, they concluded "About the (complementary) gap function one can conclude that it is useless at least in the current context". Clearly, the piecewise linear function is not objective and cannot store energy; while for those quadratic functions $\bar{F}(u)$ they listed, the dual variable $u^* = \partial \bar{F}(u)$ depends on the configuration u. Such force field is called *follower force*. In this case, the system is not conservative and traditional variational methods do not apply. Unfortunately, similar counterexamples and conclusions are repeatedly presented in many other papers (see [39, 42] and references cited therein).

Actually, in order to study nonconvex variational problems in dissipative systems subjected to follower force field, a so-called rate variational method and the associated dual extremum principle were proposed in 1990 [18]. Also, Gao and Strang's work has been extended to general nonconvex dynamical systems to allow $\bar{F}(u)$ as a quadratic function, but notations were changed (see [12, 13]). In fact, if we let $\bar{\Lambda}(\boldsymbol{x}) = \{\Lambda(\boldsymbol{x}), \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x}\rangle\}$ and $\bar{W}(\bar{\Lambda}(\boldsymbol{x})) = V(\Lambda(\boldsymbol{x})) + \frac{1}{2}\langle \boldsymbol{x}, A\boldsymbol{x}\rangle$, the general nonconvex problem $(\mathscr{P})$ studied in this paper is simply a finite-dimensional version of the Gao and Strang's general work in large deformation theory. This method has been repeatedly used in many Gao's papers (see [17, 44]). Particularly, if $\bar{\Lambda}(u)$ is a Cauchy–Riemann strain measure, then

$$\Xi(u, e^*) = \langle \bar{\Lambda}(u); e^*\rangle - \bar{W}^*(e^*) + \bar{F}(u) \tag{71}$$

is the well-known *Hellinger–Reissner complementary energy* in finite deformation theory.[10] Furthermore, if the complementary energy $\bar{W}^*(e^*)$ is replaced by $\langle e; e^*\rangle - \bar{W}(e)$, the total complementary energy $\Xi(u, e^*)$ can be written in the so-called *pseudo-Lagrangian* (it was denoted as $L_p(u, e^*, e)$ in [22])

$$\Xi_{hw}(u, e^*, e) = \bar{W}(e) + \langle \bar{\Lambda}(u) - e; e^*\rangle + \bar{F}(u), \tag{72}$$

and we have

$$\Xi(u, e^*) = \text{sta}\{\Xi_{hw}(u, e^*, e) |\ e \in \mathscr{E}_a\}.$$

In large deformation mechanics, $\Xi_{hw}(u, e^*, e)$ is called the Hu-Washizu generalized potential energy, proposed independently by Hai-Chang Hu in 1954 and K. Washizu in 1955. The associated variational statement is the well-known *Hu-Washizu principle*, which has important applications in computational mechanics of thin-walled structures, where the geometrical equation $e = \bar{\Lambda}(u)$ is usually proposed by certain geometrical hypothesis [16, 21].

It has been emphasized in many papers that the key step in the canonical duality theory is to choose a geometrically reasonable measure $\boldsymbol{\xi} = \Lambda(\boldsymbol{x})$. It was shown in [26] that for a given nonconvex variational problem, the choice of $\Lambda(\boldsymbol{x})$ may not be unique and different geometrically admissible operators could lead to different canonical dual problems. But all these canonical dual problems must be equivalent in the sense that they have the same set of solutions. Also for complex systems, two types of sequential canonical transformations were proposed (see Chap. 4, [10]). By the fact that the objectivity and canonical duality are fundamental to all natural systems, for any given real problem, as long as the geometrical operator $\Lambda(\boldsymbol{x})$ can be chosen correctly such that the nonconvex objective function(al) can be recast by adopting a canonical form $W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x}))$, the canonical duality theory can be used to establish elegant theoretical results and to develop efficient algorithms

[10]The Hellinger–Reissner energy was first proposed by Hellinger in 1914. After the external energy $\bar{F}(u)$ and the boundary conditions in the statically admissible space $\mathscr{U}_k = \{u \in \mathscr{U}_a | e = \bar{\Lambda}(u) \in \mathscr{E}_a\}$ were fixed by Reissner in 1953, the associated variational statement has been known as the Hellinger–Reissner principle. However, the extremality condition of this principle was an open problem, and also the existence of pure complementary variational principles has been a well-known debate existing for over several decades in large deformation mechanics (see [31]). This open problem was partially solved by Gao and Strang's work and completely solve by the triality theory. While the pure complementary energy principle was formulated by Gao in 1999 [9].

for robust computations. The triality theory reveals an intrinsic duality pattern in nonconvex systems and should play important roles not only for solving a large class of challenging problems in nonconvex analysis and global optimization, but also for understanding, modeling, and simulation of complex systems.

7 Conclusion Remarks

Motivated by an open problem on the double-min duality in the triality theory that was left unaddressed since 2003, we have presented a mathematically rigorous proof for this theory based on the elementary linear algebra. Our results show that the triality theory holds strongly in the tri-duality form if the primal and its dual problems have the same dimension. Otherwise, both the canonical min-max and double-max duality statements hold strongly, but double-min duality statement holds weakly in a super-symmetric form. Additionally, a weak saddle duality theory is proposed, which shows that when the complementary gap function $G_{ap}(\boldsymbol{x}, \boldsymbol{\varsigma})$ is negative, either the primal problem $(\mathscr{P})$ (only if $m < n$) or its canonical dual $(\mathscr{P}^d)$ (only if $m > n$) could have saddle critical solutions. Therefore, this 7-year-old open problem is now solved completely and the triality theory is presented in an elegant form as expected.

The method adopted in this paper can be generalized for more general constrained global optimization problems. As it is mentioned in Remark 3 that the primal problem $(\mathscr{P})$ could be NP-hard if its canonical dual has no critical point in $\mathscr{S}_a^+$. Also, the extremality conditions for those critical points $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$ are still unknown if the Hessian matrix $G(\bar{\boldsymbol{\varsigma}})$ of the gap function is indefinite. Although a general theorem on the existence and uniqueness of the canonical dual solution in $\mathscr{S}_a^+$ was proposed in [15], and some perturbation methods were discussed in [19], detailed quantitative study on these topics is fundamentally important and critical for understanding and solving NP-hard problems.

Acknowledgements The main results of this paper were announced at the 2nd World Congress of Global Optimization, July 3–7, 2011, Chania, Greece. The paper was posted online on April 15, 2011 at https://arXiv.org/abs/1104.2970. The authors are gratefully indebted with Professor Hanif Sherali at Virginia Tech for his detailed remarks and important suggestions. This paper has benefited from three anonymous referees' constructive comments. David Gao's research is supported by US Air Force Office of Scientific Research under the grants FA2386-16-1-4082 and FA9550-17-1-0151.

References

1. Arnold, V.I.: On teaching mathematics. Uspekhi Mat. Nauk. 53(1), 229–234 (1998); English translation. Russ. Math. Surv. **53**(1), 229–236 (1998)
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**, 125–142 (2008)

4. Filly, J.A., Donniell, E.F.: The Moore-Penrose generalized inverse for sums of matrices. SIAM J. Matrix. Anal. Appl. **21**(2), 629–635 (1999)
5. Floudas, C.A.: Deterministic Global Optimization: Theory. Methods and Applications. Kluwer Academic, Dordrecht (2000)
6. Floudas, C.A.: Systems approaches in bioinformatics and computational genomics. In: Challenges for the Chemical Sciences in the 21th Century, Information and Communications Workshop, National Research Council of the National Academies, National Academies Press, pp. 116–125 (2003)
7. Gao, D.Y.: Post-buckling analysis and anomalous dual variational problems in nonlinear beam theory. In: Godoy, L.A., Suarez, L.E. (eds.) Applied Mechanics in Americans, Proceedings of the Fifth Pan American Congress of Applied Mechanics, vol. 4. The University of Iowa, Iowa city (1996)
8. Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory. Appl. Mech. Rev. **50**(11), S64–S71 (1997)
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. Kluwer Academic, Dordrecht (2000)
11. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
12. Gao, D.Y.: Perfect duality theory and complete solutions to a class of global optimization problems. Optimization **52**(4–5), 467–493 (2003)
13. Gao, D.Y.: Nonconvex semi-linear problems and canonical dual solutions. In: Gao, D.Y., Ogden, R.W. (eds.) Advances in Mechanics and Mathematics, vol. II, p. 261312. Kluwer Academic, Dordrecht (2003)
14. Gao, D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. J. Indust. Manag. Optim. **3**(2), 293–304 (2007)
15. Gao, D.Y.: Canonical duality theory: theory, method, and applications in global optimization. Comput. Chem. **33**, 1964–1972 (2009)
16. Gao, D.Y., Cheung, Y.K.: On the extremum complementary energy principles for nonlinear elastic shells. Int. J. Solids Struct. **26**, 683–693 (1989)
17. Gao, D.Y., Ogden, R.W.: Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation. Q. J. Mech. Appl. Math **61**(4), 497–522 (2008)
18. Gao, Y., Onat, E.T.: Rate variational extremum principles for finite elasto- plasticity. Appl. Math. Mech. **11**(7), 659–667 (1990)
19. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. **47**(3), 463–484 (2010)
20. 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)
21. Gao, D.Y. Sherali, H.D.: Canonical duality: connection between nonconvex mechanics and global optimization. In: Advances in Applied Mathematics and Global Optimization, pp. 249–316. Springer, Berlin (2009)
22. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Quart. Appl. Math. **47**(3), 487–504 (1989)
23. Gao, D.Y., Watson, L.T., Easterling, D.R., Thacker, W.I., Billups, S.C.: Solving the canonical dual of box- and integer-constrained nonconvex quadratic programs via a deterministic direct search algorithm. Optim. Methods Softw. **28**(2), 313–326 (2013). doi:10.1080/10556788.2011.641125
24. Gao, D.Y. Wu, C.Z.: On the triality theory in global optimization (2010). arXiv:1104.2970v1

25. 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)
26. Gao, D.Y., Yang, W.-H.: Multi-duality in minimal surface type problems. Stud. Appl. Math. **95**, 127–146 (1995)
27. 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)
28. Holzapfel, G.A.: Nonlinear Solid Mechanics: A Continuum Approach for Engineering. Wiley, New York (2000). ISBN 978-0471823193
29. Holzapfel, G.A., Ogden, R.W. (eds.): Mechanics of Biological Tissue, vol. 12. Springer, Heidelberg (2006)
30. Landau, L.D., Lifshitz, E.M.: Mechanics, vol. 1, 3rd edn. Butterworth-Heinemann, London (1976)
31. Li, S.F., Gupta, A.: On dual configuration forces. J. Elast. **84**, 13–31 (2006)
32. Moreau, J.J.: La notion de sur-potentiel et les liaisons unilatérales en élastostatique. C. R. Acad. Sc. Paris **267A**, 954–957 (1968)
33. Ogden, R.W.: Non-Linear Elastic Deformations, 1st edn, p. 544. Dover Publications, New York (1997)
34. 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)
35. Santos, H.A.F.A., Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Nonlinear Mech. **47**, 240–247 (2012). doi:10.1016/j.ijnonlinmec.2011.05.012
36. Sewell, M.J.: Maximum and Minimum Principles: A Unified Approach, with Applications. Cambridge University Press, Cambridge (1987)
37. Silva, D.M.M., Gao, D.Y.: Complete solutions and triality theory to a nonconvex optimization problem with double-well potential in $\mathbb{R}^n$ (2011). arXiv:1110.0285v1
38. Strang, G.: Introduction to Applied Mathematics. Wellesley-Cambridge Press, Wellesley (1986)
39. Strugariu, R., Voisei, M.D., Zalinescu, C.: Counter-examples in bi-duality, triality and tri-duality. Discrete and Continuous Dynamical Systems - Series A, vol. 31, No. 4, December 2011
40. Tonti, E.: On the mathematical structure of a large class of physical theories, pp. 49–56. Accad. Naz. dei Lincei, Serie VIII, LII (1972)
41. Voisei, M.D., Zalinescu, C.: Some remarks concerning Gao-Strang's complementary gap function. Appl. Anal. (2010). doi:10.1080/00036811.2010.483427
42. Voisei, M.D., Zalinescu, C.: Counterexamples to some triality and tri-duality results. J. Global Optim. **49**, 173–183 (2011)
43. Wang, Z.B., Fang, S.C., Gao, D.Y., and Xing, W.X.: Canonical dual approach to solving the maximum cut problem, to appear in, J. Glob. Optim
44. Yau, S.-T., Gao, D.Y.: Obstacle problem for von Karman equations. Adv. Appl. Math. **13**, 123–141 (1992)
45. 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

Canonical Duality Theory for Solving Non-monotone Variational Inequality Problems

Guoshan Liu, David Yang Gao and Shouyang Wang

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

1 Problems and Motivation

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

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

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

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

G. Liu (✉)
School of Business, Renmin University of China, Beijing 100872, China
e-mail: liuguoshan@gmail.com

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

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

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_7

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

$$\mathscr{K} = \left\{\mathbf{x} \in \mathbb{R}^n \mid \phi(\mathbf{x}) \leq 0\right\},$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

2 Optimization Problem and Its Canonical Dual

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

$$(OP): \min_{\mathbf{x}\in\mathbb{R}^n} \left\{ P(\mathbf{x}) = W(\mathbf{x}) + \frac{1}{2}\langle \mathbf{x}, Q\mathbf{x}\rangle - \langle \mathbf{x}, \mathbf{f}\rangle \right\} \quad (4)$$
$$s.t. \quad \phi(\mathbf{x}) \leq 0.$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

where

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

We let

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

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

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

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

$$\Xi(\mathbf{x}, \boldsymbol{\varsigma}) = \langle \Lambda(\mathbf{x}), \boldsymbol{\varsigma} \rangle - \bar{V}^{\sharp}(\boldsymbol{\varsigma}) - U(\mathbf{x}).$$

For a given $\boldsymbol{\varsigma} \in \mathscr{S}_a$, the canonical dual function can be obtained as

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

where $U^{\Lambda}(\boldsymbol{\varsigma})$ is called the Λ-conjugate of U, defined by

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

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

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

$$(\mathscr{P}^d): \quad \max\{P^d(\boldsymbol{\varsigma}): \quad \boldsymbol{\varsigma} \in \mathscr{S}_c\}. \tag{9}$$

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

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

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

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

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

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

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

$$\mathscr{S}_c = \{\boldsymbol{\varsigma} = \{\boldsymbol{\sigma}, \lambda) \in \mathscr{S}_a | \ \ \boldsymbol{\tau}(\boldsymbol{\varsigma}) \in \mathscr{C}_{ol}\mathbf{G}(\boldsymbol{\varsigma})\}.$$

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

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

It follows that

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

and

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

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

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

where

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

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

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

Similar to Lemma 1, we have the following result.

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

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

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

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

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

Proof. First, assume that $\bar{\boldsymbol{\varsigma}} \in \mathbb{R}^p \times \mathbb{R}^q_+$ is a solution of (DVI), it is obvious that $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_c \cap \mathrm{dom}(\bar{V}^{\sharp})$ and we have that

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

Let

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

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

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

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

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

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

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

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

3 Global and Local Extremalities

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

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

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

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

we know that

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

We also have that

$$\nabla^2_{\mathbf{xx}} L(\bar{\mathbf{x}}, \bar{\lambda}) = \mathscr{B}^\top(\bar{\mathbf{x}})\nabla^2_{\varepsilon\varepsilon}V(\bar{\varepsilon})\mathscr{B}(\bar{\mathbf{x}}) + \mathbf{G}(\bar{\boldsymbol{\varsigma}}).$$

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

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

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

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

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

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

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

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

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

and

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

We also have that

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

Then, we have

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

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

4 Existence of the Solution

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

$$\mathscr{S}_c^+ = \{\bar{\boldsymbol{\varsigma}} \in \mathbb{R}^p \times \mathbb{R}_+^q |\ \ \gamma_{\mathbf{G}}(\bar{\boldsymbol{\varsigma}}) > 0\},$$

$$\bar{\mathscr{S}_c} = \{\bar{\boldsymbol{\varsigma}} \in \mathbb{R}^p \times \mathbb{R}_+^q |\ \ \gamma_{\mathbf{G}}(\bar{\boldsymbol{\varsigma}}) = 0\},$$

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

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

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

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

and

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

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

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

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

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

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

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

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

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

$$\Omega(\delta) = \left\{\boldsymbol{\varsigma} \in \mathscr{S}_c |\ \ \|\boldsymbol{\varsigma}\| \le \delta, \ \ \gamma(\mathbf{G}(\boldsymbol{\varsigma})) \ge \frac{1}{\delta}\right\}.$$

Let

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

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

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

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

$$\lim_{i\to+\infty} P^d(\boldsymbol{\varsigma}_i) = M, \tag{15}$$

where $M \in \mathbb{R}$. If $\{\boldsymbol{\varsigma}_i\}_{i=1,2,\cdots}$ is unbounded, then there is $\mathscr{K}_1 \subseteq \{1, 2, \cdots, \}$ such that

$$\lim_{i\to+\infty,\ i\in\mathscr{K}_1} \|\boldsymbol{\varsigma}_i\| = \infty.$$

Then, we have that

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

and

$$\lim_{i\to+\infty,\ i\in\mathscr{K}_1} \langle d,\ \sigma_i\rangle + \langle e,\ \lambda_i\rangle + \bar{V}^\sharp(\boldsymbol{\varsigma}_i) = +\infty.$$

It follows that

$$\lim_{i\to+\infty,\ i\in\mathscr{K}_1} P^d(\boldsymbol{\varsigma}_i) = -\infty,$$

which contradicts (15).

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

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

Now, we have that

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

Let

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

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

$$\xi_i \in \mathscr{T}_{\mathbf{G}}(\boldsymbol{\varsigma}_i), \quad i = 1, 2, \cdots,$$

we have

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

Therefore,

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

Now, we have

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

for $i \in \mathscr{K}_2$. Then, we have

$$\lim_{i\to+\infty,\ i\in\mathscr{K}_2} U^{\Lambda}(\boldsymbol{\varsigma}_i) = -\infty.$$

Therefore,

$$\lim_{i\to+\infty,\ i\in\mathscr{K}_2} P^d(\boldsymbol{\varsigma}_i) = -\infty,$$

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

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

Choose a $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_c^+$, there must be a $\bar{\delta} > 0$ such that

$$\Gamma(\bar{\delta}) \le P^d(\bar{\boldsymbol{\varsigma}}).$$

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

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

It follows that

$$P^d(\tilde{\boldsymbol{\varsigma}}) = \max_{\boldsymbol{\varsigma}\in\mathscr{S}_c^+} P^d(\boldsymbol{\varsigma}).$$

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

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

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

Proof. It is similar with the Theorem 5. □

5 Some Examples

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

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

and

$$\mathscr{K} = \left\{\mathbf{x} | \frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) \le e\right\}.$$

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

$$\begin{array}{ll}\min_{\mathbf{x}} P(\mathbf{x}) = \frac{1}{2}(\frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) - \mu)^2 + (\mathbf{x}_1^2 + \mathbf{x}_2\mathbf{x}_2 + \mathbf{x}_2^2) - (2\mathbf{x}_1 + 2\mathbf{x}_2) \\ s.t. \qquad \frac{1}{2}(\mathbf{x}_1^2 + \mathbf{x}_2^2) \le e.\end{array}$$

For the dual problem, we have that

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

It follows that

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

Then the critical condition of the dual problem becomes

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

We consider three cases with various values of μ and e.

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

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

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

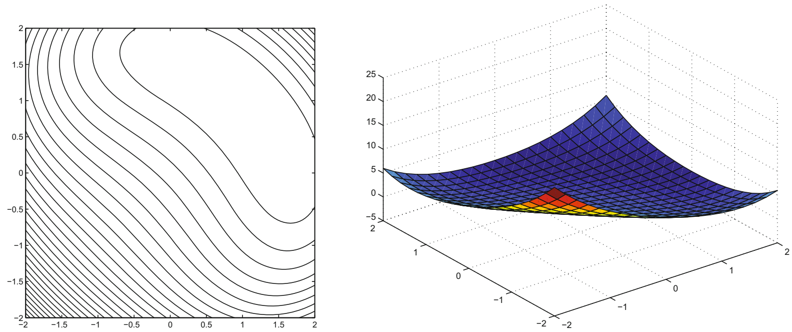


Fig. 1 Contours and graph of $P(\mathbf{x})$ in Example 1

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

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

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

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

By the fact that

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

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

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

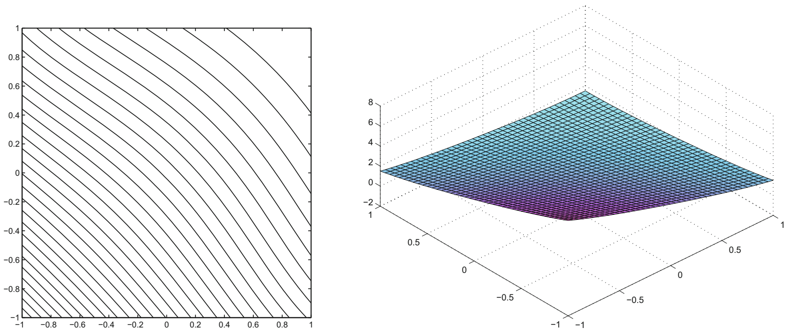


Fig. 2 Contours and graph for Example 2

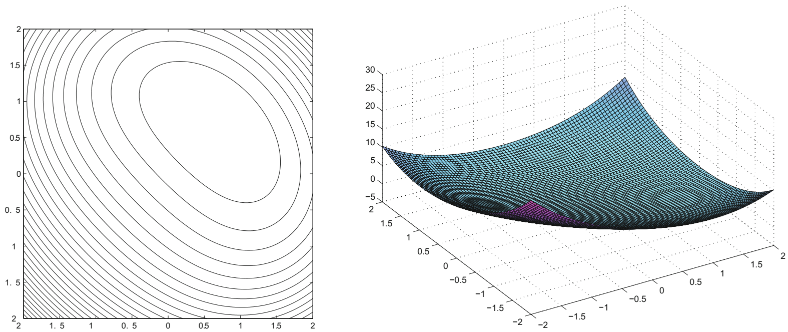


Fig. 3 Contours and graph of Example 3

$$\frac{4}{(3+(\sigma+\lambda))^2} - \sigma = \frac{4}{9}$$
$$\frac{4}{(3+(\sigma+\lambda))^2} - 2 \in \partial \Psi^\sharp(\lambda).$$

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

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

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

6 Conclusions

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

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

References

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

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

Canonical Dual Approach for Contact Mechanics Problems with Friction

Vittorio Latorre, Simone Sagratella and David Yang Gao

Abstract This paper presents an application of Canonical duality theory to the solution of contact problems with Coulomb friction. The contact problem is formulated as a quasi-variational inequality which solution is found by solving its Karush–Kuhn–Tucker system of equations. The complementarity conditions are reformulated by using the Fischer–Burmeister complementarity function, obtaining a non-convex global optimization problem. Then canonical duality theory is applied to reformulate the non-convex global optimization problem and define its optimality conditions, finding a solution of the original quasi-variational inequality. We also propose a methodology for finding the solutions of the new formulation, and report the results on well-known instances from literature.

1 Introduction

Contact mechanics provides many challenging problems in both engineering and mathematics. The problem generally consists in analyzing the forces created when an elastic body comes in contact with a rigid obstacle and search for an equilibrium of such forces. In the moment the two bodies come in contact there are not only normal forces that prevent interpenetration between the two bodies, but also friction forces that prevent the elastic body to slide on the rigid obstacle.

One of the most popular application of this class of problems is the automate planning of tasks carried out by robots. Contact problems arise in such applications

V. Latorre (✉) · S. Sagratella
Department of Computer, Control and Management Engineering "Sapienza",
University of Rome, Via Ariosto 25, 00185 Rome, Italy
e-mail: latorre@dis.uniroma1.it

S. Sagratella
e-mail: sagratella@dis.uniroma1.it

D.Y. Gao
Faculty of Science and Technology, Federation University Australia,
Mt Helen, VIC 3353, Australia
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_8

when a robotic arm has to come in contact with objects in its surrounding. In such cases it is necessary to find an equilibrium in the strength of the robotic arm so that the friction is sufficient to have a solid grip on the object without damaging it [1].

Early research focused on the frictionless contact between two or more bodies [2, 3], where a quadratic programming optimization problem or a variational inequality (like in [4]) is solved. However, in most cases, this formulation does not completely reflect the physical reality. For this reason, contact problems with Coulomb friction are studied and solved. Generally several mathematical programming methods are used for solving such problems(for more information refer to [5] and citations therein), and one of the most popular formulations for contact problems are Quasi-Variational Inequalities as reported in [6]. In the same book, the authors use a non-smooth newton method in order to compute a solution, finding satisfying results only for small values of the friction.

Quasi-variational inequalities (QVIs) are a powerful modeling tool capable of describing complex equilibrium situations that can appear in different fields such as generalized Nash games, mechanics, economics, statistics and so on (see e.g. [7–10]). For what regards QVIs there are a few works devoted to the numerical solution of finite-dimensional QVIs (see e.g. [10–16]), in particular in the recent paper [17] a solution method for QVIs based on solving their Karush–Kuhn–Tucker (KKT) conditions is proposed.

In this work we propose a novel Canonical Duality approach for solving the QVI associated with the contact problem with Coulomb friction by presenting a deeper insight on said application of the theory already presented in [18]. In particular we show that the QVI associated with the problem belongs to a particular class of QVIs called Affine Quasi-Variational Inequalities (AQVI). We search for a solution of the AQVI by determining a point that satisfies its KKT conditions. In order to find such point we reformulate the KKT conditions of the AQVI by using the Fisher–Burmeister complementarity function, obtaining a non-convex global optimization problem [19, 20]. By using Canonical Duality Theory it is possible to reformulate the obtained non-convex optimization problem and to find the conditions for a critical point to be a global solution.

The principal aim of this paper is to show the potentiality of canonical duality theory for this class of mechanics problems and propose a new methodology to solve them.

Canonical duality theory, developed from non-convex analysis and global optimization [21, 22], is a potentially powerful methodology, which has been successfully used for solving a large class of challenging problems in biology, engineering, sciences [23–25], and recently in network communications [26–28], radial basis neural networks [29] and constrained optimization [30]. In this paper we use a canonical dual transformation methodology in order to formulate the Total Complementarity Function of the original problem which stationary points do not have any duality gap in respect to the corresponding solutions of the primal problem. With the proprieties of the total complementarity function it is also possible to find the optimality conditions of the original problem.

The paper is organized as follows: In Sect. 2 we present the problem from mechanics and then report its formulation as a quasi-variational inequality. In Sect. 3 we use the dual canonical transformation to reformulate the global optimization problem as a total complementarity function and analyze its proprieties. Finally in Sect. 4 we report an optimization procedure based on the results obtained in the previous sections and numerical results on some instances of the contact friction problem.

We use the following notation: $(a, b) \in \mathbb{R}^{n_a+n_b}$ indicates the column vector comprised by vectors $a \in \mathbb{R}^{n_a}$ and $b \in \mathbb{R}^{n_b}$; $\mathbb{R}^n_+ \subset \mathbb{R}^n$ denotes the set of nonnegative numbers; $\mathbb{R}^n_{++} \subset \mathbb{R}^n$ is the set of positive numbers; $\text{sta}\{f(x) : x \in \mathscr{X}\}$ denotes the set of stationary points of function f in $\mathscr{X}$; given a matrix $Q \in \mathbb{R}^{a\times b}$ we indicate with Q_{i*} its i-th row and with Q_{*i} its i-th column; $\text{diag}(a)$ denotes the (square) diagonal matrix whose diagonal entries are the elements of the vector a; $\circ$ denotes the Hadamard (component-wise) product operator; $\mathbf{0}_n$ indicates the origin in $\mathbb{R}^n$. The double dots product $e : \sigma$ indicates $\text{trace}(e^T\sigma)$ and is a standard notation in solid mechanics where $e \in \mathbb{R}^n$ and $\sigma \in \mathbb{R}^{n\times n}$.

2 Problem Formulation

Generally in contact problems, the Coulomb friction between the body and obstacle should be considered in order to have the most realistic representation of the mathematical modeling. A simple way to define this problem is to restrict the normal displacement of the boundary points of the elastic body by means of unilateral contact constraints.

Let us consider an elastic body which occupies a smooth, bounded simply-connected domain $\Omega \subset \mathbb{R}^3$ with boundary $\Gamma = \Gamma_g \cup \Gamma_u \cup \Gamma_c$, where

1. Γ_g: associated the Neumann boundary condition. The force that moves the object is applied on this surface;
2. Γ_u: associated to the homogeneous Dirichlet boundary conditions. This part of the body is considered fixed;
3. Γ_c: associated with the unilateral boundary conditions. This is the part of the object in contact with the rigid obstacle.

The displacement of the elastic body Ω is a field function $u : \Omega \to \mathbb{R}^3$ that belongs to the following set of kinetically admissible space:

$$K = \left\{ v \in \mathscr{U}(\Omega) | \sum_{i=1}^{3} v_i n_i \leq 0 \text{ on } \Gamma_c \right\}, \tag{1}$$

where n_i for $i = 1, 2, 3$ denotes the outer normal to $\partial\Omega$ and $\mathscr{U}(\Omega)$ is a Sobolev space defined on Ω such as $v = 0$ on Γ_u. The function u solves a contact problem with given friction if it is the solution of the following variational inequality:

$$\begin{aligned}&\textit{Find } u \textit{ such that:}\\&a(u, v-u) + \int_{\Gamma_c} \gamma(|v_t| - |u_t|)d\Gamma_c \geq \int_{\Gamma_g} F(v-u)d\Gamma_g \quad \forall v \in K,\end{aligned} \tag{2}$$

where the bilinear form $a(\cdot,\cdot)$ is defined as

$$a(u, v) = \int_{\Omega} e(u) : \mathscr{H} : e(v)d\xi,$$

where $e(w)$ is the infinitesimal strain tensor and $\mathscr{H}$ is a bounded, symmetric and elliptic mapping that expresses Hooke's law. The function F represent the external force on Γ_g and γ is a given friction function on Γ_c such that $\gamma \geq 0$.

In order to solve problem (2) it is possible to use a finite element method by means of a discretization parameter h and obtain:

$$\textit{Find } u \textit{ such that: } \langle Cu, v-u\rangle - \langle f, v-u\rangle \quad \forall v \in K, \tag{3}$$

where C is the stiffness matrix of the r nodes considered on the contact surface Γ_c, f is the right-hand side vector associated with the external force F and u and $v \in \mathbb{R}^N$, where N is the number of variables of the problem that is composed by the normal and tangent components of the nodes on the contact surface Γ_c, that is $N = 2r$.

It is possible to show that the solution of problem (2) satisfies the following contact conditions with given friction:

$$\begin{aligned}&u_n \leq 0, \quad T_n \leq 0, \quad u_n T_n = 0 && \text{on } \Gamma_c\\&|T_t| \leq \gamma, \quad (\gamma - |T_t|)u_t = 0, \quad u_t T_t \leq 0 && \text{on } \Gamma_c,\end{aligned} \tag{4}$$

where u_n and u_t indicate the tangent and the normal components of vector u, T indicates the boundary stress vector on Γ_c with normal component T_n and tangent component T_t. The first relation in (4) indicates the standard unilateral contact conditions, and the second relation indicates that if the tangent component of the stress vector is lower than γ, then there is no displacement and once its value reaches γ the object begins to slide on the obstacle.

For Coulomb friction, the function γ depends linearly on the normal force, i.e., $\gamma = \Phi|T_n|$, where Φ is the coefficient of friction characterizing the physical properties of the surfaces in contact.

The solution of problem (3) can be found by solving a difficult fixed-point problem which convergence has not been proven [6]. Moreover the solution of this problem yields the displacements, while a practitioner is generally interested to find the stress on the nodes of the contact surface. Computing the contact stress from the displacement is also considered a difficult task.

To avoid these issues, it is possible to use the reciprocal variational formulation of problem (3) with variables $\mu_j, j = 1, \ldots, N$. We assume that for $j = 1, 2, \ldots, N/2$ the components corresponding to the odd number $2j-1$ are associated with the tangential components of the nodes on the contact surface and the even components are associated with the normal components of the nodes on the contact surface. Given a stress vector $\tau \in \mathbb{R}^N$, the following condition must be satisfied:

$$\tau \in \tilde{K}(\tau) = \{\mu \in \mathbb{R}^N | \ \mu_{2j} \leq 0, \ \ |\mu_{2j-1}| \leq \Phi |\tau_{2j}|, \ \ j = 1, \ldots, N/2\}. \quad (5)$$

It is easy to notice that conditions (5) represent the conditions reported in (4) and if we assume that the even components of μ are lower bounded by a value l, the bound constraints (5) can be rewritten in the following way:

$$\tilde{K}(\tau) = \{\mu \in \mathbb{R}^N | \ A\mu + B\tau - c \leq 0\}$$

with $A, B \in \mathbb{R}^{m \times N}$, $c \in \mathbb{R}^m$, $m = 2N$ and

$$(A)_{ij} = \begin{cases} 1 & \text{if } i \text{ odd}, j = (i+1)/2 \\ -1 & \text{if } i \text{ even}, j = (i)/2 \\ 0 & otherwise \end{cases},$$

$$(B)_{ij} = \begin{cases} -\Phi & \text{if } \lfloor (i+1)/2 \rfloor \text{ odd}, j = \lfloor (i+1)/2 \rfloor + 1 \\ 0 & otherwise \end{cases},$$

$$c_i = \begin{cases} l & \text{if } mod(i,4) = 0 \\ 0 & otherwise \end{cases},$$

The stress vector is the solution of the following quasi-variational inequality:

$$Find\ u\ such\ that\ \langle D\tau, \mu - \tau \rangle + \langle e, \mu - \tau \rangle \geq 0 \quad \forall \mu \in \tilde{K}(\tau), \quad (6)$$

where $D = C^{-1}$ and $e = -C^{-1}f$.

Problem (6) is an Affine Quasi-Variational Inequality AQVI(A, B, c, D, e) which equilibrium point can be found by satisfying its KKT conditions. We say that a point $\tau \in \mathbb{R}^N$ satisfies the KKT conditions if multipliers $\lambda \in \mathbb{R}^m$ exist such that

$$D\tau + e + A^T\lambda \ = \ \mathbf{0}_N, \quad \mathbf{0}_m \ \leq \ \lambda \ \perp \ A\tau + B\tau - c \ \leq \ \mathbf{0}_m. \quad (7)$$

It is quite easy to show the following result, whose proof we omit.

Theorem 1. *If a point τ, together with a suitable vector $\lambda \in \mathbb{R}^m$ of multipliers, satisfies the KKT system (7), then τ is a solution of the AQVI (A, B, c, D, e). Vice versa, if τ is a solution of the AQVI (A, B, c, D, e) then multipliers $\lambda \in \mathbb{R}^m$ exist such that the pair (τ, λ) satisfies the KKT conditions (7).*

Generally it is difficult to deal with the complementarity conditions in (7). Such conditions can be replaced by using complementarity functions. A *complementarity function* is a function $\theta : \mathbb{R}^2 \to \mathbb{R}$ such that $\theta(a, b) = 0$ if and only if $a \geq 0, b \geq 0$, and $ab = 0$. One of the most prominent complementarity functions is the Fischer–Burmeister function:

$$\theta_{\text{FB}}(a,b) = \sqrt{a^2+b^2} - (a+b).$$

We can then consider the following problem equivalent to the solution of system (7):

$$(\mathscr{P}): \quad \min_{x,\lambda} \; P(x,\lambda) = \frac{1}{2}\left\| \begin{matrix} D\tau + e + A^T\lambda \\ [\theta_i(\lambda, g(\tau))]_{i=1}^m \end{matrix} \right\|^2 = \\ \frac{1}{2}\sum_{i=1}^m \theta_{\text{FB}}(\lambda_i, -g_i(\tau))^2 + \frac{1}{2}(\tau,\lambda)^T M(\tau,\lambda) - f^T(\tau,\lambda) = \\ W(\tau,\lambda) + \frac{1}{2}(\tau,\lambda)^T M(\tau,\lambda) - f^T(\tau,\lambda) \tag{8}$$

where for all $i = 1, \ldots, m$:

$$g_i(\tau) = A_{i*}\tau + B_{i*}\tau - c_i,$$

and

$$M = \begin{pmatrix} D^T \\ A \end{pmatrix} \begin{pmatrix} D \; A^T \end{pmatrix}, \quad f = -\begin{pmatrix} D^T \\ A \end{pmatrix} e.$$

It is easy to see that problem $\mathscr{P}$ is non-convex, furthermore, since only the global minima of problem (8) correspond to a stress vector solutions, it is very hard to solve. In fact it is well known that not all critical points of P are solutions of the AQVI [19, 20].

3 Canonical Dual Transformation and Proprieties

The first step of a canonical dual transformation for problem (8) is the introduction of operator $\xi = \Lambda(\tau,\lambda) : \mathbb{R}^{N+m} \to \mathscr{E}_0 \equiv \mathbb{R}^m$, which is defined as

$$\xi_i = \Lambda_i(\tau,\lambda_i) = \sqrt{\lambda_i^2 + g_i(\tau)^2} - \lambda_i + g_i(\tau), \qquad i = 1, \ldots, m, \tag{9}$$

note that each ξ_i is convex since it is defined as a composition of a convex function and a linear function. Furthermore we introduce a convex function $V_0 : \mathscr{E}_0 \to \mathbb{R}$ (associated with ξ), i.e., defined as

$$V_0(\xi) = \frac{1}{2}\sum_{i=1}^m \xi_i^2. \tag{10}$$

It is easy to see that

$$W(\tau,\lambda) = V_0(\Lambda(x,\lambda)) = V_0(\xi). \tag{11}$$

Furthermore, we introduce a dual variable

$$\sigma = \nabla V_0(\xi) = \xi, \tag{12}$$

which is defined on the range $\mathscr{S}_0 \equiv \mathbb{R}^m$ of $\nabla V_0(\cdot)$. Since the (duality) mapping (12) is invertible, i.e., ξ can be expressed as a function of σ, then the function $V_0(\xi)$ is said to be a canonical function on $\mathscr{E}_0$, see [21].

In order to define the total complementarity function in both primal and dual variables (τ, λ, σ) we use a Legendre transformation [21]. Specifically the Legendre conjugate $V_0^*(\sigma) : \mathscr{S}_0 \to \mathbb{R}$ is defined in the following way

$$V_0^*(\sigma) = \text{sta}\left\{\xi^T\sigma - V_0(\xi) \ : \ \xi \in \mathscr{E}_0\right\},$$

which is equal to the function $\xi^T\sigma - V_0(\xi)$ in which ξ is fixed to a stationary point. Since $\xi^T\sigma - V_0(\xi)$ is a quadratic strictly concave function in ξ, then it is easy to see that its (unique) stationary point is $\bar{\xi} = \sigma$, and then

$$V_0^*(\sigma) = \bar{\xi}^T\sigma - V_0(\bar{\xi}) = \sigma^T\sigma - V_0(\bar{\xi}(\sigma)) \overset{(10)}{=} \frac{1}{2}\sum_{i=1}^{n}\sigma_i^2, \tag{13}$$

moreover we obtain that

$$V_0(\xi) = \xi^T\sigma - V_0^*(\sigma). \tag{14}$$

Since

$$W(\tau, \lambda) \overset{(11)}{=} V_0(\xi) \overset{(14)}{=} \xi^T\sigma - V_0^*(\sigma)$$

we obtain the total complementarity function:

$$\begin{aligned} \Xi_0(\tau, \lambda, \sigma) = \sum_{i=1}^{m}\left[\sigma_i\left(\sqrt{\lambda_i^2 + g_i(\tau)^2} - \lambda_i + g_i(\tau)\right) - \frac{1}{2}\sigma_i^2\right] + \\ \frac{1}{2}(\tau, \lambda)^T M(\tau, \lambda) - f^T(\tau, \lambda), \end{aligned} \tag{15}$$

where

$$\bar{f}(\sigma) = f + \begin{pmatrix} -(A^T + B^T)\sigma \\ \sigma \end{pmatrix}. \tag{16}$$

It is easy to see that the total complementarity function Ξ_0 is strictly concave in σ for all (τ, λ). Moreover Ξ_0 is convex in (τ, λ) (although non-smooth but only semi-smooth) for all $\sigma \in \mathbb{R}^m_+$, since $M \succeq 0$ and each function $\sqrt{\lambda_i^2 + g_i(\tau)^2}$ is convex in (x, λ).

Function (15) has some interesting properties that can be exploited to find a global solution of problem $\mathscr{P}$. In the following we report these properties omitting their proofs. The interested reader can refer to [18] for a detailed discussion on such properties. The first propriety shows the relations between the critical point of problem (8) and (15).

Theorem 2. *(Complementarity dual principle) Let* $(\bar{\tau}, \bar{\lambda}, \bar{\sigma})$ *be a critical point for* Ξ_0*, then* $(\bar{\tau}, \bar{\lambda})$ *is critical point for* $P(\bar{\tau}, \bar{\lambda})$ *and*

$$P(\bar{\tau}, \bar{\lambda}) = \Xi_0(\bar{\tau}, \bar{\lambda}, \bar{\sigma}) \tag{17}$$

Theorem 2 proves that every critical point of Ξ_0 has a corresponding critical point in $P(\bar{\tau}, \bar{\lambda})$, furthermore they have the same value of the objective function. From now on we will indicate with:

$$\mathscr{S}_a^+ = \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+^m.$$

The next theorem characterizes the critical points of Ξ_0 in a subset of the dual space:

Theorem 3. *Let a point* $(\bar{\tau}, \bar{\lambda}, \bar{\sigma}) \in \mathscr{S}_a^+$ *be critical for* Ξ_0 *then it is a saddle point for* Ξ_0.

Theorem 4. *Suppose that* (τ^*, λ^*) *exists such that* $P(\tau^*, \lambda^*) = -\frac{1}{2}e^T e$*, that is* (τ^*, λ^*) *is a global minimum of the primal problem, then*

1. $(\tau^*, \lambda^*, \mathbf{0}_m)$ *is a critical point for* Ξ_0 *and* $\Xi_0(\tau^*, \lambda^*, \mathbf{0}_m) = -\frac{1}{2}e^T e$;
2. *all points* $(\tau, \lambda, \sigma) \in \mathbb{R}^n \times \mathbb{R}^m \times \left\{\mathbb{R}_+^m \setminus \{\mathbf{0}_m\}\right\}$ *are not critical for* Ξ_0.

From Theorem 4 the critical points of Ξ_0 corresponding to the solutions of the Coulomb friction problem are located in $\mathscr{S}_a^+$. The property that the stationary points must all have $\|\sigma^*\| = 0$ means that the point (τ^*, λ^*) satisfies the KKT complementarity conditions as $\sigma_i^* = \theta_{FB}(\lambda_i^*, g_i(\tau^*))$, for $i = 1, \ldots, m$.

4 Results

In this section two instances of the contact problem with friction are solved. The two instances are taken from Problem 11.1 in [6] and are called CPCF31 and CPCF41. In the case the Coulomb friction $\Phi = 10$ the two instances correspond to problems OutKZ31and OutKZ41 of QVILIB [31], a collection of test problems from diverse sources that gives a uniform basis on which algorithms for the solution of QVIs can be tested and compared. The two problems have the same rigid obstacle and elastic body, but the segmentation of the obstacle is different. In CPCF31 the obstacle is divided into 30 segments, while in CPCF41 the obstacle is divided in 40 segments.

The approach we use in order to find a solution of (6) is based on the results of Theorem 4. In particular we search a solution of the following problem:

$$(\mathscr{S}P): \quad \min_{(\tau,\lambda)} \max_{\sigma \in \mathscr{S}_+^a} \Xi_0(\tau, \lambda, \sigma). \tag{18}$$

Since the total complementarity function is non-smooth because of the term due to the Fisher–Burmeister, we apply a simple smoothing procedure and obtain the following smoothed total complementarity function:

$$\Xi_\varepsilon(\tau, \lambda, \sigma) = \sum_{i=1}^{m} \left[\sigma_i \left(\sqrt{\lambda_i^2 + g_i(\tau)^2 + \varepsilon^2} - \lambda_i + g_i(\tau) \right) - \frac{1}{2}\sigma_i^2 \right] + \\ \frac{1}{2}(\tau, \lambda)^T M(\tau, \lambda) - f^T(\tau, \lambda),$$

$\Xi_\varepsilon(\tau, \lambda, \sigma)$ still retains its properties of convexity in respect to (τ, λ) for all $\sigma \in \mathbb{R}^m_+$ and concavity in respect to σ for all (τ, λ), but differently from Ξ_0 it is continuously differentiable in (τ, λ).

If we define the following operator:

$$H_\varepsilon(\tau, \lambda, \sigma) = \begin{pmatrix} \nabla_{\tau,\lambda} \Xi_\varepsilon(\tau, , \lambda, \sigma) \\ -\nabla_\sigma \Xi_\varepsilon(\tau, \lambda, \sigma) \end{pmatrix}, \tag{19}$$

It is easy to see that any point $(\tau^*, \lambda^*, \sigma^*)$ such that $H_\varepsilon(\tau^*, \lambda^*, \sigma^*) = \mathbf{0}_{n+2m}$ in $\mathscr{S}_a^+$ is an approximate solution of (6) for small values of ε. Furthermore this operator has some favorable properties, as it is a monotone operator on $\mathscr{S}_a^+$ that is a convex set. The Jacobian of operator H_ε is bisymmetric and has the following structure:

$$JH_\varepsilon(\tau, \lambda, \sigma) = \begin{pmatrix} \nabla^2_{(\tau,\lambda),(\tau,\lambda)} \hat{\Xi}_\varepsilon(\tau, \lambda, \sigma) & \nabla^2_{(\tau,\lambda),\sigma} \hat{\Xi}_\varepsilon(\tau, \lambda, \sigma) \\ -\nabla^2_{(\tau,\lambda),\sigma} \hat{\Xi}_\varepsilon(\tau, \lambda, \sigma)^T & I_m \end{pmatrix}. \tag{20}$$

In the following we describe an heuristic based on the presented theory.

Algorithm 1: Canonical Duality VI approach for AQVI

`(S.0)`: Choose $(x^0, \lambda^0, \sigma^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, $\delta^0 > 0$, $\{\varepsilon^k\} \to 0$, $\gamma \in (0, 1)$, and set $k = 0$.

`(S.1)`: If $(x^k, \lambda^k, \sigma^k)$ is an approximate solution of the AQVI: STOP.

`(S.2)`: Find a solution $(x^*, \lambda^*, \sigma^*)$ of the VI$(H_{\varepsilon^k}, \mathscr{S}^+_{a,\delta^k})$, where

$$\mathscr{S}^+_{a,\delta^k} = \{(x, \lambda, \sigma) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m : \sigma_i \geq -\delta^k, i = 1, \ldots, m\},$$

using an iterative method starting from $(x^k, \lambda^k, \mathbf{0}_m)$.

`(S.3)`: Set $(x^{k+1}, \lambda^{k+1}, \sigma^{k+1}) = (x^*, \lambda^*, \sigma^*)$, $\delta^{k+1} = \gamma\delta^k$, $k \leftarrow k + 1$, and go to (S.1).

All the computations in this paper are done using Matlab 7.6.0 on a Ubuntu 12.04 64 bits PC with Intel Core i3 CPU M 370 at 2.40GHz × 4 and 3.7 GiB of RAM. In our implementation, in order to compute a solution of the VI at step `(S.2)`, we used a C version of the PATH solver with a Matlab interface downloaded from `http://pages.cs.wisc.edu/~ferris/path/` and whose detailed description can be found in [32]. We set PATH convergence tolerance equal to 1e-3. The stopping criterion at step `(S.1)` is based on the following equation reformulation of the KKT conditions of the AQVI

$$Y(x,\lambda) = \begin{pmatrix} Dx + e + A^T\lambda \\ \phi_{\mathrm{FB}}(\lambda_i, -g_i(x))_{i=1}^m \end{pmatrix}.$$

Then the main termination criterion is $\|Y(x^k,\lambda^k)\|_\infty \le 1e-4$. In the case the algorithm stops to a value that does not satisfy the termination criterion it is labelled as failure. Starting points are taken with τ^0, λ^0 and σ^0 with all zero entries. The sequence $\{\varepsilon^k\}$ is defined by $\varepsilon^0 = 1e-4$ and $\varepsilon^{k+1} = 10^{-(k+1)}\varepsilon^k$, and we set $\delta^0 = 0.1$, $\gamma = 0.1$.

In order to have an exhaustive analysis on the problem, we run several tests on the considered problems varying the coefficient of friction Φ from $1e{-}3$ to $1e5$. The coefficient of friction is important because it is the parameter that determines the difficulty of the problem. As a matter of facts in several works [33, 34] the existence of a solution for contact problems with Coulomb friction has been proved only for small values of the friction coefficient. Furthermore the convergence of the algorithm proposed in [17] has been proved, when applied to this kind of problems, only for small value of the friction. In other words the analyzed examples with the value of the Coulomb friction $\Phi \ge 10$ can be considered difficult friction contact problem instances. In Table 1 we list

- the value of the friction Φ;
- the number of iterations, which is equal to the number of VIs solved;
- the number of crash, major and minor iterations of the PATH solver;
- the number of evaluations of H;
- the number of evaluations of JH;
- elapsed CPU time in seconds;
- the value of the *KKT* violation measure $\|Y(x,\lambda)\|_\infty$ at termination.

From Table 1 it is possible to see that the method based on canonical duality reaches a good approximation of the stress vector solution in 16 instances on 18. The only instances in which the algorithm fails are those of CPCF41 with really big values of the coefficient of friction. The solution is reached in less than a second in all the proposed instances, and it is possible to notice that the running time substantially increases when the value of the coefficient exceeds 10, showing that the instances with big values of the friction coefficient are indeed difficult to solve.

Table 1 Numerical results of Algorithm 1 for the contact problems

Problem	Φ	Iter	(crash, maj, min)	H	JH	Time	$\|Y\|_\infty$
CPCF31	1e-3	2	(2, 4, 4)	8	8	0.1376	5.11300e-07
CPCF31	1e-2	2	(2, 4, 4)	8	8	0.1338	5.03590e-07
CPCF31	1e-1	1	(1, 5, 5)	7	7	0.1116	4.19531e-07
CPCF31	1e0	2	(2, 7, 7)	11	11	0.1689	1.10312e-05
CPCF31	1e1	1	(1, 5, 5)	7	7	0.1262	6.81677e-07
CPCF31	1e2	2	(2, 10, 10)	32	14	0.2547	3.54464e-07
CPCF31	1e3	2	(2, 7, 7)	35	11	0.2235	1.50271e-05
CPCF31	1e4	1	(0, 13, 310)	44	14	0.3597	3.19657e-08
CPCF31	1e5	2	(1, 20, 26)	96	23	0.4406	4.48908e-06
CPCF41	1e-3	1	(1, 6, 6)	8	8	0.2001	1.14243e-05
CPCF41	1e-2	1	(1, 6, 6)	8	8	0.1705	1.30477e-06
CPCF41	1e-1	1	(1, 6, 6)	8	8	0.1788	1.09641e-06
CPCF41	1e0	1	(1, 8, 8)	10	10	0.2215	4.32535e-05
CPCF41	1e1	1	(1, 6, 6)	8	8	0.2136	4.78402e-05
CPCF41	1e2	2	(2, 11, 11)	20	15	0.4600	2.01671e-07
CPCF41	1e3	2	(2, 14, 15)	28	18	0.5676	9.80390e-05
CPCF41	1e4	Failure					
CPCF41	1e5	Failure					

5 Conclusions

In this paper we presented a canonical duality approach to the solution of contact problem in mechanics with Coulomb friction. We formulated the contact friction problem as a quasi-variational inequality and then exploited the Fisher–Burmeister complementarity function in order to obtain a global optimization problem. Such global optimization problem is non-convex, but with canonical duality theory it is possible to define the optimality conditions of the problem and create a simple strategy that converges to a stress vector solution of the contact problem. We also presented results on some instances of such problems varying the values of the friction coefficient on a vast range, obtaining encouraging results.

In our future research we will improve both the theory and the algorithms in order to extend such approach to other problems in mechanics and improve the methods used to find a solution.

References

1. Panagiotopoulos, R.D., Fahed, A.M.: Robot hand grasping and related problems: optimal control and identification. Int. J. Robot. Res. **13**, 127–136 (1994)
2. Conry, T.F., Seireg, A.: A mathematical programming method for design for elastic bodies in contact. J. Appl. Mech. **38**, 1293–1307 (1971)
3. Friedman, V.M., Chernina, V.S.: Iterative methods applied to the solution of contact problems between bodies. Mekh. Tverd. Tela **1**, 116–120 (1967). (in Russian)
4. Duvaut, G., Lions, J.L.: Inequalities in Mechanics and Physics. Springer, Berlin (1976)
5. Ferris, M.C., Pang, J.S.: Engineering and economic applications of complementarity problems. SIAM Rev. **39**, 669–713 (1997)
6. Outrata, J., Kocvara, M., Zowe, J.: Nonsmooth approach to optimization problems with equilibrium constraints. Kluwer Academic Publishers, Dordrecht and Boston (1998)
7. Baiocchi, C., Capelo, A.: Variational and Quasivariational Inequalities: Applications to Free Boundary Problems. Wiley, New York (1984)
8. Mosco, U.: Implicit variational problems and quasi variational inequalities. In: Gossez, J., Lami Dozo, E., Mawhin, J., Waelbroeck, L. (eds.) Nonlinear Operators and the Calculus of Variations. Lecture Notes in Mathematics, vol. 543, pp. 83–156. Springer, Berlin (1976)
9. Outrata, J., Kocvara, M.: On a class of quasi-variational inequalities. Optim. Methods Softw. **5**, 275–295 (1995)
10. Pang, J.S., Fukushima, M.: Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games. Comp. Manag. Sci. **2**, 21–56 (2005). (Erratum: ibid 6, 373–375 (2009))
11. Chan, D., Pang, J.S.: The generalized quasi-variational inequality problem. Math. Oper. Res. **7**, 211–222 (1982)
12. Fukushima, M.: A class of gap functions for quasi-variational inequality problems. J. Ind. Manag. Optim. **3**, 165–171 (2007)
13. Harms, N., Kanzow, C., Stein, O.: Smoothness Properties of a Regularized Gap Function for Quasi-Variational Inequalities. Preprint 313, Institute of Mathematics, University of Würzburg, Würzburg, March 2013
14. Nesterov, Y., Scrimali, L.: Solving strongly monotone variational and quasi-variational inequalities. CORE Discussion Paper 2006/107, Catholic University of Louvain, Center for Operations Research and Econometrics (2006)
15. Noor, M.A.: On general quasi-variational inequalities. J. King Saud Univ. **24**, 81–88 (2012)
16. Ryazantseva, I.P.: First-order methods for certain quasi-variational inequalities in Hilbert space. Comput. Math. Math. Phys. **47**, 183–190 (2007)
17. Facchinei, F., Kanzow, C., Sagratella, S.: Solving quasi-variational inequalities via their KKT conditions. Math. Prog. Ser. A (2013). doi:10.1007/s10107-013-0637-0
18. Latorre, V., Sagratella, S.: A canonical duality approach for the solution of affine quasi-variational inequalities. J. Glob. Optim. (2014). doi:10.1007/s10898-014-0236-5
19. Facchinei, F., Kanzow, C., Sagratella, S., et al.: On the solution of the KKT conditions of generalized Nash equilibrium problems. SIAM J. Optim. **21**, 1082–1108 (2011)
20. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. I and II. Springer, New York (2003)
21. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory. Methods and Applications. Springer, New York (2000)
22. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
23. Gao, D.Y.: Canonical duality theory: theory, method, and applications in global optimization. Comput. Chem. **33**, 1964–1972 (2009)
24. Wang, Z.B., Fang, S.C., Gao, D.Y., et al.: Canonical dual approach to solving the maximum cut problem. J. Glob. Optim. **54**, 341–352 (2012)
25. 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)

26. 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. Springer (2010)
27. Ruan, N., Gao, D.Y.: Global optimal solutions to a general sensor network localization problem. To appear in Perform. Eval. 2013 published online at arXiv:654731
28. Latorre, V.: A potential reduction method for canonical duality, with an application to the sensor network localization problem. J. Glob. Optim. (2014). arXiv:1403.5991
29. Latorre, V., Gao, D.Y.: Canonical dual solution to nonconvex radial basis neural network optimization problem. Neurocomputing **134**, 189–197 (2014)
30. Latorre, V., Gao, D.Y.: Canonical duality for solving general nonconvex constrained problems. Opt, Lett. **10**(8), 1763–1779 (2016). arXiv:1310.2014
31. Facchinei, F., Kanzow, C., Sagratella, S.: QVILIB: a library of quasi-variational inequality test problems. Pac. J. Optim. **9**, 225–250 (2013)
32. Dirkse, S.P., Ferris, M.C.: The PATH solver: a non-monotone stabilization scheme for mixed complementarity problems. Optim. Methods Softw. **5**, 123–156 (1995)
33. Jarušek. J.: Contact problems with bounded friction. Coercive case. Czechoslov. Math. J. **33**, 237–261 (1983)
34. Nečas, J., Jarušek, J., Haslinger, J.: On the solution of the variational inequality to the Signorini-problem with small friction. Bolletino deal UMI **17B**, 796–811 (1980)

Canonical Duality Theory for Solving Nonconvex/Discrete Constrained Global Optimization Problems

Ning Ruan and David Yang Gao

Abstract This paper presents a canonical duality theory for solving general nonconvex/discrete constrained minimization problems. By using the *canonical dual transformation*, these challenging problems can be reformulated as a unified canonical dual problem (i.e., with zero duality gap) in continuous space, which can be solved easily to obtain global optimal solution. Some basic concepts and general theory in canonical systems are reviewed. Applications to Boolean least squares problems are illustrated.

1 Introduction

We start with the following general nonlinear programming problem:

$$(\mathscr{P}): \quad \min \{P(\mathbf{x}) \quad : \quad \mathbf{x} \in \mathscr{X}_k\}, \tag{1}$$

where $P(\mathbf{x})$ is a given differentiable nonconvex function, the feasible space $\mathscr{X}_k \subset \mathbb{R}^n$ is defined as

$$\mathscr{X}_k = \{\mathbf{x} \in \mathbb{R}^n | \ \mathbf{g}(\mathbf{x}) \leq 0 \in \mathbb{R}^m\},$$

where $\mathbf{g}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ is a given vector-valued differentiable (not necessarily convex) function.

The problem $(\mathscr{P})$ involves minimizing a nonconvex function over a nonconvex feasible space [23]. 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 $\mathscr{X}_k$, the classical Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m_+ \to \mathbb{R}$ is given by

N. Ruan (✉) · D.Y. Gao
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia
e-mail: n.ruan@federation.edu.au

D.Y. Gao
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics
and Mathematics 37, DOI 10.1007/978-3-319-58017-3_9

$$L(\mathbf{x}, \boldsymbol{\sigma}) = P(\mathbf{x}) + \boldsymbol{\sigma}^T \mathbf{g}(\mathbf{x}). \tag{2}$$

If $P(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$ are convex functions, 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 (i.e., Lagrange multipliers) $\boldsymbol{\sigma}$, and the Lagrangian dual problem can be defined by the Fenchel–Moreau–Rockafellar transformation

$$P^*(\boldsymbol{\sigma}) = \inf_{\mathbf{x} \in \mathbb{R}^n} L(\mathbf{x}, \boldsymbol{\sigma}). \tag{3}$$

Under certain constraint qualifications that ensure the existence of a Karush–Kuhn–Tucker (KKT) solution, we have the following strong min–max duality relation [7]:

$$\inf_{\mathbf{x} \in \mathscr{X}_k} P(\mathbf{x}) = \sup_{\boldsymbol{\sigma} \in \mathbb{R}^m_+} P^*(\boldsymbol{\sigma}). \tag{4}$$

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

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

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

The slack in the inequality (5) is called the *duality gap* in global optimization.

In order to close the duality gap inherent in the classical Lagrange duality theory, a *canonical duality theory* has been developed, first in nonconvex mechanics [19] and analysis [9], then in global optimization [10, 18]. This new theory is composed mainly of a potentially useful *canonical dual transformation* and an associated *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 the duality gap, while the triality theory can be used to identify both global and local extrema [16, 18, 19]. This theory has been used for solving quadratic minimization problems with nonconvex constraints [21].Recently, after an open problem on the double-min duality left in 2003 [11, 12]has been solved completely [5, 20], the canonical duality–triality is recognized as a powerful methodological theory in nonconvex analysis and global optimization, by several review experts, with successful applications for solving a large class of challenging problems in nonlinear dynamical systems [28], sensor localization problems [29], and finite element method for post-buckling analysis in nonconvex mechanics [4, 31].

The purpose of the present paper is to illustrate application of the canonical duality theory for solving the foregoing general minimization problem with nonconvex constraints. In the next section, some preliminary definitions are presented. In the Sect. 3, we will show how to use the canonical dual transformation to convert the nonconvex problem into a canonical dual problem. Certain particular cases are illustrated in Sect. 4. Finally, in Sect. 5 the concluding remark is presented.

2 Canonical System and Definitions

Canonical system was introduced in Gao's book [9], which provides a unified modeling for a large class of mathematical problems from real-world applications. This section presents some basic concepts and definitions needed for this paper. The original and detailed terminologies were given in [32] and the celebrated textbook by Gil Strang [33] for linear systems, and in [9] for nonlinear systems.

Let $\mathscr{X}$, $\mathscr{X}^*$ and $\mathscr{E}$, $\mathscr{E}^*$ be two pairs of real linear spaces, finite- or infinite-dimensional, in duality by the bilinear forms $\langle \mathbf{x}, \mathbf{x}^* \rangle : \mathscr{X} \times \mathscr{X}^* \to \mathbb{R}$ and $\langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle : \mathscr{E} \times \mathscr{E}^* \to \mathbb{R}$, respectively. By introducing a so-called *geometric operator* $\Lambda : \mathscr{X} \to \mathscr{E}_a \subset \mathscr{E}$ and a *balance operator* $\Upsilon : \mathscr{E}^*_a \subset \mathscr{E}^* \to \mathscr{X}^*$, we have the so-called *primal system* $\mathbb{S}_p := \{\mathscr{X}, \mathscr{E}; \Lambda\}$ and the *dual system* $\mathbb{S}_d := \{\mathscr{X}^*, \mathscr{E}^*; \Upsilon\}$. The duality relation between $\mathscr{E}_a$ and $\mathscr{E}^*_a$ is linked by a *constitutive (or physical) mapping* $C : \mathscr{E}_a \to \mathscr{E}^*_a$. Thus, the composition of $\mathbb{S}_p$ and $\mathbb{S}_d$ integrants a *first-order system* $\mathbb{S} = \{\langle \mathscr{X}, \mathscr{X}^* \rangle, \langle \mathscr{E}; \mathscr{E}^* \rangle; \Lambda, C, \Upsilon\}$ (see Sect. 4.3.2, [9]).

Definition 1. The system $\mathbb{S}$ is called a *canonical system* if the constitutive mapping C is invertible; The system $\mathbb{S}$ is called *physically nonlinear (reps. linear)* if C is nonlinear (resp. linear); The system $\mathbb{S}$ is called *geometrically nonlinear (reps. linear)* if Λ is nonlinear (resp. linear). The system $\mathbb{S}$ is called *fully nonlinear (resp. linear)* if it is both physically and geometrically nonlinear (resp. linear).

As indicated in [9] (Sect. 4.3.2) that the geometrical operator describes "topological property" of the system such that $\boldsymbol{\xi}$ can be used to measure the internal response of the system. The canonical duality relation $\boldsymbol{\xi}^* = C(\boldsymbol{\xi})$ reveals the internal physical (constitutive) behavior of the system. If Λ is an $m \times n$ matrix, then $\mathbb{S}$ is a finite-dimensional *algebraic system*. Optimization in such systems is known as *mathematical programming*. If Λ is a continuous (partial) differential operator, then $\mathbb{S}$ is an infinite-dimensional *(partial) differential system*, and optimization problems fall into the *calculus of variations*. For geometrically linear systems, the balance operator $\Upsilon = \Lambda^*$ is the adjoint operator of Λ, defined by $\langle \Lambda \mathbf{x}; \boldsymbol{\xi}^* \rangle = \langle \mathbf{x}, \Lambda^* \boldsymbol{\xi}^* \rangle$. However, for geometrically nonlinear systems, the balance operator Υ depends on a (generalized) Gâteau derivative $\Lambda_t(\mathbf{x})$ of the operator $\Lambda(\mathbf{x})$, due to the so-called *virtual work principle* [9], i.e. for any virtual variation $\delta \mathbf{x} \in \mathscr{X}$, we have

$$\langle \delta \Lambda(\mathbf{x}); \boldsymbol{\xi}^* \rangle = \langle \Lambda_t(\mathbf{x}) \delta \mathbf{x}; \boldsymbol{\xi}^* \rangle = \langle \delta \mathbf{x}, \Lambda_t^*(\mathbf{x}) \boldsymbol{\xi}^* \rangle = \langle \delta \mathbf{x}, \mathbf{x}^* \rangle.$$

Thus, the governing equations of a canonical system are

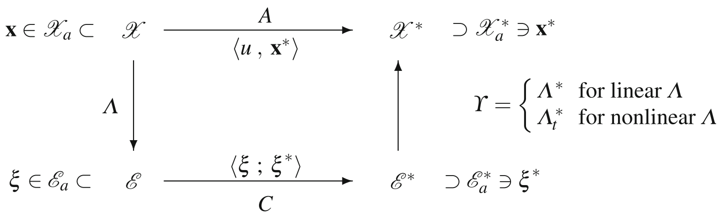


Fig. 1 Diagrammatic representation for canonical systems

(a) Geometrical equation: $\boldsymbol{\xi} = \Lambda(\mathbf{x})$
(b) Constitutive equation: $\boldsymbol{\xi}^* = C(\boldsymbol{\xi})$
(c) Balance equation: $\mathbf{x}^* = \Lambda_t^*(\mathbf{x})\boldsymbol{\xi}^*$.

A diagrammatic representation of the canonical system is shown in Fig. 1.

For conservative systems, there exists a Gâteau differentiable function $V : \mathscr{V}_a \to \mathbb{R}$ such that the constitutive relation $\boldsymbol{\xi}^* = C(\boldsymbol{\xi}) = \delta V(\boldsymbol{\xi}) : \mathscr{V}_a \to \mathscr{V}_a^*$ is invertible, where $\delta V(\boldsymbol{\xi})$ represents the Gâteau derivative of V at $\boldsymbol{\xi}$. In mathematical programming, $\delta V(\boldsymbol{\xi})$ is simply the gradient of V, denoted as ∇V. The Legendre conjugate $V^*(\boldsymbol{\xi}^*) : \mathscr{V}_a^* \to \mathbb{R}$ of V is defined by the *Legendre transformation*

$$V^*(\boldsymbol{\xi}^*) = \text{sta}\{\langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle - V(\boldsymbol{\xi}) \quad : \quad \boldsymbol{\xi} \in \mathscr{V}_a\}.$$

The notation $\text{sta}\{\mathbf{f}(\mathbf{x}) : \mathbf{x} \in \mathscr{X}_a\}$ denotes for finding stationary points of $\mathbf{f}(\mathbf{x})$ subjected to $x \in \mathscr{X}_a$.

Definition 2. A real-valued function $V : \mathscr{E}_a \subset \mathscr{E} \to \mathbb{R}$ is called a canonical function on $\mathscr{V}_a$ if its Legendre conjugate $V^*(\boldsymbol{\xi}^*)$ can be uniquely defined on $\mathscr{E}_a^* \subset \mathscr{E}^*$ such that the following canonical duality relations hold on $\mathscr{E}_a \times \mathscr{E}_a^*$:

$$\boldsymbol{\xi}^* = \delta V(\boldsymbol{\xi}) \quad \Leftrightarrow \quad \boldsymbol{\xi} = \delta V^*(\boldsymbol{\xi}^*) \quad \Leftrightarrow \quad \langle \boldsymbol{\xi} \; ; \; \boldsymbol{\xi}^* \rangle = V(\boldsymbol{\xi}) + V^*(\boldsymbol{\xi}^*). \tag{6}$$

The canonical duality lays a foundation for canonical dual transformation. For a given source (input) $\mathbf{f} = \mathbf{x}^* \in \mathscr{X}_a^*$ and necessary geometrical constraints (such as boundary–initial conditions, etc.) in $\mathscr{X}_a \subset \mathscr{X}$, the primal problem associated with the canonical system can be formulated by the following canonical form:

$$(\mathscr{P}) : \quad \min\{P(\mathbf{x}) = V(\Lambda(\mathbf{x})) - F(\mathbf{x}) | \;\; \mathbf{x} \in \mathscr{X}_a\}, \tag{7}$$

where $F(\mathbf{x}) = \langle \mathbf{x}, \mathbf{f} \rangle$ is the *external energy* [19] and the feasible set $\mathscr{X}_k = \{\mathbf{x} \in \mathscr{X}_a | \;\; \Lambda(\mathbf{x}) \in \mathscr{E}_a\}$ is the so-called *kinetically admissible space* [9]. The criticality condition $\delta P(\mathbf{x}) = 0$ leads to a general equilibrium equation

$$A(\mathbf{x}) = \Lambda_t^*(\mathbf{x}) C(\Lambda(\mathbf{x})) = \mathbf{f}. \tag{8}$$

Clearly, this equilibrium equation has a solution only if the input $\mathbf{f}$ is in the range of the equilibrium mapping $A : \mathscr{X}_a \to \mathscr{X}_a^* \subset \mathscr{X}^*$.

For dissipative systems, although the nonlinear operator $A(\mathbf{x})$ may be not a potential operator, by the least squares method $\min\{\|A(\mathbf{x}) - \mathbf{f}\|^2\}$, the equilibrium problem $A(\mathbf{x}) = \mathbf{f}$ can still be written in the canonical form $(\mathscr{P})$ with $\Lambda(\mathbf{x}) = A(\mathbf{x})$ (see [30]). Therefore, the canonical form $(\mathscr{P})$ provides a unified modeling for general nonlinear systems.

Remark 1. The geometrical nonlinearity is a well-known concept in mechanics, which means large (or finite) deformation (see Sect. 6.3 [9]). The physical nonlinearity is governed by constitutive laws of the system, which could cover many natural phenomena, such as hyperelasticity, plasticity, hysteresis, locking effects, etc. (see Sect. 3.1, [9]). By Definition 1, the physical nonlinearity in a canonical system must be monotone, i.e., the stored energy $V(\boldsymbol{\xi})$ is a convex function. Therefore, the geometrical nonlinearity is a main challenge in canonical systems, which leads to bifurcation in static systems, chaotic phenomena in dynamical systems, and the NP-hard problems in global optimization. This is the reason why the "geometrical nonlinearity" was emphasized in the title of Gao and Strang's paper [19], wherein, the stored energy $V(\boldsymbol{\xi})$ is an *objective function*, which must be nonlinear (at least quadratic), while the external energy $F(\mathbf{x})$ must be linear such that its Gâteauderivative is the external force. Objectivity is also a fundamental concept in continuum physics and nonlinear analysis (see Definition 6.1.2 [9, 15]).

The canonical systems and duality theory have been well studied in mathematical physics [32, 33] and convex analysis [7] for geometrically linear systems, where, the equilibrium operator $A = \Lambda^* C(\Lambda)$ is symmetrical. For geometrically nonlinear systems, the duality theory was first studied by Gao and Strang [19]. By using the Fenchel–Young equality $V(\boldsymbol{\xi}) = \langle \boldsymbol{\xi}; \boldsymbol{\xi}^* \rangle - V^*(\boldsymbol{\xi}^*)$ to replace $V(\Lambda(\mathbf{x}))$ in $P(\mathbf{x})$, the Gao–Strang total complementary function $\Xi : \mathscr{X}_a \times \mathscr{E}_a^* \to \mathbb{R}$ can be formulated as [19]

$$\Xi(\mathbf{x}, \boldsymbol{\xi}^*) = \langle \Lambda(\mathbf{x}); \boldsymbol{\xi}^* \rangle - V^*(\boldsymbol{\xi}^*) - F(\mathbf{x}). \tag{9}$$

By this total complementary function, the canonical dual function can be obtained by

$$P^d(\boldsymbol{\xi}^*) = \text{sta}\{\Xi(\mathbf{x}, \boldsymbol{\xi}^*) |\ \mathbf{x} \in \mathscr{X}_a\} = F^\Lambda(\boldsymbol{\xi}^*) - V^*(\boldsymbol{\xi}^*), \tag{10}$$

where $F^\Lambda(\boldsymbol{\xi}^*)$ is the Λ-transformation of $F(\xi)$ defined by

$$F^\Lambda(\boldsymbol{\xi}^*) = \text{sta}\{\langle \Lambda(\mathbf{x}); \boldsymbol{\xi}^* \rangle - F(\mathbf{x}) |\ \boldsymbol{\xi} \in \mathscr{X}_a\}.$$

If $\Lambda(\mathbf{x})$ is a quadratic operator and let $\mathbf{G}(\boldsymbol{\xi}^*) = \nabla^2_{\mathbf{x}} \Xi(\mathbf{x}, \boldsymbol{\xi}^*)$ be the Hessian of Ξ, then the canonical dual function P^d can be explicitly formulated as

$$P^d(\boldsymbol{\xi}^*) = -\frac{1}{2}\mathbf{f}^T \mathbf{G}(\boldsymbol{\xi}^*)^{-1}\mathbf{f} - V^*(\boldsymbol{\xi}^*) \tag{11}$$

which is well-defined on $\mathscr{S}_a = \{\boldsymbol{\xi}^* \in \mathscr{E}_a^* | \ \det \mathbf{G}(\boldsymbol{\xi}^*) \neq 0\}$. In order to identify global and local extremality conditions of $P^d(\boldsymbol{\xi}^*)$, we need the following two subsets:

$$\mathscr{S}_a^+ = \{\boldsymbol{\xi}^* \in \mathscr{S}_a | \ \mathbf{G}(\boldsymbol{\xi}^*) \succ 0\}, \quad \mathscr{S}_a^- = \{\boldsymbol{\xi}^* \in \mathscr{S}_a | \ \mathbf{G}(\boldsymbol{\xi}^*) \prec 0\}. \tag{12}$$

Then the canonical duality–triality theory can be presented as [9]:

Theorem 1 (**Complementary-Dual Principle**)
The problem $(\mathscr{P}^d)$ *is canonically dual to* $(\mathscr{P})$ *in the sense that if* $\bar{\boldsymbol{\xi}}^* \in \mathscr{S}_a$ *is a stationary point of* $P^d(\boldsymbol{\xi}^*)$*, then*

$$\bar{\mathbf{x}} = \mathbf{G}(\bar{\boldsymbol{\xi}}^*)^{-1}\mathbf{f} \tag{13}$$

is a stationary point of $P(\mathbf{x})$ *on* $\mathscr{X}_c$ *and* $P(\bar{\mathbf{x}}) = P^d(\bar{\boldsymbol{\xi}}^*)$.

Theorem 2 (**Triality Theory**)
Suppose that $\bar{\boldsymbol{\xi}}^* \in \mathscr{S}_a$ *is a stationary point of* $P^d(\boldsymbol{\xi}^*)$ *and* $\bar{\mathbf{x}} = \mathbf{G}(\bar{\boldsymbol{\xi}}^*)^{-1}\mathbf{f}$.

If $\bar{\boldsymbol{\xi}}^* \in \mathscr{S}_a^+$*, then it is a global maximizer of* $P^d(\boldsymbol{\xi}^*)$ *on* $\mathscr{S}_a^+$ *if and only if* $\bar{\mathbf{x}}$ *is a global minimizer of* $P(\mathbf{x})$ *on* $\mathscr{X}_c$*, i.e.,*

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathscr{X}_c} P(\mathbf{x}) \quad \Leftrightarrow \quad \max_{\boldsymbol{\xi}^* \in \mathscr{S}_a^+} P^d(\boldsymbol{\xi}^*) = P^d(\bar{\boldsymbol{\xi}}^*). \tag{14}$$

If $\bar{\boldsymbol{\xi}}^* \in \mathscr{S}_a^-$*, then on the neighborhood* $\mathscr{X}_o \times \mathscr{S}_o$ *of* $(\bar{\mathbf{x}}, \bar{\boldsymbol{\xi}}^*)$*, we have either*

$$P(\bar{\mathbf{x}}) = \max_{\mathbf{x} \in \mathscr{X}_o} P(\mathbf{x}) \quad \Leftrightarrow \quad \max_{\boldsymbol{\xi}^* \in \mathscr{S}_o} P^d(\boldsymbol{\xi}^*) = P^d(\bar{\boldsymbol{\xi}}^*), \tag{15}$$

or (only if $\dim \mathscr{X}_c = \dim \mathscr{S}_a$*)*

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathscr{X}_o} P(\mathbf{x}) \quad \Leftrightarrow \quad \min_{\boldsymbol{\xi}^* \in \mathscr{S}_o} P^d(\boldsymbol{\xi}^*) = P^d(\bar{\boldsymbol{\xi}}^*). \tag{16}$$

The Complementary-Duality Principle was originally proposed in geometrically nonlinear mechanics, which solved a 50-year old open problem in finite deformation theory and is known as the Gao principle (see [25]). The triality theory was first discovered in post-bifurcation of a large deformed beam in 1996 [8]. The condition $\dim \mathscr{X}_k = \dim \mathscr{S}_a$ for the double-min duality (16) was an open problem first discovered in 2003 [11], which was solved recently for general global optimization problems in canonical systems [20]. The canonical duality–triality theory has been used successfully for solving a large class of challenging problems in global optimization and nonconvex analysis, including the recent complete set of solutions to 3-dimensional nonlinear partial differential equations in finite deformation theory [15].

3 Canonical Dual Problem and Strong Duality

Since both $P(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$ are nonconvex functions, we need to put this problem in the framework of the canonical systems [9]. We assume that there exists Gâteaux differentiable operators $\Lambda_i,\ i = 0, 1, \cdots, m$, such that

$$\begin{aligned} \xi_0 &= \Lambda_0(\mathbf{x}), \\ \xi_i &= \Lambda_i(\mathbf{x}),\ i = 1, \cdots, m, \end{aligned}$$

and the *canonical function* $V_i,\ i = 0, \cdots, m$ such that

$$P(\mathbf{x}) = V_0(\Lambda_0(\mathbf{x})), \quad \mathbf{g}(\mathbf{x}) = \{V_i(\Lambda_i(\mathbf{x}))\}. \tag{17}$$

Then, the Lagrangian (2) can be written in the canonical form:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = V_0(\xi_0(\mathbf{x})) + \sum_{i=1}^{m} \lambda_i V_i(\xi_i(\mathbf{x})), \tag{18}$$

where, $\boldsymbol{\lambda} \in \mathscr{S}_\lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^m |\ \lambda_i \geq 0, i = 1, \cdots, m\}$ is the Lagrange multiplier, and $\boldsymbol{\xi} = [\xi_0, \xi_1, \cdots, \xi_m]^T$.

Define

$$\begin{aligned} \sigma_0 &= \nabla V_0(\xi_0), \\ \sigma_i &= \nabla V_i(\xi_i),\ i = 1, \cdots, m \end{aligned}$$

and $\boldsymbol{\sigma} = [\sigma_1, \cdots, \sigma_m]^T$, $\boldsymbol{\xi}^* = [\sigma_0, \boldsymbol{\sigma}]^T$. Then Legendre conjugate can be uniquely defined by

$$\begin{aligned} V_0^*(\sigma_0) &= \text{sta}\{\xi_0\sigma_0 - V_0(\xi_0)\ :\ \ \xi_0 \in \mathbb{R}\}, \\ V_i^*(\sigma_i) &= \text{sta}\{\xi_i\sigma_i - V_i(\xi_i)\ :\ \xi_i \in \mathbb{R}\},\ i = 1, \cdots, m. \end{aligned}$$

Let $\mathscr{S}_\sigma$ be the feasible domain of $V_0^*(\sigma_0)$ and $V_i^*(\sigma_i)$, and $\mathbf{V}^* = [V_1^*, \cdots, V_m^*]^T$, $\boldsymbol{\Lambda}(\mathbf{x}) = [\Lambda_1(\mathbf{x}), \ldots, \Lambda_m(\mathbf{x})]^T$. By the canonical dual transformation, the total complementary function can be written as

$$\Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*) = \Lambda_0(\mathbf{x})\sigma_0 - V_0^*(\sigma_0) + \boldsymbol{\lambda}^T(\boldsymbol{\Lambda}(\mathbf{x}) \circ \boldsymbol{\sigma} - \mathbf{V}^*(\boldsymbol{\sigma})), \tag{19}$$

where the symbol $\circ$ denotes the Hadamard product of two vectors, i.e.,

$$s \circ t = [s_1 t_1, \ldots, s_m t_m]^T.$$

Thus the canonical dual function can be obtained by

$$P^d(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) = U^\Lambda(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) - V_0^*(\sigma_0) - \boldsymbol{\lambda}^T \mathbf{V}^*(\boldsymbol{\sigma}),$$

where $U^\Lambda(\lambda, \boldsymbol{\xi}^*)$ is defined by

$$U^\Lambda(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) = \text{sta}\{\Lambda_0(\mathbf{x})\sigma_0 + \boldsymbol{\lambda}^T(\Lambda(\mathbf{x}) \circ \boldsymbol{\sigma}) \quad : \quad \mathbf{x} \in \mathbb{R}^n\}. \tag{20}$$

Let $\mathscr{S}_a \subset \mathscr{S}_\lambda \times \mathscr{S}_\sigma$ be the canonical dual feasible space such that $U^\Lambda(\boldsymbol{\lambda}, \boldsymbol{\xi}^*)$ is well-defined, the dual problem of $(\mathscr{P})$ can be proposed as the following:

$$(\mathscr{P}^d): \quad \max\{P^d(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \quad : \quad (\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a\}. \tag{21}$$

Theorem 3 *Suppose that the point* $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$ *is a KKT point for the total complementary function (19), then* $\bar{\mathbf{x}}$ *is a KKT point of the primal problem* $(\mathscr{P})$*, the vector* $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$ *is a KKT point of the dual problem* $(\mathscr{P}^d)$*, and*

$$P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*) = P^d(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$$

Proof. If $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$ is a KKT point of Ξ, then we have the following first-order optimality conditions [2]

$$\Xi_{\mathbf{x}}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*) = \nabla\Lambda_0(\bar{\mathbf{x}})\bar{\sigma}_0 + \bar{\boldsymbol{\lambda}}^T(\nabla\boldsymbol{\Lambda}(\bar{\mathbf{x}}) \circ \bar{\boldsymbol{\sigma}}), \tag{22}$$

$$\Xi_{\sigma_0}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*) = \Lambda_0(\bar{\mathbf{x}}) - \nabla V_0^*(\bar{\sigma}_0) = 0, \tag{23}$$

$$\Xi_{\boldsymbol{\sigma}}(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*) = \bar{\boldsymbol{\lambda}} \circ (\boldsymbol{\Lambda}(\bar{\mathbf{x}}) - \nabla\mathbf{V}^*(\bar{\boldsymbol{\sigma}})) = 0 \tag{24}$$

and the KKT conditions:

$$\bar{\boldsymbol{\lambda}} \geq 0, \quad \boldsymbol{\Lambda}(\bar{\mathbf{x}}) \circ \bar{\boldsymbol{\sigma}} - \mathbf{V}^*(\bar{\boldsymbol{\sigma}}) \leq 0, \quad \bar{\boldsymbol{\lambda}}^T(\boldsymbol{\Lambda}(\bar{\mathbf{x}}) \circ \bar{\boldsymbol{\sigma}} - \mathbf{V}^*(\bar{\boldsymbol{\sigma}})) = 0 \tag{25}$$

By the canonical duality, Eqs. (23) and (24) are equivalent to

$$\bar{\sigma}_0 = \nabla V_0(\Lambda_0(\bar{\mathbf{x}})), \;\; \bar{\boldsymbol{\sigma}} = \nabla\mathbf{V}(\boldsymbol{\Lambda}(\bar{\mathbf{x}})). \tag{26}$$

Substituting condition (26) in (22) and (25), using the chain rule of derivation on f_0, and $g_i, \; i = 1, \cdots, m$, we obtain

$$\nabla P(\bar{\mathbf{x}}) + \boldsymbol{\lambda}^T\nabla\mathbf{g}(\bar{\mathbf{x}}) = 0,$$

$$\bar{\boldsymbol{\lambda}} \geq 0, \quad \mathbf{g}(\bar{\mathbf{x}}) \leq 0, \quad \bar{\boldsymbol{\lambda}}^T\mathbf{g}(\bar{\mathbf{x}}) = 0. \tag{27}$$

This shows that $\bar{\mathbf{x}}$ is a KKT point of problem (1). Furthermore, by the complementary condition $\bar{\boldsymbol{\lambda}}^T(\boldsymbol{\Lambda}(\bar{\mathbf{x}}) \circ \bar{\boldsymbol{\sigma}} - \mathbf{V}^*(\bar{\boldsymbol{\sigma}})) = 0$ we obtain $P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$.

Dually, the Eq. (22) leads to the stationarity condition (20):

$$U^{\Lambda}(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*) = \Lambda_0(\bar{\mathbf{x}})\bar{\sigma}_0 + \bar{\boldsymbol{\lambda}}^T(\Lambda(\bar{\mathbf{x}}) \circ \bar{\boldsymbol{\sigma}})$$

Combining with conditions (23), (24), and (25), it prove that $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$ is a KKT point of dual problem (21) and $\Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*) = P^d(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$. □

In order to identify global optimal solutions to the original problem (1), we let

$$\mathscr{S}_a^+ = \{(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a | \ \mathbf{G}(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) = \nabla_x^2 \Xi_{\mathbf{x}}(\bar{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*) \succ 0, \ \boldsymbol{\lambda}_i > 0\}.$$

Theorem 4 *Suppose that canonical functions $V_i(\xi_i)$, $i = 0, 1, \cdots, m$, are convex and $\mathscr{S}_a^+$ is convex. If $(\bar{\mathbf{x}}, \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$ is a KKT point of $\Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*)$ and $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*) \in \mathscr{S}_a^+$, then $(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*)$ is a global maximizer of $P^d(\boldsymbol{\lambda}, \boldsymbol{\xi}^*)$ on $\mathscr{S}_a^+$, and $\bar{\mathbf{x}}$ is a global minimizer of $P(\mathbf{x})$ on $\mathscr{X}_k$, that is,*

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathscr{X}_k} P(\mathbf{x}) = \max_{(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a^+} P^d(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) = P^d(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\xi}}^*).$$

Proof. Since functions $V_i(\xi_m)$, $i = 0, 1, \cdots, m$ are convex, their Legendre conjugates are also convex. Due to the positivity of $\boldsymbol{\lambda}$, the total complementary function $\Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*)$ is concave in the dual variables $\boldsymbol{\xi}^* = (\sigma_0, \boldsymbol{\sigma})$, and these variables are decoupled. We have

$$\max_{(\sigma_0, \boldsymbol{\sigma})} \Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*) = \max_{\sigma_0} \max_{\boldsymbol{\sigma}} \Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*).$$

Therefore, for any given $\mathbf{x}$, the Fenchel duality leads to

$$\max_{(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a^+} \Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*) = \max_{\boldsymbol{\lambda} \in \mathbb{R}_+^m} L(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} P(\mathbf{x}) & \text{if } \mathbf{x} \in \mathscr{X}_a, \\ +\infty & \text{otherwise.} \end{cases}$$

Since $\Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*)$ is linear in $\boldsymbol{\lambda}$, if $\boldsymbol{\xi}^* \in \mathscr{S}_a^+$, then the total complementary function is convex in $\mathbf{x}$ and concave in $\boldsymbol{\xi}^*$, therefore, by the saddle min-max duality, we have

$$\begin{aligned} \min_{\mathbf{x} \in \mathscr{X}_k} P(\mathbf{x}) &= \min_{\mathbf{x} \in \mathbb{R}^n} \max_{(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a^+} \Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*) = \max_{(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a^+} \min_{\mathbf{x} \in \mathbb{R}^n} \Xi(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\xi}^*) \\ &= \max_{(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a^+} P^d(\boldsymbol{\lambda}, \boldsymbol{\xi}^*). \end{aligned}$$

This proves the theorem. □

Applications to some challenging problems will be given in the next section.

4 Boolean Least Squares Problems

The canonical duality theory can be applied to solve the following *Boolean least squares problem* :

$$(\mathscr{P}_b): \quad \min_{\mathbf{x}\in\mathscr{X}_{ip}} \left\{ P_b(\mathbf{x}) = -\frac{1}{2}\|A\mathbf{x}-\mathbf{f}\|^2 \;:\; B\mathbf{x}=\mathbf{b} \right\}. \tag{28}$$

where

$$\mathscr{X}_{ip} = \{\mathbf{x}\in\mathbb{R}^n \mid -\mathbf{1}\le\mathbf{x}\le\mathbf{1}, \;\; \mathbf{x}\in\{-1,1\}^n\},$$

and $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{m\times n}$ are two given matrices; $\mathbf{f}\in\mathbb{R}^p$ and $\mathbf{b}\in\mathbb{R}^m$ are given vectors. We assume that $m < n$ and rank $B = m$ so that the problem $(\mathscr{P}_b)$ is not overconstrained. This problem arises from a large number of applications in communication systems such as the channel decoding, multiuser detection, resource allocation in wireless systems, etc. [6, 13].

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

$$(\mathscr{P}_{bo}): \quad \min_{\mathbf{x}\in\mathscr{X}_{ip}} \left\{ P_{bo}(\mathbf{x}) = -\frac{1}{2}\|A\mathbf{x}-\mathbf{f}\|^2 \right\}. \tag{29}$$

Due to the nonconvex target function and integer constraints, traditional direct methods for solving either $(\mathscr{P}_b)$ or $(\mathscr{P}_{bo})$ are fundamentally difficult. Indeed, integer programming problems are considered to be NP-hard in global optimization and computer science.

The key step for solving integer programming problems is to reformulate the problems in the canonical form. By the fact that the integer constraints are governed by the "physical behavior" of the system, which must be written in the constitutive duality form. Therefore, by introducing a canonical measure $\boldsymbol{\xi} = \Lambda(\mathbf{x}) = \mathbf{x}\circ\mathbf{x}$: $\mathbb{R}^n \to \mathscr{E} = \mathbb{R}^n_+$ and a convex, lower semi-continuous function

$$V(\boldsymbol{\xi}) = \begin{cases} 0 & \text{if } \boldsymbol{\xi} = \mathbf{1}\in\mathbb{R}^n, \\ +\infty & \text{otherwise,} \end{cases}$$

the integer constrained problem $(\mathscr{P}_{bo})$ can be equivalently written in the canonical form

$$(\mathscr{P}_{bo}): \quad \min \left\{ V(\Lambda(\mathbf{x})) - \frac{1}{2}\|A\mathbf{x}-\mathbf{f}\|^2 \;:\; \mathbf{x}\in\mathbb{R}^n \right\}. \tag{30}$$

Since the canonical function $V(\boldsymbol{\xi})$ is not differentiable, the canonical duality relations (6) should be replaced by the generalized sub-differential forms [13]

$$\boldsymbol{\sigma}\in\partial V(\boldsymbol{\xi}) \;\;\Leftrightarrow\;\; \boldsymbol{\xi}\in\partial V^*(\boldsymbol{\sigma}) \;\;\Leftrightarrow\;\; \boldsymbol{\xi}^T\boldsymbol{\sigma} = V(\boldsymbol{\xi}) + V^*(\boldsymbol{\sigma}), \tag{31}$$

where $V^*(\boldsymbol{\sigma})$ is the Fenchel conjugate of $V(\boldsymbol{\xi})$ defined by

$$V^*(\boldsymbol{\sigma}) = \sup_{\boldsymbol{\xi}\in\mathscr{E}}\{\boldsymbol{\xi}^T\boldsymbol{\sigma} - V(\boldsymbol{\xi})\} = \begin{cases} \mathbf{1}^T\boldsymbol{\sigma} & \text{if } \boldsymbol{\sigma} \neq 0 \in \mathbb{R}^n, \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

By theory of convex analysis, the generalized canonical duality relations (31) are equivalent to

$$\boldsymbol{\xi} = \mathbf{1}, \quad \boldsymbol{\sigma} \neq 0, \quad \boldsymbol{\sigma}^T(\boldsymbol{\xi} - \mathbf{1}) = 0. \tag{33}$$

Clearly, for $\boldsymbol{\sigma} \neq 0$, the complementarity condition $\boldsymbol{\sigma}^T(\boldsymbol{\xi} - \mathbf{1}) = 0$ leads to the integer condition $\mathbf{x} \circ \mathbf{x} = \mathbf{1}$. Thus, replacing $V(\Lambda(\mathbf{x}))$ in (30) by $\Lambda(\mathbf{x})^T\boldsymbol{\sigma} - V^*(\boldsymbol{\sigma})$, the total complementary function $\Xi(\mathbf{x}, \boldsymbol{\sigma})$ of the problem $(\mathscr{P}_{bo})$ can be obtained as

$$\begin{aligned} \Xi(\mathbf{x}, \boldsymbol{\sigma}) &= \Lambda(\mathbf{x})^T\boldsymbol{\sigma} - V^*(\boldsymbol{\sigma}) - \frac{1}{2}\|A\mathbf{x} - \mathbf{f}\|^2 \\ &= \frac{1}{2}\mathbf{x}^T\mathbf{G}_a(\boldsymbol{\sigma})\mathbf{x} - \boldsymbol{\sigma}^T\mathbf{1} + \mathbf{x}^T A^T\mathbf{f} - \frac{1}{2}\|\mathbf{f}\|^2 \quad \forall \boldsymbol{\sigma} \in \mathbb{R}^n_+ \end{aligned} \tag{34}$$

where

$$\mathbf{G}_a(\boldsymbol{\sigma}) = \nabla^2_{\mathbf{x}}\Xi(\mathbf{x}, \boldsymbol{\sigma}) = -A^T A + 2\text{Diag}\,(\boldsymbol{\sigma}). \tag{35}$$

By the fact that Ξ is a quadratic function of $\mathbf{x}$, the stationarity condition $\nabla_{\mathbf{x}}\Xi(\mathbf{x}, \boldsymbol{\sigma}) = 0$ leads to $\mathbf{G}_a(\boldsymbol{\sigma})\mathbf{x} = -A^T\mathbf{f}$. Then, on the dual feasible space

$$\mathscr{S}_b = \{\boldsymbol{\sigma} \in \mathbb{R}^n | \ \ \boldsymbol{\sigma} \neq 0, \ \ \det \mathbf{G}_a(\boldsymbol{\sigma}) \neq 0\}, \tag{36}$$

the canonical dual function of $P_{bo}(\mathbf{x})$ can be formulated as

$$P^d_{bo}(\boldsymbol{\sigma}) = -\frac{1}{2}\mathbf{f}^T A[\mathbf{G}_a(\boldsymbol{\sigma})]^{-1}A^T\mathbf{f} - \sum_{i=1}^n \sigma_i - \frac{1}{2}\|\mathbf{f}\|^2. \tag{37}$$

Theorem 5 *If $\bar{\boldsymbol{\sigma}} \in \mathscr{S}_b$ is a KKT point the canonical dual function $P^d_{bo}(\boldsymbol{\sigma})$, then $\bar{\mathbf{x}} = -[\mathbf{G}_a(\bar{\boldsymbol{\sigma}})]^{-1}A^T\mathbf{f}$ is a KKT point of the Boolean least squares problem $(\mathscr{P}_{bo})$. If*

$$\bar{\boldsymbol{\sigma}} \in \mathscr{S}^+_b = \{\boldsymbol{\sigma} \in \mathscr{S}_b | \ \boldsymbol{\sigma} > 0, \ \ \mathbf{G}_a(\boldsymbol{\sigma}) \succ 0\}, \tag{38}$$

then $\bar{\mathbf{x}}$ is a global minimizer of $P_{bo}(\mathbf{x})$ on $\mathscr{X}_{ip}$ and

$$P_{bo}(\bar{\mathbf{x}}) = \min_{\mathbf{x}\in\mathscr{X}_{ip}} P_{bo}(\mathbf{x}) = \max_{\boldsymbol{\sigma}\in\mathscr{S}^+_b} P^d_{bo}(\boldsymbol{\sigma}) = P^d_{bo}(\bar{\boldsymbol{\sigma}}). \tag{39}$$

This theorem shows that the integer programming problem $(\mathscr{P}_{bo})$ is equivalent to a maximizing concave dual problem in a convex continuous space $\mathscr{S}^+_b$. If $P^d_{bo}(\boldsymbol{\sigma})$ has a stationary point in $\mathscr{S}^+_b$, the global optimal solution to the integer primal problem can be easily obtained by solving its canonical dual problem. Detailed study on the canonical duality theory for solving general integer programming problems are given in [13, 17].

Now let us consider the Boolean least squares problem $(\mathscr{P}_b)$. The equality constraint in $(\mathscr{P}_b)$ can be easily relaxed by letting

$$\mathbf{x} = \mathbf{x}_b + N_B \mathbf{x}^o, \tag{40}$$

where $\mathbf{x}_b \in \mathbb{R}^n$ is a particular solution of $B\mathbf{x} = \mathbf{b}$, i.e., $B\mathbf{x}_b = \mathbf{b}$, the matrix $N_B \in \mathbb{R}^{n\times r}$ $(r = n - m)$ is the null space of B, i.e., $BN_B\mathbf{x}^o = 0 \in \mathbb{R}^m\ \forall \mathbf{x}^o \in \mathbb{R}^r$. Thus, substituting $\mathbf{x} = \mathbf{x}_b + N_B\mathbf{x}^o$ in $\Xi(\mathbf{x}, \boldsymbol{\sigma})$, the canonical dual of the Boolean least squares problem can be formulated as [13]

$$(\mathscr{P}_b^d):\quad \max_{\boldsymbol{\sigma}\in\mathscr{S}_b^+} \left\{ P_b^d(\boldsymbol{\sigma}) = \frac{1}{2}\mathbf{x}_b^T \mathbf{G}_a(\boldsymbol{\sigma})\mathbf{x}_b - \frac{1}{2}\|\mathbf{f}\|^2 + \mathbf{x}_b^T A^T\mathbf{f} - \sum_{i=1}^n \sigma_i - G_b(\boldsymbol{\sigma}) \right\}, \tag{41}$$

where

$$G_b(\boldsymbol{\sigma}) = \frac{1}{2}(A^T\mathbf{f} + \mathbf{G}_a(\boldsymbol{\sigma})\mathbf{x}_b)^T N_B[N_B^T\mathbf{G}_a(\boldsymbol{\sigma})N_B]^{-1}N_B^T(A^T\mathbf{f} + \mathbf{G}_a(\boldsymbol{\sigma})\mathbf{x}_b).$$

Similarly to Theorem 4, we can obtain the following theorem.

Theorem 6 ([13]) *The primal problem* $(\mathscr{P}_b)$ *is canonically dual to* $(\mathscr{P}_b^d)$ *in the sense that if* $\bar{\boldsymbol{\sigma}}$ *is a KKT point of* $(\mathscr{P}_b^d)$ *and* $\bar{\boldsymbol{\sigma}} > 0$*, then the vector*

$$\bar{\mathbf{x}} = \mathbf{x}_b - N_B[N_B^T\mathbf{G}_a(\bar{\boldsymbol{\sigma}})N_B]^{-1}N_B^T(A^T\mathbf{f} + \mathbf{G}_a(\bar{\boldsymbol{\sigma}})\mathbf{x}_b) \tag{42}$$

is a KKT point of $(\mathscr{P}_b)$ *and*

$$P_b(\bar{\mathbf{x}}) = P_b^d(\bar{\boldsymbol{\sigma}}). \tag{43}$$

Moreover, if $\mathbf{G}_a(\bar{\boldsymbol{\sigma}})$ *is positive definite, then* $\bar{\boldsymbol{\sigma}}$ *is a global maximizer of* $(\mathscr{P}_b^d)$ *on* $\mathscr{S}_b^+$ *and* $\bar{\mathbf{x}}$ *is a global minimizer of* $(\mathscr{P}_b)$ *on* $\mathscr{X}_{ip}$*, i.e.,*

$$P_b(\bar{\mathbf{x}}) = \min_{\mathbf{x}\in\mathscr{X}_{ip}} P_b(\mathbf{x}) = \max_{\boldsymbol{\sigma}\in\mathscr{S}_b^+} P_b^d(\boldsymbol{\sigma}) = P_b^d(\bar{\boldsymbol{\sigma}}). \tag{44}$$

Example 1
We now consider the following 2-dimensional problem:

$$\min\ P_{bo}(x_1, x_2) = -\frac{1}{2}\|A\mathbf{x} - \mathbf{f}\|^2, \tag{45}$$

$$s.t.\ x_i^2 \le 1,\quad i = 1, 2, \tag{46}$$

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

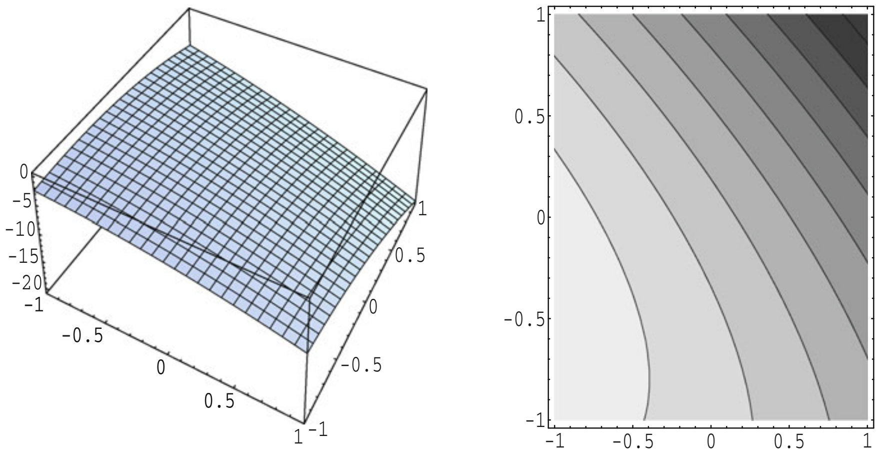


Fig. 2 Graph of the least square function P_{bo} and its contour

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

has four stationary points:

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

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

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

It is easy to check that we have only one stationary point $\boldsymbol{\sigma}_1 \in \mathscr{S}^+_b$. By Theorem 5, we know that $\mathbf{x}_1$ is a global minimizer (see Fig. 2). It is easy to verify that $P_{bo}(\mathbf{x}_k) = P^d_{bo}(\boldsymbol{\sigma}_k)\ \ k = 1, 2, 3, 4$ and

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

We note that if the inequality constraints in (46) are replaced by $x_i^2 = 1,\ \ i = 1, 2$, then the problem (45) is a Boolean least squares problem. Since all the four dual solutions $\boldsymbol{\sigma}_k \neq 0$ $(k = 1, 2, 3, 4)$, it turns out that the primal solutions $\mathbf{x}_k$ are all integer vectors.

5 Conclusions and Open Problems

We have presented applications of the canonical duality theory to several nonconvex constrained optimization problems. Our results show that by using the canonical dual transformation, these global optimization problems with nonconvex and integer constraints can be reformulated uniformly as a concave maximization dual problem in continuous space, which can be solved easily if the canonical dual has a stationary point in its convex domain $\mathscr{S}_a^+$. In this case, the global minimizer to the primal problems is unique. On the other hand, if the primal problem has a unique global minimizer, it does not ensure its canonical dual has a stationary point in $\mathscr{S}_a^+$. The existence conditions were discussed in [14]. If the canonical dual has no stationary point in $\mathscr{S}_a^+$, the primal problem could be really NP-hard, which is equivalent to a nonconvex minimal stationary point problem [14]

$$\min \operatorname{sta}\{P^d(\boldsymbol{\lambda}, \boldsymbol{\xi}^*) : \quad (\boldsymbol{\lambda}, \boldsymbol{\xi}^*) \in \mathscr{S}_a\}. \tag{47}$$

To solve this nonconvex canonical dual problem is still a challenging task. Generally speaking, if the primal problem has multiple global minimizers, its canonical dual could have multiple stationary points on the boundary of the open set $\mathscr{S}_a^+$. In this case, the canonical dual problem can be solved efficiently by perturbation methods [28, 29, 34].

Interested readers are suggested to use the idea and method presented in this article to solve many other difficult problems in global optimization [24, 27], nonconvex mechanics [1, 3], network communication [6], and scientific computations [22].

References

1. An, L.T.H., Tao, P.D.: The DC (difference of convex functions) programming and DCA revisited with DC models with real world nonconvex optimization problems. Ann. Oper. Res. **133**, 23–46 (2005)
2. Bazaraa, M.S., Sherali, H.D.: Nonlinear Programming: Theory and Algorithms, third edn. Wiley, New York (2006)
3. Burke, J.V., Lewis, A.S., Overton, M.L.: A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. SIAM J. Optim. **15**(3), 751–779 (2005)
4. Cai, K., Gao, D.Y., Qin, Q.H.: Post-buckling solutions of hyper-elastic beam by canonical dual finite element method. Math. Mech. Solids **19**(6), 659–671 (2014)
5. Chen, Y., Gao, D.Y.: Global solutions to nonconvex optimization of 4th-order polynomial and log-sum-exp functions. J. Global Optim. **64**, 417–431 (2016)
6. Chiang, M.: Geometric Programming for Communication Systems. Now Publishers Inc., Hanover (2005)
7. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland, Amsterdam (1976)
8. Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory. Appl. Mech. Rev. **50**, S64–S71 (1997)
9. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory Methods and Applications. Kluwer Academic Publishers, Dordrecht (2000)

10. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
11. Gao, D.Y.: Perfect duality theory and complete set of solutions to a class of global optimization. Optimization **52**, 467–493 (2003)
12. Gao, D.Y.: Nonconvex semi-linear problems and canonical duality solutions. In: Advances in Mechanics and Mathematics II, pp. 261–311. Kluwer Academic Publishers, Dordrecht (2003)
13. Gao, D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. J. Ind. Manag. Optim. **3**(2), 293–304 (2007)
14. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solutions for global optimization. Comput. Chem. Eng. **33**, 1964–1972 (2009)
15. Gao, D.Y.: Analytical solutions to general anti-plane shear problems in finite elasticity. Continuum Mech. Thermodyn. **28**, 175–194 (2016)
16. Gao, D.Y., Ogden, R.W.: Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem. ZAMP **59**, 498–517 (2008)
17. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. **47**(3), 463–484 (2010)
18. Gao, D.Y., Sherali, H.D.: Canonical Duality Theory: Connections between nonconvex mechanics and global optimization. In: Advance in Mechanics and Mathematics, vol. 3, pp. 257–326. Springer, (2008)
19. Gao, D.Y., Strang, G.: Geometric nonlinearity: Potential energy, complementary energy, and the gap function. Q. Appl. Math. **47**(3), 487–504 (1989)
20. 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)
21. Gao, D.Y., Ruan, N., Sherali, H.: Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality. J. Glob. Optim. **45**, 473–497 (2009)
22. Hajela, P.: Genetic search-an approach to the nonconvex optimization problem. AIAA J. **28**(7), 1205–1210 (1990)
23. Hansen, P., Jaumard, B., Ruiz, M., Xiong, J.: Global minimization of indefinite quadratic functions subjects to box constraints. Nav. Res. Logist. **40**(3), 373–392 (1993)
24. Horst, R., Pardalos, P.M., Romeijn, H.E.: Handbook of Global Optimization. Kluwer Academic Publishers, Dordrecht (2000)
25. Li, S.F., Gupta, A.: On dual configuration forces. J. Elast. **84**, 13–31 (2006)
26. Nesterov, Y.: Squared functional systems and optimization problems. In: Frenk, H., et al. (eds.) High Performance Optimization, pp. 405–440. Kluwer Academic Publishers, Dordrecht (2000)
27. Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization. Kluwer Academic Publishers, Dordrecht (2000)
28. Ruan, N., Gao, D.Y.: Canonical duality approach for nonlinear dynamical systems. IMA J. Appl. Math. **79**(2), 313–325 (2014)
29. Ruan, N., Gao, D.Y.: Global optimal solutions to a general sensor network localization problem. Perform. Eval. **75**(76), 1–16 (2014)
30. 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)
31. Santos, H.A.F.A., Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Nonlinear Mech. **47**, 240–247 (2012)
32. Tonti, E.: On the mathematical structure of a large class of physical theories. Accad. Naz. dei Lincei, Serie **VIII**, 49–56 (1972)
33. Strang, G.: Introducation to Applied Mathematics. Wellesley-Cambridge Press, Wellesley (1986)
34. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Canonical dual approach to solving the maximum cut problem. J. Glob. Optim. **54**, 341–351 (2012)

On D.C. Optimization Problems

Zhong Jin and David Yang Gao

Abstract A canonical d.c. (difference of canonical and convex functions) programming problem is proposed, which can be used to model general global optimization problems in complex systems. It shows that by using canonical duality theory, a large class of nonconvex minimization problems can be equivalently converted to a unified concave maximization problem over a convex domain, which can be solved easily under certain conditions. Additionally, a detailed proof for triality theory is provided, which can be used to identify local extremal solutions. Applications are illustrated and open problems are presented.

1 Mathematical Modeling and Objectivity

It is known that in Euclidean space every continuous global optimization problem on a compact set can be reformulated as a d.c. optimization problem, i.e., a nonconvex problem which can be described in terms of *d.c. functions* (difference of convex functions) and *d.c. sets* (difference of convex sets) [19]. By the fact that any constraint set can be equivalently relaxed by a nonsmooth indicator function, general nonconvex optimization problems can be written in the following standard d.c. programming form

$$\min\{f(x) = g(x) - h(x) \mid \forall x \in \mathscr{X}\}, \tag{1}$$

Z. Jin (✉)
Department of Mathematics, Shanghai Maritime University, 1550 Haigang Ave, Shanghai, People's Republic of China
e-mail: asdjz1234@163.com

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

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_10

where $\mathscr{X} = \mathbb{R}^n$, $g(x)$, $h(x)$ are convex proper lower-semicontinuous functions on $\mathbb{R}^n$, and the d.c. function $f(x)$ to be optimized is usually called the "objective function" in mathematical optimization. A more general model is that $g(x)$ can be an arbitrary function [19]. Clearly, this d.c. programming problem is artificial. Although it can be used to "model" a very wide range of mathematical problems [15] and has been studied extensively during the last thirty years (cf. [16, 18]), it comes at a price: it is impossible to have elegant theory and powerful algorithms for solving this problem without detailed structures on these arbitrarily given functions. As the result, even some very simple d.c. programming problems are considered as NP-hard. This dilemma is mainly due to the existing gap between mathematical optimization and mathematical physics.

The real-world applications show a simple fact, i.e., the functions $g(x)$ and $h(x)$ in the standard d.c. programming problem (1) cannot be arbitrarily given, they must obey certain fundamental laws in physics in order to model real-world systems. In Lagrange mechanics and continuum physics, a real-valued function $W : \mathscr{X} \to \mathbb{R}$ is said to be objective if and only if (see [6], Chap. 6)

$$W(x) = W(Rx) \quad \forall x \in \mathscr{X}, \quad \forall R \in \mathscr{R}, \tag{2}$$

where $\mathscr{R}$ is a special rotation group such that $R^{-1} = R^T, \quad \det R = 1, \quad \forall R \in \mathscr{R}$. Based on the original concept of objectivity, a general multi-scale mathematical model was proposed by Gao in [6]:

$$(\mathscr{P}) : \qquad \inf\{\Pi(x) = W(Dx) - F(x) \mid \forall x \in \mathscr{X}\}, \tag{3}$$

where $D : \mathscr{X} \to \mathscr{Y}$ is a linear operator; $W : \mathscr{Y} \to \mathbb{R} \cup \{+\infty\}$ is an objective function on its effective domain $\mathscr{Y}_a \subset \mathscr{Y}$, in which, certain physical constraints (such as constitutive laws, etc.) are given; correspondingly, $F : \mathscr{X} \to \mathbb{R} \cup \{-\infty\}$ is a so-called *subjective function*, which must be linear on its effective domain $\mathscr{X}_a \subset \mathscr{X}$, wherein, certain "geometrical constraints" (such as boundary/initial conditions, etc.) are given. By Riesz representation theorem, the subjective function can be written as $F(x) = \langle x, \bar{x}^* \rangle$, where $\bar{x}^* \in \mathscr{X}^*$ is a given input (or source), the bilinear form $\langle x, x^* \rangle : \mathscr{X} \times \mathscr{X}^* \to \mathbb{R}$ puts $\mathscr{X}$ and $\mathscr{X}^*$ in duality. Therefore, the extremality condition $0 \in \partial \Pi(x)$ leads to the equilibrium equation [6]

$$0 \in D^* \partial W(Dx) - \partial F(x) \quad \Leftrightarrow \quad D^* y^* - x^* = 0 \;\; \forall x^* \in \partial F(x), \quad y^* \in \partial W(y). \tag{4}$$

In this model, the objective duality relation $y^* \in \partial W(y)$ is governed by the constitutive law, which depends on mathematical modeling of the system; the subjective duality relation $x^* \in \partial F(x)$ leads to the input $\bar{x}^*$ of the system, which depends only on each given problem. Thus, the problem $(\mathscr{P})$ can be used to model general real-world applications.

Canonical duality-triality is a breakthrough theory which can be used not only for modeling complex systems within a unified framework, but also for solving real-world problems with a unified methodology. This theory was developed originally from Gao and Strang's work in nonconvex mechanics [11] and has been applied successfully for solving a large class of challenging problems in both nonconvex analysis/mechanics and global optimization, such as phase transitions in solids [12], post-buckling of large deformed beam [17], nonconvex polynomial minimization problems with box and integer constraints [8, 10, 13], Boolean and multiple integer programming [3, 20], fractional programming [4], mixed integer programming [14], polynomial optimization [9], high-order polynomial with log-sum-exp problem [1].

The goal of this paper is to apply the canonical duality theory for solving the challenging d.c. programming problem (1). The rest of this paper is arranged as follows. Based on the concept of objectivity, a canonical d.c. optimization problem and its canonical dual are formulated in the next section. Analytical solutions and triality theory for a general d.c. minimization problem with sum of nonconvex polynomial and exponential functions are discussed in Sects. 3 and 4. Four special examples are illustrated in Sect. 5. Some conclusions and future work are given in Sect. 6.

2 Canonical D.C. Problem and Its Canonical Dual

It is known that the linear operator $D : \mathscr{X} \to \mathscr{Y}$ can't change the nonconvex $W(Dx)$ to a convex function. According to the definition of the objectivity, a nonconvex function $W : \mathscr{Y} \to \mathbb{R}$ is objective if and only if there exists a function $V : \mathscr{Y} \times \mathscr{Y} \to \mathbb{R}$ such that $W(y) = V(y^T y)$. Based on this fact, a canonical transformation was proposed by Gao in 2000 [7].

Definition 1 (Canonical Transformation and Canonical Measure).
For a given nonconvex function $g : \mathscr{X} \to \mathbb{R} \cup \{\infty\}$, if there exists a nonlinear mapping $\Lambda : \mathscr{X} \to \mathscr{E}$ and a convex, l.s.c function $V : \mathscr{E} \to \mathbb{R} \cup \{\infty\}$ such that

$$g(x) = V(\Lambda(x)), \tag{5}$$

then, the nonlinear transformation (5) is called the canonical transformation and $\xi = \Lambda(x)$ is called a canonical measure.

The canonical measure $\xi = \Lambda(x)$ is also called the *geometrically admissible measure* in the canonical duality theory [7], which is not necessarily to be objective. But the most simple canonical measure in $\mathbb{R}^n$ is the quadratic function $\xi = x^T x$, which is clearly objective. Therefore, the canonical function can be viewed as a generalized objective function.

According to the canonical duality theory, the subjective function $F(x) = \langle x, \bar{x}^* \rangle$ is necessary for any given real-world system in order to have non-trivial solutions (states or outputs). Since the function $g(x)$ in the standard d.c. programming (1) could be nonconvex, it is reasonable to assume the convex function $h(x)$ in (1) is a quadratic function

$$Q(x) = \frac{1}{2}\langle x, Cx\rangle + \langle x, f\rangle, \tag{6}$$

where $C : \mathscr{X} \to \mathscr{X}^*$ is a given symmetrical positive definite operator (or matrix) and $f \in \mathscr{X}^*$ is a given input. Thus, a canonical d.c. (CDC for short) minimization problem can be proposed as the following

$$(CDC) : \quad \min\{\Pi(x) = V(\Lambda(x)) - Q(x) |\ \ x \in \mathscr{X}\} \tag{7}$$

Since the canonical measure $\xi = \Lambda(x) \in \mathscr{E}$ is nonlinear and $V(\xi)$ is convex on $\mathscr{E}$, the composition $V(\Lambda(x))$ has a higher order nonlinearity than $Q(x)$. Therefore, the coercivity for the target function $\Pi(x)$ should naturally satisfied, i.e.,

$$\lim_{\|x\|\to\infty} \{\Pi(x) = V(\Lambda(x)) - Q(x)\} = \infty \tag{8}$$

which is a necessary condition for the existence of the global minimal solution to (CDC). Clearly, this generalized d.c. minimization problem can be used to model a reasonably large class of real-world systems.

By the fact that $V(\xi)$ is convex, l.s.c. on $\mathscr{E}$, its conjugate can be uniquely defined by the Fenchel transformation

$$V^*(\xi^*) = \sup\{\langle \xi; \xi^*\rangle - V(\xi) |\ \ \xi \in \mathscr{E}\}. \tag{9}$$

The bilinear form $\langle \xi; \xi^*\rangle$ puts $\mathscr{E}$ and $\mathscr{E}^*$ in duality. According to convex analysis (cf. [2]), $V^* : \mathscr{E}^* \to \mathbb{R} \cup \{+\infty\}$ is also convex, l.s.c. on its domain $\mathscr{E}^*$ and the following generalized canonical duality relations [7] hold on $\mathscr{E} \times \mathscr{E}^*$

$$\xi^* \in \partial V(\xi) \quad \Leftrightarrow \quad \xi \in \partial V^*(\xi^*) \quad \Leftrightarrow \quad V(\xi) + V^*(\xi^*) = \langle \xi; \xi^*\rangle. \tag{10}$$

Replacing $V(\Lambda(x))$ in the target function $\Pi(x)$ by the Fenchel-Young equality $V(\xi) = \langle \xi; \xi^*\rangle - V^*(\xi^*)$, Gao and Strang's total complementary function (see [7]) $\Xi : \mathscr{X} \to \mathscr{E}^* \to \mathbb{R} \cup \{-\infty\}$ for this (CDC) can be obtained as

$$\Xi(x, \xi^*) = \langle \Lambda(x); \xi^*\rangle - V^*(\xi^*) - Q(x). \tag{11}$$

By this total complementary function, the canonical dual of $\Pi(x)$ can be obtained as

$$\Pi^d(\xi^*) = \inf\{\Xi(x, \xi^*) |\ \ x \in \mathscr{X}\} = Q^\Lambda(\xi^*) - V^*(\xi^*), \tag{12}$$

where $Q^\Lambda : \mathscr{E}^* \to \mathbb{R} \cup \{-\infty\}$ is the so-called Λ-conjugate of $Q(x)$ defined by (see [7])

$$Q^\Lambda(\xi^*) = \inf\{\langle \Lambda(x); \xi^*\rangle - Q(x) \ |\ \ x \in \mathscr{X}\}. \tag{13}$$

If this Λ-conjugate has a non-empty effective domain, the following canonical duality

$$\inf_{x\in\mathscr{X}} \Pi(x) = \sup_{\xi^*\in\mathscr{E}^*} \Pi^d(\xi^*) \tag{14}$$

holds under certain conditions, which will be illustrated in the next section.

3 Application and Analytical Solution

Let us consider a special application in $\mathbb{R}^n$ such that

$$g(x) = \sum_{i=1}^{p} \exp\left(\frac{1}{2}x^T A_i x - \alpha_i\right) + \sum_{j=1}^{r} \frac{1}{2}\left(\frac{1}{2}x^T B_j x - \beta_j\right)^2, \tag{15}$$

where $\{A_i\}_{i=1}^p \in \mathbb{R}^{n\times n}$ are symmetric matrices and $\{B_j\}_{j=1}^r \in \mathbb{R}^{n\times n}$ are symmetric positive definite matrices, α_i and β_j are real numbers. Clearly, $g : \mathbb{R}^n \to \mathbb{R}$ is nonconvex and highly nonlinear. This type of nonconvex function covers many real applications.

The canonical measure in this application can be given as

$$\xi = \begin{pmatrix} \theta \\ \eta \end{pmatrix} = \Lambda(x) = \begin{pmatrix} \left\{\frac{1}{2}x^T A_i x\right\}_{i=1}^p \\ \left\{\frac{1}{2}x^T B_j x\right\}_{j=1}^r \end{pmatrix} \; : \; \mathbb{R}^n \to \mathscr{E}_a \subseteq \mathbb{R}^m$$

where $m = p + r$. Therefore, a canonical function can be defined on $\mathscr{E}_a$:

$$V(\xi) = V_1(\theta) + V_2(\eta)$$

where

$$V_1(\theta) = \sum_{i=1}^{p} \exp\left(\theta_i - \alpha_i\right),$$

$$V_2(\eta) = \sum_{j=1}^{r} \frac{1}{2}(\eta_j - \beta_j)^2.$$

Here θ_i and η_j denote the ith component of θ and the jth component of η, respectively. Since $V_1(\theta)$ and $V_2(\eta)$ are convex, $V(\xi)$ is a convex function. By Legendre transformation, we have the following equation

$$V(\xi) + V^*(\zeta) = \xi^T \zeta, \tag{16}$$

where

$$\zeta = \begin{pmatrix} \tau \\ \sigma \end{pmatrix} = \begin{pmatrix} \nabla V_1(\theta) \\ \nabla V_2(\eta) \end{pmatrix} = \begin{pmatrix} \{\exp(\theta_i - \alpha_i)\}_{i=1}^{p} \\ \{\eta_j - \beta_j\}_{j=1}^{r} \end{pmatrix} \ : \ \mathscr{E}_a \to \mathscr{E}_a^* \subset \mathbb{R}^m$$

and $V^*(\zeta)$ is the conjugate function of $V(\xi)$, defined as

$$V^*(\zeta) = V_1^*(\tau) + V_2^*(\sigma) \tag{17}$$

with

$$V_1^*(\tau) = \sum_{i=1}^{p} (\alpha_i + \ln(\tau_i) - 1)\,\tau_i,$$
$$V_2^*(\sigma) = \frac{1}{2}\sigma^T\sigma + \beta^T\sigma,$$

where $\beta = \{\beta_j\}$.

Since the canonical measure in this application is a quadratic operator, the total complementary function $\Xi : \mathbb{R}^n \times \mathscr{E}_a^* \to \mathbb{R}$ has the following form

$$\Xi(x,\zeta) = \frac{1}{2}x^T G(\zeta)x - f^T x - V_1^*(\tau) - V_2^*(\sigma), \tag{18}$$

where

$$G(\zeta) = \sum_{i=1}^{p} \tau_i A_i + \sum_{j=1}^{r} \sigma_j B_j - C.$$

Notice that for any given ζ, the total complementary function $\Xi(x,\zeta)$ is a quadratic function of x and its stationary points are the solutions of the following equation

$$\nabla_x \Xi(x,\zeta) = G(\zeta)x - f = 0. \tag{19}$$

If $\det(G(\zeta)) \neq 0$ for a given ζ, then (19) can be solved analytically to have a unique solution $x = G(\zeta)^{-1}f$. Let

$$\mathscr{S}_a = \left\{ \zeta \in \mathscr{E}_a^* | \ \ \det(G(\zeta)) \neq 0 \right\}. \tag{20}$$

Thus, on $\mathscr{S}_a$ the canonical dual function $\Pi^d(\zeta)$ can then be written explicitly as

$$\Pi^d(\zeta) = -\frac{1}{2}f^T G(\zeta)^{-1} f - V_1^*(\tau) - V_2^*(\sigma). \tag{21}$$

Clearly, both $\Pi^d(\zeta)$ and its domain $\mathscr{S}_a$ are nonconvex. The canonical dual problem is to find all stationary points of $\Pi^d(\zeta)$ on its domain, i.e.,

$$(\mathscr{P}^d): \quad \text{sta}\left\{ \Pi^d(\zeta) \mid \zeta \in \mathscr{S}_a \right\}. \tag{22}$$

Theorem 1 (Analytic Solution and Complementary-Dual Principle). *Problem ($\mathscr{P}^d$) is canonical dual to the problem ($\mathscr{P}$) in the sense that if $\bar{\zeta} \in \mathscr{S}_a$ is a stationary point of $\Pi^d(\zeta)$, then*

$$\bar{x} = G(\bar{\zeta})^{-1} f \tag{23}$$

is a stationary point of $\Pi(x)$, the pair $(\bar{x}, \bar{\zeta})$ is a stationary point of $\Xi(x, \zeta)$, and we have

$$\Pi(\bar{x}) = \Xi(\bar{x}, \bar{\zeta}) = \Pi^d(\bar{\zeta}). \tag{24}$$

The proof of this theorem is analogous with that in [6]. Theorem 1 shows that there is no duality gap between the primal problem ($\mathscr{P}$) and the canonical dual problem ($\mathscr{P}^d$).

4 Triality Theory

In this section we will study the global optimality conditions for the critical solutions of the primal and dual problems. In order to identify both global and local extrema of both two problems, we let

$$\mathscr{S}_a^+ = \{ \zeta \in \mathscr{S}_a \mid G(\zeta) \succ 0 \},$$
$$\mathscr{S}_a^- = \{ \zeta \in \mathscr{S}_a \mid G(\zeta) \prec 0 \}.$$

where $G \succ 0$ means that G is a positive definite matrix and where $G \prec 0$ means that G is a negative definite matrix. It is easy to prove that both $\mathscr{S}_a^+$ and $\mathscr{S}_a^-$ are convex sets and

$$Q^\Lambda(\zeta) = \inf\{\langle \Lambda(x); \zeta \rangle - Q(x) | \ \ x \in \mathbb{R}^n\} = \begin{cases} -\frac{1}{2} f^T G(\zeta)^{-1} f & \text{if } \zeta \in \mathscr{S}_a^+ \\ -\infty & \text{otherwise} \end{cases} \tag{25}$$

This shows that $\mathscr{S}_a^+$ is an effective domain of $Q^\Lambda(\zeta)$.

For convenience, we first give the first and second derivatives of functions $\Pi(x)$ and $\Pi^d(\zeta)$:

$$\nabla \Pi(x) = Gx - f, \tag{26}$$

$$\nabla^2 \Pi(x) = G + Z_0 H Z_0^T, \tag{27}$$

$$\nabla \Pi^d(\zeta) = \begin{pmatrix} \left\{ \frac{1}{2} f^T G^{-1} A_i G^{-1} f - \alpha_i - \ln(\tau_i) \right\}_{i=1}^{p} \\ \left\{ \frac{1}{2} f^T G^{-1} B_j G^{-1} f - \sigma_j - \beta_j \right\}_{j=1}^{r} \end{pmatrix}, \tag{28}$$

$$\nabla^2 \Pi^d(\zeta) = -Z^T G^{-1} Z - H^{-1}, \tag{29}$$

where $Z_0, Z \in \mathbb{R}^{n \times m}$ and $H \in \mathbb{R}^{m \times m}$ are defined as

$$Z_0 = \left[A_1 x, \ldots, A_p x, B_1 x, \ldots, B_r x\right],$$
$$Z = \left[A_1 G^{-1} f, \ldots, A_p G^{-1} f, B_1 G^{-1} f, \ldots, B_r G^{-1} f\right],$$
$$H = \begin{bmatrix} \mathrm{diag}(\tau) & 0 \\ 0 & E_n \end{bmatrix},$$

where E_n is a $n \times n$ identity matrix. By the fact that $\tau > 0$, the matrix H^{-1} is positive definite.

Next we can get the lemma as follows whose proof is trivial.

Lemma 1. *If $M_1, M_2, \ldots, M_N \in \mathbb{R}^{n\times n}$ are symmetric positive semi-definite matrices, then $M = M_1 + M_2 + \ldots + M_N$ is also a positive semi-definite matrix.*

Lemma 2. *If λ_G is an arbitrary eigenvalue of G, it follows that*

$$\lambda_G \geq \sum_{i=1}^{p} \tau_i \lambda_{min}^{A_i} + \sum_{j=1}^{r} \sigma_j \bar{\lambda}^{B_j} - \lambda_{max}^{C},$$

in which $\lambda_{min}^{A_i}$ is the smallest eigenvalue of A_i, $\lambda_{max}^{C_i}$ is the largest eigenvalue of C_i, and

$$\bar{\lambda}^{B_j} = \begin{cases} \lambda_{min}^{B_j}, & \sigma_j > 0 \\ \lambda_{max}^{B_j}, & \sigma_j \leq 0, \end{cases} \tag{30}$$

where $\lambda_{min}^{B_j}$ and $\lambda_{max}^{B_j}$ are the smallest eigenvalue and the largest eigenvalue of B_j respectively.

Proof. Firstly, we need prove $\tau_i(A_i - \lambda_{min}^{A_i} E_n)$, $\lambda_{max}^{C} E_n - C$ and $\sigma_j(B_j - \bar{\lambda}^{B_j} E_n)$ are all symmetric positive semi-definite matrices.

(a) As $\lambda_{min}^{A_i}$ is the smallest eigenvalue of A_i, then $A_i - \lambda_{min}^{A_i} E_n$ is symmetric positive semi-definite, so $\tau_i(A_i - \lambda_{min}^{A_i} E_n)$ is symmetric positive semi-definite with $\tau_i = \exp(\theta_i - \alpha_i) > 0$.

(b) As λ_{max}^{C} is the largest eigenvalue of C, then $\lambda_{max}^{C} E_n - C$ is a symmetric positive semi-definite matrix.

(c)(c.1) As $\lambda_{min}^{B_j}$ is the smallest eigenvalue of B_j, then $B_j - \lambda_{min}^{B_j} E_n$ is symmetric positive semi-definite, so when $\sigma_j > 0$ it holds that $\sigma_j(B_j - \lambda_{min}^{B_j} E_n)$ is symmetric positive semi-definite.

(c.2) As $\lambda_{max}^{B_j}$ is the largest eigenvalue of B_j, then $B_j - \lambda_{max}^{B_j} E_n$ is symmetric negative semi-definite, so when $\sigma_j \leq 0$ it holds that $\sigma_j(B_j - \lambda_{max}^{B_j} E_n)$ is symmetric positive semi-definite.

From (c.1) and (c.2), we know $\sigma_j(B_j - \bar{\lambda}^{B_j} E_n)$ is always symmetric positive semi-definite.

Then by (a), (b), (c) and Lemma 1, we have

$$\sum_{i=1}^{p} \tau_i(A_i - \lambda_{min}^{A_i} E_n) + \sum_{j=1}^{r} \sigma_j(B_j - \bar{\lambda}^{B_j} E_n) + \lambda_{max}^{C} E_n - C$$

is a positive semi-definite matrix, which is equivalent to

$$G - \left(\sum_{i=1}^{p} \tau_i \lambda_{min}^{A_i} + \sum_{j=1}^{r} \sigma_j \bar{\lambda}^{B_j} E_n - \lambda_{max}^{C} \right) E_n$$

is a positive semi-definite matrix, which implies that for every eigenvalue of G, it is greater than or equal to $\sum_{i=1}^{p} \tau_i \lambda_{min}^{A_i} + \sum_{j=1}^{r} \sigma_j \bar{\lambda}^{B_j} - \lambda_{max}^{C}$. □

Based on the above lemma, the following assumption is given for the establishment of solution method.

Assumption 1 *There is a critical point $\zeta = (\tau, \sigma)$ of $\Pi^d(\zeta)$, satisfying $\Delta > 0$ where*

$$\Delta = \sum_{i=1}^{p} \tau_i \lambda_{min}^{A_i} + \sum_{j=1}^{r} \sigma_j \bar{\lambda}^{B_j} - \lambda_{max}^{C}.$$

Lemma 3. *If $\bar{\zeta}$ is a stationary point of $\Pi^d(\zeta)$ satisfying Assumption 1, then $\bar{\zeta} \in \mathcal{S}_a^+$.*

Proof. From Lemma 3, we know if λ_G is an arbitrary eigenvalue of G, it holds that $\lambda_G \geq \Delta$. If $\bar{\zeta}$ is a critical point satisfying Assumption 1, then $\Delta > 0$, so for every eigenvalue of G, we have $\lambda_G \geq \Delta > 0$, then G is a positive definite matrix, i.e., $\bar{\zeta} \in \mathcal{S}_a^+$. □

The following lemma is needed here. Its proof is omitted, which is similar to that of Lemma 6 in [5].

Lemma 4. *Suppose that $P \in \mathbb{R}^{n\times n}$, $U \in \mathbb{R}^{m\times m}$ and $W \in \mathbb{R}^{n\times m}$ are given symmetric matrices with*

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \prec 0, \quad U = \begin{bmatrix} U_{11} & 0 \\ 0 & U_{22} \end{bmatrix} \succ 0, \text{ and } W = \begin{bmatrix} W_{11} & 0 \\ 0 & 0 \end{bmatrix},$$

where P_{11}, U_{11} and W_{11} are $r \times r$-dimensional matrices, and W_{11} is nonsingular. Then,

$$-W^T P^{-1} W - U^{-1} \preceq 0 \Leftrightarrow P + WUW^T \preceq 0. \tag{31}$$

Now, we give the main result of this paper, triality theorem, which illustrates the relationships between the primal and canonical dual problems on global and local solutions under Assumption 1.

Theorem 2. (**Triality Theorem**) *Suppose that $\bar{\zeta}$ is a critical point of $\Pi^d(\zeta)$, and $\bar{x} = G(\bar{\zeta})^{-1} f$.*

1. *Min–max duality: If $\bar{\zeta}$ is the critical point satisfying Assumption 1, then the canonical min–max duality holds in the form of*

$$\Pi(\bar{x}) = \min_{x\in\mathbb{R}^n} \Pi(x) = \max_{\zeta\in\mathcal{S}_a^+} \Pi^d(\zeta) = \Pi^d(\bar{\zeta}). \tag{32}$$

2. *Double-max duality: If $\bar{\zeta} \in \mathscr{S}_a^-$, the double-max duality holds in the form that if $\bar{x}$ is a local maximizer of $\Pi(x)$ or $\bar{\zeta}$ is a local maximizer of $\Pi^d(\zeta)$, we have*

$$\Pi(\bar{x}) = \max_{x \in \mathscr{X}_0} \Pi(x) = \max_{\zeta \in \mathscr{S}_0} \Pi^d(\zeta) = \Pi^d(\bar{\zeta}) \tag{33}$$

where $\bar{x} \in \mathscr{X}_0 \subset \mathbb{R}^n$ and $\bar{\zeta} \in \mathscr{S}_0 \subset \mathscr{S}_a^-$.

3. *Double-min duality: If $\bar{\zeta} \in \mathscr{S}_a^-$, then the double-min duality holds in the form that when $m = n$, if $\bar{x}$ is a local minimizer of $\Pi(x)$ or $\bar{\zeta}$ is a local minimizer of $\Pi^d(\zeta)$, we have*

$$\Pi(\bar{x}) = \min_{x \in \mathscr{X}_0} \Pi(x) = \min_{\zeta \in \mathscr{S}_0} \Pi^d(\zeta) = \Pi^d(\bar{\zeta}) \tag{34}$$

where $\bar{x} \in \mathscr{X}_0 \subset \mathbb{R}^n$ and $\bar{\zeta} \in \mathscr{S}_0 \subset \mathscr{S}_a^-$.

Proof. 1. Because $\bar{\zeta}$ is a critical point satisfying Assumption 1, by Lemma 4 it holds $\bar{\zeta} \in \mathscr{S}_a^+$, i.e., $G(\bar{\zeta}) \succ 0$. As $G(\bar{\zeta}) \succ 0$ and $H \succ 0$, by (29) we know the Hessian of the dual function is negative definitive, i.e., $\nabla^2 \Pi^d(\zeta) \prec 0$, which implies that $\Pi^d(\zeta)$ is strictly concave over $\mathscr{S}_a^+$. Hence, we get

$$\Pi^d(\bar{\zeta}) = \max_{\zeta \in \mathscr{S}_a^+} \Pi^d(\zeta).$$

By the convexity of $V(\xi)$, we have $V(\xi) - V(\bar{\xi}) \geq (\xi - \bar{\xi})^T \nabla V(\bar{\xi}) = (\xi - \bar{\xi})^T \bar{\zeta}$ (see [11]), so

$$V(\Lambda(x)) - V(\Lambda(\bar{x})) \geq (\Lambda(x) - \Lambda(\bar{x}))^T \bar{\zeta},$$

which implies

$$\begin{aligned}\Pi(x) - \Pi(\bar{x}) &\geq (\Lambda(x) - \Lambda(\bar{x}))^T \bar{\zeta} - \frac{1}{2} x^T C x + \frac{1}{2} \bar{x}^T C \bar{x} + f^T (x - \bar{x}) \\ &= \frac{1}{2} x^T G(\bar{\zeta}) x - \frac{1}{2} \bar{x}^T G(\bar{\zeta}) \bar{x} - (x - \bar{x})^T G(\bar{\zeta}) \bar{x},\end{aligned} \tag{35}$$

Because $G(\bar{\zeta}) \succ 0$, the convexity of $\frac{1}{2} x^T G(\bar{\zeta}) x$ with respect to x in $\mathbb{R}^n$ leads to

$$\frac{1}{2} x^T G(\bar{\zeta}) x - \frac{1}{2} \bar{x}^T G(\bar{\zeta}) \bar{x} \geq (x - \bar{x})^T G(\bar{\zeta}) \bar{x}$$

Then by (35), $\Pi(x) \geq \Pi(\bar{x})$ for any $x \in \mathbb{R}^n$, which with Theorem 1 and (4) shows that the Eq. (32) is true.

2. If $\bar{\zeta}$ is a local maximizer of $\Pi^d(\zeta)$ over $\mathscr{S}_a^-$, it is true that $\nabla^2 \Pi^d(\bar{\zeta}) = -Z^T G^{-1} Z - H^{-1} \preceq 0$ and there exists a neighborhood $\mathscr{S}_0 \subset \mathscr{S}_a^-$ such that for all $\zeta \in \mathscr{S}_0$, $\nabla^2 \Pi^d(\zeta) \preceq 0$. Since the map $x = G^{-1} f$ is continuous over $\mathscr{S}_a$, the image of the map over $\mathscr{S}_0$ is a neighborhood of $\bar{x}$, which is denoted by $\mathscr{X}_0$.

Now we prove that for any $x \in \mathscr{X}_0$, $\nabla^2\Pi(x) \preceq 0$, which plus the fact that $\bar{x}$ is a critical point of $\Pi(x)$ implies $\bar{x}$ is a maximizer of $\Pi(x)$ over $\mathscr{X}_0$. By singular value decomposition, there exist orthogonal matrices $J \in \mathbb{R}^{n\times n}$, $K \in \mathbb{R}^{m\times m}$ and $R \in \mathbb{R}^{n\times m}$ with

$$R_{ij} = \begin{cases} \delta_i, & i = j \text{ and } i = 1, \ldots, r, \\ 0, & \text{otherwise}, \end{cases} \tag{36}$$

where $\delta_i > 0$ for $i = 1, \ldots, r$ and $r = \text{rank}(F)$, such that $ZH^{\frac{1}{2}} = JRK$, then

$$Z = JRKH^{-\frac{1}{2}}. \tag{37}$$

For any $x \in \mathscr{X}_0$, let ζ be a point satisfying $x = G^{-1}f$. Therefore, $\nabla^2\Pi^d(\zeta) = -Z^TG^{-1}Z - H^{-1} \preceq 0$, then it holds that

$$-H^{-\frac{1}{2}}K^TR^TJ^TG^{-1}JRKH^{-\frac{1}{2}} - H^{-1} \preceq 0. \tag{38}$$

Multiplying above inequality by $KH^{\frac{1}{2}}$ from the left and $H^{\frac{1}{2}}K^T$ from the right, it can be obtained that

$$-R^TJ^TG^{-1}JR - E_m \preceq 0, \tag{39}$$

which, by Lemma 4, is further equivalent to

$$J^TGJ + RR^T \preceq 0, \tag{40}$$

then it follows that

$$-G \succeq JRR^TJ^T = JRKH^{-\frac{1}{2}}HH^{-\frac{1}{2}}K^TR^TJ^T = ZHZ^T. \tag{41}$$

Thus, $\nabla^2\Pi(x) = G + ZHZ^T \preceq 0$, then $\bar{x}$ is a maximizer of $\Pi(x)$ over $\mathscr{X}_0$. Similarly, we can prove that if $\bar{x}$ is a maximizer of $\Pi(x)$ over $\mathscr{X}_0$, then $\bar{\zeta}$ is a maximizer of $\Pi^d(\zeta)$ over $\mathscr{S}_0$. By the Theorem 1, the Eq. (33) is proved.

3. Now we prove the double-min duality. Suppose that $\bar{\zeta}$ is a local minimizer of $\Pi^d(\zeta)$ in $\mathscr{S}_a^-$, then there exists a neighborhood $\mathscr{S}_0 \subset \mathscr{S}_a^-$ of $\bar{\zeta}$ such that for any $\zeta \in \mathscr{S}_0$, $\nabla^2\Pi^d(\zeta) \succeq 0$. Let $\mathscr{X}_0$ denote the image of the map $x = G^{-1}f$ over $\mathscr{S}_0$, which is a neighborhood of $\bar{x}$. For any $x \in \mathscr{X}_0$, let ζ be a point that satisfies $x = G^{-1}f$. It follows from $\nabla^2\Pi^d(\zeta) = -Z^TG^{-1}Z - H^{-1} \succeq 0$ that $-Z^TG^{-1}Z \succeq H^{-1} \succ 0$, which implies the matrix F is invertible. Then it is true that

$$-G^{-1} \succeq (Z^T)^{-1}H^{-1}Z^{-1}, \tag{42}$$

which is further equivalent to

$$-G \preceq ZHZ^T. \tag{43}$$

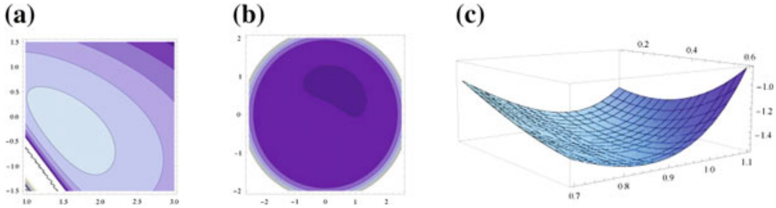


Fig. 1 The min–max duality in Example 1: **a** contour plot of function $\Pi^d(\tau,\sigma)$ near $(\bar{\tau}_1,\bar{\sigma}_1)$; **b** contour plot of function $\Pi(x,y)$; **c** graph of function $\Pi(x,y)$ near $(\bar{x}_1,\bar{y}_1)$

Thus, $\nabla^2\Pi(x) = G + ZHZ^T \succeq 0$ and x is a local minimizer of $\Pi(x)$. The converse can be proved similarly. By Theorem 1, the Eq. (34) is then true.

The theorem is proved. □

5 Examples

In this section, let $p = r = 1$. From the definition of (CDC) problem, A_1 is a symmetric matrix, B_1 and C_1 are two positive definite matrices. According to different cases of A_1, following five motivating examples are provided to illustrate the proposed canonical duality method in our paper. By examining the critical points of the dual function, we will show how the dualities in the triality theory are verified by these examples.

Example 1

We consider the case that A_1 is positive definite. Let $\alpha_1 = \beta_1 = 1$ and

$$A_1 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1.5 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix}, \text{ and } f = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

then the primal problem:

$$\min_{(x,y)\in\mathbb{R}^2} \Pi(x,y) = \exp\left(x^2 + 1.5y^2 - 1\right) + 0.5\left(0.5x^2 + 0.75y^2 - 1\right)^2 - 0.25x^2$$
$$-y^2 - x - 2y.$$

The corresponding canonical dual function is

$$\Pi^d(\tau,\sigma) = -0.5\left(\frac{1}{2\tau+\sigma-0.5} + \frac{4}{3\tau+1.5\sigma-2}\right) - \tau\ln(\tau) - 0.5\sigma^2 - \sigma.$$

so there is no duality gap, then $(\bar{x}_1,\bar{y}_1)$ is the global solution of the primal problem, which demonstrates the min–max duality(see Fig. 1).

Example 2

We consider the case that A_1 is negative definite. Let $\alpha_1 = -4$, $\beta_2 = 0.5$ and

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & -1.5 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \text{ and } f = \begin{bmatrix} 5 \\ 2 \end{bmatrix},$$

then the primal problem:

$$\min_{(x,y)\in\mathbb{R}^2} \Pi(x,y) = \left(-0.5x^2 - 0.75y^2 + 4\right) + 0.5\left(x^2 + 0.5y^2 - 0.5\right)^2 - x^2$$
$$-1.5y^2 - 5x - 2y.$$

The corresponding canonical dual function is

$$\Pi^d(\tau,\sigma) = -0.5\left(\frac{25}{-\tau + 2\sigma - 2} + \frac{4}{-1.5\tau + \sigma - 3}\right) - \tau\ln(\tau) + 5\tau - 0.5\sigma^2 - 0.5\sigma.$$

In this problem, $\lambda_{min}^{A_1} = -1.5$, $\lambda_{min}^{B_1} = 1$, $\lambda_{max}^{B_1} = 2$, and $\lambda_{max}^{C_1} = 3$. It is noticed that $(\bar{\tau}_1, \bar{\sigma}_1) = (0.145563, 3.95352)$ is a critical point of the dual function $\Pi^d(\tau,\sigma)$(see Fig. 2a). As $\bar{\sigma}_1 > 0$, we have $\bar{\lambda}^{B_1} = \lambda_{min}^{B_1}$ and

$$\Delta = \bar{\tau}_1\lambda_{min}^{A_1} + \bar{\sigma}_1\lambda_{min}^{B_1} - \lambda_{max}^{C_1} = 0.7352 > 0,$$

so Assumption 1 is satisfied, then $(\bar{\tau}_1, \bar{\sigma}_1)$ is in $\mathscr{S}_a^+$. By Theorem 1, we get $(\bar{x}_1, \bar{y}_1) = (0.867833, 2.72044)$. Moreover, we have

$$\Pi(\bar{x}_1, \bar{y}_1) = \Pi^d(\bar{\tau}_1, \bar{\sigma}_1) = -13.6736,$$

so there is no duality gap, then $(\bar{x}_1, \bar{y}_1)$ is the global solution of the primal problem, which demonstrates the min–max duality(see Fig. 2).

For showing the double-max duality of Example 2, we find a local maximum point of $\Pi^d(\tau,\sigma)$ in $\mathscr{S}_a^-$: $(\bar{\tau}_2, \bar{\sigma}_2) = (54.3685, -0.492123)$. By Theorem 1, we get $(\bar{x}_2, \bar{y}_2) = (-0.0871798, -0.023517)$. Moreover, we have

$$\Pi(\bar{x}_2, \bar{y}_2) = \Pi^d(\bar{\tau}_2, \bar{\sigma}_2) = 54.9641,$$

and $(\bar{x}_2, \bar{y}_2)$ is also a local maximum point of $\Pi(x,y)$, which demonstrates the double-max duality(see Fig. 3).

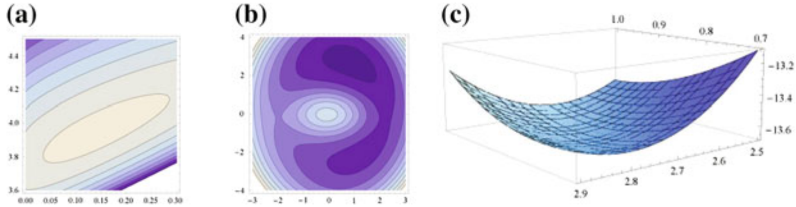


Fig. 2 The min–max duality in Example 2: **a** contour plot of function $\Pi^d(\tau,\sigma)$ near $(\bar{\tau}_1,\bar{\sigma}_1)$; **b** contour plot of function $\Pi(x,y)$; **c** graph of function $\Pi(x,y)$ near $(\bar{x}_1,\bar{y}_1)$

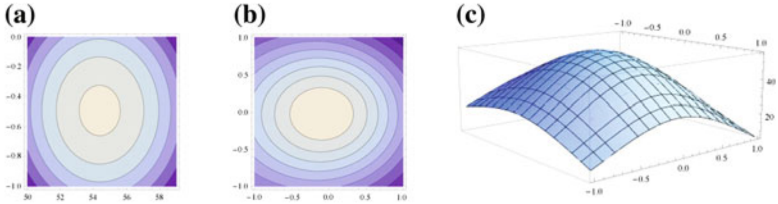


Fig. 3 The double-max duality in Example 2: **a** contour plot of function $\Pi^d(\tau,\sigma)$ near $(\bar{\tau}_2,\bar{\sigma}_2)$; **b** contour plot of function $\Pi(x,y)$ near $(\bar{x}_2,\bar{y}_2)$; **c** graph of function $\Pi(x,y)$ near $(\bar{x}_2,\bar{y}_2)$

Example 3

We consider the case that A_1 is indefinite. Let $\alpha_1=\beta_1=1$ and

$$A_1=\begin{bmatrix}1&0\\0&-2\end{bmatrix},\ \ B_1=\begin{bmatrix}1&0\\0&1\end{bmatrix},\ \ C_1=\begin{bmatrix}1.5&0\\0&1\end{bmatrix},\ \text{and}\, f=\begin{bmatrix}1\\1\end{bmatrix},$$

then the primal problem:

$$\min_{(x,y)\in\mathbb{R}^2}\Pi(x,y)=\exp\left(0.5x^2-y^2-1\right)+0.5\left(0.5x^2+0.5y^2-1\right)^2$$
$$-0.75x^2-0.5y^2-x-y.$$

The corresponding canonical dual function is

$$\Pi^d(\tau,\sigma)=-0.5\left(\frac{1}{\tau+\sigma-0.5}+\frac{1}{-2\tau+\sigma-1}\right)-\tau\ln(\tau)-0.5\sigma^2-\sigma.$$

In this problem, $\lambda_{min}^{A_1}=-2$, $\lambda_{min}^{B_1}=\lambda_{max}^{B_1}=1$, and $\lambda_{max}^{C_1}=1.5$. It is noticed that $(\bar{\tau}_1,\bar{\sigma}_1)=(0.143473, 1.91093)$ is a critical point of the dual function $\Pi^d(\tau,\sigma)$(see Fig. 4a). As $\bar{\sigma}_1>0$, we have $\bar{\lambda}^{B_1}=\lambda_{min}^{B_1}$ and

$$\Delta=\bar{\tau}_1\lambda_{min}^{A_1}+\bar{\sigma}_1\lambda_{min}^{B_1}-\lambda_{max}^{C_1}=0.1240>0,$$

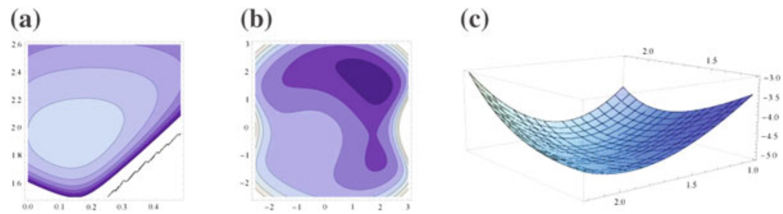


Fig. 4 The min–max duality in Example 3: **a** contour plot of function $\Pi^d(\tau, \sigma)$ near $(\bar{\tau}_1, \bar{\sigma}_1)$; **b** contour plot of function $\Pi(x, y)$; **c** graph of function $\Pi(x, y)$ near $(\bar{x}_1, \bar{y}_1)$

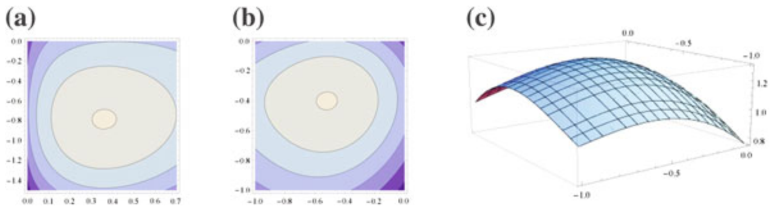


Fig. 5 The double-max duality in Example 3: **a** contour plot of function $\Pi^d(\tau, \sigma)$ near $(\bar{\tau}_2, \bar{\sigma}_2)$; **b** contour plot of function $\Pi(x, y)$ near $(\bar{x}_2, \bar{y}_2)$; **c** graph of function $\Pi(x, y)$ near $(\bar{x}_2, \bar{y}_2)$

so Assumption 1 is satisfied, then $(\bar{\tau}_1, \bar{\sigma}_1)$ is in $\mathscr{S}_a^+$. By Theorem 1, we get $(\bar{x}_1, \bar{y}_1) = (1.80375, 1.60261)$. Moreover, we have

$$\Pi(\bar{x}_1, \bar{y}_1) = \Pi^d(\bar{\tau}_1, \bar{\sigma}_1) = -5.16136,$$

so there is no duality gap, then $(\bar{x}_1, \bar{y}_1)$ is the global solution of the primal problem, which demonstrates the min–max duality(see Fig. 4).

For showing the double-max duality of Example 3, we find a local maximum point of $\Pi^d(\tau, \sigma)$ in $\mathscr{S}_a^-$: $(\bar{\tau}_2, \bar{\sigma}_2) = (0.358833, -0.785507)$. By Theorem 1, we get $(\bar{x}_2, \bar{y}_2) = (-0.519029, -0.399493)$. Moreover, we have

$$\Pi(\bar{x}_2, \bar{y}_2) = \Pi^d(\bar{\tau}_2, \bar{\sigma}_2) = 1.30402,$$

and $(\bar{x}_2, \bar{y}_2)$ is also a local maximum point of $\Pi(x, y)$, which demonstrates the double-max duality(see Fig. 5).

Example 4

We also consider the case that A_1 is indefinite. Let $\alpha_1 = 1$, $\beta_1 = 2$ and

$$A_1 = \begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 4 & 0 \\ 0 & 4.4 \end{bmatrix}, \quad \text{and } f = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

then the primal problem:

$$\min_{(x,y)\in\mathbb{R}^2} \Pi(x,y) = \exp\left(-1.5x^2+0.5y^2-1\right)+0.5\left(0.5x^2+0.5y^2-2\right)^2$$
$$-2x^2-2.2y^2-x-y.$$

The corresponding canonical dual function is

$$\Pi^d(\tau,\sigma) = -0.5\left(\frac{1}{-3\tau+\sigma-4}+\frac{1}{\tau+\sigma-4.4}\right)-\tau\ln(\tau)-0.5\sigma^2-2\sigma.$$

In this problem, $\lambda_{min}^{A_1}=-3$, $\lambda_{min}^{B_1}=\lambda_{max}^{B_1}=1$, and $\lambda_{max}^{C_1}=4.4$. It is noticed that $(\bar{\tau}_1,\bar{\sigma}_1)=(0.0612941, 4.67004)$ is a critical point of the dual function $\Pi^d(\tau,\sigma)$(see Fig. 6a). As $\bar{\sigma}_1>0$, we have $\bar{\lambda}^{B_1}=\lambda_{min}^{B_1}$ and

$$\Delta = \bar{\tau}_1\lambda_{min}^{A_1}+\bar{\sigma}_1\lambda_{min}^{B_1}-\lambda_{max}^{C_1}=0.0862>0,$$

so Assumption 1 is satisfied, then $(\bar{\tau}_1,\bar{\sigma}_1)$ is in $\mathscr{S}_a^+$. By Theorem 1, we get $(\bar{x}_1,\bar{y}_1)=(2.05695, 3.01812)$. Moreover, we have

$$\Pi(\bar{x}_1,\bar{y}_1)=\Pi^d(\bar{\tau}_1,\bar{\sigma}_1)=-22.6111,$$

so there is no duality gap, then $(\bar{x}_1,\bar{y}_1)$ is the global solution of the primal problem, which demonstrates the min–max duality(see Fig. 6).

For showing the double-max duality of Example 4, we find a local maximum point of $\Pi^d(\tau,\sigma)$ in $\mathscr{S}_a^-$: $(\bar{\tau}_2,\bar{\sigma}_2)=(0.361948,-1.97615)$. By Theorem 1, we get $(\bar{x}_2,\bar{y}_2)=(-0.141603,-0.166273)$. Moreover, we have

$$\Pi(\bar{x}_2,\bar{y}_2)=\Pi^d(\bar{\tau}_2,\bar{\sigma}_2)=2.52149,$$

and $(\bar{x}_2,\bar{y}_2)$ is also a local maximum point of $\Pi(x,y)$, which demonstrates the double-max duality(see Fig. 7).

For showing the double-min duality of Example 4, we find a local minimum point of $\Pi^d(\tau,\sigma)$ in $\mathscr{S}_a^-$: $(\bar{\tau}_3,\bar{\sigma}_3)=(0.149286, 3.90584)$. By Theorem 1, we get $(\bar{x}_3,\bar{y}_3)=(-1.84496,-2.89962)$. Moreover, we have

$$\Pi(\bar{x}_3,\bar{y}_3)=\Pi^d(\bar{\tau}_3,\bar{\sigma}_3)=-12.7833,$$

and $(\bar{x}_3,\bar{y}_3)$ is also a a local minimum point of $\Pi(x,y)$, which demonstrates the double-min duality(see Fig. 8).

From above double-min duality in Example 4, we can find our proposed canonical dual method can avoids a local minimum point $(\bar{x}_3,\bar{y}_3)$ of the primal problem. In fact, by the canonical dual method, the global solution is obtained, so any local minimum point is avoided. For instance, the point $(1.29672,-2.09209)$ is a local minimum point of the primal problem in Example 2 (see Fig. 9a), and the minimum value is -3.98411, but our proposed canonical dual method obtains the global minimum value

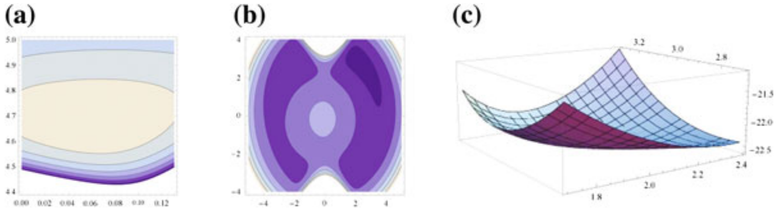


Fig. 6 The min–max duality in Example 4: **a** contour plot of function $\Pi^d(\tau, \sigma)$ near $(\bar{\tau}_1, \bar{\sigma}_1)$; **b** contour plot of function $\Pi(x, y)$; **c** graph of function $\Pi(x, y)$ near $(\bar{x}_1, \bar{y}_1)$

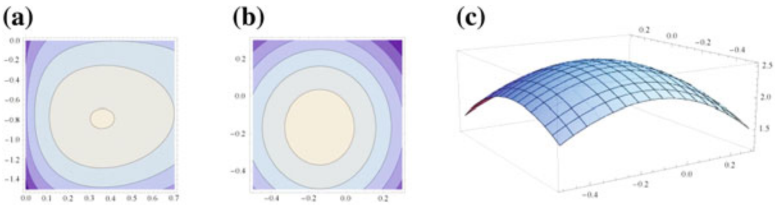


Fig. 7 The double-max duality in Example 4: **a** contour plot of function $\Pi^d(\tau, \sigma)$ near $(\bar{\tau}_2, \bar{\sigma}_2)$; **b** contour plot of function $\Pi(x, y)$ near $(\bar{x}_2, \bar{y}_2)$; **c** graph of function $\Pi(x, y)$ near $(\bar{x}_2, \bar{y}_2)$

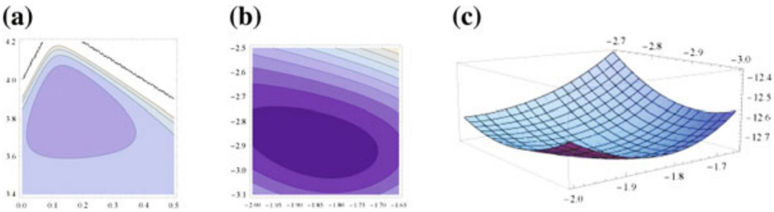


Fig. 8 The double-min duality in Example 4: **a** contour plot of function $\Pi^d(\tau, \sigma)$ near $(\bar{\tau}_3, \bar{\sigma}_3)$; **b** contour plot of function $\Pi(x, y)$ near $(\bar{x}_3, \bar{y}_3)$; **c** graph of function $\Pi(x, y)$ near $(\bar{x}_3, \bar{y}_3)$

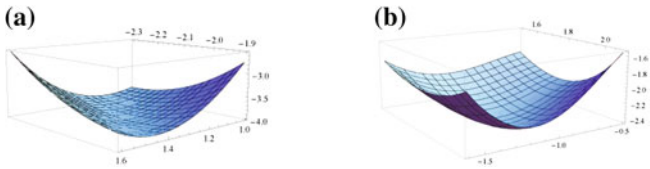


Fig. 9 graph of the primal problem near a local minimum point: **a** in Example 2; **b** in Example 3

-13.6736; the point $(1.88536, -1.10196)$ is a local minimum point of the primal problem in Example 3 (see Fig. 9b), and the minimum value is -2.45219, but our proposed canonical dual method obtains the global minimum value -22.6111.

6 Conclusions

Based on the original definition of objectivity in continuum physics, a canonical d.c. optimization problem is proposed, which can be used to model general nonconvex optimization problems in complex systems. Detailed application is provided by solving a challenging problem in $\mathbb{R}^n$. By the canonical duality theory, this nonconvex problem is able to reformulated as a concave maximization dual problem in convex domain. A detailed proof for the triality theory is provided under a reasonable assumption. This theory can be used to identify both global and local extrema, and to develop a powerful algorithm for solving this general d.c. optimization problem. Several examples are given to illustrate detailed situations. All these examples support the Assumption 1. However, we should emphasize that this assumption is only a sufficient condition for the existence of a canonical dual solution in $\mathcal{S}_a^+$. How to relax this assumption and to obtain a necessary condition for $\mathcal{S}_a^+ \neq \emptyset$ are open questions and deserve detailed study.

References

1. Chen, Y., Gao, D.Y.: Global solutions to nonconvex optimization of 4th-order polynomial and log-sum-exp functions. J. Global Optim. **64**(3), 1–15 (2016)
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 (2007)
4. Fang, S.-C., Gao, D.Y., Sheu, R.-L., Xing, W.X.: Global optimization for a class of fractional programming problems. J. Global Optim. **45**(3), 337–353 (2009)
5. Gao, D., Wu, C.: On the triality theory for a quartic polynomial optimization problem. J. Ind. Manag. Optim. **8**, 229–242 (2012)
6. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods, and Applications. Kluwer, Dordrecht (2000)
7. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Global Optim. **17**(1/4), 127–160 (2000)
8. Gao, D.Y.: Sufficient conditions and perfect duality in nonconvex minimization with inequality constraints. J. Ind. Manag. Optim. **1**(1), 59–69 (2005)
9. Gao, D.Y.: Complete solutions and extremality criteria to polynomial optimization problems. J. Global Optim. **35**, 131–143 (2006)
10. Gao, D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. J. Ind. Manag. Optim. **3**(2), 293–304 (2007)
11. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Q. J. Appl. Math. **47**(3), 487–504 (1989)
12. 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)
13. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Global Optim. **47**(3), 463–484 (2010)
14. Gao, D.Y., Ruan, N., Sherali, H.D.: Canonical duality solutions for fixed cost quadratic program. Optim. Optimal Control **39**, 139–156 (2010)
15. Hiriart-Urruty, J.B.: Generalized differentiability, duality and optimization for problems dealing with differences of convex functions. Lecture Notes in Economics and Mathematical Systems, vol. 256, pp. 37–70. Springer, Berlin (1985)

16. Horst, R., Thoai, N.V.: DC Programming: overview. J. Opt. Theory Appl. **103**, 1–43 (1999)
17. Santos, H.A.F.A., Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Nonlinear Mech. **47**, 240–247 (2011)
18. Tao, P.D., An, L.T.H.: Recent advances in DC programming and DCA. Trans. Comput. Collect. Intell. **13**, 1–37 (2014)
19. Tuy, H.: D.C. Optimization: Theory, Methods and Algorithms. In: Horst, R., Pardalos, P.M. (eds.) Handbook of Global Optimization, pp. 149–216. Kluwer Academic Publishers, Dordrecht (1995)
20. Wang, Z., Fang, S.-C., Gao, D.Y., Xing, W.: Global extremal conditions for multi-integer quadratic programming. J. Ind. Manag. Optim. **4**(2), 213–225 (2008)

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:

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

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, $\boldsymbol{f} = \{f_i\} \in \mathbb{R}^n$ is a given vector, $\langle *, *\rangle$ denotes a bilinear form in $\mathbb{R}^n \times \mathbb{R}^n$; the feasible space $\mathscr{X}_a$ is an open convex subset of $\mathbb{R}^n$ such that on which the nonconvex function $W : \mathscr{X}_a \to \mathbb{R}$ is well-defined.

Due to the nonconvexity, Problem $(\mathscr{P}_o)$ 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

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics
and Mathematics 37, DOI 10.1007/978-3-319-58017-3_11

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 ($\mathscr{P}_o$).

Generally speaking, a powerful algorithm should 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 ($\mathscr{P}_o$).

Assumption 1

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

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

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

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

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

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

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

The canonical primal problem $(\mathscr{P})$ 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 $\mathscr{V}_a$, therefore, the canonical dual mapping $\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) : \mathscr{V}_a \to \mathscr{V}_a^* \subset \mathbb{R}^m$ is one-to-one onto the convex set $\mathscr{V}_a^* \subset \mathbb{R}^m$ such that the following canonical duality relations hold on $\mathscr{V}_a \times \mathscr{V}_a^*$

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

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

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

By convex analysis, we have

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

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

$$\min_{\mathbf{x}} \max_{\boldsymbol{\varsigma}} \left\{ \Xi(\mathbf{x}, \boldsymbol{\varsigma}) \mid (\mathbf{x}, \boldsymbol{\varsigma}) \in \mathscr{X}_a \times \mathscr{V}_a^* \right\}, \tag{8}$$

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

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

in which

$$\mathbf{G}(\boldsymbol{\varsigma}) = \mathbf{A} + \sum_{k=1}^{m} \varsigma_k \mathbf{A}_k, \tag{10}$$

and

$$\boldsymbol{\tau}(\boldsymbol{\varsigma}) = \boldsymbol{f} + \sum_{k=1}^{m} \varsigma_k \boldsymbol{b}_k. \tag{11}$$

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

$$\mathbf{G}(\boldsymbol{\varsigma})\mathbf{x} = \boldsymbol{\tau}(\boldsymbol{\varsigma}). \tag{12}$$

Let

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

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

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

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

$$\mathscr{S}_a^+ = \left\{ \boldsymbol{\varsigma} \in \mathscr{V}_a^* \mid \mathbf{G}(\boldsymbol{\varsigma}) \succeq 0 \right\}, \tag{14}$$

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

$$(\mathscr{P}^d): \qquad \max\{P^d(\boldsymbol{\varsigma}) \mid \boldsymbol{\varsigma} \in \mathscr{S}_a^+\}. \tag{15}$$

The following result is due to the canonical duality theory.

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

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

is a stationary point to $(\mathscr{P})$ *and* $P(\bar{\boldsymbol{x}}) = P^d(\bar{\boldsymbol{\varsigma}})$.

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

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

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

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

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

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 $\mathscr{X}_a$ *is a bounded convex polytope subset of* $\mathbb{R}^n$*. Since* $\mathscr{X}_a$ *contains only linear constraints, both* $\mathscr{V}_a$ *and* $\mathscr{S}_a$ *are also close and bounded. Let* χ *be the smallest eigenvalue of* $\sum_{k=1}^m \varsigma_k \mathbf{A}_k$*, where* $\boldsymbol{\varsigma} = [\varsigma_1, \ldots, \varsigma_m]^T \in \mathscr{S}_a$*. Since* $\mathscr{S}_a$ *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 $\frac{1}{2}\langle \mathbf{x}, \mathbf{A}\mathbf{x} \rangle$ 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.

[1] In fact, Problem $(\mathscr{P})$ 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:

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

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 $\{\boldsymbol{a}_k\}$ 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 $\Lambda_k(\mathbf{x})=\|\mathbf{x}_k-\boldsymbol{a}_k\|^2$. 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 $\boldsymbol{\xi}=\Lambda(\bar{\boldsymbol{x}})$ and $\bar{\boldsymbol{x}}$ is an optimal solution of problem (18). It is easy to check that the dual variable $\bar{\boldsymbol{\varsigma}}=0$. Thus, $\det\mathbf{G}(\bar{\boldsymbol{\varsigma}})=0$. 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 $(\mathscr{P})$ is relaxed to the following canonical saddle point problem:

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

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

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

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

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

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

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

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

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

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

$$\Xi(\bar{\boldsymbol{x}}, \boldsymbol{\varsigma}) \leq \Xi(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}}) \;\; \forall\, \boldsymbol{\varsigma} \in \mathscr{V}_a^* \subset \mathbb{R}^m. \tag{20}$$

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

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

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

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

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

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

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

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

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

has a unique solution $\bar{\boldsymbol{\varsigma}}$ such that $\mathbf{G}(\bar{\boldsymbol{\varsigma}}) \succ 0$. Since $\mathscr{X}_a = \mathbb{R}^n$, the convex minimization problem (22) is equivalent to Problem $(\mathscr{S}_p)$ 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 $\bar{\Lambda}(\mathbf{x}) = \{\Lambda(\mathbf{x}), \frac{1}{2}\langle \mathbf{x}, \mathbf{A}\mathbf{x}\rangle\}$, then, the Gao–Strang complementary gap function is simply defined as

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

Clearly, this gap function is strictly positive for any nonzero $\mathbf{x} \in \mathscr{X}_a$ if and only if $\mathbf{G}(\boldsymbol{\varsigma}) \succ 0$. Then by Theorem 2 in [7] we know that the primal problem has a unique solution if the problem $(\mathscr{S}_p)$ is non-degenerate. By Theorem 2 and Proposition 1 we know that the nonconvex problem $(\mathscr{P})$ 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 $(\mathscr{S}_p)$ is non-degenerate. By the fact that $g = \frac{1}{2}\langle \mathbf{G}^{-1}(\boldsymbol{\varsigma})\boldsymbol{\tau}(\boldsymbol{\varsigma}), \boldsymbol{\tau}(\boldsymbol{\varsigma})\rangle$ 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 $(\mathscr{P}^d)$ defined by (15). Moreover, the canonical duality theory can also be used to find the biggest local extrema of the nonconvex problem $(\mathscr{P})$ (see [10]).

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

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

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

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

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

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

Furthermore,

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

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

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

Thus,

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

Therefore,

$$\Xi(\bar{\boldsymbol{x}},\boldsymbol{\varsigma}) \le \Xi(\bar{\boldsymbol{x}},\bar{\boldsymbol{\varsigma}}) \le \Xi(\mathbf{x},\bar{\boldsymbol{\varsigma}}),\ \ \forall\,(\mathbf{x},\boldsymbol{\varsigma}) \in \mathscr{X}_a \times \mathscr{S}_a^+.$$

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

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

$$P(\bar{\boldsymbol{x}}) = \Xi(\bar{\boldsymbol{x}},\bar{\boldsymbol{\varsigma}}) \le \Xi(\mathbf{x},\bar{\boldsymbol{\varsigma}}),\ \ \forall\,(\mathbf{x},\boldsymbol{\varsigma}) \in \mathscr{X}_a \times \mathscr{S}_a^+.$$

On the other hand,

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

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

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

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

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

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

$$\mathbf{G}(\boldsymbol{\varsigma}) = \begin{bmatrix} \varsigma_1+\varsigma_2 & \varsigma_1-\varsigma_2 \\ \varsigma_1-\varsigma_2 & \varsigma_1+\varsigma_2 \end{bmatrix},$$

$V^*(\boldsymbol{\varsigma}) = \frac{1}{2}\boldsymbol{\varsigma}^T\boldsymbol{\varsigma}$. Thus, $\mathbf{G}(\boldsymbol{\varsigma}) \succeq 0 \Leftrightarrow \varsigma_1 \ge 0$ and $\varsigma_2 \ge 0$. Clearly, $(\bar{\boldsymbol{x}},\bar{\boldsymbol{\varsigma}})$ is a saddle point of Problem $(\mathscr{S}_p)$ if and only if $(\bar{\boldsymbol{x}},\bar{\boldsymbol{\varsigma}})$ is the solution of the following variational inequality:

$$\mathbf{G}(\bar{\boldsymbol{\varsigma}})\bar{\boldsymbol{x}} = 0, \tag{25}$$

$$\langle \nabla V^*(\bar{\boldsymbol{\varsigma}}) - \Lambda(\bar{\boldsymbol{x}});\boldsymbol{\varsigma}-\bar{\boldsymbol{\varsigma}}\rangle \ge 0,\ \forall\boldsymbol{\varsigma} \ge 0. \tag{26}$$

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 $\bar{\boldsymbol{\varsigma}} = 0$. Thus, $(\bar{\boldsymbol{x}},\bar{\boldsymbol{\varsigma}})$ is a saddle point of Problem $(\mathscr{S}_p)$ if and only if $\bar{\boldsymbol{\varsigma}} = 0$ and $\bar{\boldsymbol{x}}$ satisfies

$$(x_1 + x_2)^2 \le 1,$$
$$(x_1 - x_2)^2 \le 1.$$

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

Now we turn our attention to the saddle point problem $(\mathscr{S}_p)$. 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:

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

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* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$ *of the associated Problem* $(\mathscr{S}_{p1})$ *is on the boundary of* $\mathscr{S}_a^+$, *then for any given* $\varepsilon > 0$, *there exists a* $\Delta \boldsymbol{f} \in \mathbb{R}^n$ *such that* $\|\Delta \boldsymbol{f}\| \le \varepsilon$ *and the perturbed saddle point Problem* $(\mathscr{S}_{p1})$

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

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

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

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

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

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

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 C such that

$$\mathbf{C}^T\mathbf{A}_1\mathbf{C} = \text{diag}(a_1^1, \cdots, a_n^1),$$
$$\mathbf{C}^T\mathbf{A}_2\mathbf{C} = \text{diag}(a_1^2, \cdots, a_n^2).$$

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

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

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

From Proposition 1 we know that if the solution $\bar{\boldsymbol{x}}$ of Problem $(\mathscr{P}_1)$ satisfies $\mathbf{G}(\bar{\boldsymbol{\varsigma}}) \succ 0$, then it can be obtained by simply solving the concave maximization dual problem $(\mathscr{P}^d)$. Otherwise, Proposition 2 shows that this solution can be obtained under a small perturbation. Thus, the nonconvex optimization problem $(\mathscr{P}_1)$ 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 $(\mathscr{S}_p)$. 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 $(\mathscr{S}_p)$. Clearly, Problem $(\mathscr{S}_p)$ is strictly concave with respect to $\boldsymbol{\varsigma}$. However, if Problem $(\mathscr{S}_p)$ is degenerate, i.e., $\bar{\boldsymbol{\varsigma}} \in \partial \mathscr{S}_a^+$, then Problem $(\mathscr{S}_p)$ is convex but not strictly in terms of $\mathbf{x}$. In this case, Problem $(\mathscr{S}_p)$ 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 $(\mathscr{S}_p)$, a regularized saddle point problem can be proposed as

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

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

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

where $\mu_k < \rho_k$. Note that

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

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

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

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

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

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

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

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* $\boldsymbol{\varsigma}_{k+1} = \arg\max_{\boldsymbol{\varsigma} \in \mathscr{S}^+_{\mu_k}} P^d_{\rho_k}(\boldsymbol{\varsigma})$ *and* $\mathbf{x}_{k+1} = (\mathbf{G}(\boldsymbol{\varsigma}_{k+1}) + \rho_k I)^{-1}(\rho_k \mathbf{x}_k + \boldsymbol{\tau}(\boldsymbol{\varsigma}_{k+1}))$*.*
Step 3 *If* $\|\boldsymbol{\varsigma}_{k+1} - \boldsymbol{\varsigma}_k\| \le \varepsilon$*, stop. Otherwise, set* $k = k + 1$ *and go to Step 2.*

Theorem 4. *Suppose that*

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

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

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

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

That is

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

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

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

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

Now we adopt the following steps to prove our results.

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

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

Clearly,

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

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

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

Combining (32) and (33), we obtain

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

Thus,

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

Repeating the above process, we get

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

On the other hand,

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

Substituting (36) into (35) gives rise to

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

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

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

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

On the other hand,

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

Summing the above inequalities together yields that

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

By Assumption (2) and $\mathbf{x}^k \to \bar{\boldsymbol{x}}$, we know that $\boldsymbol{\varsigma}_k$ is a bounded sequence. Now we suppose that there are two subsequences $\{\boldsymbol{\varsigma}_k^1\}$ and $\{\boldsymbol{\varsigma}_k^2\}$ of $\{\boldsymbol{\varsigma}_k\}$ such that $\{\boldsymbol{\varsigma}_k^1\} \to \boldsymbol{\varsigma}^1$ and $\{\boldsymbol{\varsigma}_k^2\} \to \boldsymbol{\varsigma}^2$. Denote $\{\mathbf{x}_k^1\}$ and $\{\mathbf{x}_k^2\}$ are two subsequences of $\{\mathbf{x}_k\}$ associated with $\{\boldsymbol{\varsigma}_k^1\}$ and $\{\boldsymbol{\varsigma}_k^2\}$. Clearly, $\boldsymbol{\varsigma}^1, \boldsymbol{\varsigma}^2 \in \mathscr{S}_a^+$. Note that

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

Thus,

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

Taking limit on both sides of the above inequality yields to

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

In a similar way, we can show that

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

Therefore,

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

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

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

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

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

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

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 $(\mathscr{P})$ is suggested as the following.

Algorithm 2

Step 1 *Adopt Algorithm 1 to solve Problem* $(\mathscr{S}_p)$*. Denote the obtained solution as* $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$.

Step 2 *If* $\|\Lambda(\bar{\boldsymbol{x}}) - \nabla V^*(\bar{\boldsymbol{\varsigma}})\| \leq \varepsilon$*, output* $\bar{\boldsymbol{x}}$ *is a global minimizer of Problem* $(\mathscr{P})$*, where* ε *is the tolerance. Otherwise, a gradient-based optimization method is used to refine Problem* $(\mathscr{P})$ *with initial condition* $\bar{\boldsymbol{x}}$.

Remark 2. Since Problem $(\mathscr{S}_p)$ is a convex–concave saddle point problem, many exact and inexact proximal point methods can be adapted [12, 14, 30]. In fact, solving Problem $(\mathscr{S}_p)$ is an easy task since it is essentially a convex optimization problem. However, to obtain a solution of Problem $(\mathscr{P})$ from the solution set of Problem $(\mathscr{S}_p)$ 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 $\mathbb{R}^n \times \mathscr{S}_a^+$ 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 $\mathbb{R}^n \times \mathscr{S}_a^+$. By the fact that $\bigcap_k \mathscr{S}_{\mu_k}^+ = \mathscr{S}_a^+$, our algorithm generates a convergent sequence and its clustering point is a saddle point of Problem $(\mathscr{S}_p)$ under certain conditions. Since $[\nabla_{\mathbf{x}}\Xi(\mathbf{x}, \boldsymbol{\varsigma}), -\nabla_{\boldsymbol{\varsigma}}\Xi(\mathbf{x}, \boldsymbol{\varsigma})]$ in $\mathbb{R}^n \times \mathscr{S}_a^+$ is not monotone for each subproblem, it is natural to approximate an optimal solution of Problem $(\mathscr{P})$ 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 $(\mathscr{P})$, 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 $\rho_k \to 0$. For classical proximal point methods, this condition was not required. In fact, this condition is adopted for simple proof that of clustering point $(\bar{x}, \bar{\varsigma})$ of the sequence $\{\mathbf{x}_k, \varsigma_k\}$ being a saddle point of Problem $(\mathscr{P})$. Our simulations show that $\rho_k \to 0$ 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 $\rho_k \to 0$.

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 $\mathbf{G}(\bar{\varsigma}) \approx 0$, 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 $\mathbf{A}_i, i = 1, \cdots, m$, are randomly generated semi-definite matrix and $d_i, i = 1, \cdots, m$, are chosen such that the optimal function value of $P(\mathbf{x})$ is 0. In fact, $\mathbf{G}(\bar{\varsigma}) = 0$ 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	$\bar{x}$	$\bar{\varsigma}$	$P(\bar{x}) = \frac{1}{2}\|\bar{\varsigma} - \Lambda(\bar{x})\|^2$
$\begin{pmatrix} 0.81472369 \\ 0.90579194 \end{pmatrix}$	$\begin{pmatrix} -1.12001364 \times 10^{-14} \\ 1.00004756 \end{pmatrix}$	$\begin{pmatrix} -3.48372378 \\ -3.48372376 \end{pmatrix} \times 10^{-9}$	$0.93735607 \times 10^{-8}$
$\begin{pmatrix} 0.60684258 \\ 0.48598247 \end{pmatrix}$	$\begin{pmatrix} 1.00004756 \\ 5.39453096 \times 10^{-14} \end{pmatrix}$	$\begin{pmatrix} -3.48358490 \\ -3.48358548 \end{pmatrix} \times 10^{-9}$	$0.93735508 \times 10^{-8}$
$\begin{pmatrix} -0.61543234 \\ -0.79193703 \end{pmatrix}$	$\begin{pmatrix} 0.56709252 \times 10^{-14} \\ -1.00004840 \end{pmatrix}$	$\begin{pmatrix} -3.48379359 \\ -3.48379378 \end{pmatrix} \times 10^{-9}$	$0.93735627 \times 10^{-8}$
$\begin{pmatrix} -0.92181297 \\ -0.73820724 \end{pmatrix}$	$\begin{pmatrix} -1.00004756 \\ 0.12834042 \times 10^{-13} \end{pmatrix}$	$\begin{pmatrix} -3.48370090 \\ -3.48370051 \end{pmatrix} \times 10^{-9}$	$0.93735602 \times 10^{-8}$

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 $\Delta \boldsymbol{f}$, the corresponding saddle point problem $(\mathscr{S}_p)$ has no solution $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$ such that $\mathbf{G}(\bar{\boldsymbol{\varsigma}}) \succ 0$ 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 $(\mathscr{P})$. 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 $(\mathscr{P})$. In fact, for all the cases in Table 2, if we set $\varepsilon = 10^{-4}$, then all the obtained results are satisfied with $\max_i |\bar{x}_i^* - x_i^{true}| \leq \varepsilon, i = 1, \cdots, n$, where $\mathbf{x}^{true} = [x_1^{true}, \cdots, x_n^{true}]^T$ is one of exact optimal solutions of Problem $(\mathscr{P})$. 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 $P(\bar{\boldsymbol{x}}) \approx 10^{-8}$ for all cases in Table 2.

It is obvious that Problem (39) has at least two solutions because of its symmetry, i.e., if $\bar{\boldsymbol{x}}$ is its solution, so is $-\bar{\boldsymbol{x}}$. 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, $a_2, \cdots, a_M \in \mathbb{R}^d$. The locations of N sensor points $x_1, x_2, \cdots, x_N \in \mathbb{R}^d$ 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 ith and the jth 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 ith sensor point and the kth 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 $i = 1, 2, \cdots, N$, 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:

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

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

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

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

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

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, $i = 1, \cdots, N$, where $\widehat{x}_i$ and x_i^* are the estimated position and true positions, respectively, $i = 1, \cdots, N$. The software package SFSDP [17] is applied for generating test problems and comparison. During our simulation, all of sensors are placed in [0, 1]×[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.

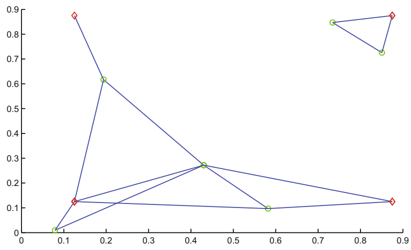


Fig. 1 Network topology of six sensors and four anchors

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

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

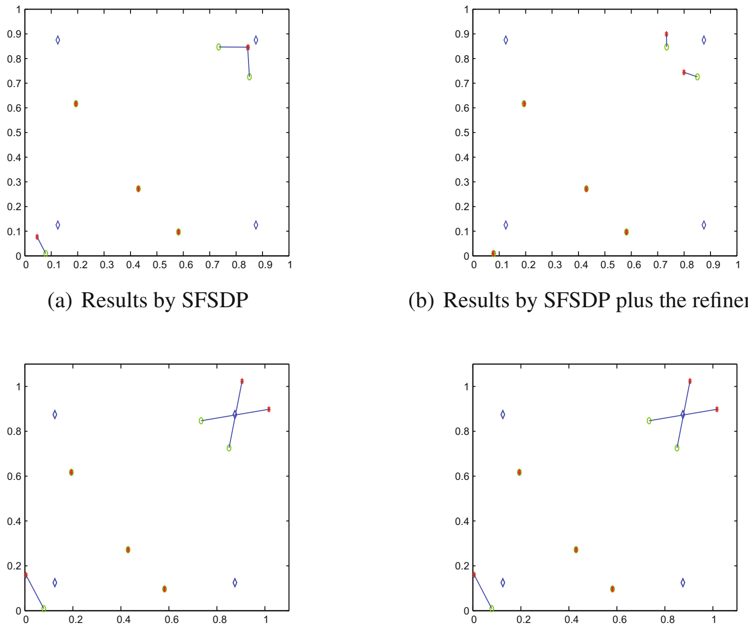
$$\widehat{e_{ik}} = \max\{(1+\xi_{ik}), 0.1\}e_{ik} \quad ((i,k) \in N_a), \tag{45}$$

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



(a) Results by SFSDP

(b) Results by SFSDP plus the refinement

(c) Results by Algorithm 1

(d) Results by Algorithm 5

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

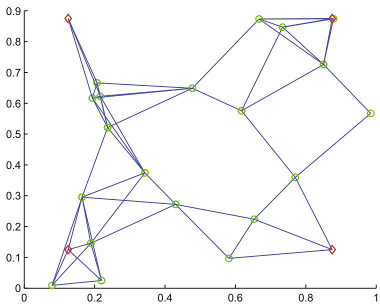


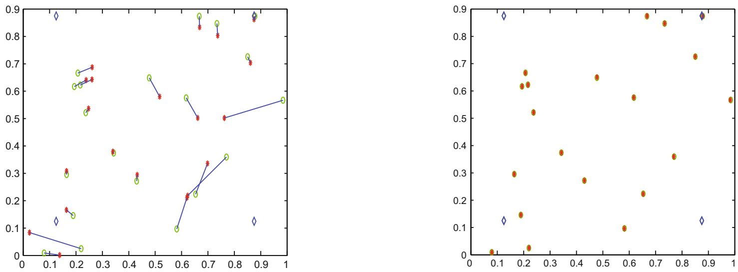
Fig. 3 Network topology of 20 sensors and 4 anchors

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.



(a) Results by SFSDP plus the refinement

(b) Results by Algorithm 5

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

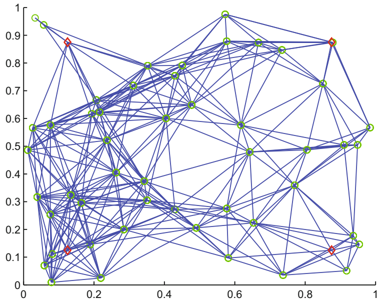
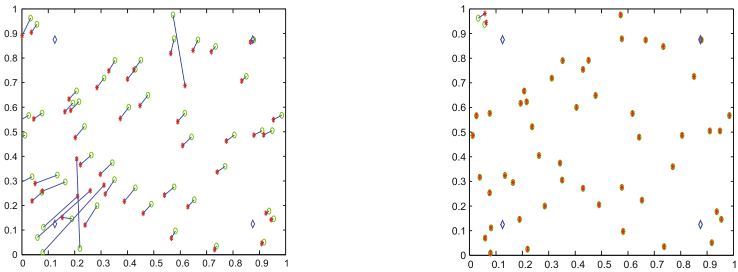


Fig. 5 Network topology of 50 sensors and 4 anchors



(a) Results by SFSDP plus the refinement

(b) Results by Algorithm 2

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 nonconvex optimization problems. Using the canonical duality theory, the original nonconvex optimization problem is first relaxed to a convex–concave saddle point optimization problem. Depending on the singularity of the matrix **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 $\|\Lambda(\bar{\boldsymbol{x}}) - \nabla V^*(\bar{\boldsymbol{\varsigma}})\| \leq \varepsilon$ within a given error tolerance, then the solution of the primal problem is also recovered exactly. Otherwise, $\bar{\boldsymbol{x}}$ 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 SFSDP: 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
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

Unified Interior Point Methodology for Canonical Duality in Global Optimization

Vittorio Latorre

Abstract We propose an interior point method to solve instances of the nonconvex optimization problems reformulated with canonical duality theory. To this aim we propose an interior point potential reduction algorithm based on the solution of the primal–dual total complementarity function. We establish the global convergence result for the algorithm under mild assumptions. Our methodology is quite general and can be applied to several problems which dual has been formulated with canonical duality theory and shows the possibility of devising efficient interior points methods for nonconvex duality.

1 Introduction

We want to introduce a framework to solve the following saddle point problem:

$$\min_{x\in\mathbb{R}^n}\max_{\sigma\in\mathbb{R}^m} \Xi(x,\sigma) = \frac{1}{2}x^T G(\sigma)x - F(\sigma)^T x - V^*(\sigma), \quad s.t. \quad G(\sigma)\succeq 0, \tag{1}$$

where $\succeq$ indicates that G is positive semidefinite, $G(\sigma)$ is a $n\times n$ symmetric matrix such that the map $G(\sigma):\mathbb{R}^m\rightarrow\mathbb{R}^{n\times n}$ is positive semidefinite convex, that is,

$$G(t\sigma_1+(1-t)\sigma_2)\succeq tG(\sigma_1)+(1-t)G(\sigma_2), \quad \forall\sigma_1,\sigma_2\in\mathbb{R}^m, \forall t\in(0,1).$$

$V^*(\sigma)$ is a convex and two times continuously differentiable function in σ. it is easy to notice that Problem (1) is convex in x for every σ such that $G(\sigma)\succeq 0$ and it is concave for every σ.

Such problem arises from the reformulation of nonconvex optimization problems in Canonical Duality Theory. Canonical duality is a methodology to formulate the dual of nonconvex optimization problems without any duality gap between the

V. Latorre (✉)
Department of Computer, Control and Management Engineering,
“Sapienza”, University of Rome, Via Ariosto 25, 00185 Rome, Italy
e-mail: latorre@dis.uniroma1.it

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_12

stationary points of the primal problem and the stationary points of the dual problem. The interest in canonical duality is not only due to the absence of duality gap, but also for the possibility to define global optimality conditions for many of such nonconvex optimization problems. In the recent years, canonical duality theory has been applied in biology, engineering, sciences [6, 16], and recently in network communications [7, 15], radial basis neural networks [10] and constrained optimization [9].

In spite of its theoretical prowess and range of applications, there are few results regarding the numerical solution of problems formulated with canonical duality theory. In [16] several mid-sized instances of the maximum cut problem are solved, to a maximum of 500 variables, with good performances in terms of speed; however, no convergence result is given. A convergence result is given in [17]; however, the assumptions on the convergence are rather strong. In a more recent work on the application of canonical duality theory to Quasi-Variational Inequalities [11], the authors reformulate problem (1) as a monotone Variational Inequality (VI) and are able to solve high-dimensional problems with several thousand of variables, without giving any convergence result, but suggesting that the methodology could have some interesting proprieties.

In this paper we partially resume the approach presented in [11]. We consider the Karush–Kunt–Tucker conditions of the monotone variational inequality associated with (1), reformulate the problem as a system of constrained equations and then prove the convergence of a potential reduction interior point method to the desired solution under mild assumptions.

The approach we consider is a potential reduction algorithm based on the damped Newton method reported in [3, 13]. The framework of this algorithm rests on six main assumptions on the operator, the feasible set, and the potential reduction merit function. The convergence result easily follows once it is proved that the proposed methodology satisfies these assumptions. The same framework has been applied to Generalized Nash Equilibrium Problems [1] and more recently to Quasi-Variational Inequalities [2], providing in both cases new important benchmarks to solve these problems.

The paper is organized as follows. In the next section we briefly show how problem (1) is obtained from general nonconvex optimization problem. In Sect. 3 we reformulate problem (1) as a system of equations, while in Sect. 4 we briefly report the key assumptions of the framework introduced in [13] and present the interior point method together with its convergence proprieties and the boundedness of the generated sequence. In Sect. 5 we report the conclusions.

Notation. For a given subset of S of $\mathbb{R}^n$ we let int S, cl S, and bd S denote, respectively, the interior, the closure, and the boundary of S; Given a set $\mathscr{A}$ we indicate with $|\mathscr{A}|$ the number of elements in $\mathscr{A}$. If the mapping $H : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable in a point x in its domain, the Jacobian matrix of H at x is denoted $JH(x)$.

The set of real matrices with n rows and m columns is defined as $\mathbb{R}^{n\times m}$; the set of $n - dimensional$ squared and symmetric matrices is denoted as $\mathscr{S}^n$; given a matrix A, we denote with a_{ij} its element on the i^{th} row and j^{th} column. The inner product defined on the set $\mathbb{R}^{n\times n}$ of squared matrices is given by

$$X \bullet Y = tr(X^T Y), \quad (X; Y) \in \mathbb{R}^{n \times n},$$

where "tr" denotes the trace of a matrix. This inner product induces the Frobenius norm for matrices given by

$$\|X\|_F = \sqrt{tr(X^T X)}, \quad X \in \mathbb{R}^{n \times n}.$$

Given a mapping $F(x, Y) : \mathbb{R}^n \times \mathscr{S}^n \to \mathbb{R}^n \times \mathscr{S}^n$ defined as

$$F(x, Y) = \begin{pmatrix} g(x, Y) \\ h(x, Y) \end{pmatrix},$$

with $g(x, Y) : \mathbb{R}^n \times \mathscr{S}^n \to \mathbb{R}^n$ and $h(x, Y) : \mathbb{R}^n \times \mathscr{S}^n \to \mathscr{S}^n$, a vector $\bar{x} \in \mathbb{R}^n$ and a matrix $\bar{Y} \in \mathscr{S}^n$, with a small abuse of notation we define the product between the mapping and the elements of $\mathbb{R}^n \times \mathscr{S}^n$ as:

$$F(x, Y) \bullet (\bar{x}, \bar{Y}) = g(x, Y)^T \bar{x} + h(x, Y) \bullet \bar{Y}.$$

The subsets of $\mathscr{S}^n$ consisting of the positive semidefinite and positive definite matrices are denoted by $\mathscr{S}^n_+$ and $\mathscr{S}^n_{++}$, respectively. For two matrices A and B in $\mathscr{S}^n$, we write $A \succeq B$ if $A - B \in \mathscr{S}^n_+$; similarly, $A \succ B$ means $A - B \in \mathscr{S}^n_{++}$; furthermore we define $\preceq$ and $\prec$ such that $A \preceq B$ if $-A \succeq -B$ and $A \prec B$ if $-A \succ -B$. $\mathbb{R}^n_+ \subset \mathbb{R}^n$ denotes the set of nonnegative numbers in $\mathbb{R}^n$; $\mathbb{R}^n_{++} \subset \mathbb{R}^n$ denotes the set of positive numbers in $\mathbb{R}^n$; sta$\{f(x) : x \in \mathscr{X}\}$ denotes the set of stationary points of function f in $\mathscr{X}$; diag(a) denotes the (square) diagonal matrix whose diagonal entries are the elements of the vector a; $vect\{A\}$ denotes the vector $\in \mathbb{R}^{n^2}$ such that the first n elements are the elements in the first column of A, the elements from $n + 1$ to $2n$ are the elements in the second column of A and so on till the last n elements that correspond to the elements in the n^{th} column of A; $\circ$ denotes the Hadamard (component-wise) product operator; and $\mathbf{0}_n$ denotes the origin in $\mathbb{R}^n$, likewise $\mathbf{0}_{n \times m}$ denotes the origin in $\mathbb{R}^{n \times m}$. If no index is indicated, the dimension of $\mathbf{0}$ is deduced from the context; $\mathbf{1}_n$ denotes the vectors of all ones in $\mathbb{R}^n$; I_n denotes the identity matrix in $\mathbb{R}^{n \times n}$.

2 Problem Description

Canonical duality theory is applied to the following general nonconvex problem:

$$(\mathscr{P}) : \min_{x \in \mathbb{R}^n} \left\{ \Pi(x) = W(x) + \frac{1}{2} x^T A x - c^T x \right\},$$

where $W(x)$ is a nonconvex term in the objective function, $A \in \mathscr{S}^n$ and $c \in \mathbb{R}^n$. The canonical dual transformation can be applied if the following assumption is satisfied:

Assumption 1 *There exists a nonlinear operator*

$$\xi = \Lambda(x) : \mathbb{R}^n \to \mathscr{E}_a \subseteq \mathbb{R}^m$$

function of x, such that the nonconvex functional $W(x)$ can be rewritten as

$$W(x) = V(\Lambda(x)) = V(\xi) : \mathscr{E}_a \to \mathbb{R}, \tag{2}$$

where V is a convex and differentiable function in ξ.

If Assumption 1 is satisfied, the primal problem can be rewritten in the following form:

$$\min_{x \in \mathbb{R}^n} \left\{ \Pi(x) = V(\Lambda(x)) + \frac{1}{2} x^T A x - c^T x \right\}.$$

As $V(\xi)$ is convex and differentiable, it is possible to apply the Legendre transformation, and write the total complementarily function in the primal variable x and dual variable $\sigma \in \mathscr{S}_a \subseteq \mathbb{R}^m$:

$$\Xi(x, \sigma) = \Lambda(x)^T \sigma - V^*(\sigma) + \frac{1}{2} x^T A x - c^T x,$$

where $V^*(\sigma)$ is the Fenchel conjugate of $V(\xi)$.

In many real-world applications, the geometrically nonlinear operator $\Lambda(x)$ is usually a quadratic function, say

$$\Lambda(x) = \left\{ \frac{1}{2} x^T C_k x - x^T b_k \right\}^m : \mathbb{R}^n \to \mathscr{E}_a \subset \mathbb{R}^m. \tag{3}$$

In the following we focus on the transformation for a general quadratic operator. With operator (3) the total complementarity function can be reformulated as

$$\begin{aligned} \Xi(x, \sigma) &= \frac{1}{2} x^T G(\sigma) x - F(\sigma)^T x - V^*(\sigma), \\ G(\sigma) &= A + \sum_{k=1}^m C_k \sigma_k, \quad F(\sigma) = c + \sum_{k=1}^m \sigma_k b_k. \end{aligned} \tag{4}$$

The dual is obtained by exploiting the stationarity conditions of (4) in the primal variable:

$$\nabla_x \Xi(x, \sigma) = \mathbf{0}_n \Rightarrow x = G(\sigma)^{-1} F(\sigma),$$

and substituting the newfound value in the total complementarity function:

$$\Pi^d(\sigma) = -\frac{1}{2}F(\sigma)^T G(\sigma)^{-1} F(\sigma) - V^*(\sigma). \tag{5}$$

Note that the feasible set $\mathscr{S}_a$ is not convex; then, in order to identify the global optimality conditions, we need to introduce the following subset of $\mathscr{S}_a$:

$$\mathscr{S}_a^+ = \{\sigma \in \mathscr{S}_a | \ G(\sigma) \succeq 0\}.$$

Theorem 1. (**Global Optimality** *[5]*) *Given a critical point* $(\bar{x}, \bar{\sigma})$ *of* $\Xi(x, \sigma)$, $\bar{x}$ *is the unique global minimizer of* $\Pi(x)$ *if* $\bar{\sigma} \in \mathscr{S}_a^+$ *is the global maximizer of* $\Pi^d(\sigma)$ *on* $\mathscr{S}_a^+$*, and there is no duality gap between the primal, dual, and total complementarity functions, i.e.,*

$$\min_{x \in \mathbb{R}^n} \Pi(x) = P(\bar{x}) = \Xi(\bar{x}, \bar{\sigma}) = \Pi^d(\bar{\sigma}) = \max_{\sigma \in \mathscr{S}_a^+} \Pi^d(\sigma). \tag{6}$$

The result reported in equation (6) clearly shows the global optimality conditions. The original nonconvex primal problem is reduced to the maximization of the dual function $\Pi^d(\sigma)$ on the convex set $\mathscr{S}_a^+$. Furthermore it easy to notice from the (5) that the dual is concave on $\mathscr{S}_a^+$, and therefore the resulting problem is convex. Finally, we want to underline that there is no duality gap between the solution of the dual and the global minimum in the primal.

3 Reformulation of the Problem as a System of Constrained Equations

By the results of Theorem 1, it is possible to find the global solution of Problem ($\mathscr{P}$) by different approaches. One approach is to directly solve the dual formulation on $\mathscr{S}_a^+$, but this method has several faults:

- It is necessary to calculate the inverse of matrix $G(\sigma)$ every time the objective function is evaluated, and such operation could be necessary several times per iteration;
- The inverse matrix operation can become even more time expensive or generate errors in the case $G(\sigma)$ is ill-conditioned or it is not full rank;
- If the algorithm that solves the dual problem fails to converge to a good enough approximation of a stationary point, it is difficult to retrieve informations on the corresponding point in the primal problem.

For these reasons we propose a method that exploits the information available on both the primal and dual problems and search for a saddle point of the total complementarity function in $\mathscr{S}_a^+$, that is exactly the problem in the form of (1). As a matter of facts, it is easy to notice that finding the maximum of $\Pi^d(\sigma)$ in $\mathscr{S}_a^+$ is equivalent to solve the following canonical saddle point problem:

$$\min_{x\in\mathbb{R}^n}\max_{\sigma\in\mathbb{R}^m} \Xi(x,\sigma) = \frac{1}{2}x^T G(\sigma)x - F(\sigma)^T x - V^*(\sigma) \quad \text{s.t.,} \quad G(\sigma) \succeq 0, \tag{7}$$

that is the same problem presented in the introduction. The solution of (7) can be found by solving a monotone variational inequality on a convex set [3]:

$$\Gamma(x,\sigma) = 0, \quad G(\sigma) \succeq 0, \tag{8}$$

where $\Gamma : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ is defined as

$$\Gamma(x,\sigma) = \begin{pmatrix} \nabla_x \Xi(x,\sigma) \\ -\nabla_\sigma \Xi(x,\sigma) \end{pmatrix}.$$

The operator Γ is monotone because $\Xi(x,\sigma)$ is convex in the primal variables for $\sigma \in \mathscr{S}_a^+$ and it is concave for all $\sigma \in \mathscr{S}_a$ [14], while the set of positive definite matrices is a convex cone. We want to find a solution of (8) by solving the Karush–Kunt–Tucker (KKT) conditions associated with the problem, that is,

$$\begin{gathered} \Gamma_L(x,\sigma,L) = \begin{pmatrix} \nabla_x \Xi(x,\sigma) \\ -\nabla_\sigma \Xi(x,\sigma) - \nabla_\sigma (L \bullet G(\sigma)) \end{pmatrix} = \mathbf{0}_{n+m} \\ L \bullet G(\sigma) = 0, \quad L \succeq 0, \quad G(\sigma) \succeq 0, \end{gathered} \tag{9}$$

where $L \in \mathscr{S}_+^n$ is the matrix of the Lagrangian multipliers. The mapping $\Gamma_L(x,\sigma,L)$ is monotone as a result of Lemma 7 in [12]. Problems can arise when searching for the solution of (8) when there are KKT points located on the boundary of the feasible set. As a matter of facts, a point satisfying conditions (9) with $L \neq 0$ does not correspond to a saddle point of the total complementarity function $\Xi(x,\sigma)$ (in fact they generally correspond to stationary points of the primal problem). In other words we are interested in KKT points which matrix of multipliers L is equal to $\mathbf{0}_{n\times n}$.

To this aim, we reformulate the conditions (9) as a system of Constrained Equations (CE) and propose an interior point method specifically designed to solve this system of Constrained Equations and send the matrix of Lagrange multipliers to zero. We introduce the matrix $W \in \mathscr{S}_+^n$ of slack variables and consider the $CE(H,\Omega)$ system:

$$H(z) = \mathbf{0}, \quad z = (x,\sigma,L,W) \in \Omega, \tag{10}$$

where $H : \Omega \to S$ with $\Omega = \mathbb{R}^{n+m} \times \mathscr{S}_+^n \times \mathscr{S}_+^n$ and $S = \mathbb{R}^{n+m} \times \mathscr{S}_+^n \times \mathscr{S}_+^n \times \mathscr{S}_+^n$, is defined as

$$H(x, \sigma, L, W) = \begin{pmatrix} \Gamma_L(x, \sigma, L) \\ \Phi(\sigma, L, W) \\ L \end{pmatrix} \tag{11}$$

with $\Phi(\sigma, L, W)$ defined as

$$\Phi(\sigma, L, W) = \begin{pmatrix} W - G(\sigma) \\ (LW + WL)/2 \end{pmatrix}.$$

The last set of equations in (11) forces the matrix of Lagrange multipliers to go to zero when the algorithm reaches convergence, assuring that the solution of $CE(\Omega, H)$ is a saddle point of (7).

4 Key Assumptions and Convergence Result

In this section we present the conditions which the operator H and the feasible set Ω must satisfy together with a suitable potential reduction function in order to assure the convergence to a solution of the (10). The framework we use is the same as the one presented in [3] and [13]. This framework is based on six main assumptions that we report here for convenience.

Given the set Ω, operator H, and a potential function $p : \text{int } S \to \mathbb{R}$, the following assumptions must be satisfied by a potential reduction method in order to assure convergence to a solution of the $CE(\Omega, H)$.
(A1) the closed set Ω has a nonempty interior.
(A2) there exists a closed set $S \subseteq \mathbb{R}^{n+m} \times \mathscr{S}^n_+ \times \mathscr{S}^n_+ \times \mathscr{S}^n_+$ such that

1. $\mathbf{0} \in S$;
2. the open set $\Omega_I = H^{-1}(\text{int } S) \cap \text{int } \Omega$ is nonempty;
3. the set $H^{-1}(\text{int } S) \cap \text{bd } \Omega$ is empty.

(A3) H is continuously differentiable on Ω_I, and $JH(x)$ is full rank for all $x \in \Omega_I$
(A4) for every sequence $\{u^k\} \subset \text{int } S$ such that

$$\text{either } \lim_{k\to\infty} \|u^k\| = \infty \text{ or } \lim_{k\to\infty} u^k = \bar{u} \in \text{bd } S\backslash\{0\}$$

we have

$$\lim_{k\to\infty} p(u^k) = \infty.$$

(A5) p is continuously differentiable in its domain and $u \bullet \nabla p(u) > 0$ for all nonzero $u \in \text{int } S$.
(A6) there exists a nonzero vector $o \in S$ and a scalar $\bar{\beta} \in (0, 1]$ such that

$$u \bullet \nabla p(u) \geq \bar{\beta} \frac{(o \bullet u)(o \bullet \nabla p(u))}{\|o\|^2}, \quad \forall u \in \text{int } S.$$

In the following theorems we show that operator H and the feasible set Ω satisfy the aforementioned assumptions with the choice of a suitable potential reduction function.

Theorem 2. *Suppose that $\Xi(x, \lambda)$ is twice differentiable in x and σ, then the set Ω and the operator H in (11) satisfy conditions (A1)–(A3).*

Proof. Condition $(A1)$ is trivially satisfied, also condition $(A2).1$ holds. The point $(\mathbf{0}_{n+m}, I_n, I_n)$ belongs to both Ω_I and $\text{int}\,\Omega$, therefore condition $(A2).2$ holds. From condition

$$(LW + WL)/2,$$

we can define the following set:

$$\mathscr{U} = \{(L, W) \in \mathscr{S}^n_{++} \times \mathscr{S}^n_{++} : LW + WL \in \mathscr{S}^n_{++}\}.$$

It has been proved in lemma 1 of [12] that

$$\mathscr{U} = \{(L, W) \in \mathscr{S}^n_{+} \times \mathscr{S}^n_{+} : LW + WL \in \mathscr{S}^n_{++}\}.$$

This alternative representation implies the $(A2).3$. Finally condition $(A3)$ is satisfied because of the assumption on $\Xi(x, \lambda)$ and the monotonicity of the operator $\Gamma_L(x, \sigma, L)$. □

Theorem 3. *the potential function $p : S \to \mathbb{R}$ defined as*

$$p(a, B, C, D) = \eta \log(\|a\|^2 + \|B\|_F^2 + \|C\|_F^2 + \|D\|_F^2) - \\ \log(\det(B)) - \log(\det(C)) - \log(\det(D)), \tag{12}$$

where $\eta \geq 2n$, satisfies assumptions (A4)-(A6), with $o = (\mathbf{0}_{n+m}, \mathbf{0}_{n\times n}, I_n, \mathbf{0}_{n\times n})$ and $\bar{\beta} < 1/3$

Proof. It can be easily noticed that the value of p goes to ∞ as the sequence $\{a_k, B_k, C_k, D_k\}$ approaches the boundary of the feasible set. Considering that $\|Z\|_F = \sqrt{tr(Z^T Z)}$, then $\|Z\|_F^2$ is the sum of the squares of the n eigenvalues of Z and that $\det(Z)$ is the product of said eigenvalues, we have

$$p(a, B, C, D) = \eta \log\left(\sum_{i=1}^{n+m} \|a\|^2 + \sum_{i=1}^{n} b_i^2 + \sum_{i=1}^{n} c_i^2 + \sum_{i=1}^{n} d_i^2\right) - \\ \sum_{i=1}^{n} \log b_i - \sum_{i=1}^{n} \log c_i - \sum_{i=1}^{n} \log d_i,$$

where $b_i = 1, \dots, n$, $c_i = 1, \dots, n$ and $d_i = 1, \dots, n$ are the eigenvalues of B, C, and D respectively. Also considering that $n \log\left(\sum_{i=1}^{n} u_i\right) \geq \sum_{i=1}^{n} \log u_i + n \log n$ it is possible to write

$$p(a, B, C, D) > \left(\frac{2\eta}{3n} - 1\right)\left(\sum_{i=1}^{n} \log b_i + \sum_{i=1}^{n} \log c_i + \sum_{i=1}^{n} \log d_i\right),$$

therefore assumption (A4) is satisfied for $\eta > \frac{3}{2}n$.
If we define

$$\tau = \|a\|^2 + \|B\|_F^2 + \|C\|_F^2 + \|D\|_F^2,$$

it is possible to write the derivative of the potential function p as

$$\nabla p(a, B, C, D) = \begin{pmatrix} \frac{2\eta}{\tau} a \\ \frac{2\eta}{\tau} B - B^{-1} \\ \frac{2\eta}{\tau} C - C^{-1} \\ \frac{2\eta}{\tau} D - D^{-1} \end{pmatrix},$$

we have

$$(a, B, C, D) \bullet \nabla p(a, B, C, D) = 2\eta - 3n > 0,$$

and thus Assumption (A5) holds. For Assumption (A6), considering that $tr(Z)^2 \leq n\|Z\|_F^2$ and $n^2 \leq tr(Z^{-1})tr(Z)$ (for the arithmetic geometric mean inequality) we have

$$\begin{aligned} &\frac{[\nabla p(a, B, C, D) \bullet (\mathbf{0}_{n+m}, \mathbf{0}_{n\times n}, I_n, \mathbf{0}_{n\times n})][(a, B, C, D) \bullet (\mathbf{0}_{n+m}, \mathbf{0}_{n\times n}, I_n, \mathbf{0}_{n\times n})]}{\|(\mathbf{0}_n, I_n, \mathbf{0}_{n\times n}, \mathbf{0}_{n\times n})\|_F^2} = \\ &\frac{2\eta}{n}\frac{tr(C)^2}{\tau} - \frac{tr(C^{-1})tr(C)}{n} \leq \\ &\frac{2\eta}{n}\frac{tr(C)^2}{\|C\|_F^2} - \frac{tr(C^{-1})tr(C)}{n} \leq \\ &2\eta - n < \frac{1}{\beta}(2\eta - 3n) = \frac{1}{\beta}[(a, B, C, D) \bullet \nabla p(a, B, C, D)]. \end{aligned}$$

□

We let

$$z = (x, \sigma, L, W), \quad \psi(z) = p(H(z)),$$

and report the following method that follows the same scheme of the interior point method presented in [13]:

Algorithm 1: CPRA: Complementarity Potential Reduction Algorithm

`(S.0)` : Choose $z^0 = (x^0, \sigma^0, L^0, W^0) \in \Omega$, $\gamma \in (0, 1)$, $\bar{\beta} < 1/3$, $\varepsilon > 0$, and set $k := 0$.

`(S.1)` : If $\|\Gamma(x, \sigma)\|^2 < \varepsilon$: STOP

`(S.2)` : Choose a scalar $\beta_k \in (0, \bar{\beta})$ and find a solution $d^k = (dx^k, d\sigma^k, dL^k, dW^k)$ of the following linear least squares problem:

$$\min_d \left\{ \frac{1}{2} \left\| Q(z^k, d) + H(z^k) - \beta_k \frac{o^T H(z^k)}{\|o\|^2} o \right\|^2 \right\}.$$

where

$$Q\ (z^k, d) = \begin{pmatrix} \nabla^2_{xx} \Xi(x^k, \sigma^k) dx + \nabla^2_{x\sigma} \Xi(x^k, \sigma^k) d\sigma \\ -\nabla^2_{x\sigma} \Xi(x^k, \sigma^k)^T dx - \nabla^2_{\sigma\sigma} \Xi(x^k, \sigma^k) d\sigma + \nabla_{\sigma L}(L^k \bullet G(\sigma^k)) dL \\ dW - G(d\sigma) \\ (dL) W^k + W^k (dL) + L^k (dW) + (dW) L^k \\ dL \end{pmatrix}$$

`(S.3)` : find a step size α_k such that

$$z^k + \alpha_k d^k \in \Omega$$

and

$$\psi(z^k + \alpha_k d^k) \le \psi(z^k) + \gamma \nabla \psi(z^k) \bullet d^k$$

`(S.4)` : Set $z^{k+1} = z^k + \alpha_k d^k$, $k \leftarrow k + 1$, and go to (S.1).

Algorithm 1 is a modified, damped version of the Newton method. At Step `(S.0)` the initial values of the variables and parameters are set. In order to assure the feasibility of z^0, it generally suffices to put a large enough positive value of σ^0, such that $G(\sigma^0) \succ 0$. At Step `(S.1)` there is the stopping criterion that assures the final point is a good enough approximation of a stationary point of $\Xi(x, \sigma)$. At Step `(S.2)` the modified newton direction is calculated. As the linear system is not squared, the least squares solution to the system of equations is returned. One of the main features of the algorithm is the presence of the vector o that bends the direction toward the interior of the feasible set. It is important to underline that the calculated direction at every iteration is unique for Assumption ($A3$) and always a descent direction of $\psi(\cdot)$ in z_k as shown in the following theorem:

Theorem 4. *Suppose that conditions* ($A5$) *and* ($A6$) *hold. Assume also that* $z \in \Omega_I$, $d^k = (dx^k, d\sigma^k, dL^k, dW^k) \in \mathbb{R}^{n+m} \times \mathcal{S}^n_+ \times \mathcal{S}^n_+$ *and* $\beta \in \mathbb{R}$ *are such that*

$$
\begin{aligned}
&H(z) \neq 0, \qquad 0 \leq \beta < \bar{\beta}, \\
&d^k = arg\min_d \left\{ \tfrac{1}{2} \left\| Q(z,d) + H(z) - \beta_k \tfrac{o^T H(z)}{\|o\|^2} o \right\|^2 \right\},
\end{aligned} \tag{13}
$$

where $o \in S$ and $\bar{\beta} \in [0, 1]$ are as in condition (A6). Then d^k is a descent direction for $\psi(\cdot)$ in z, that is $\nabla\psi(z) \bullet d^k < 0$

Proof. We introduce the following vector in $\mathbb{R}^{n+m+3n^2}$:

$$
\hat{H}(z) = \begin{pmatrix} \Gamma_L(x, \sigma, L) \\ vect\{W - G(\sigma)\} \\ vect\{(LW + WL)/2\} \\ vect\{L\} \end{pmatrix}. \tag{14}
$$

The Jacobian of $\hat{H}(z)$ is the following $(n + m + 3n^2) \times (n + m + 2n^2)$ matrix:

$$
J\hat{H}(z) = \begin{pmatrix} \nabla^2_{xx}\Xi(x,\sigma) & \nabla^2_{x\sigma}\Xi(x,\sigma) & \mathbf{0}_{n\times n^2} & \mathbf{0}_{n\times n^2} \\ -\nabla^2_{x\sigma}\Xi(x,\sigma) & \nabla^2_{\sigma\sigma}\Xi(x,\sigma) & C^T & \mathbf{0}_{m\times n^2} \\ \mathbf{0}_{n^2\times n} & C & \mathbf{0}_{n^2\times n^2} & I_{n^2} \\ \mathbf{0}_{n^2\times n} & \mathbf{0}_{n^2\times m} & W_{en} & L_{en} \\ \mathbf{0}_{n^2\times n} & \mathbf{0}_{n^2\times m} & I_{n^2} & \mathbf{0}_{n^2\times n^2} \end{pmatrix}. \tag{15}
$$

where

$$
W_{en} = \begin{pmatrix} W + I_n w_{11} & I_n w_{12} & \cdots & I_n w_{1n} \\ I_n w_{21} & W + I_n w_{22} & \cdots & I_n w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ I_n w_{n1} & I_n w_{n2} & \cdots & W + I_n w_{nn} \end{pmatrix}, \tag{16}
$$

$$
L_{en} = \begin{pmatrix} L + I_n l_{11} & I_n l_{12} & \cdots & I_n l_{1n} \\ I_n l_{21} & L + I_n l_{22} & \cdots & I_n l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ I_n l_{n1} & I_n l_{n2} & \cdots & L + I_n l_{nn} \end{pmatrix}, \tag{17}
$$

and $C \in \mathbb{R}^{n^2\times m}$ is $\nabla_{\sigma L}(L \bullet G(\sigma^k))^T$. Let $u \equiv \hat{H}(z)$, if we consider $\hat{d}^k \in \mathbb{R}^{n+m+2n^2}$, solution of the following least squares problem:

$$
\hat{d}^k = arg\min_d \left\{ \frac{1}{2} \left\| (Ju)d + u - \beta_k \frac{\hat{o}^T u}{\|\hat{o}\|^2} \hat{o} \right\|^2 \right\}, \tag{18}
$$

where $\hat{o}$ has been suitably changed from o to match the dimension of $\hat{H}(z)$, it is easy to notice that $\hat{d}^k$ is equivalent to d^k, solution of the least squares problem in (13), in the following sense:

$$\hat{d}^k = \begin{pmatrix} dx^k \\ d\sigma^k \\ vect\{dL^k\} \\ vect\{dW^k\} \end{pmatrix}.$$

Furthermore, if we define

$$\nabla\hat{\psi}(z) = \begin{pmatrix} \nabla_x \psi(z) \\ \nabla_\sigma \psi(z) \\ vect\{\nabla_L \psi(z)\} \\ vect\{\nabla_W \psi(z)\} \end{pmatrix}, \quad \nabla\hat{p}(u) = \begin{pmatrix} \nabla_x p(H(z)) \\ \nabla_\sigma p(H(z)) \\ vect\{\nabla_L p(H(z))\} \\ vect\{\nabla_W p(H(z))\} \end{pmatrix},$$

for the symmetry of the matrices involved in the calculations, we have

$$\nabla\psi(z^k) \bullet d^k = \nabla\hat{\psi}(z)^T \hat{d}^k, \quad \nabla\hat{\psi}(z) = Ju^T \nabla\hat{p}(u).$$

Another propriety of $\hat{d}^k$ is that it satisfies the normal equations of (18)

$$\hat{d}^k = \left(Ju^T Ju\right)^{-1} Ju^T \left(\beta_k \frac{\hat{o}^T u}{\|\hat{o}\|^2}\hat{o} - u\right). \tag{19}$$

Therefore, from the assumptions of the theorem and by exploiting the (19) it is possible to obtain

$$\begin{aligned} \nabla\hat{\psi}(z)^T \hat{d}^k &= \nabla\hat{p}(u)^T (Ju)\hat{d}^k \\ &\overset{(19)}{=} \nabla\hat{p}(u)^T Ju \left(Ju^T Ju\right)^{-1} Ju^T \left(\beta_k \tfrac{\hat{o}^T u}{\|\hat{o}\|^2}\hat{o} - u\right) \\ &= \nabla\hat{p}(u)^T Ju Ju^{-1} (Ju^T)^{-1} Ju^T \left(\beta_k \tfrac{\hat{o}^T u}{\|\hat{o}\|^2}\hat{o} - u\right) \\ &= \nabla\hat{p}(u)^T \left(\beta_k \tfrac{\hat{o}^T u}{\|\hat{o}\|^2}\hat{o} - u\right) \leq -\nabla\hat{p}(u)^T u (1 - \tfrac{\beta_k}{\bar{\beta}}) \\ &= -\nabla p(H(z)) \bullet H(z)(1 - \tfrac{\beta_k}{\bar{\beta}}) \overset{(A5)}{<} 0, \end{aligned}$$

where with Ju^{-1} and $(Ju^T)^{-1}$ are the Moore Penrose pseudo inverses of Ju and Ju^T, respectively. The third equality derives from the propriety

$$(AB)^{-1} = B^{-1}A^{-1},$$

valid for the Moore Penrose pseudo inverse in the case we are considering (interested readers can refer to [8]). The last equality follows from the definition of $\hat{H}(z)$ and $\hat{p}(u)$. □

At step `(S.3)` the potential function (12) is used to measure the progress of the algorithm. Finally at Step `(S.4)` the value of k is updated and the loop is completed.

It is possible to observe that the sequence generated by Algorithm 1 necessarily belongs to Ω. We now present the convergence result:

Theorem 5. *Assume that $CE(\Omega, H)$ has a solution. Let $\{z^k\}$ be the sequence generated by Algorithm 1, then*

(a) the sequence $\{H(z^k)\}$ is bounded;
(b) any accumulation point of $\{z^k\}$, if it exists, solves CE(Ω, H);
(c) $\lim_{k\to\infty} H(z^k) = 0$*;*
(d) the sequence $\{z^k\} = \{(x^k, \sigma^k, L^k, W^k)\}$ is bounded.

Proof. The proof of statements (a) and (b) follows from Theorem 3 of [13].

In order to prove the (c) we first have to prove the (d), that is the boundedness of $\{z^k\}$. To prove the boundless of $\{z^k\}$ we have to prove the boundedness of the sequences $\{x^k\}$, $\{\sigma^k\}$, $\{L^k\}$, and $\{W^k\}$. The boundedness of $\{L^k\}$ is a direct consequence of the boundedness of $\{H(z^k)\}$.

To prove the boundedness of the sequences $\{x^k\}$ and $\{\sigma^k\}$ we use the operator Γ. In detail, from the (4) we obtain

$$\nabla_x \Xi(x, \sigma) = \quad G(\sigma)x - F(\sigma), \tag{20}$$

$$-\nabla_\sigma \Xi(x, \sigma) = \nabla V^*(\sigma) - \nabla V(\Lambda(x)). \tag{21}$$

It is easy to see that if one of the two sequences goes to infinity while the other converges, $\|\Gamma(x^k, \sigma^k)\| \to \infty$ contradicting the (a).
We consider the case in which $\{x^k\}$ and $\{\sigma^k\}$ go to infinity simultaneously. It is possible to notice from the (4) that $F(\sigma)$ is linear in σ, and therefore if both the variables go to infinity we have $\|\nabla_x \Xi(x^k, \sigma^k)\| \to \infty$. Finally if we suppose that $\{W^k\} \to \infty$, from the boundedness of $\{\sigma^k\}$ and constraint $W - G(\sigma)$ we obtain the desired contradiction with the (a).

The (c) is a direct consequence of conditions (b) and (d). □

5 Conclusions

We presented an interior points method framework for canonical duality theory that converges under mild assumptions. The framework in this paper not only has really favorable convergence proprieties, but it is also general and potentially able to handle large-sized problems efficiently with a good level of reliability.

In our view, these results constitute an important step for several topics in optimization. The new findings of this paper indicate that it is possible to adapt interior points methods to the problems reformulated with canonical duality. Therefore, other popular interior points methods such as primal–dual methods could be used to solve problem (1) and find the global solution of many nonconvex optimization problems efficiently.

There are also several applications that can be investigated with the presented framework. In detail, the maximum cut problem and the radial basis function neural networks problems can also be solved with canonical duality [10, 16], and the proposed algorithm could be useful to find their global solutions for large-sized instances.

Acknowledgements The Author would like to thank Simone Sagratella for his help. Without his suggestions in the initial conception of this method, it would have been quite difficult to understand the right path to take in order to create the presented framework. The author would also like to thank Professor Stefano Lucidi for his suggestions for improving the paper.

References

1. Dreves, A., Facchinei, F., Kanzow, C., Sagratella, S.: On the solution of the KKT conditions of generalized Nash equilibrium problems. SIAM J. Optim. **21**, 1082–1108 (2011)
2. Facchinei, F., Kanzow, C., Sagratella, S.: Solving quasi-variational inequalities via their KKT conditions. Math. Prog. Ser. A **144**, 369–412 (2014). doi:10.1007/s10107-013-0637-0
3. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. Volume. I and II. Springer, New York (2003)
4. Gao, D.Y.: Duality principles in nonconvex systems: Theory, methods and applications. Kluwer Academic Publishers, Dordrecht (2000)
5. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
6. Gao, D.Y.: Canonical duality theory: theory, method, and applications in global optimization. Comput. Chem. **33**, 1964–1972 (2009)
7. 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 (2012)
8. Greville, T.N.E.: Note on the generalized inverse of a matrix product. SIAM Review **8**(4), 518–521 (1966)
9. Latorre, V., Gao, D.Y.: Canonical duality for solving general nonconvex constrained problems. Optim. Lett. **10**(8), 1763–1779 (2016). doi:10.1007/s11590-015-0860-0
10. Latorre, V., Gao, D.Y.: Canonical dual solution to nonconvex radial basis neural network optimization problem. Neurocomputing **134**, 189–197 (2014)
11. Latorre, V., Sagratella, S.: A Canonical Duality Approach for the Solution of Affine Quasi-Variational Inequalities. J. Glob. Opt. (2014). doi:10.1007/s10898-014-0236-5
12. Monteiro, R.D.C., Pang, J.S.: On two interior-point mappings for nonlinear semidefinite complementarity problems. Math. Oper. Res. **23**, 39–60 (1998)
13. Monteiro, R.D.C., Pang, J.S.: A potential reduction Newton method for constrained equations. SIAM J. Optim. **9**(3), 729–754 (1999)
14. Rockafellar, R.T.: Monotone operators associated with saddle-functions and minimax problems. Proc. Pure Math **18**, 241–250 (1970)
15. Ruan, N., Gao, D.Y.: Global optimal solutions to a general sensor network localization problem. Perform. Eval. **75–76**, 1–16 (2014)
16. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Canonical dual approach to solving the maximum cut problem. J. Glob. Optim. **54**, 341–352 (2012)
17. Wu, C., Gao, D. Y.: Canonical primal-dual method for solving non-convex minimization problems. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Canonical Duality Theory. Springer (2017)

Canonical Duality Theory for Topology Optimization

David Yang Gao

Abstract This paper presents a canonical duality approach for solving a general topology optimization problem of nonlinear elastic structures. Based on the principle of minimum total potential energy, this most challenging problem can be formulated as a bi-level mixed integer nonlinear programming problem (MINLP), i.e., for a given deformation, the first-level optimization is a typical linear constrained 0–1 programming problem, while for a given structure, the second-level optimization is a general nonlinear continuous minimization problem in computational nonlinear elasticity. It is discovered that for linear elastic structures, first-level optimization is a typical Knapsack problem, which is considered to be NP-complete in computer science. However, by using canonical duality theory, this well-known problem can be solved analytically to obtain exact integer solution. A perturbed canonical dual algorithm (CDT) is proposed and illustrated by benchmark problems in topology optimization. Numerical results show that the proposed CDT method produces desired optimal structure without any gray elements. The checkerboard issue in traditional methods is much reduced. Additionally, an open problem on NP-hardness of the Knapsack problem is proposed.

1 General Topology Optimization Problem and Challenges

Topology optimization is a mathematical method that optimizes material layout within a given design space, for a given set of loads, boundary conditions, and constraints with the goal of maximizing the performance of the system. Due to its broad applications, the topology optimization has been subjected to extensively study since the seminal paper by Bendsoe and Kikuch [4]. Generally speaking, a typical topology optimization problem involves both continuous-state variable and discrete density distribution that can take either the value 0 (void) or 1 (solid material) at any point in the design domain. Thus, numerical discretization methods (say FEM)

D.Y. Gao (✉)
Faculty of Science and Technology, Federation University Australia,
Mount Helen, VIC, Australia
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics
and Mathematics 37, DOI 10.1007/978-3-319-58017-3_13

for solving topology optimization problems lead to a so-called mixed integer nonlinear programming (MINLP) problem, which appears extensively in computational engineering, decision and management sciences, operations research, industrial, and systems engineering [10].

Let us consider an elastically deformable body that in an undeformed configuration occupies an open domain $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ with boundary $\Gamma = \partial\Omega$. We assume that the body is subjected to a body force $\mathbf{f}$ (per unit mass) in the reference domain Ω and a given surface traction $\mathbf{t}(\mathbf{x})$ of dead-load type on the boundary $\Gamma_t \subset \partial\Omega$, while the body is fixed on the remaining boundary $\Gamma_u = \partial\Omega \cap \Gamma_t$. Based on the minimal potential principle in continuum mechanics, the topology optimization of this elastic body can be formulated in the following coupled minimization problem.

$$(\mathscr{P}): \quad \min_{\mathbf{u}\in\mathscr{U}_a} \min_{\rho\in\mathscr{Z}} \left\{ \Pi(\mathbf{u}, \rho) = \int_\Omega W(\nabla\mathbf{u})\rho \mathrm{d}\Omega + \int_\Omega \mathbf{u}\cdot\mathbf{f}\rho\mathrm{d}\Omega - \int_{\Gamma_t} \mathbf{u}\cdot\mathbf{t}\mathrm{d}\Gamma \right\}, \tag{1}$$

where the unknown $\mathbf{u} : \Omega \to \mathbb{R}^d$ is a displacement vector field, the design variable $\rho(\mathbf{x}) \in \{0, 1\}$ is a discrete scalar field, and the stored energy per unit reference volume $W(\mathbf{D})$ is a nonlinear differentiable function of the deformation gradient $\mathbf{D} = \nabla\mathbf{u}$. The notation $\mathscr{U}_a$ identifies a *kinematically admissible space* of deformations, in which, certain geometrical/boundary conditions are given, and

$$\mathscr{Z} = \left\{ \rho(\mathbf{x}) : \Omega \to \{0, 1\} | \ \int_\Omega \rho(\mathbf{x})\mathrm{d}\Omega \le V_c \right\}$$

is a design feasible space, in which, $V_c > 0$ is the desired volume.

Mathematically speaking, the topology optimization $(\mathscr{P})$ is a coupled nonlinear-discrete minimization problem in infinite-dimensional space. For large deformation problems, the stored energy $W(\mathbf{D})$ is usually nonconvex. It is fundamentally difficult to analytically solve this type of problems. Numerical methods must be adopted.

Finite element method is the most popular numerical approach for topology optimization, by which the domain Ω is divided into n disjointed elements $\{\Omega_e\}$ and in each element, the unknown fields can be numerically discretized as

$$\mathbf{u}(\mathbf{x}) = \mathbf{N}_e(\mathbf{x})\mathbf{u}_e, \quad \rho(\mathbf{x}) = \rho_e \in \{0, 1\} \ \ \forall \mathbf{x} \in \Omega_e, \tag{2}$$

where $\mathbf{N}_e$ is an interpolation matrix, $\mathbf{u}_e$ is a nodal displacement vector, the binary design variable $\rho_e \in \{0, 1\}$ is used for determining whether the element Ω_e is a void $(\rho_e = 0)$ or a solid $(\rho_e = 1)$. Thus, by substituting (2) into $\Pi(\mathbf{u}, \rho)$ and let $\mathscr{U}_a^m \subset \mathbb{R}^m$ be an admissible nodal displacement space,

$$\mathscr{Z}_a = \left\{ \boldsymbol{\rho} = \{\rho_e\} \in \{0, 1\}^n | V(\boldsymbol{\rho}) = \sum_{e=1}^n \rho_e \Omega_e \le V_c \right\}, \tag{3}$$

the variational problem ($\mathscr{P}$) can be numerically reformulated the following global optimization problem:

$$(\mathscr{P}_h): \quad \min_{\mathbf{u}\in\mathscr{U}_a^m} \min_{\boldsymbol{\rho}\in\mathscr{Z}_a} \left\{\Pi_h(\mathbf{u}, \boldsymbol{\rho}) = C(\boldsymbol{\rho}, \mathbf{u}) - \mathbf{u}^T\mathbf{f}(\boldsymbol{\rho})\right\}, \tag{4}$$

where

$$C(\boldsymbol{\rho}, \mathbf{u}) = \boldsymbol{\rho}^T\mathbf{c}(\mathbf{u}), \quad \mathbf{c}(\mathbf{u}) = \left\{\int_{\Omega_e} W(\nabla\mathbf{N}_e(\mathbf{x})\mathbf{u}_e)\mathrm{d}\Omega\right\} \in \mathbb{R}^n, \tag{5}$$

$$\mathbf{f}(\boldsymbol{\rho}) = \left\{\int_{\Omega_e} \rho_e\mathbf{N}_e(\mathbf{x})^T\mathbf{b}_e(\mathbf{x})\mathrm{d}\Omega\right\} + \left\{\int_{\Gamma_t^e} \mathbf{N}_e(\mathbf{x})^T\mathbf{t}(\mathbf{x})\mathrm{d}\Gamma\right\} \in \mathbb{R}^m. \tag{6}$$

Clearly, this discretized topology optimization involves both the continuous variable $\mathbf{u} \in \mathscr{U}_a^m$ and the integer variable $\boldsymbol{\rho} \in \mathscr{Z}_a$; it is the so-called *mixed integer nonlinear programming problem* (MINLP) in mathematical programming. Since $\rho_e^p = \rho_e \ \ \forall\rho_e \in \{0, 1\}, \ \ \forall p \in \mathbb{R}$, we have

$$C_p(\boldsymbol{\rho}, \mathbf{u}) := \sum_{e=1}^{n} \rho_e^p c_e(\mathbf{u}) = (\underbrace{\boldsymbol{\rho} \circ \ldots \circ \boldsymbol{\rho}}_{p \text{ times}})^T\mathbf{c}(\mathbf{u}) = C(\boldsymbol{\rho}, \mathbf{u}) \quad \forall p \in \mathbb{R}, \tag{7}$$

where $\boldsymbol{\rho} \circ \mathbf{c} = \{\rho_e c_e\}$ represents the Hadamard product. Particularly, for $p = 2$, we write

$$C_2(\boldsymbol{\rho}, \mathbf{u}) := \frac{1}{2}\boldsymbol{\rho}^T\mathbf{A}(\mathbf{u})\boldsymbol{\rho}, \quad \mathbf{A}(\mathbf{u}) = 2\mathrm{Diag}\{\mathbf{c}(\mathbf{u})\}. \tag{8}$$

Clearly, $C_2(\boldsymbol{\rho}, \mathbf{u})$ is a convex function of $\boldsymbol{\rho}$ since $\mathbf{A}(\mathbf{u}) \succeq 0 \ \ \forall\mathbf{u} \in \mathscr{U}_a^m$. By the facts that $\boldsymbol{\rho} \in \mathscr{Z}_a$ is the main design variable and the displacement $\mathbf{u}$ depends on each given domain Ω, the problem ($\mathscr{P}_h$) is actually a so-called bi-level programming problem:

$$(\mathscr{P}_{bl}): \qquad \min_{\boldsymbol{\rho}\in\mathscr{Z}_a} \min_{\mathbf{u}\in\mathscr{U}_a^m} \{C_p(\boldsymbol{\rho}, \mathbf{u}) - \mathbf{u}^T\mathbf{f}(\boldsymbol{\rho})\} \tag{9}$$

$$s.t. \ \ \mathbf{u} = \arg\min_{\mathbf{v}\in\mathscr{U}_a^m} \Pi_h(\mathbf{v}, \boldsymbol{\rho}). \tag{10}$$

In this formulation, $C_p(\boldsymbol{\rho}, \mathbf{u}) - \mathbf{u}^T\mathbf{f}(\boldsymbol{\rho})$ represents the upper level cost function and the total potential energy $\Pi_h(\mathbf{u}, \boldsymbol{\rho})$ represents the lower level cost function. For large deformation problems, the total potential energy Π_h is usually a nonconvex function of $\mathbf{u}$. Therefore, this bi-level optimization could be the most challenging problem in global optimization.

For linear elastic structures, the total potential energy Π_h is a quadratic function of $\mathbf{u}$

$$\Pi_h(\mathbf{u}, \boldsymbol{\rho}) = \frac{1}{2}\mathbf{u}^T\mathbf{K}(\boldsymbol{\rho})\mathbf{u} - \mathbf{u}^T\mathbf{f}(\boldsymbol{\rho}), \tag{11}$$

where $\mathbf{K}(\boldsymbol{\rho}) = \{\rho_e \mathbf{K}_e\} \in \mathbb{R}^{m\times m}$ is the overall stiffness matrix, which is obtained by assembling the sub-matrix $\rho_e \mathbf{K}_e$ for each element Ω_e. In this case, the lower level optimization (10) is a convex minimization and for each given upper level design variable $\boldsymbol{\rho}$, the lower level solution is simply governed by the linear equilibrium equation $\mathbf{K}(\boldsymbol{\rho})\mathbf{u} = \mathbf{f}(\boldsymbol{\rho})$. Therefore, the topology optimization for linear elasticity is mathematically a linearly constrained integer programming problem:

$$(\mathscr{P}_{le}): \quad \min_{\boldsymbol{\rho}\in\mathscr{Z}_a} \min_{\mathbf{u}\in\mathscr{U}_a^m} \left\{ -\frac{1}{2}\mathbf{u}^T\mathbf{K}(\boldsymbol{\rho})\mathbf{u} | \quad \mathbf{K}(\boldsymbol{\rho})\mathbf{u} = \mathbf{f}(\boldsymbol{\rho}) \right\}. \tag{12}$$

Due to the integer constraint, to solve this mixed integer quadratic minimization problem is fundamentally difficult. In order to overcome the combinatorics complexity in this problem, various approximations were proposed during the last decades, including homogenization [4], density-based approximations [3], level set method [21], and topological derivative [19]. These approaches generally relax the MINLP problem into a continuous parameter optimization problem by using size, density, or shape, and then solve it based on the traditional Newton-type (gradient-based) or evolutionary optimization algorithms. A comprehensive survey on these approaches was given in [18].

The so-called Simplified Isotropic Material with Penalization (SIMP) is one of the most popular approaches in topology optimization:

$$(SIMP): \quad \min_{\boldsymbol{\rho}\in\mathbb{R}^N} C_p(\boldsymbol{\rho}, \mathbf{u}(\boldsymbol{\rho})) \tag{13}$$

$$s.t. \;\; \mathbf{K}(\boldsymbol{\rho}^p)\mathbf{u} = \mathbf{f}(\boldsymbol{\rho}), \quad V(\boldsymbol{\rho}) \le V_c, \tag{14}$$

$$0 < \rho_e \le 1, \; e = 1, \ldots, n \tag{15}$$

where p is the so-called penalization parameter in topology optimization. The SIMP formulation has been studied extensively in topology optimization and numerous research papers have been produced during the past decades. By the fact that $\boldsymbol{\rho}^p = \boldsymbol{\rho} \;\; \forall p \in \mathbb{R}, \;\; \forall \boldsymbol{\rho} \in \{0, 1\}^n$, we can see that the integer constraint $\boldsymbol{\rho} \in \{0, 1\}^n$ in $(\mathscr{P}_{le})$ is simply replaced by the box constraint $\boldsymbol{\rho} \in (0, 1]^n$. Although it was discovered by engineers that the "magic number" $p = 3$ can ensure good convergence to almost 0-1 solutions, the SIMP formulation is not mathematically equivalent to the topology optimization problem $(\mathscr{P}_{le})$. Actually, in many real-world applications, most SIMP solutions $\{\rho_e\}$ are only approximate to 0 or 1 but never be exactly 0 or 1. Correspondingly, these elements are in grayscale which have to be filtered or interpreted artificially. Additionally, this method suffers some key limitations such as the unsure global optimization, many grayscale elements, checkerboard patterns, etc.

2 Canonical Dual Problem and Analytical Solution

Canonical dual finite element methods for solving elasto-plastic structures and large deformation problems have been studied since 1988 [5, 6]. Applications to nonconvex mechanics are given recently for post-buckling problems [1, 15]. This paper will address the canonical duality theory for solving the challenging integer programming problem in $(\mathscr{P}_u)$.

Let $\mathbf{a} = \{a_e = \mathrm{Vol}(\Omega_e)\} \in \mathbb{R}^n$, where $\mathrm{Vol}(\Omega_e)$ represents the volume of each element Ω_e. Then we have $\mathscr{Z}_a = \{\boldsymbol{\rho} \in \{0, 1\}^n | \ \boldsymbol{\rho}^T \mathbf{a} \le V_c\}$. By the fact that $\min_{\boldsymbol{\rho}} \min_{\mathbf{u}} = \min_{\mathbf{u}} \min_{\boldsymbol{\rho}}$, the alternative iteration can be adopted for solving the topology optimization problem. Since $C_1(\boldsymbol{\rho}, \mathbf{u}) = \frac{1}{2}\mathbf{u}^T \mathbf{K}(\boldsymbol{\rho})\mathbf{u} = \boldsymbol{\rho}^T \mathbf{c}(\mathbf{u})$, for a given solution of (10), the energy vector $\mathbf{c}_u = \mathbf{c}(\mathbf{u}) \in \mathbb{R}^n_+$ is nonnegative. Thus, the iterative method for linear elastic topology optimization $(\mathscr{P}_{le})$ can be proposed for solving the following linear 0–1 programming problem ($(\mathscr{P})$ for short):

$$(\mathscr{P}): \quad \min \left\{ P_u(\boldsymbol{\rho}) = -\mathbf{c}_u^T \boldsymbol{\rho} \ | \ \boldsymbol{\rho} \in \{0, 1\}^n, \quad \boldsymbol{\rho}^T \mathbf{a} \le V_c \right\}. \tag{16}$$

This is the well-known Knapsack problem. Due to the 0–1 constraint, even this most simple linear integer programming is listed as one of Karp's 21 NP-complete problems [13].

The canonical duality theory for general integer programming was first proposed by Gao in 2007 [9]. The key idea of this theory is the introduction of a canonical measure

$$\boldsymbol{\xi} = \Lambda(\boldsymbol{\rho}) = \{\boldsymbol{\rho} \circ \boldsymbol{\rho} - \boldsymbol{\rho}, \quad \boldsymbol{\rho}^T \mathbf{a} - V_c\} : \quad \mathbb{R}^n \to \mathscr{E} = \mathbb{R}^{n+1}. \tag{17}$$

Let

$$\mathscr{E}_a := \{\boldsymbol{\xi} = \{\boldsymbol{\varepsilon}, \nu\} \in \mathbb{R}^{n+1} | \ \boldsymbol{\varepsilon} \le 0, \quad \nu \le 0\} \tag{18}$$

be a convex cone in $\mathbb{R}^{n+1}$. Its indicator $\Psi(\boldsymbol{\xi})$ is defined by

$$\Psi(\boldsymbol{\xi}) = \begin{cases} 0 & \text{if } \boldsymbol{\xi} \in \mathscr{E}_a \\ +\infty & \text{otherwise} \end{cases}$$

which is a convex and lower semi-continuous (l.s.c) function in $\mathbb{R}^{n+1}$. By this function, the primal problem can be relaxed in the following unconstrained minimization form:

$$\min \left\{ \Phi(\boldsymbol{\rho}) = P_u(\boldsymbol{\rho}) + \Psi(\Lambda(\boldsymbol{\rho})) \ | \ \boldsymbol{\rho} \in \mathbb{R}^n \right\}. \tag{19}$$

Due to the convexity of $\Psi(\boldsymbol{\xi})$, its conjugate function can be defined uniquely by the Fenchel transformation:

$$\Psi^*(\boldsymbol{\zeta}) = \sup_{\boldsymbol{\xi}\in\mathbb{R}^{n+1}} \{\boldsymbol{\xi}^T\boldsymbol{\zeta} - \Psi(\boldsymbol{\xi})\} = \begin{cases} 0 & \text{if } \boldsymbol{\zeta} \in \mathscr{E}_a^* \\ +\infty & \text{otherwise} \end{cases} \tag{20}$$

where $\mathscr{E}_a^* = \{\boldsymbol{\zeta} = \{\boldsymbol{\sigma}, \varsigma\} \in \mathbb{R}^{n+1} |\ \boldsymbol{\sigma} \geq 0,\quad \varsigma \geq 0\}$ is the dual space of $\mathscr{E}_a$. Thus, by using the Fenchel-Young equality $\Psi(\boldsymbol{\xi}) + \Psi^*(\boldsymbol{\zeta}) = \boldsymbol{\xi}^T\boldsymbol{\zeta}$, the function $\Phi(\boldsymbol{\rho})$ can be written in the Gao–Strang total complementary function [12]

$$\Xi(\boldsymbol{\rho}, \boldsymbol{\zeta}) = P_u(\boldsymbol{\rho}) + \Lambda(\boldsymbol{\rho})^T\boldsymbol{\zeta} - \Psi^*(\boldsymbol{\zeta}). \tag{21}$$

Based on this function, the canonical dual of $\Phi(\boldsymbol{\rho})$ can be defined by

$$\Phi^d(\boldsymbol{\zeta}) = \text{sta}\,\{\Xi(\boldsymbol{\rho}, \boldsymbol{\zeta}) |\ \boldsymbol{\rho} \in \mathbb{R}^m\} = P_u^{\Lambda}(\boldsymbol{\zeta}) - \Psi^*(\boldsymbol{\zeta}), \tag{22}$$

where sta $\{f(x) |\ x \in X\}$ stands for finding a stationary value of $f(x)\ \ \forall x \in X$, and

$$P_u^{\Lambda}(\boldsymbol{\zeta}) = \text{sta}\,\{\Lambda(\boldsymbol{\rho})^T\boldsymbol{\zeta} + P_u(\boldsymbol{\rho})\} = -\frac{1}{4}\boldsymbol{\tau}_u^T(\boldsymbol{\zeta})\mathbf{G}^{-1}(\boldsymbol{\zeta})\boldsymbol{\tau}_u(\boldsymbol{\zeta}) - \varsigma V_c \tag{23}$$

is the Λ-conjugate of $P_u(\boldsymbol{\rho})$, in which,

$$\mathbf{G}(\boldsymbol{\zeta}) = \text{ Diag}\{\boldsymbol{\sigma}\}, \quad \boldsymbol{\tau}_u(\boldsymbol{\zeta}) = \boldsymbol{\sigma} - \varsigma\mathbf{a} + \mathbf{c}_u.$$

Clearly, $P_u^{\Lambda}(\boldsymbol{\zeta})$ is well defined if $\det \mathbf{G} \neq 0$, i.e., $\boldsymbol{\sigma} \neq 0 \in \mathbb{R}^n$. Let $\mathscr{S}_a = \{\boldsymbol{\zeta} \in \mathscr{E}_a^* |\ \det \mathbf{G} \neq 0\}$. We have the following standard result in the canonical duality theory:

Theorem 1 (Complementary-Dual Principle). *For a given* $\mathbf{u} \in \mathscr{U}_a^m$, *if* $(\bar{\boldsymbol{\rho}}, \bar{\boldsymbol{\zeta}})$ *is a KKT point of* Ξ, *then* $\bar{\boldsymbol{\rho}}$ *is a KKT point of* Φ, $\bar{\boldsymbol{\zeta}}$ *is a KKT point of* Φ^d, *and*

$$\Phi(\bar{\boldsymbol{\rho}}) = \Xi(\bar{\boldsymbol{\rho}}, \bar{\boldsymbol{\zeta}}) = \Phi^d(\bar{\boldsymbol{\zeta}}). \tag{24}$$

Proof. By the convexity of $\Psi(\boldsymbol{\xi})$, we have the following canonical duality relations:

$$\boldsymbol{\zeta} \in \partial\Psi(\boldsymbol{\xi}) \quad \Leftrightarrow \quad \boldsymbol{\xi} \in \partial\Psi^*(\boldsymbol{\zeta}) \quad \Leftrightarrow \quad \Psi(\boldsymbol{\xi}) + \Psi^*(\boldsymbol{\zeta}) = \boldsymbol{\xi}^T\boldsymbol{\zeta}, \tag{25}$$

where

$$\partial\Psi(\boldsymbol{\xi}) = \begin{cases} \boldsymbol{\zeta} & \text{if } \boldsymbol{\zeta} \in \mathscr{E}_a^* \\ \emptyset & \text{otherwise} \end{cases}$$

is the sub-differential of Ψ. Thus, in terms of $\boldsymbol{\xi} = \Lambda(\boldsymbol{\rho})$ and $\boldsymbol{\zeta} = \{\boldsymbol{\sigma}, \varsigma\}$, the canonical duality relations (25) can be equivalently written as

$$\boldsymbol{\rho} \circ \boldsymbol{\rho} - \boldsymbol{\rho} \leq 0 \quad \Leftrightarrow \quad \boldsymbol{\sigma} \geq 0 \quad \Leftrightarrow \quad \boldsymbol{\sigma}^T(\boldsymbol{\rho} \circ \boldsymbol{\rho} - \boldsymbol{\rho}) = 0 \tag{26}$$

$$\boldsymbol{\rho}^T\mathbf{a} - V_c \leq 0 \quad \Leftrightarrow \quad \varsigma \geq 0 \quad \Leftrightarrow \quad \varsigma(\boldsymbol{\rho}^T\mathbf{a} - V_c) = 0. \tag{27}$$

These are exactly the KKT conditions for the inequality constraints $\boldsymbol{\rho} \circ \boldsymbol{\rho} - \boldsymbol{\rho} \leq 0$ and $\boldsymbol{\rho}^T \mathbf{a} - V_c \leq 0$. Thus, $(\bar{\boldsymbol{\rho}}, \bar{\boldsymbol{\zeta}})$ is a KKT point of Ξ if and only if $\bar{\boldsymbol{\rho}}$ is a KKT point of Φ, $\bar{\boldsymbol{\zeta}}$ is a KKT point of Φ^d. The equality (24) holds due to the canonical duality relations in (25). □

Indeed, on the effective domain $\mathscr{E}_a^*$ of $\Psi^*(\boldsymbol{\zeta})$, the total complementary function Ξ can be written as

$$\Xi(\boldsymbol{\rho}, \boldsymbol{\sigma}, \varsigma) = P_u(\boldsymbol{\rho}) + \boldsymbol{\sigma}^T(\boldsymbol{\rho} \circ \boldsymbol{\rho} - \boldsymbol{\rho}) + \varsigma(\boldsymbol{\rho}^T \mathbf{a} - V_c), \tag{28}$$

which can be considered as the Lagrangian of $(\mathscr{P})$ for the canonical constraint $\Lambda(\boldsymbol{\rho}) \leq 0 \in \mathbb{R}^{n+1}$. The Lagrange multiplier $\boldsymbol{\zeta} = \{\boldsymbol{\sigma}, \varsigma\} \in \mathscr{E}_a^*$ must satisfy the KKT conditions in (26) and (27). By the complementarity condition $\boldsymbol{\sigma}^T(\boldsymbol{\rho} \circ \boldsymbol{\rho} - \boldsymbol{\rho}) = 0$ we know that $\boldsymbol{\rho} \circ \boldsymbol{\rho} = \boldsymbol{\rho}$ if $\boldsymbol{\sigma} > 0$. Let

$$\mathscr{S}_a^+ = \{\boldsymbol{\zeta} = \{\boldsymbol{\sigma}, \varsigma\} \in \mathscr{E}_a^* | \ \boldsymbol{\sigma} > 0\}. \tag{29}$$

Then for any given $\boldsymbol{\zeta} = \{\boldsymbol{\sigma}, \varsigma\} \in \mathscr{S}_a^+$, the function $\Xi(\cdot, \boldsymbol{\zeta}) : \mathbb{R}^m \to \mathbb{R}$ is strictly convex, the canonical dual function of P_u can be well defined by

$$P_u^d(\boldsymbol{\zeta}) = \min_{\boldsymbol{\rho} \in \mathbb{R}^m} \Xi(\boldsymbol{\rho}, \boldsymbol{\zeta}) = -\frac{1}{4} \boldsymbol{\tau}_u^T(\boldsymbol{\zeta}) \mathbf{G}^{-1}(\boldsymbol{\zeta}) \boldsymbol{\tau}_u(\boldsymbol{\zeta}) - \varsigma V_c. \tag{30}$$

Thus, the canonical dual problem of $(\mathscr{P})$ can be proposed as follows:

$$(\mathscr{P}^d): \quad \max\{P_u^d(\boldsymbol{\sigma}, \varsigma) | \ (\boldsymbol{\sigma}, \varsigma) \in \mathscr{S}_a^+\}. \tag{31}$$

Theorem 2 (Analytical Solution). *For any given* $\mathbf{u} \in \mathscr{U}_a^m$*, if* $\bar{\boldsymbol{\zeta}}$ *is a solution to* $(\mathscr{P}^d)$*, then*

$$\bar{\boldsymbol{\rho}} = \frac{1}{2} \mathbf{G}^{-1}(\bar{\boldsymbol{\zeta}}) \boldsymbol{\tau}_u(\bar{\boldsymbol{\zeta}}) \tag{32}$$

is a global optimal solution to $(\mathscr{P})$ *and*

$$P_u(\bar{\boldsymbol{\rho}}) = \min_{\boldsymbol{\rho} \in \mathscr{Z}_a} P_u(\boldsymbol{\rho}) = \max_{\boldsymbol{\zeta} \in \mathscr{S}_a^+} P_u^d(\boldsymbol{\zeta}) = P_u^d(\bar{\boldsymbol{\zeta}}). \tag{33}$$

Proof. It is easy to prove that for any given $\mathbf{u} \in \mathscr{U}_a^m$, the canonical dual function $P_u^d(\boldsymbol{\zeta})$ is concave on the open convex set $\mathscr{S}_a^+$. If $\bar{\boldsymbol{\zeta}}$ is a KKT point of $P_u^d(\boldsymbol{\zeta})$, then it must be a unique global maximizer of $P_u^d(\boldsymbol{\zeta})$ on $\mathscr{S}_a^+$. By Theorem 1 we know that if $\bar{\boldsymbol{\zeta}} = \{\bar{\boldsymbol{\sigma}}, \bar{\varsigma}\} \in \mathscr{S}_a^+$ is a KKT point of $\Phi^d(\boldsymbol{\zeta})$, then $\bar{\boldsymbol{\rho}} = \boldsymbol{\rho}(\bar{\boldsymbol{\zeta}})$ defined by (32) must be a KKT point of $\Phi(\boldsymbol{\rho})$. Since $\Xi(\boldsymbol{\rho}, \boldsymbol{\zeta})$ is a saddle function on $\mathbb{R}^n \times \mathscr{S}_a^+$, we have

$$\min_{\rho \in \mathbb{R}^n} \Phi(\rho) = \min_{\rho \in \mathbb{R}^n} \max_{\zeta \in \mathscr{S}_a^+} \Xi(\rho, \zeta) = \max_{\zeta \in \mathscr{S}_a^+} \min_{\rho \in \mathbb{R}^n} \Xi(\rho, \zeta)$$
$$= \max_{\zeta \in \mathscr{S}_a^+} \Phi^d(\zeta) = \max_{\zeta \in \mathscr{S}_a^+} P_u^d(\zeta),$$

Since $\bar{\sigma} > 0$, the complementarity condition in (26) leads to

$$\bar{\rho} \circ \bar{\rho} - \bar{\rho} = 0 \quad \text{i.e. } \bar{\rho} \in \{0, 1\}^n.$$

Thus, we have

$$P_u(\bar{\rho}) = \min_{\rho \in \mathscr{Z}_a} P_u(\rho) = \max_{\zeta \in \mathscr{S}_a^+} P_u^d(\zeta) = P_u^d(\bar{\zeta})$$

as required.

Remark 1. Theorem 2 shows that although the canonical dual problem is a concave maximization in continuous space, it produces the analytical solution (32) to the well-known integer Knapsack problem $(\mathscr{P}_u)$! This analytical solution was first obtained by Gao in 2007 for general quadratic integer programming problems (see Theorem 3, [9]). The indicator function of a convex set and its sub-differential were first introduced by J.J. Moreau in 1968 in his study on unilateral constrained problems in contact mechanics [14]. His pioneering work laid a foundation for modern analysis and the canonical duality theory. In solid mechanics, the indicator of a plastic yield condition is also called a *super-potential*. Its sub-differential leads to a general constitutive law and a unified pan-penalty finite element method in plastic limit analysis [5]. In mathematical programming, the canonical duality leads to a unified framework for nonlinear constrained optimization problems in multiscale systems [7, 8, 10, 11].

3 Perturbed Canonical Duality Method and Algorithm

Numerically speaking, although the global optimal solution of the integer programming problem $(\mathscr{P})$ can be obtained by solving the canonical dual problem $(\mathscr{P}^d)$, the rate of convergence is very slow since $P_u^d(\sigma, \varsigma)$ is nearly a linear function of $\sigma \in \mathscr{S}_a^+$ when σ is far from its origin. In order to overcome this problem, a so-called β-perturbed canonical dual method has been proposed by Gao and Ruan in integer programming [11], i.e., by introducing a perturbation parameter $\beta > 0$, the problem $(\mathscr{P}^d)$ is replaced by

$$(\mathscr{P}_\beta^d): \quad \max\left\{ P_\beta^d(\sigma, \varsigma) = P_u^d(\sigma, \varsigma) - \frac{1}{4}\beta^{-1}\sigma^T\sigma \mid \{\sigma, \varsigma\} \in \mathscr{S}_a^+ \right\} \qquad (34)$$

which is strictly concave on $\mathscr{S}_a^+$.

Theorem 3. *For a given* $\mathbf{u} \neq 0 \in \mathbb{R}^m$ *and* $V_c > 0$*, there exists a* $\beta_c > 0$ *such that for any given* $\beta \geq \beta_c$*, the problem* $(\mathscr{P}^d_\beta)$ *has a unique solution* $\boldsymbol{\zeta}_\beta \in \mathscr{S}^+_a$*. If* $\boldsymbol{\rho}_\beta = \frac{1}{2}\mathbf{G}^{-1}(\boldsymbol{\zeta}_\beta)\boldsymbol{\tau}_u(\boldsymbol{\zeta}_\beta) \in \{0, 1\}^n$*, then* $\boldsymbol{\rho}_\beta$ *is a global optimal solution to* $(\mathscr{P})$*.*

Proof. It is easy to show that for any given $\beta > 0$, $P^d_\beta(\boldsymbol{\zeta})$ is strictly concave on the open convex set $\mathscr{S}^+_a$, i.e., $(\mathscr{P}^d_\beta)$ has a unique solution. Particularly, the criticality condition $\nabla P^d_\beta(\boldsymbol{\zeta}) = 0$ leads to the following canonical dual algebraic equations:

$$2\beta^{-1}\sigma_e^3 + \sigma_e^2 = (\varsigma a_e - c_e)^2, \quad e = 1, \ldots, n, \tag{35}$$

$$\sum_{e=1}^{n} \frac{1}{2}\frac{a_e}{\sigma_e}(\sigma_e - a_e\varsigma + c_e) - V_c = 0. \tag{36}$$

It was proved in [8] that for any given $\beta > 0$ and $\theta_e = \varsigma a_e - c_e \neq 0, \quad e = 1, \ldots, n$, the canonical dual algebraic equation (35) has a unique positive real solution

$$\sigma_e = \frac{1}{6}\beta[-1 + \phi_e(\varsigma) + \phi_e^c(\varsigma)] > 0, \quad e = 1, \ldots, n \tag{37}$$

where

$$\phi_e(\varsigma) = \eta^{-1/3}\left[2\theta_e^2 - \eta + 2i\sqrt{\theta_e^2(\eta - \theta_e^2)}\right]^{1/3}, \quad \eta = \frac{\beta^2}{27},$$

and ϕ_e^c is the complex conjugate of ϕ_e, i.e., $\phi_e\phi_e^c = 1$. Thus, the canonical dual algebraic equation (36) has a unique solution

$$\varsigma = \frac{\sum_{e=1}^{n} a_e(1 + c_e/\sigma_e) - 2V_c}{\sum_{e=1}^{n} a_e^2/\sigma_e}. \tag{38}$$

This shows that the perturbed canonical dual problem $(\mathscr{P}^d_\beta)$ has a unique solution in $\mathscr{S}^+_a$, which can be analytically obtained by (37) and (38). The rest proof of this theorem is similar to that given in [11]. □

Theoretically speaking, for any given $V_c < V_o$, the perturbed canonical duality method can produce desired optimal solution to the integer constrained problem $(\mathscr{P})$. However, if $V_c \ll V_o$, to reduce the initial volume V_o directly to V_c by solving the bi-level topology optimization problem $(\mathscr{P}_{bl})$ may lead to unreasonable solutions. In order to resolve this problem, a volume decreasing control parameter $\mu \in (V_c/V_o, 1)$ is introduced to slowly reduce the volume in the iteration. Thus, based on the above strategies, the canonical duality algorithm (CDT) for solving the general topology optimization problem $(\mathscr{P}_{bl})$ can be proposed below.

Algorithm 1. (Canonical Dual Algorithm for Topology Optimization (CDT))

(I) Initialization. Let $\boldsymbol{\rho}^0 = \{1\} \in \mathbb{R}^n$. Find $\mathbf{u}^0$ by solving the sublevel optimization problem

$$\mathbf{u}^0 = \arg\min\{\Pi_h(\mathbf{u}, \boldsymbol{\rho}^0)|\ \ \mathbf{u} \in \mathscr{U}_a\}. \tag{39}$$

Compute $\mathbf{c}^0 = \mathbf{c}(\mathbf{u}^0)$ according to (5). Define an initial value $\varsigma_0 > 0$ and an initial volume $V_\gamma \in [V_c, V_o)$. Let $\gamma = 0,\ k = 1$.

(II) Find $\boldsymbol{\sigma}_k = \{\sigma_e^k\} \in \mathbb{R}^n$ by

$$\sigma_e^k = \frac{1}{6}\beta[-1 + \phi(\varsigma^{k-1}) + \phi^c(\varsigma^{k-1})],\quad e = 1, \ldots, n.$$

(III) Find ς^k by

$$\varsigma^k = \frac{\sum_{e=1}^n a_e(1 + c_e^\gamma/\sigma_e^k) - 2V_\gamma}{\sum_{e=1}^n a_e^2/\sigma_e^k.}$$

(IV) If

$$|P_\beta^d(\boldsymbol{\sigma}^k, \varsigma^k) - P_\beta^d(\boldsymbol{\sigma}^{k-1}, \varsigma^{k-1})| \le \omega_1,$$

compute $\boldsymbol{\rho}^\gamma$ by

$$\rho_e^\gamma = \frac{1}{2}[1 - (\varsigma^k a_e - c_e^\gamma)/\sigma_e^k],\quad e = 1, \ldots, n.$$

then go to (V); otherwise, let $k = k + 1$, go to (II).

(V) Find $\mathbf{u}^\gamma$ by solving

$$\mathbf{u}^\gamma = \arg\min\{\Pi_h(\mathbf{u}, \boldsymbol{\rho}^\gamma)|\ \ \mathbf{u} \in \mathscr{U}_a\} \tag{40}$$

(VI) Convergence test: If

$$|C(\boldsymbol{\rho}^\gamma, \mathbf{u}^\gamma) - C(\boldsymbol{\rho}^{\gamma-1}, \mathbf{u}^{\gamma-1})| \le \omega_2,\quad V_\gamma \le V_c$$

then stop; otherwise, let $V_{\gamma+1} = \mu V_\gamma \ge V_o$ and computing $\mathbf{c}^{\gamma+1} = \mathbf{c}(\mathbf{u}^\gamma)$, …, n. Let $\gamma = \gamma + 1$, $k = 1$, go to (II).

The penalty parameter in this algorithm is usually taken $\beta > 10$. For linear elastic materials, the lower level optimization (40) in the algorithm (CDT) can be simply replaced by $\mathbf{u}^\gamma = \mathbf{K}^{-1}(\boldsymbol{\rho}^\gamma)\mathbf{f}(\boldsymbol{\rho}^\gamma)$.

Table 1 The comparison between the SIMP and CDT

Method	Structures	Steps	Compliance
SIMP		41	169.2908
CDT		28	164.7108

4 Numerical Examples for Linear Elastic Structures

The proposed semi-analytic method is implemented in Matlab. For the purpose of illustration, the applied load and geometry data are chosen as dimensionless. Young's modulus and Poisson's ratio of the material are taken as $E = 1$ and $\nu = 0.3$, respectively. The volume fraction is $\mu_c = V_c/V_0 = 0.6$. The stiffness matrix of the structure in CDT algorithm is given by $\mathbf{K}(\boldsymbol{\rho}) = \sum_{e=1}^{n}[E_{min} + (E - E_{min})\rho_e]\mathbf{K}_e$ where $E_{min} = 10^{-9}$ in order to avoid singularity in computation. The evolutionary rate used in the CDT is $\mu = 0.975$. To compare with the SIMP approach, the well-known 88-line algorithm proposed by Andreassen et al. [2] is used with the parameters penal = 3, rmin = 1.5, ft = 1.

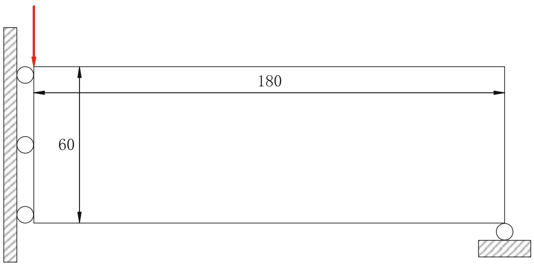


Fig. 1 The design domain, boundary conditions, and external load for a MBB beam

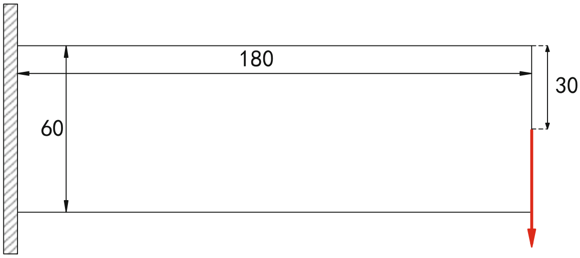


Fig. 2 A test example of the benchmark Cantilever problem

4.1 MBB Beam Problem

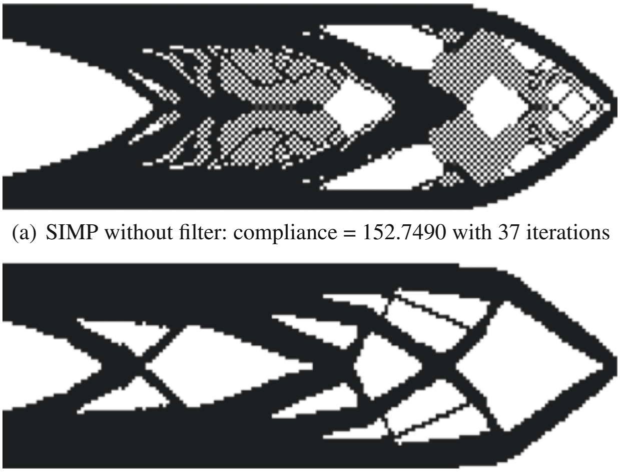
The well-known benchmark Messerschmitt–Bölkow–Blohm (MBB) beam problem in topology optimization is selected as the first test example (see Fig. 1). The design domain is discretized with 180×60 square mesh elements. Computational results obtained by both CDT and SIMP are reported in Table 1.

4.2 Cantilever Beam

The second test example is the classical Cantilever problem (see Fig. 2). The beam is fixed along its left side with a downward traction applied at its right middle point. The example consists of 180×60 quad meshes and the target volume fraction is $\mu_c = 0.6$. Numerical results by both the CDT and SIMP are shown in Fig. 3.

4.3 Summary of Computational Results

The computational results for the above benchmark problems show clearly that without filter, the SIMP produces a large range of checkerboard patterns and gray elements, while by the CDT method, precise void-solid optimal structure can be obtained with very few checkerboard patterns. By the fact that the optimal density distribution $\boldsymbol{\rho}$ can be obtained analytically at each iteration, the CDT method produces desired optimal structure within much less computing time. The convergence of the CDT method depends mainly on the parameter $\mu \in [\mu_c, 1)$. Generally speaking, the smaller μ produces fast convergent but less optimal results. Detailed study on this issue will be addressed in the future research. From the proof of Theorem 3 we know that if $\theta_e = 0$, the canonical dual algebraic equation (32) has two zero solutions, which are located on the boundary of $\mathscr{S}_a^+$. Correspondingly, the density ρ_e can't be analytically given by equation (35). In this case, the primal problem $(\mathscr{P})$



(a) SIMP without filter: compliance = 152.7490 with 37 iterations

(b) CDT: compliance = 153.6767 with 23 iterations

Fig. 3 Topology optimization for the cantilever beam by the SIMP (**a**) and CDT (**b**) methods

could be really NP-hard, which is a conjecture proposed in [10]. This open problem deserves theoretically study in order to completely solve the Knapsack problem.

Acknowledgements Matlab code for the CDT algorithm was helped by Professor M. Li from Zhejiang University. The research is supported by US Air Force Office of Scientific Research under grants FA2386-16-1-4082 and FA9550-17-1-0151.

References

1. Ali, E.J., Gao, D.Y.: Improved canonical dual finite element method and algorithm for post buckling analysis of nonlinear gao beam. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Canonical Duality-Triality: Unified Theory and Methodology for Multidisciplinary Study. Springer, Berlin (2016)
2. Andreassen, E., Clausen, A., Schevenels, M., Lazarov, B.S., Sigmund, O.: Efficient topology optimization in MATLAB using 88 lines of code. Struct. Multidiscip. Optim. **43**(1), 1–16 (2011)
3. Bendsoe, M.P.: Optimal shape design as a material distribution problem. Struct. Optim. **1**, 193C202 (1989)
4. Bendsoe, M.P., Kikuchi, N.: Generating optimal topologies in structural design using a homogenization method. Comput. Methods Appl. Mech. Eng. **72**(2), 197–224 (1988)
5. Gao, D.Y.: Panpenalty finite element programming for limit analysis. Comput. Struct. **28**, 749–755 (1988)
6. Gao, D.Y.: Complementary finite element method for finite deformation nonsmooth mechanics. J. Eng. Math. **30**, 339–353 (1996)
7. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solutions for global optimization. Comput. Chem. Eng. **33**, 1964–1972 (2009)

8. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications, pp. xviii + 454. Springer, New York (2000)
9. Gao, D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. J. Indust. Manage. Optim. **3**(2), 293–304 (2007)
10. Gao, D.Y.: On unified modeling, theory, and method for solving multi-scale global optimization problems. AIP Conf. Proc. **1776**, 020005 (2016). doi:10.1063/1.4965311
11. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. **47**, 463–484 (2010)
12. Gao, D.Y., Strang, G.: Geometric nonlinearity: Potential energy, complementary energy, and the gap function. Quart. Appl. Math. **47**(3), 487–504 (1989)
13. Karp, R.K.: Reducibility among combinatorial problems. In: Miller, R.E., Thatcher, J.W. (eds.) Complexity of Computer Computations, pp. 85–103. Plenum, New York (1972)
14. Moreau, J.J.: La notion de sur-potentiel et les liaisons unilatérales en élastostatique. C.R. Acad. Sci. Paris **267 A**, 954–957 (1968)
15. Santos, H.A.F.A., Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Nonlinear Mech. **7**, 240–247 (2011)
16. Sigmund, O.: A 99 line topology optimization code written in matlab. Struct. Multidiscip. Optim. **21**(2), 120–127 (2001)
17. Sigmund, O., Petersson, J.: Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima. Struct. Optim. **16**(1), 68–75 (1998)
18. Sigmund, O., Maute, K.: Topology optimization approaches: a comparative review. Struct. Multidiscip. Optim. **48**(6), 1031–1055 (2013)
19. Sokolowski, J., Zochowski, A.: On the topological derivative in shape optimization. Struct. Optim. **37**, 1251–1272 (1999)
20. Stolpe, M., Bendsoe, M.P.: Global optima for the Zhou-Rozvany problem. Struct. Multidiscip. Optim. **43**(2), 151–164 (2011)
21. van Dijk, N.P., Maute, K., Langelaar, M., van Keulen, F.: Level-set methods for structural topology optimization: a review. Struct. Multidiscip. Optim. **48**(3), 437–472 (2013)

Improved Canonical Dual Finite Element Method and Algorithm for Post-Buckling Analysis of Nonlinear Gao Beam

Elaf Jaafar Ali and David Yang Gao

Abstract This paper deals a study on post-buckling problem of a large deformed elastic beam by using a canonical dual mixed finite element method (CD-FEM). The nonconvex total potential energy of this beam can be used to model post-buckling problems. To verify the triality theory, different types of dual stress interpolations are used. Applications are illustrated with different boundary conditions and different external loads using semi-definite programming (SDP) algorithm. The results show that the global minimizer of the total potential energy is stable buckled configuration, the local maximizer solution leads to the unbuckled state, and both of these two solutions are numerically stable. While the local minimizer is unstable buckled configuration and very sensitive.

1 Introduction

Nonconvex variational problems have always presented serious challenges not only in numerical analysis, but also in computational mechanics and engineering sciences. By numerical discretization techniques, nonconvex variational problems are linked with certain nonconvex global optimization minimization problems. Due to the lack of global optimality condition, conventional numerical methods and direct approaches cannot solve these problems deterministically. The popular primal–dual interior point methods suffer from uncertain error bounds in nonconvex analysis because of the intrinsic duality gaps produced by traditional duality theories. Therefore, most nonconvex minimization problems are considered as **NP-hard** in global

E.J. Ali (✉) · D.Y. Gao
Faculty of Science and Technology, Federation University Australia, Mt Helen, Victoria 3353, Australia
e-mail: elafali@students.federation.edu.au

D.Y. Gao
e-mail: d.gao@federation.edu.au

E.J. Ali
University of Basrah College of Science, Basra, Iraq

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_14

optimization and computer sciences. Unfortunately, this fundamental difficulty is not fully recognized in computational mathematics and mechanics due to the significant gap between these fields.

Canonical duality theory is a newly developed, potentially powerful methodological theory which can transfer general multi-scale nonconvex problems in R^n to a unified convex dual problem in continuous space R^m with $m \le n$ and without duality gap. The associated *triality theory* provides extremality criteria for both global and local optimal solutions, which can be used to develop powerful algorithms for solving general nonconvex variational problems. This talk will present a canonical dual finite element method (CD-FEM) for solving general nonconvex variational problems. Using Gao–Strang's complementary–dual principle and mixed finite element discretization, the general nonconvex variational problem can be reformulated as a min–max optimization problem of a saddle function. Based on the triality theory and the SDP method, a canonical primal–dual algorithm is proposed. Detailed application will be illustrated by post-buckling problem of a large elastic deformations of beam, which is governed by a fourth-order nonlinear differential equation. The total potential energy of this beam is a double-welled nonconvex functional with two local minimizers, representing the two buckled states, and one local maximizer representing the unbuckled state.

The purpose of the present work is to verify the triality theory to find all solutions of the post-buckling problem of a large deformation nonlinear beam. Mixed finite element method with mixed meshes of different dual stress interpolations are used to get a closed dimensions between the discretized displacement and discretized stress. Numerical results show that the our algorithm can produce a stable solutions for the global minimizer and local maximizer. However, the local minimizer is very sensitive to numerical discretization and external loads.

2 Nonconvex Problem and Canonical Dual–Complementary Principle

Let us consider an elastic beam subjected to a vertical distributed lateral load $q(x)$ and compressive external axial force F at the right end as shown in Fig. 1. It was

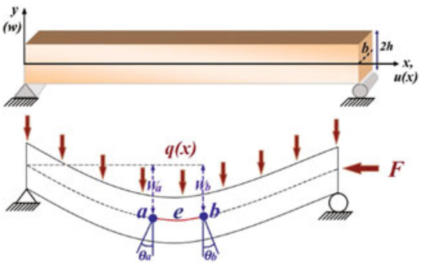


Fig. 1 Simply supported beam model

discovered by Gao in 1996 that the well-known von Karman nonlinear plate model in one dimension is actually equivalent to a linear differential equation and therefore, it cannot be used for studying post-buckling phenomena [4]. The main reason for this "paradox" is due to the fact that the stress in lateral direction of large deformed plate was ignored by von Karman. Therefore, von Karman equation works only for thin plate and cannot be used as a beam model. For a relatively thick beam such that $h/L \sim w(x) \in O(1)$, the deformation in the lateral direction can not be ignored. Based on the finite deformation theory for Hooke' material and EulerBernoulli hypothesis (i.e., straight lines normal to the mid-surface remain straight and normal to the mid-surface after deformation), a nonlinear beam model was proposed by Gao [4]:

$$EI.w_{,xxxx} - \alpha E w_{,x}^2 w_{,xx} + E\lambda w_{,xx} - f(x) = 0, \ \forall x \in [0\ L], \tag{1}$$

where E is the elastic modulus of material, $I = 2h^3/3$ is the second moment of area of the beam's cross section, w is the transverse displacement field of the beam, $\alpha = 3h(1 - \nu^2) > 0$ with ν as the Poisson's ratio, $\lambda = (1 + \nu)(1 - \nu^2)F/E > 0$ is an integral constant, $f(x) = (1 - \nu^2)q(x)$ depends mainly on the distributed lateral load $q(x)$; $2h$ and L represent to the height and length of the beam, respectively. The axial displacement $u(x)$ is governed by the following differential equation [4]:

$$u_x = -\frac{1}{2}(1 + \nu)w_{,x}^2 - \frac{\lambda}{2h(1 + \nu)}\ , \tag{2}$$

which shows that $u(x) \sim w_{,x}(x) \in O(\epsilon),\ \ u_{,x}(x) \sim w_{,xx}(x) \in O(\epsilon^2)$. The total potential energy attendant of this problem is the function $\Pi(w) : \mathscr{U}_a \to R$ define by

$$\Pi(w) = \int_0^L \left(\frac{1}{2}EIw_{,xx}^2 + \frac{1}{12}E\alpha w_{,x}^4 - \frac{1}{2}E\lambda w_{,x}^2 - f(x)\ w\right)dx = 0, \tag{3}$$

where $\mathscr{U}_a$ is the kinematically admissible space, in which certain necessary boundary conditions are given. Thus, for the given external loads $f(x)$ and λ, the primal variational problem is to find $\bar{w} \in \mathscr{U}_a$ such that

$$(\mathscr{P}) : \quad \Pi(\bar{w}) = \inf\,\{\Pi(w) | w \in \mathscr{U}_a\}. \tag{4}$$

It is easy to prove that the stationary condition $\delta\Pi(w) = 0$ leads to the governing equation (1). From the classic beam theory, the Euler buckling load can be determined by

$$\lambda_{cr} = \inf_{w \in \mathscr{U}_a} \frac{\int_0^L EIw_{,xx}^2 dx}{\int_0^L Ew_{,x}^2 dx}\ . \tag{5}$$

Clearly, before the axial load λ reaches to the Euler buckling load λ_{cr}, the total potential energy $\Pi(w)$ is convex on $\mathscr{U}_a$ and the nonlinear differential equation (1) has only one solution. When $\lambda > \lambda_{cr}$, the beam is in a post-buckling state. In this case, the total potential energy Π is nonconvex and Eq. (1) may have at most three (strong) solutions [6] at each material point $x \in [0\ L]$: two minimizers corresponding to the two possible buckled states, one maximizer corresponding to the possible unbuckled state. Clearly, these solutions are sensitive to both the axial load λ and the distributed lateral force field $f(x)$. By Eq. (2) we know that the axial deformation could be relatively larger, the Gao beam model can be used for studying both pre- and post-buckling problems in engineering and sciences [2, 11]. Mathematically, due to the fact that traditional numerical methods and convex optimization techniques cannot identify the global minimizer at each numerical iteration, most of nonconvex optimization problems are considered to be NP-hard in global optimization and computer science [7]. the Gao–Strang total complementary energy $\Xi : \mathscr{U}_a \times \mathscr{S}_a \to R$ [8] in nonlinear elasticity can be defined as

$$\begin{aligned}\Xi(w,\sigma) &= \int_0^L \left(\frac{1}{2} EI w_{,xx}^2 + \frac{1}{2}\sigma w_{,x}^2 - \frac{3}{4E\alpha}(\sigma + E\lambda)^2 - f(x)w \right) dx \\ &= G(w,\sigma) - \int_0^L [V^*(\sigma) - f(x)w] dx, \end{aligned} \tag{6}$$

where $\mathscr{S}_a = \{\sigma \in C[0, L] |\ \ \sigma(x) \geq -\lambda E \quad \forall x \in [0, L]\}$ and

$$G(w,\sigma) = \int_0^L \left(\frac{1}{2} EI w_{,xx}^2 + \frac{1}{2}\sigma w_{,x}^2 \right) dx$$

is the generalized Gao–Strang complementary gap function [8].

3 Mixed Finite Element Method and Triality Theory

In order to apply FEM, the domain of the beam is discretized into m elements $[0, L] = \bigcup_{e=1}^m \Omega^e$. In each element $\Omega^e = [x_a, x_b]$, the deflection, rotating angular, and dual stress for the node x_a are marked as w_a, θ_a, and σ_a, respectively, and similar for the node x_b. Then, we have the nodal displacement vector $w_e^T = [w_a \ \ \theta_a \ \ w_b \ \ \theta_b]$ of the e-th element and the nodal dual stress element $\sigma_e^T = [\sigma_a \ \ \sigma_b]$. In each element, we use mixed finite element interpolations for both $w(x)$ and $\sigma(x)$, i.e.,

$$w_e^h(x) = N_w^T(x) w_e \quad , \qquad \sigma_e^h(x) = N_\sigma^T(x) \boldsymbol{\sigma}_e \quad \forall x \in \Omega^e.$$

Thus, the spaces $\mathscr{U}_a$ and $\mathscr{S}_a$ can be numerically discretized to the finite-dimensional spaces $\mathscr{U}_a^h$ and $\mathscr{S}_a^h$, respectively. The shape function for $w(x)$ is based on piecewise-cubic polynomial, i.e.,

$$N_w = \begin{bmatrix} \frac{1}{4}\ (1-\xi)^2\ (2+\xi) \\ \frac{L_e}{8}(1-\xi)^2\ (1+\xi) \\ \frac{1}{4}\ (1+\xi)^2\ (2-\xi) \\ \frac{L_e}{8}(1+\xi)^2\ (\xi-1) \end{bmatrix},$$

where $\xi = 2x/L_e - 1$ and L_e is the length of e-th beam element. The shape function for σ is based on different dual stress interpolations; piecewise-linear stresses (PLS, $\delta = 1$), piecewise-quadratic stresses (PQS, $\delta = 2$), and piecewise-cubic stresses (PCS, $\delta = 3$) as follows:

$$N_\sigma|_{\delta=1} = \frac{1}{2}\begin{bmatrix} 1-\xi \\ 1+\xi \end{bmatrix}, \qquad N_\sigma|_{\delta=2} = \frac{1}{2}\begin{bmatrix} \xi^2-\xi \\ 1-\xi^2 \\ \xi^2+\xi \end{bmatrix},$$

and

$$N_\sigma|_{\delta=3} = \frac{1}{16}\begin{bmatrix} -1+\xi+9\xi^2-9\xi^3 \\ 9-27\xi-9\xi^2+27\xi^3 \\ 9+27\xi-9\xi^2-27\xi^3 \\ -1-\xi+9\xi^2+9\xi^3 \end{bmatrix},$$

where δ refers to the number of straight lines inside the element e as shown in Fig. 2.

Thus, on the discretized feasible deformation space $\mathscr{U}_a^h$, the Gao–Strang total complementary energy can be expressed in the following discretized form:

$$\begin{aligned} \Xi^h(\mathbf{w}, \boldsymbol{\sigma}) &= \sum_{e=1}^{m} \left(\frac{1}{2} w_e^T\ G^e(\sigma_e)\ w_e - \frac{1}{2}\sigma_e^T\ K_e\ \sigma_e - \lambda_e^T\ \sigma_e - f_e^T\ w_e - c_e \right) \\ &= \frac{1}{2}\mathbf{w}^T\ \mathbf{G}(\boldsymbol{\sigma})\ \mathbf{w} - \frac{1}{2}\boldsymbol{\sigma}^T\ \mathbf{K}\ \boldsymbol{\sigma} - \boldsymbol{\lambda}^T\ \boldsymbol{\sigma} - \mathbf{f}^T\ \mathbf{w} - c, \end{aligned} \tag{7}$$

where $\mathbf{w} \in \mathscr{U}_a^h \subset R^{2(m+1)}$ and $\boldsymbol{\sigma} \in \mathscr{S}_a^h \subset R^{\delta m+1}$ are nodal deflection and dual stress vectors, respectively. We let

$$\mathscr{S}_a^h = \{\boldsymbol{\sigma} \in R^{\delta m+1} |\ \ \det \mathbf{G}(\boldsymbol{\sigma}) \neq 0\}. \tag{8}$$

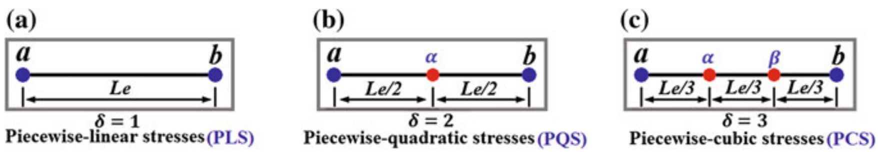


Fig. 2 Dual stress nodes in an element

The Hessian matrix of the gap function $\mathbf{G}(\boldsymbol{\sigma}) \in R^{2(m+1)} \times R^{2(m+1)}$ is obtained by assembling the following symmetric matrices $G^e(\sigma_e)$:

$$G^e(\sigma_e) = \int_{\Omega_e} \left(EI\ N_w''\ (N_w'')^T + (N_\sigma)^T\ \sigma_e\ N_w'\ (N_w')^T \right) dx. \tag{9}$$

The matrix $\mathbf{K} \in R^{\delta m+1} \times R^{\delta m+1}$ is obtained by assembling the following positive-definite matrices K_e

$$K_e = \int_{\Omega_e} \left(\frac{3}{2E\alpha} N_\sigma\ N_\sigma^T \right) dx.$$

Also, $\boldsymbol{\lambda} = \{\lambda_e\} \in R^{\delta m+1}$, $\mathbf{f} = \{f_e\} \in R^{2(m+1)}$ are defined by assembling the corresponding element components $\lambda_e = \int_{\Omega_e} \left(\frac{3}{2\alpha} \lambda N_\sigma \right) dx,\ \ f_e = \int_{\Omega_e} f\big(x\big)\ N_w\ dx$, and $c = \sum_{e=1}^{m} c_e \in R$, where $c_e = \int_{\Omega_e} \left(\frac{3E}{4\alpha} \lambda^2 \right) dx = \frac{3}{4\alpha} E L_e \lambda^2$.

By the critical condition $\delta \Xi^h(\mathbf{w}, \boldsymbol{\sigma}) = 0$, we obtain the two equations $\mathbf{G}(\boldsymbol{\sigma})\ \mathbf{w} - \mathbf{f} = 0$, and $\frac{1}{2} \mathbf{w}^T\ \mathbf{G}_{,\boldsymbol{\sigma}}(\boldsymbol{\sigma})\ \mathbf{w} - \mathbf{K}\ \boldsymbol{\sigma} - \boldsymbol{\lambda} = 0$, where $\mathbf{G}_{,\boldsymbol{\sigma}}(\boldsymbol{\sigma})$ is gradient of $\mathbf{G}$ respect to $\boldsymbol{\sigma}$. The discretized pure complementary energy $\Pi_d^h : \mathscr{S}_a^h \to R$ can be obtained by the following canonical dual transformation:

$$\Pi_d^h(\boldsymbol{\sigma}) = -\frac{1}{2} \mathbf{f}^T\ \mathbf{G}^{-1}(\boldsymbol{\sigma})\ \mathbf{f} - \frac{1}{2} \boldsymbol{\sigma}^T\ \mathbf{K}\ \boldsymbol{\sigma} - \boldsymbol{\lambda}^T\ \boldsymbol{\sigma} - c \tag{10}$$

Suppose $(\bar{\mathbf{w}}, \bar{\boldsymbol{\sigma}})$ is a stationary point of $\Xi^h(\mathbf{w}, \boldsymbol{\sigma})$, and let $\mathscr{S}_a^+ = \{\boldsymbol{\sigma} \in \mathscr{S}_a^h \mid \mathbf{G}(\boldsymbol{\sigma}) \succ 0\}$, and $\mathscr{S}_a^- = \{\boldsymbol{\sigma} \in \mathscr{S}_a^h \mid \mathbf{G}(\boldsymbol{\sigma}) \prec 0\}$. Then, by "*Complementary–duality Principle theorem*" [5], we have the following theorem.

Theorem 1. *Suppose* $(\bar{\mathbf{w}}, \bar{\boldsymbol{\sigma}})$ *is a stationary point of* $\Xi^h(\mathbf{w}, \boldsymbol{\sigma})$*, then* $\Pi_p^h(\bar{\mathbf{w}}) = \Xi^h(\bar{\mathbf{w}}, \bar{\boldsymbol{\sigma}}) = \Pi_d^h(\bar{\boldsymbol{\sigma}})$*. Moreover, if* $\bar{\boldsymbol{\sigma}} \in \mathscr{S}_a^+$*, then we have*

Canonical Min–Max Duality: *The stationary point* $\bar{\mathbf{w}}$ *is a global minimizer of* $\Pi_p^h(\mathbf{w})$ *on* $\mathscr{U}_a^h$ *if and only if* $\bar{\boldsymbol{\sigma}}$ *is a global maximizer of* $\Pi_d^h(\boldsymbol{\sigma})$ *on* $\mathscr{S}_a^+$*, i.e.,*

$$\Pi_p^h(\bar{\mathbf{w}}) = \min_{\mathbf{w} \in \mathscr{U}_a^h} \Pi_p^h(\mathbf{w}) \quad \Leftrightarrow \quad \max_{\boldsymbol{\sigma} \in \mathscr{S}_a^+} \Pi_d^h(\boldsymbol{\sigma}) = \Pi_d^h(\bar{\boldsymbol{\sigma}}). \tag{11}$$

If $\bar{\boldsymbol{\sigma}} \in \mathscr{S}_a^-$*, then on a neighborhood* $\mathscr{U}_o \times \mathscr{S}_o \subset \mathscr{U}_a^h \times \mathscr{S}_a^-$ *of* $(\bar{\mathbf{w}}, \bar{\boldsymbol{\sigma}})$ *we have*

Canonical Double-max Duality: *The stationary point* $\bar{\mathbf{w}}$ *is a local maximizer of* $\Pi_p^h(\mathbf{w})$ *on* $\mathscr{U}_o$ *if and only if the stationary point* $\bar{\boldsymbol{\sigma}}$ *is a local maximizer of* $\Pi_d^h(\boldsymbol{\sigma})$ *on* $\mathscr{S}_o$*, i.e.,*

$$\Pi_p^h(\bar{\mathbf{w}}) = \max_{\mathbf{w} \in \mathscr{U}_o} \Pi_p^h(\mathbf{w}) \quad \Leftrightarrow \quad \max_{\boldsymbol{\sigma} \in \mathscr{S}_o} \Pi_d^h(\boldsymbol{\sigma}) = \Pi_d^h(\bar{\boldsymbol{\sigma}}) \tag{12}$$

Canonical Double-min Duality: *The stationary point* $\bar{\mathbf{w}}$ *is a local minimizer of* $\Pi_p^h(\mathbf{w})$ *on* $\mathscr{U}_o$ *if and only if the stationary point* $\bar{\boldsymbol{\sigma}}$ *is a local minimizer of* $\Pi_d^h(\boldsymbol{\sigma})$ *on* $\mathscr{S}_o$*, i.e.,*

$$\Pi_p^h(\bar{\mathbf{w}}) = \min_{\mathbf{w}\in\mathcal{U}_o} \Pi_p^h(\mathbf{w}) \quad \Leftrightarrow \quad \min_{\boldsymbol{\sigma}\in\mathcal{S}_o} \Pi_d^h(\boldsymbol{\sigma}) = \Pi_d^h(\bar{\boldsymbol{\sigma}}). \tag{13}$$

The proof of this theorem follows from the general results in global optimization [3, 9, 10]. The canonical min–max duality can be used to find global minimizer of the nonconvex problem by the canonical dual problem $\max\{\Pi_d^h(\boldsymbol{\sigma})|\ \boldsymbol{\sigma} \in \mathcal{S}_a^+\}$, which is a concave maximization problem and can be solved easily by well-developed convex analysis and optimization techniques. The canonical double-max and double-min duality statements can be used to find the biggest local maximizer and a local minimizer of the nonconvex primal problem, respectively. It was proved in [3, 9, 10] that both the canonical min–max and double-max duality statements hold strongly regardless the dimensions of $\mathcal{U}_a^h$ and $\mathcal{S}_a^h$, while the canonical double-min duality statement (13) holds weakly for $\dim \mathcal{U}_a^h \neq \dim \mathcal{S}_a^h$, but it holds strongly if $\dim \mathcal{U}_a^h = \dim \mathcal{S}_a^h$. This case is within our reach in the following applications.

4 Semi-definite Programming Algorithm

According to Schur complement lemma [12], the global optimization problem $\min_{\mathbf{w}\in\mathcal{U}_a^h} \Pi_p^h(\mathbf{w}) = \min_{\mathbf{w}\in\mathcal{U}_a^h} \max_{\boldsymbol{\sigma}\in\mathcal{S}_a^h} \Xi(\mathbf{w},\boldsymbol{\sigma})\ \ s.t.\ \ \mathbf{G}(\boldsymbol{\sigma}) \succeq 0$, can be relaxed to the following SDP problem [1]:

$$\begin{aligned} &\max_{\boldsymbol{\sigma},t}\ t \\ &s.t.\ \mathbf{G}(\boldsymbol{\sigma}) \succeq 0, \quad \begin{bmatrix} 2\mathbf{K}^{-1} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma}^T & \frac{1}{2}\mathbf{w}^T\mathbf{G}(\boldsymbol{\sigma})\mathbf{w} - \boldsymbol{\lambda}^T\boldsymbol{\sigma} - \mathbf{f}^T\mathbf{w} - c - t \end{bmatrix} \succeq 0, \end{aligned} \tag{14}$$

where $\mathbf{w} = \mathbf{w}(\boldsymbol{\sigma}) = \mathbf{G}^{-1}(\boldsymbol{\sigma})\mathbf{f}$. By the fact that $\mathbf{K} \succeq 0$, the second inequality constraint implies to; $t(\mathbf{w},\boldsymbol{\sigma}) \leq \frac{1}{2}\mathbf{w}^T\mathbf{G}(\boldsymbol{\sigma})\mathbf{w} - \boldsymbol{\lambda}^T\boldsymbol{\sigma} - \mathbf{f}^T\mathbf{w} - c$.

By the same way, the SDP relaxation for the canonical double-max duality statement, $\max_{\mathbf{w}\in\mathcal{U}_a^h} \Pi_p^h(\mathbf{w}) = \max_{\mathbf{w},\boldsymbol{\sigma}} \Xi(\mathbf{w},\boldsymbol{\sigma}) = \max \Pi_d^h(\boldsymbol{\sigma})\ \ \text{s.t.}\ \ \boldsymbol{\sigma} \in \mathcal{S}_a^-$ should be equivalent to [1]:

$$\begin{aligned} &\max_{\boldsymbol{\sigma},t}\ t \\ &s.t.\ -\mathbf{G}(\boldsymbol{\sigma}) \succeq 0, \quad \begin{bmatrix} 2\mathbf{K}^{-1} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma}^T & \frac{1}{2}\mathbf{w}^T\mathbf{G}(\boldsymbol{\sigma})\mathbf{w} - \boldsymbol{\lambda}^T\boldsymbol{\sigma} - \mathbf{f}^T\mathbf{w} - c - t \end{bmatrix} \succeq 0. \end{aligned} \tag{15}$$

which leads to a local maximum solution to the post-buckling problem.

To find the local minimum for the beam post-buckling problem, it is appropriate to use the following new formula of pure complementary energy [1]:

$$\widehat{\Pi}^d(\boldsymbol{\sigma},\mathbf{w}) = -\frac{1}{2}\mathbf{f}^T\ \mathbf{G}^{-1}(\boldsymbol{\sigma})\ \mathbf{f} - \frac{1}{2}\mathbf{w}^T\ \mathbf{M}(\boldsymbol{\sigma})\ \mathbf{w} - \frac{1}{2}\boldsymbol{\lambda}^T\ \boldsymbol{\sigma} - c. \tag{16}$$

The SDP relaxation for the canonical double-min duality statement $\min_{\mathbf{w}} \Pi_p^h(\mathbf{w}) = \min_{\mathbf{w},\boldsymbol{\sigma}} \Xi(\mathbf{w},\boldsymbol{\sigma}) = \min_{\mathbf{w},\boldsymbol{\sigma}} \widehat{\Pi}^d(\boldsymbol{\sigma},\mathbf{w})$ s.t. $\boldsymbol{\sigma} \in \mathscr{S}_a^-$ and for $\mathbf{w} = \mathbf{w}(\boldsymbol{\sigma})$ should be equivalent to

$$\min_{\boldsymbol{\sigma},t} \; t$$
$$s.t. \; -\mathbf{G}(\boldsymbol{\sigma}) \succ 0, \quad \begin{bmatrix} -2\mathbf{G}(\boldsymbol{\sigma}) & \mathbf{f} \\ \mathbf{f}^T & \frac{1}{2}\mathbf{w}^T\mathbf{M}(\boldsymbol{\sigma})\mathbf{w} + \frac{1}{2}\boldsymbol{\lambda}^T\boldsymbol{\sigma} + c + t \end{bmatrix} \succeq 0. \tag{17}$$

Where $\mathbf{M}(\boldsymbol{\sigma})$ is obtained by assembling the following symmetric matrices $M^e(\sigma_e)$:

$$M^e(\sigma_e) = \int_{\Omega_e} \frac{1}{2}\Big((N_\sigma)^T \; \sigma_e \; N_w' \; (N_w')^T \Big) dx \tag{18}$$

The post-buckling configurations of a large deformed nonlinear beam can be found by the following steps:

1. With an initial point $\mathbf{w}^{(k=1)}$, the next steps are repeated as $\mathbf{w}^{(k+1)}$ converges to the solution.
2. Find $\boldsymbol{\sigma}^{(k+1)}$ by applying SDP algorithm for global maximizer and local minimizer problems in (15) and (17), respectively.
3. Compute $\mathbf{w}^{(k+1)} = \mathbf{G}^{-1}(\boldsymbol{\sigma}^{(k+1)})\mathbf{f}$.
4. Check convergence; if $\|\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\|/\|\mathbf{w}^{(k)}\| \le \epsilon$, stop with optimal solution $\mathbf{w}^* = \mathbf{w}^{(k+1)}$, where ϵ is a small positive real number. Otherwise, put $k = k + 1$ and return to step 2.

For applying SDP algorithm, a software package named SeDuMi [13] is used to solve the problems (15) and (17) via the interior point method.

5 Numerical Solutions with Different Dual Stress Interpolations

According to the triality theory, the canonical double-min duality statement (13) holds strongly if $\dim \mathscr{U}_a^h = \dim \mathscr{S}_a^h$. So, the piecewise-quadratic stress ($\delta = 2$) is the most convenient to verify this theory to obtain closed dimensions between the discretized displacement $\mathbf{w} \in R^{2(m+1)}$ and discretized stress $\boldsymbol{\sigma} \in R^{\delta m+1}$. But these two dimensions are still not equal. However, it is possible to make these dimensions equal if we use mixed different dual stress interpolations on the elements of the same beam. So, many mixed meshes of dual stress interpolations are used in this paper beside to the "PLS mesh" and "PQS mesh" in order to improve the local unstable buckled configuration solution of a large deformed beam.

We present four different types of beams which are controlled by different boundary conditions. Some geometrical data are kept fixed for all computations; $E = 1000\,\text{Pa}$, $\nu = 0.3$, $L = 1\,\text{m}$, $h = 0.05\,\text{m}$ with an odd number of beam elements

$m = 51$. Different loading conditions, including both axial and transverse arrangements, are considered in our applications.

5.1 Simply Supported Beam

A simply supported beam model is fixed in both directions at $x = 0$ and fixed only in the y-direction at $x = L$ as shown in Fig. 3-a. By applying the boundary conditions, $w(0) = w''(0) = w(L) = w''(L) = 0$, two elements of discretized displacement $\mathbf{w} = \{w_e\} \in R^{2(m+1)}$ should be zero. Then, the remaining nonzero elements of the vector w is $(2m)$. We used three types of dual stress interpolations to construct a mixed mesh of dual stress fields in order to obtain $\dim \mathscr{U}_a^h = \dim \mathscr{S}_a^h$. The PQS is applied on $(m-3)$ elements and the PCS is used for only one element that is on the central of the beam. While the PLS is applied on two beam elements which surround the central element as shown in "*Mesh-1*" in Fig. 4. So, we have $dim(\boldsymbol{\sigma}) = dim(\mathbf{w}) = 2m$, and this dimension equals 102 for $m = 51$. The critical load of the simply supported beam is $\lambda_{cr} = 0.00097m^2$, see Eq. (5). The approximate deflections with $\lambda > \lambda_{cr}$ under both of uniformly distributed load and concentrated force are shown in Figs. 5 and 6, respectively.

5.2 Doubly/Clamped Beam

Doubly/clamped beam is fixed at both ends (see Fig. 3-c). The boundary conditions, $w(0) = w'(0) = w(L) = w'(L) = 0$, force the first two and the last two elements

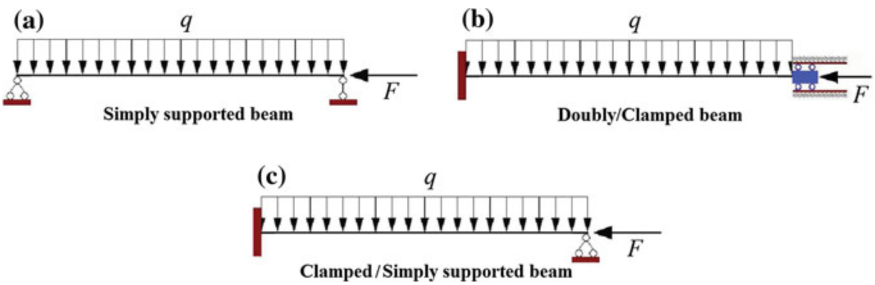


Fig. 3 Different types of beams

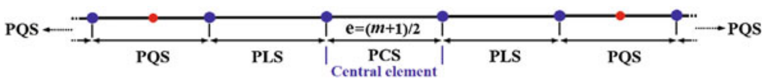


Fig. 4 **Mesh-1**: Mixed dual stress interpolations of beam elements

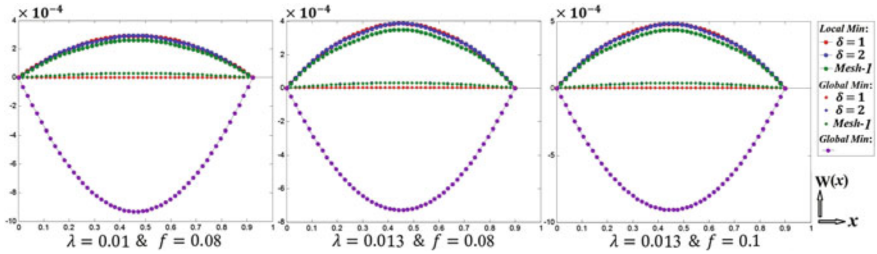


Fig. 5 Post-buckling solutions of simply supported beam under uniformly distributed load

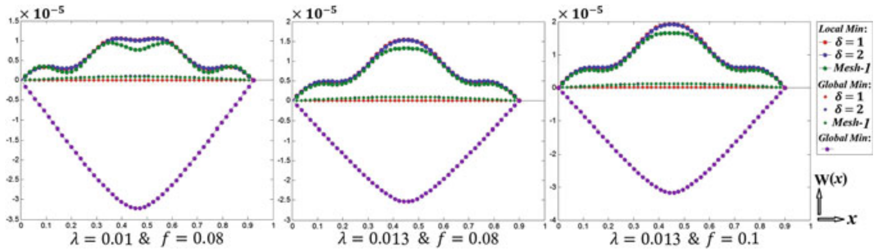


Fig. 6 Post-buckling solutions of simply supported beam under a concentrated force

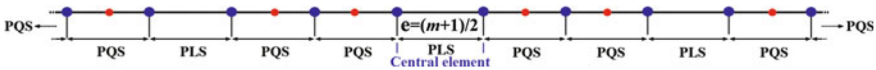


Fig. 7 **Mesh-3**: Mixed dual stress interpolations of beam elements

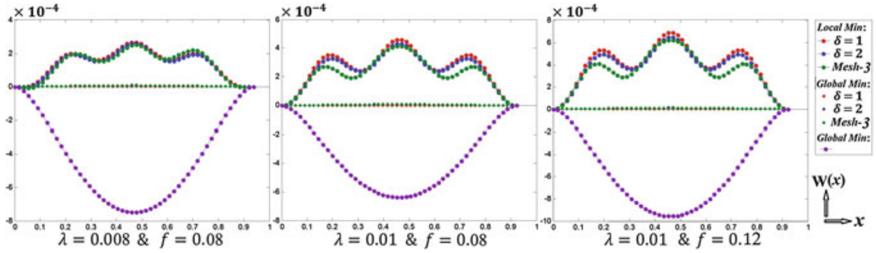


Fig. 8 Post-buckling configurations of clamped beam under uniformly distributed load

of discretized displacement w to be zero. Thus, the remaining nonzero element of displacement vector is $(2m - 2)$. The selected mixed mesh for dual stress field is "*Mesh-3*" which contains $(m - 3)$ of PQS, while PLS is used for three beam elements (see Fig. 7). For $m = 51$, the $dim(\boldsymbol{\sigma}) = dim(\mathbf{w}) = 100$. The approximate deflections for $\lambda > \lambda_{cr}$ with $\lambda_{cr} = 0.0041m^2$ under uniformly distributed load and concentrated force are shown in Figs. 8 and 9, respectively.

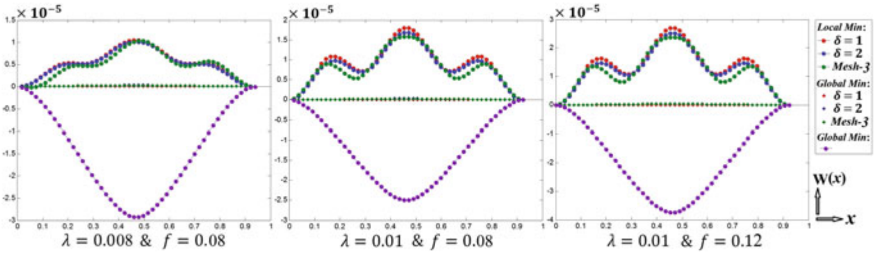


Fig. 9 Post-buckling configurations of clamped beam under a concentrated force

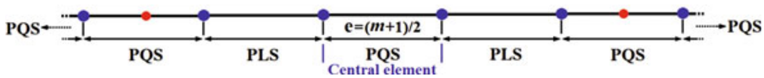


Fig. 10 **Mesh-4**: Mixed dual stress interpolations of beam elements

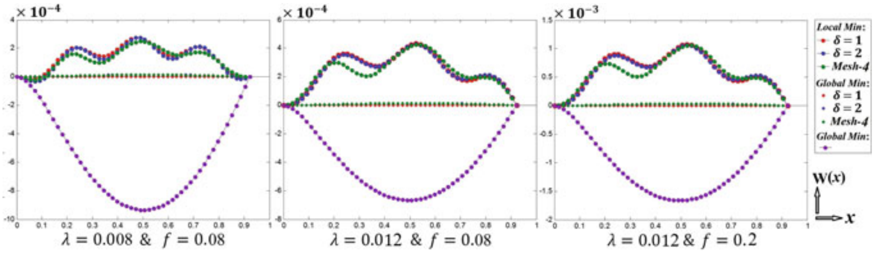


Fig. 11 Post-buckling configurations of clamped/simply supported beam under uniformly distributed load

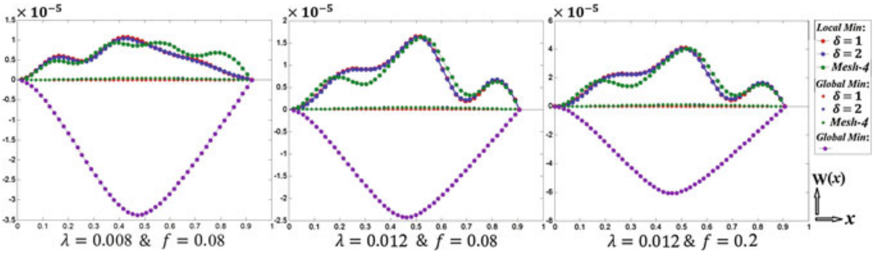


Fig. 12 Post-buckling configurations of clamped/simply supported beam under a concentrated force

5.3 Clamped/Simply Supported Beam

Clamped/simply supported beam is clamped at $x = 0$ and fixed in both directions at $x = L$ as shown in Fig. 3-d. Three elements of discretized displacement w should be zero after applying the boundary conditions; $w(0) = w'(0) = w(L) = w''(L) = 0$. The remaining nonzero element of w is $(2m - 1)$. The "*Mesh-4*" is designed by applying two different dual stress interpolations. The PQS is applied for $(m - 3)$

beam elements, while the PLS is applied on two beam elements which surround the central element as shown in Fig. 10. Thus, for $m = 51$, the $dim(\boldsymbol{\sigma}) = dim(\mathbf{w}) = 101$. The critical load of this beam is $\lambda_{cr} = 0.0034m^2$. The approximate deflections under uniformly distributed load and concentrated force are shown in Figs. 11 and 12, respectively.

6 Conclusions

This paper presents a CD-FEM for the post-buckling analysis with a large elastic deformations beam which is governed by a fourth-order nonlinear differential equation which was introduced by Gao in 1996. The generalized total complementary energy $\Xi(\mathbf{w}, \boldsymbol{\sigma})$ associated with this model is a nonconvex functional and was used to study the post-buckling problems. Combining the generalized total complementary energy and the proposed formula of pure complementary energy $\widehat{\Pi}^d(\boldsymbol{\sigma}, \mathbf{w})$ with the triality theory, a canonical duality algorithm is studied for solving post-buckling problems using SDP algorithm. According to the triality theory, the dimensions of discretized displacement and dual stress have been made equal by designing a number of mixed meshes of different dual stress interpolations. Different boundary conditions and different loading conditions, including both axial and transverse arrangements are considered in our applications. The numerical results show that the global minimizer and local maximizer of the total potential energy are stable buckled configuration for different dual stress meshes. While the local minimizer present unstable deformation states and the solutions of unstable buckled state is sensitive to both stress interpolations and external loads.

Acknowledgements The research is supported by US Air Force Office of Scientific Research under grants FA2386-16-1-4082 and FA9550-17-1-0151.

References

1. Ali, E., Gao, D.: On SDP method for solving canonical dual problem in post buckling of large deformed elastic beam. Appl. Math. Comput. (2016). arXiv:1609.04675
2. Cai, K., Gao, D., Qin, Q.: Post-buckling solutions of hyper-elastic beam by canonical dual finite element method. Math. Mech. Solids **19**(6), 659–671 (2014)
3. Chen, Y., Gao, D.: Global solutions to nonconvex optimization of 4th-order polynomial and log-sum-exp functions. J. Glob. Optim. (2014). doi:10.1007/s10898-014-0244-5
4. Gao, D.: Nonlinear elastic beam theory with application in contact problems and variational approaches. Mech. Res. Commun. **23**(1), 11–17 (1996)
5. Gao, D.: Duality Principles in Nonconvex Systems: Theory. Methods and Applications. Kluwer, London (2000)
6. Gao, D., Ogden, R.: Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem. Z. Angew. Math. Phys. **59**, 498–517 (2008)
7. Gao, D., Sherali, H.: Canonical duality: connection between nonconvex mechanics and global optimization. Springer, Advances in Applied Mathematics and Global Optimization (2009)

8. Gao, D., Strang, G.: Geometric nonlinearity: potential energy, complementary energy and the gap function. Q. Appl. Math. XLVII **3**, 487–504 (1989)
9. Gao, D., Wu, C.: On the triality theory for a quartic polynomial optimization problem. J. Ind. Manag. Optim. **8**(1), 229–242 (2012)
10. Morales, D., Gao, D.: Canonical duality theory and triality for solving general nonconstrained global optimization problems. Math. Mech, Complex Syst. **3**(2), 139–161 (2015). arXiv:1210.0180
11. Santos, H., Gao, D.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Non-Linear Mech. **47**, 240–247 (2012)
12. Stephen, B., Lieven, V.: Convex Optimization, 1st edn. Cambridge University Press, Cambridge (2004)
13. Sturm, J.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. **11**, 625–653 (1999)

Global Solutions to Spherically Constrained Quadratic Minimization via Canonical Duality Theory

Yi Chen and David Yang Gao

Abstract This paper presents a detailed study on global optimal solutions to a nonconvex quadratic minimization problem with a spherical constraint, which is well known as a trust region subproblem and has been studied extensively for decades. The main challenge is solving the so-called hard case, i.e., the problem has multiple solutions on the boundary of the sphere. By canonical duality-triality theory, this challenging problem is able to be reformulated as a one-dimensional canonical dual problem, without any duality gaps. Results show that this problem is in the hard case if and only if certain conditions are satisfied by both the direction and norm of coefficient of the linear item in the objective function. A perturbation method and associated algorithms are proposed to solve hard-case problems. Theoretical results and methods are verified by numerical examples.

1 Introduction

We consider the following quadratic minimization problem:

$$(\mathscr{P}) \quad \min \quad P(x) = x^T \boldsymbol{Q} x - 2\boldsymbol{f}^T x \qquad \text{s.t.} \quad x \in \mathscr{X}_a,$$

where the given matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is assumed to be symmetric, $\boldsymbol{f} \in \mathbb{R}^n$ is an arbitrarily given vector, and the feasible region is defined as

$$\mathscr{X}_a = \left\{ x \in \mathbb{R}^n \mid \|x\| \leq r \right\},$$

with r being a positive real number and $\|x\| = \|x\|_2$ representing ℓ_2 norm in $\mathbb{R}^n$.

Y. Chen · D.Y. Gao (✉)
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 353, Australia
e-mail: yi.chen@federation.edu.au

D.Y. Gao
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics
and Mathematics 37, DOI 10.1007/978-3-319-58017-3_15

Problem ($\mathscr{P}$) arises naturally in computational mathematical physics with extensive applications in engineering sciences. From the point view of systems theory, if the vector $\boldsymbol{f} \in \mathbb{R}^n$ is considered as an input (or source), then the solution $x \in \mathbb{R}^n$ is referred to as the output (or state) of the system. By the fact that the capacity of any given system is limited, the spherical constraint in $\mathscr{X}_a$ is naturally required for virtually every real-world system. For example, in engineering structural analysis, if the applied force $\boldsymbol{f} \in \mathbb{R}^\infty$ is big enough, the stress distribution in the structure will reach its elastic limit and the structure will collapse. For elasto-perfectly plastic materials, the well-known von Mises yield condition is a nonlinear inequality constraint $\|x\|_2 \leq r$ imposed on each material point[1] (see Chap. 7, [1]). By finite element method, the variational problem in structural limit analysis can be formulated as a large-size nonlinear optimization problem with m quadratic inequality constraints (m depends on the number of total finite elements). Such problems have been studied extensively in computational mechanics for more than fifty years and the so-called penalty-duality finite element programing [2, 3] is one of the well-developed efficient methods for solving this type of problems in engineering sciences.

In mathematical programing, the problem ($\mathscr{P}$) is known as a trust region subproblem, which arises in trust region methods [4, 5]. In literatures, two similar problems are also discussed: in [6–8], the convexity of the quadratic constraint is removed; while in [9, 10], the constraint is replaced by a two-sided (lower and upper bounded) quadratic constraint. Although the function $P(x)$ may be nonconvex, it is proved that the problem ($\mathscr{P}$) possesses the *hidden convexity*, i.e., ($\mathscr{P}$) is actually equivalent to a convex optimization problem [10], and for each optimal solution $\bar{\boldsymbol{x}}$, there exist a Lagrange multiplier $\bar{\mu}$ such that the following conditions hold [11]:

$$(\boldsymbol{Q} + \bar{\mu}\boldsymbol{I})\bar{\boldsymbol{x}} = \boldsymbol{f}, \tag{1}$$

$$\boldsymbol{Q} + \bar{\mu}\boldsymbol{I} \succeq 0, \tag{2}$$

$$\|\bar{\boldsymbol{x}}\| \leq r, \quad \bar{\mu} \geq 0, \quad \bar{\mu}(\|\bar{\boldsymbol{x}}\| - r) = 0. \tag{3}$$

Let λ_1 be the smallest eigenvalue of the matrix $\boldsymbol{Q}$. From conditions (2) and (3), we have

$$\bar{\mu} \geq \max\{0, -\lambda_1\}.$$

If the problem ($\mathscr{P}$) has no solutions on the boundary of $\mathscr{X}_a$, then $\boldsymbol{Q}$ must be positive definite, and $\|\boldsymbol{Q}^{-1}\boldsymbol{f}\| < r$, which leads to $\bar{\mu} = 0$. Now suppose the solution $\bar{\boldsymbol{x}}$ is on the boundary of $\mathscr{X}_a$. If $(\boldsymbol{Q} + \bar{\mu}\boldsymbol{I}) \succ 0$, we have $\|(\boldsymbol{Q} + \bar{\mu}\boldsymbol{I})^{-1}\boldsymbol{f}\| = r$ and the multiplier $\bar{\mu}$ can be easily found. While if $\det(\boldsymbol{Q} + \bar{\mu}\boldsymbol{I}) = 0$, it becomes very challenging to solve the problem [12–16] and the situation is referred to as 'hard case' (see [17]). Mathematically speaking, when the problem is in the hard case, there are multiple solutions for the equation $(\boldsymbol{Q} + \bar{\mu}\boldsymbol{I})x = \boldsymbol{f}$ and they are in the

[1]The well-known Tresca yield condition $\|x\|_\infty \leq r$ is equivalent to a box constraint at each material point. It was shown in the well-known experiment by Taylor and Quinney in 1931 that the von Mises yield condition is better than the Tesca yield condition for metal structures (see [1] p. 404.).

form $x = (\boldsymbol{Q} + \bar{\mu}\boldsymbol{I})^{\dagger}\boldsymbol{f} + \tau\tilde{\boldsymbol{x}}$ with $(\boldsymbol{Q} + \bar{\mu}\boldsymbol{I})\tilde{\boldsymbol{x}} = 0$. As pointed out in [12, 15, 16, 18], the hard case always implies that $\boldsymbol{f}$ is perpendicular to the subspace generated by all the eigenvectors corresponding to λ_1. We show by Theorem 3 and Example 2 in this paper that this condition is only a necessary condition for the problem being in the hard case. Many methods have been proposed for handling the problem $(\mathscr{P})$, especially focusing on the hard case: Newton type methods [17, 19], methods recasting the problem in terms of a parameterized eigenvalue problem [12, 15], methods sequential searching Krylov subspaces [18, 20], semidefinite programing methods [13, 16], and the D.C. (difference of convex functions) method [21].

Canonical duality theory is a powerful methodological theory which has been used successfully for solving a large class of difficult (nonconvex, nonsmooth, and discrete) problems in global optimization (see [22, 23]), within a unified framework. This theory is mainly comprised of (1) a *canonical dual transformation*, which can be used to reformulate nonconvex/discrete problems from different systems as a unified canonical dual problem without duality gaps; (2) a *complementary-dual principle*, which provides a unified analytical solution form in terms of the canonical dual variable; and (3) a *triality theory*, which is composed of *canonical min–max duality*, *double-min duality*, and *double-max duality*. The canonical min–max duality can be used to find a global optimal solution for the primal problem, while the double-min and double-max dualities can be used to identify the biggest local minimizer and the biggest local maximizer, respectively.

The canonical duality-triality theory was developed from Gao and Strang's original work [24], which discusses the nonconvex/nonsmooth variational problem

$$\min\{P(u) = W(\boldsymbol{D}u) + F(u)\}, \tag{4}$$

where the variational argument u is a continuous function in an infinite-dimensional space, $\boldsymbol{D}$ is a linear operator, $W(w)$ is the stored energy, which is an *objective functional* and depends only on the mathematical model, and $F(u)$ is the external energy, which is a "subjective" functional and depends on each problem (boundary-initial conditions). It is well known in nonlinear analysis [25] and continuum physics (see [1], p. 288) that a real-valued function $W(w)$ is called *objective* only if $W(w)$ satisfies the *frame-invariance principle*,[2] i.e., $W(w) = W(\boldsymbol{R}w)$ for any rotation matrices $\boldsymbol{R}$ such that $\boldsymbol{R}^T = \boldsymbol{R}^{-1}$ and det $\boldsymbol{R} = 1$. It was emphasized in [25] that the objectivity is not an assumption but an axiom. This means that the objective function depends only on the constitutive property of the system. Geometrically speaking, *the objective function should be an invariant under orthogonal transformation*. This concept lays a foundation for the canonical duality theory, i.e., instead of the design variable u (the linear operator $\boldsymbol{D}$ can not change the nonconvexity of $W(\boldsymbol{D}u)$), the *canonical dual transformation* is to choose a geometrically admissible (say objective) measure $\boldsymbol{\xi} = \Lambda(u)$ and a convex function $V(\boldsymbol{\xi})$ such that $W(\boldsymbol{D}u) = V(\Lambda(u))$ and the duality relation $\boldsymbol{\xi}^* = \nabla V(\boldsymbol{\xi})$ is invertible. Such one-to-one duality is called the

[2]See web page http://en.wikipedia.org/wiki/Objectivity_(frame_invariance).

canonical duality. The most simple objective measure is the ℓ_2 norm $\Lambda(u) = u^T u$ since $\Lambda(\boldsymbol{R}u) = \Lambda(u)$. Thus, the objective function $W(w)$ can not be linear. On the other hand, the so-called subjective function $F(u)$ depends on input (such as external force, market demanding, cost/price, etc.) and boundary-initial constraints for each problem, which must be linear. Therefore, the combination of $W(w)$ and $F(u)$ can be used to model general problems in complex systems[3] [1, 27]. Using numerical discretization (say, the finite element method) for the unknown variable $u(x)$, the general variational problem (4) becomes a very general global optimization problem in finite dimensional space (see [2, 28]). This is the basic reason why the canonical duality theory can be used for solving a large class of problems from different fields. However, the objective function in mathematical programing has been misused with other concepts such as cost, target, utility, and energy functions. It turns out that the canonical duality theory has been challenged (cf. [29]) by oppositely using linear $W(w)$ and nonlinear $F(u)$ as counterexamples (see [30]). These conceptual mistakes show a big gap between mathematical physics and optimization.

The goal of this paper is to find global solutions for the problem $(\mathscr{P})$, especially when it is in the hard case. We first show in the next section that by the canonical dual transformation, this constrained nonconvex problem can be reformulated as a one-dimensional optimization problem. The complementary-dual principle shows that this one-dimensional problem is canonically dual to $(\mathscr{P})$ in the sense that both problems have the same set of KKT solutions. While the canonical min–max duality in the triality theory provides a sufficient and necessary condition for identifying global optimal solutions. In order to solve the hard case, a perturbation method is proposed in Sect. 4 and, accordingly, a canonical primal-dual algorithm is developed in Sect. 5. Numerical results are presented in Sect. 6. The paper is ended with some conclusion remarks.

2 Canonical Dual Problem

By the fact that the condition $\|x\| \leq r$ is a physical constraint (required by mathematical model), it must be written in canonical form. Therefore, instead of the ℓ_2 norm, the canonical dual transformation is to introduce a quadratic (objective) measure $\xi = \Lambda(x) = x^T x : \mathbb{R}^n \to \mathscr{E}_a = \{\xi \in \mathbb{R} | \ \xi \geq 0\}$ and a convex function $V : \mathscr{E}_a \to \mathbb{R} \cup \{+\infty\}$

$$V(\xi) = \begin{cases} 0 & \text{if } \xi \leq r^2, \\ +\infty & \text{otherwise} \end{cases}$$

[3] Gao and Strang's model (4) has been generalized as $\min\{P(u) = W(\boldsymbol{D}u) - U(u)\}$, where $U(u)$ is a quadratic function, in order to cover more general problems in nonlinear dynamical systems and global optimization [26].

such that the constrained problem ($\mathscr{P}$) can be written equivalently in the following canonical form [22, 26, 27, 31]

$$\min\left\{\Pi(x) = V(\Lambda(x)) - U(x) \mid x \in \mathbb{R}^n\right\},$$

where $U(x) = -x^T \boldsymbol{Q} x + 2\boldsymbol{f}^T x$. By the Fenchel transformation, the conjugate of $V(\xi)$ can be uniquely defined as

$$V^*(\sigma) = \sup\{\xi\sigma - V(\xi) \mid \xi \in \mathscr{E}_a\} = \begin{cases} r^2\sigma & \text{if } \sigma \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Clearly, $V^*(\sigma)$ is convex, lower semi-continuous on $\mathscr{E}_a^* = \mathbb{R}$. According to convex analysis [32], we have the following equivalent relations on $\mathscr{E}_a \times \mathscr{E}_a^*$:

$$\sigma \in \partial V(\xi) \iff \xi \in \partial V^*(\sigma) \iff V(\xi) + V^*(\sigma) = \xi\sigma.$$

By the canonical duality theory, the pair (ξ, σ) satisfying (2) is called the (generalized) canonical duality pair (see [31] and Remark 1 in [22]). Clearly, the canonical duality (2) is equivalent to

$$\xi - r^2 \leq 0, \quad \sigma \geq 0, \quad \sigma(\xi - r^2) = 0.$$

This shows that the KKT conditions in (3) are equivalently relaxed by one of the canonical duality relations in (2). Replacing $V(\xi)$ in $\Pi(x)$ by the Fenchel-Young equality $V(\xi(x)) = \xi(x)\sigma - V^*(\sigma)$, the Gao-Strang total complementary function can be naturally obtained as [26, 27]:

$$\Xi(x, \sigma) = \xi(x)\sigma - V^*(\sigma) - U(x) = x^T \boldsymbol{G}(\sigma) x - 2\boldsymbol{f}^T x - V^*(\sigma),$$

where $\boldsymbol{G}(\sigma) = \boldsymbol{Q} + \sigma \boldsymbol{I}$. Let

$$\mathscr{S}_a = \{\sigma \in \mathbb{R} \mid \sigma \geq 0, \ \det \boldsymbol{G}(\sigma) \neq 0\}$$

be a canonical dual feasible space. Then for any given $\sigma \in \mathscr{S}_a$, the canonical dual function $P^d : \mathscr{S}_a \to \mathbb{R}$ can be defined by

$$P^d(\sigma) = \text{sta}\left\{\Xi(x, \sigma) \mid x \in \mathbb{R}^n\right\} = -\boldsymbol{f}^T \boldsymbol{G}(\sigma)^{-1} \boldsymbol{f} - r^2\sigma,$$

where the notation $\text{sta}\{\Xi(x, \sigma) \mid x \in \mathbb{R}^n\}$ stands for computing stationary points of $\Xi(x, \sigma)$ with respect to x. Therefore, the stationary canonical dual problem is to find KKT points $\bar{\sigma}$ of $P^d(\sigma)$ such that [33]

$$P^d(\bar{\sigma}) = \text{sta}\{P^d(\sigma) \mid \sigma \in \mathscr{S}_a\}.$$

We need to emphasize that $P^d(\sigma)$ is a function of a scalar variable $\sigma \in \mathscr{S}_a \subset \mathbb{R}$, regardless of the dimension of the primal problem, and the inequality $\det \boldsymbol{G}(\sigma) \neq 0$ is actually not a constraint (the Lagrange multiplier for this inequality is zero). Therefore, the KKT points for this canonical dual problem are much easier to be obtained than that for the primal problem. By the canonical duality theory, we have the following theorem.

Theorem 1. (Analytical Solution and Complementary-Dual Principle [33]**)** *Suppose that the symmetrical matrix* $\boldsymbol{Q}$ *has* $m\ (\leq n)$ *distinct eigenvalues* λ_i, $i = 1, \ldots, m$ *and* $i_d \leq m$ *of them are strictly negative such that* $\lambda_1 < \lambda_2 < \cdots < \lambda_{i_d} < 0 \leq \lambda_{i_d+1} < \cdots < \lambda_m$. *Then for a given vector* $\boldsymbol{f} \in \mathbb{R}^n$ *and a sufficiently large* $r > 0$, *the canonical dual problem (2) has at most* $2i_d + 1$ *KKT points* $\bar{\sigma}_i$ *satisfying*

$$\bar{\sigma}_1 > -\lambda_1 > \bar{\sigma}_2 \geq \bar{\sigma}_3 > -\lambda_2 > \cdots > -\lambda_{i_d} > \bar{\sigma}_{2i_d} \geq \bar{\sigma}_{2i_d+1} > 0.$$

For each $\bar{\sigma}_i, \quad i = 1, \ldots, 2i_d + 1$, *the vector*

$$\bar{x}_i = \boldsymbol{G}(\bar{\sigma}_i)^{-1}\boldsymbol{f} \tag{5}$$

is a KKT point of the primal problem ($\mathscr{P}$), *and we have*

$$P(\bar{x}_j) \geq P(\bar{x}_i) = \Xi(\bar{x}_i, \bar{\sigma}_i) = P^d(\bar{\sigma}_i) \leq P^d(\bar{\sigma}_j) \ \ \forall i, j = 1, \ldots, 2i_d + 1, \ \ i \leq j.$$

This theorem shows that the nonconvex function $P(x)$ is canonically dual (without duality gaps) to $P^d(\sigma)$ at each KKT point $(\bar{x}_i, \bar{\sigma}_i)$, and the function values of $P^d(\sigma_i)$ are in an opposite order with its critical points $\sigma_1 > \sigma_2 \geq \ldots$ (see Fig. 1). Clearly, the KKT solution $\bar{x}_1$ is a global minimizer of the primal problem ($\mathscr{P}$).

In order to identify global optimal solutions among all the critical points of $P^d(\sigma)$, a subset of $\mathscr{S}_a$ is needed:

$$\mathscr{S}_a^+ = \{\sigma \in \mathscr{S}_a \mid \boldsymbol{G}(\sigma) \succ \boldsymbol{0}\}.$$

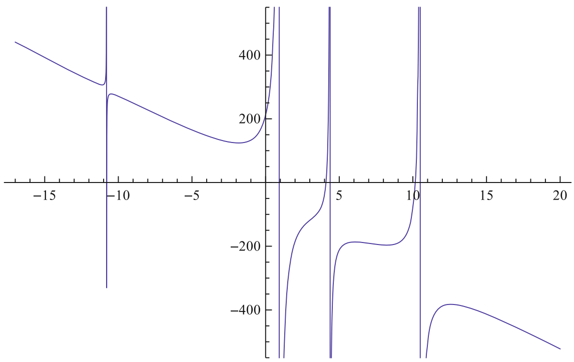


Fig. 1 The graph of canonical dual function $P^d(\sigma)$ for $n = 4$ (see Example 3 for details)

The problem canonically dual to $(\mathscr{P})$ can be proposed as the following

$$(\mathscr{P}^d) \quad \max \left\{ P^d(\sigma) \mid \sigma \in \mathscr{S}_a^+ \right\}.$$

Theorem 2. **(Global Optimality Condition** [1, 23]**)** *Suppose that $\bar{\sigma}$ is a critical point of $P^d(\sigma)$. If $\bar{\sigma} \in \mathscr{S}_a^+$, then $\bar{\sigma}$ is a global maximal solution of the problem $(\mathscr{P}^d)$ on $\mathscr{S}_a^+$ and $\bar{x} = \boldsymbol{G}(\bar{\sigma})^{-1}\boldsymbol{f}$ is a global minimal solution of the primal problem $(\mathscr{P})$, i.e.,*

$$P(\bar{x}) = \min_{x \in \mathscr{X}_a} P(x) = \max_{\sigma \in \mathscr{S}_a^+} P^d(\sigma) = P^d(\bar{\sigma}).$$

According to the triality theorem [1, 29], the global optimality condition (2) is called canonical min–max duality. By the fact that $P^d(\sigma)$ is strictly concave on the (open) convex set $\mathscr{S}_a^+$, this theorem guarantees that if there is a critical point in $\mathscr{S}_a^+$, it must be unique and the nonconvex minimization problem $(\mathscr{P})$ is equivalent to a concave maximization problem $(\mathscr{P}^d)$. Similar result is also discussed by Corollary 5.3 in [9] and Theorem 1 in [13]. Moreover, for the case when $n = 1$, the double-min duality statement in the weak-triality theory proven recently (see [29, 34, 35]) shows that the problem $(\mathscr{P})$ has at most one local minimizer, which is corresponding to a critical point $\bar{\sigma} \in \mathscr{S}_a^- = \{\sigma \in \mathscr{S}_a \mid \boldsymbol{G}(\sigma) \prec 0\}$. All these previous results show that the canonical duality-triality theory provides detailed information on a complete set of solutions to the nonconvex problem $(\mathscr{P})$.

Remark 1. Duality theory for quadratic minimization problems with ℓ_2-norm constraints was discussed extensively in plastic mechanics fifty years ago. It was shown by Gao in [3] that for the quadratic ℓ_2^2 constraint, the canonical dual can be easily formulated and a primal-dual finite element programing algorithm was first developed for solving minimal potential variational problems in infinite dimensional space [2]. By the fact that the geometrical measure $\xi(x) = x^T x$ is quadratic, the first term in $\Xi(x,\sigma)$ is the so-called (generalized) *complementary gap function* [26, 27] denoted by

$$G_{ap}(x,\sigma) = \xi(x)\sigma + x^T \boldsymbol{Q} x = x^T \boldsymbol{G}(\sigma) x.$$

Clearly, $G_{ap}(x,\sigma) \geq 0 \;\; \forall x \in \mathbb{R}^n$ if and only if $\sigma \in \mathscr{S}_a^+$. Therefore, $\Xi(x,\sigma)$ is a saddle function on $\mathbb{R}^n \times \mathbb{R}$ if $G_{ap}(x,\sigma) \geq 0 \;\; \forall x \in \mathbb{R}^n$. This result was first discovered by Gao and Strang in nonconvex mechanics [24], where they proved that this gap function recovers a broken symmetry in geometrically nonlinear systems and provides a global optimality condition for general nonconvex variational problems in mathematical physics. Particularly, the total complementary function $\Xi(x,\sigma)$ on $\mathbb{R}^n \times \mathbb{R}_+ = \{\sigma \in \mathbb{R} \mid \sigma \geq 0\}$ has a simple form

$$\Xi(x,\sigma) = x^T \boldsymbol{G}(\sigma) x - 2x^T \boldsymbol{f} - r^2\sigma = P(x) + \sigma(x^T x - r^2),$$

which can be viewed as a Lagrangian of $(\mathscr{P})$ for the ℓ_2^2-norm constraint $x^T x \leq r^2$. Indeed, the total complementary function $\Xi(x,\sigma)$ was also called nonlinear Lagrangian in [1] or extended Lagrangian in [31]. However, for nonconvex target function $P(x)$, the classical Lagrangian duality theory will produce a well-known duality gap unless the global optimality condition $G_{ap}(x,\sigma) \ge 0 \ \forall x \in \mathbb{R}^n$ is satisfied. Therefore, the Lagrangian duality theory is only a special case of the canonical duality theory for certain problems. Also, by the fact that a large class of nonconvex/discrete global optimization problems can be equivalently reformulated as a unified canonical dual form (2) (see [22, 26, 27]), which is equivalent to a convex minimization problem over a convex feasible set, the so-called “hidden convexity" is indeed a special case of the canonical min–max duality theory.

For the hard case, the matrix $\boldsymbol{G}(\sigma)$ is singular at the KKT point $\bar{\sigma}$, the canonical dual $P^d(\sigma)$ should be replaced by (see [36])

$$P^d(\sigma) = -\boldsymbol{f}^T \boldsymbol{G}(\sigma)^\dagger \boldsymbol{f} - r^2\sigma,$$

where $\boldsymbol{G}(\sigma)^\dagger$ stands for a generalized inverse of $\boldsymbol{G}(\sigma)$. In [9, 13], the dual function is also presented in discussions of the strong duality. Since this function is not strictly concave on the closure of $\mathscr{S}_a^+$, it may have multiple critical points located on the boundary of $\mathscr{S}_a^+$. In the following sections, we will first study the existence conditions of these critical points, and then study an associated algorithm for computing these solutions.

3 Existence Conditions

As $\boldsymbol{Q}$ is symmetrical, there exist a diagonal matrix Ł and an orthogonal matrix U such that $\boldsymbol{Q} = U Ł U^T$. The diagonal entities of Ł are the eigenvalues of $\boldsymbol{Q}$ and are arranged in a nondecreasing order,

$$\lambda_1 = \cdots = \lambda_k < \lambda_{k+1} \le \cdots \le \lambda_n.$$

The columns of U are corresponding eigenvectors.

Let $\hat{\boldsymbol{f}} = U^T \boldsymbol{f}$. Because $(\boldsymbol{Q} + \sigma \boldsymbol{I})^{-1} = U(Ł + \sigma \boldsymbol{I})^{-1} U^T$, we can rewrite the canonical dual function $P^d(\sigma)$ as

$$P^d(\sigma) = -\frac{\sum_{i=1}^k \hat{f}_i^2}{\lambda_1 + \sigma} - \sum_{i=k+1}^{n} \frac{\hat{f}_i^2}{\lambda_i + \sigma} - r^2\sigma,$$

where $\hat{f}_i$, $i = 1, \ldots, n$ are elements of $\hat{\boldsymbol{f}}$. It is now easy to see that as long as $\boldsymbol{f} \neq 0$, $P^d(\sigma)$ has stationary points in $\mathscr{S}_a$ and thus the canonical dual problem (2) is well defined. Whereas, for the case when $\boldsymbol{f} = 0$, a perturbation should be introduced, which is discussed in the next section.

Theorem 3. (Existence Conditions) *Suppose that for any given $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{f} \in \mathbb{R}^n$, λ_i, and $\hat{f}_i$ are defined as above.*

The canonical dual function $P^d(\sigma)$ has a critical point $\bar{\sigma}$ in $(-\lambda_1, +\infty)$ if and only if either $\sum_{i=1}^k \hat{f}_i^2 \neq 0$ or $\sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i-\lambda_1)^2} > r^2$ holds true. Furthermore, if $\lambda_1 \leq 0$, then $\bar{x} = \boldsymbol{G}(\bar{\sigma})^{-1}\boldsymbol{f}$ is the unique solution of the primal problem ($\mathscr{P}$).

If $P^d(\sigma)$ has no critical points in $(-\lambda_1, +\infty)$, the primal problem ($\mathscr{P}$) has exactly two global solutions when the multiplicity of λ_1 is $k = 1$ and has infinite number of solutions when $k > 1$.

Proof: First, we prove that the existence of a critical point of $P^d(\sigma)$ in $(-\lambda_1, +\infty)$ implies that either $\sum_{i=1}^k \hat{f}_i^2 \neq 0$ or $\sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i-\lambda_1)^2} > r^2$ holds true. It is equivalent to prove that if $\sum_{i=1}^k \hat{f}_i^2 = 0$ and $\sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i-\lambda_1)^2} \leq r^2$ the dual function $P^d(\sigma)$ will have no critical points in $(-\lambda_1, +\infty)$. The first item in the expression (3) vanishes when $\sum_{i=1}^k \hat{f}_i^2 = 0$. Then because $\sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i-\lambda_1)^2} \leq r^2$, the first-order derivative of the dual function

$$(P^d(\sigma))' = \sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i + \sigma)^2} - r^2$$

is always negative in $(-\lambda_1, +\infty)$. Therefore, the dual function $P^d(\sigma)$ will have no critical points in $(-\lambda_1, +\infty)$.

Next we will give the proof of the sufficiency, which is divided into two parts:

(1) If $\sum_{i=1}^k \hat{f}_i^2 \neq 0$, then $\sigma = -\lambda_1$ is a pole of $P^d(\sigma)$, i.e., as σ approaches $-\lambda_1$ from the right side, $P^d(\sigma)$ approaches $-\infty$. The value of $P^d(\sigma)$ also approaches $-\infty$, when σ approaches $+\infty$. Thus, $-P^d(\sigma)$ is coercive on $(-\lambda_1, +\infty)$. Since, for any $\sigma \in (-\lambda_1, +\infty)$, $\boldsymbol{G}(\sigma)$ is positive definite, $P^d(\sigma)$ is strictly concave on $(-\lambda_1, +\infty)$. Thus there exists a unique critical point in $(-\lambda_1, +\infty)$.

(2) If $\sum_{i=1}^k \hat{f}_i^2 = 0$ and $\sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i-\lambda_1)^2} > r^2$, $(P^d(\sigma))'$ is positive at $\sigma = -\lambda_1$. Moreover, $(P^d(\sigma))'$ approaches $-r^2$ as σ approaches ∞. Therefore, there exists at least one root for the equation $(P^d(\sigma))' = 0$ in $(-\lambda_1, +\infty)$, which means $P^d(\sigma)$ has at least one critical point in $(-\lambda_1, +\infty)$. Similarly, because of the strict concavity of $P^d(\sigma)$ over $(-\lambda_1, +\infty)$, the critical point is unique.

Suppose $\lambda_1 \leq 0$. The uniqueness of global solution $\bar{\boldsymbol{x}}$ will be proved, if it can be proved that $(\bar{\boldsymbol{x}}, \bar{\sigma})$ is the only pair that satisfies the KKT conditions (1)–(3). As mentioned above, the dual function $P^d(\sigma)$ is strictly concave on $(-\lambda_1, +\infty)$, which, plus the criticality of $\bar{\sigma}$, implies that $(P^d(\sigma))' = \|x\|^2 - r^2 > 0$ for $\sigma \in (-\lambda_1, \bar{\sigma})$ and < 0 for $\sigma \in (\bar{\sigma}, +\infty)$, where $x = \boldsymbol{G}(\sigma)^{-1}\boldsymbol{f}$. Thus, for any $\sigma \neq \bar{\sigma}$ in $(-\lambda_1, +\infty)$, there is no x such that (x, σ) satisfies the KKT conditions (1)–(3). Except for the interval $(-\lambda_1, +\infty)$, $\sigma = -\lambda_1$ is the last candidate. However, if $\sum_{i=1}^k \hat{f}_i^2 \neq 0$, the equation $\boldsymbol{G}(-\lambda_1)x = \boldsymbol{f}$ has no solutions, and if $\sum_{i=1}^k \hat{f}_i^2 = 0$ and $\sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i-\lambda_1)^2} > r^2$, the feasibility of any solution of $\boldsymbol{G}(-\lambda_1)x = \boldsymbol{f}$ is violated by the fact that $\|x\|^2 - r^2 = \sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i-\lambda_1)^2} - r^2 > 0$. Then, $\sigma = -\lambda_1$ can not make the KKT conditions hold true. Therefore, $(\bar{\boldsymbol{x}}, \bar{\sigma})$ is the unique pair that satisfies the KKT conditions (1)–(3).

Finally, suppose that there are no critical points in $(-\lambda_1, +\infty)$, which, from the above proof, is equivalent to $\sum_{i=1}^{k} \hat{f}_i^2 = 0$ and $\sum_{i=k+1}^{n} \frac{\hat{f}_i^2}{(\lambda_i - \lambda_1)^2} \le r^2$. Then, for any global solution, we have $\bar{\sigma} = -\lambda_1$. Let $\bar{\boldsymbol{x}}$ be a global solution and $\bar{\boldsymbol{y}} = U^T \bar{\boldsymbol{x}}$. Then the canonical equilibrium equation $\boldsymbol{G}(\bar{\sigma})\bar{\boldsymbol{x}} = \boldsymbol{f}$ can be equivalently transformed into $\mathrm{diag}(\{\lambda_i + \bar{\sigma}\})\bar{\boldsymbol{y}} = \hat{\boldsymbol{f}}$. If $k = 1$, i.e., the multiplicity of λ_1 is one, the equation uniquely determines $\bar{y}_i, i = 2, \ldots, n$, but not $\bar{y}_1$. By the fact that $\bar{\boldsymbol{y}}^T \bar{\boldsymbol{y}} = r^2$, $\bar{y}_1$ has exactly two values, corresponding to the two global solutions of $(\mathscr{P})$. While, if $k > 1$, i.e., the matrix $\boldsymbol{Q}$ has at least two repeated eigenvalues $\lambda_1 = \lambda_2 = \cdots = \lambda_k \le 0$, the equations $\mathrm{diag}(\{\lambda_i + \bar{\sigma}\})\bar{\boldsymbol{y}} = \hat{\boldsymbol{f}}$ and $\bar{\boldsymbol{y}}^T \bar{\boldsymbol{y}} = r^2$ have infinite number of solutions. □

Remark 2. The complementarity relations between the primal problem $(\mathscr{P})$ and its canonical dual problem $(\mathscr{P}^d)$ are significant. When $\lambda_1 > 0$, i.e., $\boldsymbol{Q}$ is positive definite, if $(\mathscr{P})$ has a global solution in the interior of $\mathscr{X}_a$, which must be the stationary point of $P(x)$ and can be easily calculated, its canonical dual $(\mathscr{P}^d)$ has no critical point in $\mathscr{S}_a^+ = [0, +\infty)$ due to $(P^d(0))' = \|\bar{\boldsymbol{x}}\|^2 - r^2 < 0$, where $\bar{\boldsymbol{x}} = \boldsymbol{G}(0)^{-1}\boldsymbol{f}$ is the stationary point of $P(x)$. Dually, when $\lambda_1 \le 0$, the primal function $P(x)$ is nonconvex and the global minimizer of $(\mathscr{P})$ must be on the boundary of $\mathscr{X}_a$. In this case, if the canonical dual $(\mathscr{P}^d)$ has a critical point in $\mathscr{S}_a^+ = (-\lambda_1, +\infty)$, the primal problem $(\mathscr{P})$ is then not in the hard case and has a unique solution, which can be easily obtained by solving the canonical dual problem. Whereas if $(\mathscr{P}^d)$ has no critical points in $\mathscr{S}_a^+$, i.e., $P^d(-\lambda_1) = \sup\{P^d(\sigma) | \ \sigma \in \mathscr{S}_a^+\}$, the primal problem $(\mathscr{P})$ is in the hard case, because, for any $\sigma \in \mathscr{S}_a^+$ and $x = \boldsymbol{G}(\sigma)^{-1}\boldsymbol{f}$, we have $(P^d(\sigma))' = \|x\|^2 - r^2 < 0$, which destroys the complementary condition in (3), and only $\sigma = -\lambda_1$ can make the KKT conditions (1)–(3) hold.

Therefore, combining with Theorem 3, we have the following result.

Corollary 1. *If $\lambda_1 \le 0$, the nonconvex problem $(\mathscr{P})$ is in the hard case if and only if both conditions (i) $\sum_{i=1}^{k} \hat{f}_i^2 = 0$ and (ii) $\sum_{i=k+1}^{n} \frac{\hat{f}_i^2}{(\lambda_i - \lambda_1)^2} \le r^2$ hold true.*

The condition (i) is well known: the trust region subproblem could be in the hard case only if the coefficient $\boldsymbol{f}$ is perpendicular to the subspace generated by eigenvectors of the smallest eigenvalue. The condition (ii) is new, which shows that the hard case of $(\mathscr{P})$ depends not only on the direction of $\boldsymbol{f}$, but also on its norm.

Theorem 3 and Corollary 1 show an important fact that the given vector $\boldsymbol{f}$ plays an important role to the solutions of the problem $(\mathscr{P})$. From the point of view of solid mechanics, if $\boldsymbol{f}$ is considered as an applied force, then the decision variable x is the displacement and the spherical constraint $\|x\| \le r$ is corresponding to the von Mises yield condition, which represents the capacity of the system. If the norm of $\boldsymbol{f}$ is big enough, the deformation x should reach the limit $\|x\| = r$ and the problem $(\mathscr{P})$ has a solution on the boundary of $\mathscr{X}_a$. By the canonical duality, the problem $(\mathscr{P}^d)$ must have a critical point in $\mathscr{S}_a^+$. If the norm of $\boldsymbol{f}$ is too small, the primal problem $(\mathscr{P})$ could have multiple solutions. In this case, $(\mathscr{P}^d)$ has no critical point in $\mathscr{S}_a^+$ and $(\mathscr{P})$ could be in the hard case.

To illustrate Theorem 3, let us consider a 3-dimensional problem with coefficients

$$\boldsymbol{Q} = \begin{pmatrix} [r] -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} [r]0 \\ 0 \\ -1.8 \end{pmatrix}, \text{ and } r = 2.$$

In this case, the eigenvalues of $\boldsymbol{Q}$ are $\lambda_1 = \lambda_2 = -1$, and $\lambda_3 = 1$. So we have $k = 2$ and the target function

$$P(x) = -\frac{1}{2}(x_1^2 + x_2^2) + \frac{1}{2}x_3^2 + 1.8x_3$$

is nonconvex, whose minimizers are on the boundary of the feasible region. Replacing $x_1^2 + x_2^2$ with $r^2 - x_3^2$, the target function $P(x)$ can be reformulated as a univariate function of x_3,

$$g(x_3) = x_3^2 + 1.8x_3 - 2,$$

which achieves the minimum at $x_3 = -0.9$. Then we obtain the following equation

$$x_1^2 + x_2^2 = r^2 - x_3^2 = 2^2 - (-0.9)^2 = 3.19.$$

So all $\bar{\boldsymbol{x}} \in \mathbb{R}^3$ satisfying $\bar{x}_1^2 + \bar{x}_2^2 = 3.19$ and $\bar{x}_3 = -0.9$ are global minimizers of the problem.

By the fact that $\sum_{i=1}^{2} \hat{f}_i^2 = 0$ and $\sum_{i=2+1}^{3} \frac{\hat{f}_i^2}{(\lambda_i - \lambda_1)^2} = (-1.8)^2/(1+1)^2 \le r^2 = 4$, Theorem 3 shows that $P^d(\sigma)$ has no critical point in $\mathscr{S}_a^+$, and $(\mathscr{P})$ is indeed in the hard case and has infinite number of global solutions. If we choose either a smaller r or a vector $\boldsymbol{f}$ with a larger magnitude such that $\sum_{i=2+1}^{3} \frac{\hat{f}_i^2}{(\lambda_i - \lambda_1)^2} > r^2$, the global solution will be unique. For example, let $r = 0.5$. Then $x_3 = -0.9$ is no longer the minimizer of $g(x_3)$ and the problem $\min\{g(x_3) | \; x_3^2 \le 0.5^2\}$ leads to $x_3 = -0.5$. From $x_1^2 + x_2^2 = r^2 - x_3^2 = 0.5^2 - (-0.5)^2 = 0$, we know the unique global solution of $(\mathscr{P})$ is $\bar{\boldsymbol{x}} = (0, 0, -0.5)^T$.

In [37], Martinez investigated the 'local-nonglobal minimizers' of the problem $(\mathscr{P})$, of which the main results (Theorem 3.1 in [37]) can be restated in the following theorem.

Theorem 4. *(i) If $\bar{\boldsymbol{x}}$ is a local-nonglobal minimizer of $(\mathscr{P})$, then there is a $\bar{\sigma} \in (\max\{0, -\lambda_2\}, -\lambda_1)$ such that $\boldsymbol{G}(\bar{\sigma})\bar{\boldsymbol{x}} = \boldsymbol{f}$ and $(P^d(\bar{\sigma}))'' \ge 0$.*

(ii) There exists at most one local-nonglobal minimizer of $(\mathscr{P})$.

(iii) If $\|\bar{\boldsymbol{x}}\| = r$, $\boldsymbol{G}(\bar{\sigma})\bar{\boldsymbol{x}} = \boldsymbol{f}$ for some $\bar{\sigma} \in (-\lambda_2, -\lambda_1)$, $\bar{\sigma} > 0$ and $(P^d(\bar{\sigma}))'' > 0$, then $\bar{\boldsymbol{x}}$ is a strict local minimizer of $(\mathscr{P})$.

From the point of view of the canonical duality theory, the $\bar{\sigma}$ in this theorem is actually a critical point of $P^d(\sigma)$. The case of $(\mathscr{P})$ having no local-nonglobal minimizers implies that all the local minimizers are global solutions. The situations that leads to this case include (i) the multiplicity of λ_1 being larger than one; (ii) no critical point in $(\max\{0, -\lambda_2\}, -\lambda_1)$, and (iii) $\boldsymbol{f}$ being perpendicular to the eigenvector of λ_1. The first situation results in $(-\lambda_2, -\lambda_1) = \emptyset$. The last situation violates the necessary condition $(P^d(\sigma))'' \geq 0$, which can be observed from the expression of $(P^d(\sigma))''$,

$$(P^d(\sigma))'' = -2\sum_{i=1}^{n} \frac{\hat{f}_i^2}{(\lambda_i + \sigma)^3}.$$

For any $\sigma \in (-\lambda_2, -\lambda_1)$, the only nonnegative item in $(P^d(\sigma))''$ is the first term $-2\hat{f}_1^2/(\lambda_1 + \sigma)^3$. Thus $(P^d(\sigma))''$ will be negative if $\hat{f}_1^2 = 0$. As shown in Fig. 1, there is a critical point $\bar{\sigma}_2 \in (-\lambda_2, -\lambda_1) = (4.37, 10.51)$ and the corresponding solution $\bar{x}_2$ obtained from the Eq. (5) is a local minimizer.

4 Perturbation Methods

This section is devoted to compute solutions for the problem when the canonical dual problem $(\mathscr{P}^d)$ has no critical point in $(-\lambda_1, +\infty)$. Since a necessary condition for the hard case is $\sum_{i=1}^{k} \hat{f}_i^2 = 0$, a perturbation can be introduced such that this condition does not hold true anymore. Impressively, once we obtain the critical point in $\mathscr{S}_a^+$, all the global solutions can be determined. Our approach has been applied successfully in canonical duality theory for solving nonlinear algebraic equations [38], chaotic dynamical systems [39], as well as a class of NP-hard problems in the global optimization [36, 40, 41].

In order to establish the existence conditions, a perturbation $\sum_{i=1}^{k} \alpha_i U_i$ with parameters

$$\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^{k} \neq 0$$

is introduced to $\boldsymbol{f}$. Let

$$\boldsymbol{p} = \boldsymbol{f} + \sum_{i=1}^{k} \alpha_i U_i, \quad \hat{\boldsymbol{p}} = U^T \boldsymbol{p}, \text{ and } P_\alpha(x) = x^T \boldsymbol{Q} x - 2\boldsymbol{p}^T x.$$

It is true that the existence conditions hold true for the perturbed problem

$$(\mathscr{P}_\alpha) \quad \min\{P_\alpha(x) \mid x \in \mathscr{X}_a\},$$

for $\sum_{i=1}^{k} \hat{p}_i^2 \neq 0$ is guaranteed by (4).

The following theorem states that if the parameter $\boldsymbol{\alpha}$ is chosen appropriately, the optimal solution of the perturbed problem approximates that of the primal problem $(\mathscr{P})$.

Theorem 5. *Suppose that $\lambda_1 \leq 0$, there is no critical point of $P^d(\sigma)$ in $\mathscr{S}_a^+$, and $\bar{\boldsymbol{x}}^*$ is the optimal solution of the problem $(\mathscr{P}_\alpha)$. Then, there is a global solution of* *the problem ($\mathscr{P}$), denoted as $\bar{\boldsymbol{x}}$, which is on the boundary of $\mathscr{X}_a$ and, for any $\varepsilon > 0$, if the parameter $\boldsymbol{\alpha}$ satisfies*

$$\|\boldsymbol{\alpha}\|^2 \leq (\lambda_2 - \lambda_1)^2 \left(r^2 - \sum_{i=k+1}^{n} \frac{\hat{f}_i^2}{(\lambda_i - \lambda_1)^2} \right) (1/\sqrt{2(1 - \cos(\varepsilon/r))} - 1)^{-2},$$

we have $\|\bar{\boldsymbol{x}}^* - \bar{x}\| \leq \varepsilon$.

Proof. For simplicity, the coordinate system is rotated and let $\boldsymbol{y} = U^T x$, $\boldsymbol{y}_k = \{y_i\}_{i=1}^k$ and $\boldsymbol{y}_\ell = \{y_i\}_{i=k+1}^n$. Since $\hat{f}_i = 0$ for $i = 1, \ldots, k$, variables y_i for $i = 1, \ldots, k$ appear in the target function only in the form of squares. On the boundary of $\mathscr{X}_a$, the problem ($\mathscr{P}$) is then equivalent to the following problem in $\mathbb{R}^{n-k}$:

$$\min_{\|\boldsymbol{y}_\ell\| \leq r} P^\ell(\boldsymbol{y}_\ell) = \sum_{i=k+1}^{n} (\lambda_i - \lambda_1) y_i^2 - \sum_{i=k+1}^{n} 2\hat{f}_i y_i + \lambda_1 r^2.$$

Since $P^\ell(\boldsymbol{y}_\ell)$ is a strictly convex function, it has a unique stationary point,

$$\bar{\boldsymbol{y}}_\ell = \left\{ \frac{\hat{f}_i}{\lambda_i - \lambda_1} \right\}_{i=k+1}^{n}.$$

Combining with the assumption of no critical point in $\mathscr{S}_a^+$, we know that this stationary point is the global optimal solution of the problem (4). Then, all $\bar{\boldsymbol{y}}$ that satisfies $\bar{\boldsymbol{y}}_k^T \bar{\boldsymbol{y}}_k = r^2 - \bar{\boldsymbol{y}}_\ell^T \bar{\boldsymbol{y}}_\ell$ are solutions of the problem ($\mathscr{P}$). Here we choose one particular solution with

$$\bar{\boldsymbol{y}}_k = h \bar{\boldsymbol{y}}_k^*, \quad h = \frac{1}{\|\bar{\boldsymbol{y}}_k^*\|} \sqrt{r^2 - \bar{\boldsymbol{y}}_\ell^T \bar{\boldsymbol{y}}_\ell},$$

where $\bar{\boldsymbol{y}}^* = U \bar{\boldsymbol{x}}^*$, and let $\bar{\boldsymbol{x}} = U \bar{\boldsymbol{y}}$.

By canceling variables y_i, $i = 1, \ldots, k$, the perturbed problem (4) with the equality constraint is equivalent to

$$\min_{\|\boldsymbol{y}_\ell\| \leq r} P_\alpha^\ell(\boldsymbol{y}_\ell) = \sum_{i=k+1}^{n} (\lambda_i - \lambda_1) y_i^2 - \sum_{i=k+1}^{n} 2\hat{f}_i y_i + \lambda_1 r^2 - 2\|\boldsymbol{\alpha}\| \sqrt{r^2 - \boldsymbol{y}_\ell^T \boldsymbol{y}_\ell}.$$

The function $P_\alpha^\ell(\boldsymbol{y}_\ell)$ is also strictly convex. Moreover, for any $\|\boldsymbol{y}_\ell\| < r$, we have $P_\alpha^\ell(\boldsymbol{y}_\ell) < P^\ell(\boldsymbol{y}_\ell)$, while for any $\|\boldsymbol{y}_\ell\| = r$, we have $P_\alpha^\ell(\boldsymbol{y}_\ell) = P^\ell(\boldsymbol{y}_\ell)$. The fact indicates that the unique stationary point of $P_\alpha^\ell(\boldsymbol{y}_\ell)$ is in the interior of $\|\boldsymbol{y}_\ell\| \leq r$. Thus the global solution $\bar{\boldsymbol{y}}_\ell^*$ is a stationary point of the problem (4) and then satisfies

$$\bar{y}_i^* = \frac{\hat{f}_i}{\lambda_i - \lambda_1 + \|\boldsymbol{\alpha}\| (r^2 - \bar{\boldsymbol{y}}_\ell^{*T} \bar{\boldsymbol{y}}_\ell^*)^{-\frac{1}{2}}}, i = k+1, \ldots, n.$$

and

$$|\bar{y}_i^*| < |\bar{y}_i|, i = k+1, \ldots, n.$$

We will prove that as $\|\boldsymbol{\alpha}\|$ approaches zero, $\bar{\boldsymbol{y}}^*$ will approach $\bar{\boldsymbol{y}}$. First, we have the following relationship

$$\begin{aligned}
\bar{\boldsymbol{y}}^{*T}\bar{\boldsymbol{y}} &= \sqrt{r^2 - \bar{\boldsymbol{y}}_\ell^{*T}\bar{\boldsymbol{y}}_\ell^*}\sqrt{r^2 - \bar{\boldsymbol{y}}_\ell^T\bar{\boldsymbol{y}}_\ell} + \bar{\boldsymbol{y}}_\ell^{*T}\bar{\boldsymbol{y}}_\ell \\
&\leq \frac{1}{2}\left(r^2 - \bar{\boldsymbol{y}}_\ell^{*T}\bar{\boldsymbol{y}}_\ell^* + r^2 - \bar{\boldsymbol{y}}_\ell^T\bar{\boldsymbol{y}}_\ell\right) + \bar{\boldsymbol{y}}_\ell^{*T}\bar{\boldsymbol{y}}_\ell \\
&= r^2 - \frac{1}{2}\|\bar{\boldsymbol{y}}_\ell^* - \bar{\boldsymbol{y}}_\ell\|^2,
\end{aligned}$$

where the first equality is derived from the definition of $\bar{\boldsymbol{y}}_k$ and the fact that $\bar{\boldsymbol{y}}^*$ locates on the surface of the sphere. Based on the relationship

$$\|\bar{\boldsymbol{y}}^* - \bar{\boldsymbol{y}}\| \leq r \arccos\left(\frac{\bar{\boldsymbol{y}}^{*T}\bar{\boldsymbol{y}}}{r^2}\right) \leq r \arccos\left(\frac{r^2 - \frac{1}{2}\|\bar{\boldsymbol{y}}_\ell^* - \bar{\boldsymbol{y}}_\ell\|^2}{r^2}\right),$$

we will have $\|\bar{\boldsymbol{y}}^* - \bar{\boldsymbol{y}}\| \leq \varepsilon$, if $\|\bar{\boldsymbol{y}}_\ell^* - \bar{\boldsymbol{y}}_\ell\|^2 \leq 2r^2(1 - \cos\frac{\varepsilon}{r})$. Then, it can be verified that

$$\|\bar{\boldsymbol{y}}_\ell^* - \bar{\boldsymbol{y}}_\ell\|^2 \leq \frac{r^2}{\left((\lambda_2 - \lambda_1)\|\boldsymbol{\alpha}\|^{-1}\sqrt{r^2 - \bar{\boldsymbol{y}}_\ell^{*T}\bar{\boldsymbol{y}}_\ell^*} + 1\right)^2}.$$

If let the right side of Eq. (4) be less than or equal to $2r^2(1 - \cos\frac{\varepsilon}{r})$, we obtain

$$\|\boldsymbol{\alpha}\|^2 \leq \frac{(\lambda_2 - \lambda_1)^2(r^2 - \bar{\boldsymbol{y}}_\ell^{*T}\bar{\boldsymbol{y}}_\ell^*)}{(1/\sqrt{2(1 - \cos\frac{\varepsilon}{r})} - 1)^2}.$$

Combining with relations in (4), we can state that $\|\bar{\boldsymbol{y}}^* - \bar{\boldsymbol{y}}\| \leq \varepsilon$ if the following inequality is true

$$\|\boldsymbol{\alpha}\|^2 \leq \frac{(\lambda_2 - \lambda_1)^2(r^2 - \sum_{i=k+1}^n \frac{\hat{f}_i^2}{(\lambda_i - \lambda_1)^2})}{(1/\sqrt{2(1 - \cos\frac{\varepsilon}{r})} - 1)^2}.$$

Since $\|\bar{\boldsymbol{x}}^* - \bar{\boldsymbol{x}}\| = \|\bar{\boldsymbol{y}}^* - \bar{\boldsymbol{y}}\|$, the Eq. (4) implies that $\|\bar{\boldsymbol{x}}^* - \bar{\boldsymbol{x}}\| \leq \varepsilon$. □

Theorem 5 shows that with a proper parameter $\boldsymbol{\alpha}$, the existence condition is guaranteed to hold true for the perturbed problem and the perturbation method can be used to solve the hard case approximately. As the perturbation parameters approach zero, the perturbed solutions will approach to one of the global solutions of ($\mathscr{P}$). By the projection theorem, the nearest points to $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{x}}^*$ in the subspace spanned by

$\{U_1, \ldots, U_k\}$ are $\sum_{i=1}^{k}(\bar{\boldsymbol{x}}^T U_i)U_i$ and $\sum_{i=1}^{k}(\bar{\boldsymbol{x}}^{*T} U_i)U_i$, respectively. Then we have the following relationship

$$\|\bar{\boldsymbol{x}}^* - \sum_{i=1}^{k}(\bar{\boldsymbol{x}}^{*T} U_i)U_i\|^2 < \|\bar{\boldsymbol{x}} - \sum_{i=1}^{k}(\bar{\boldsymbol{x}}^T U_i)U_i\|^2,$$

which means that the perturbed solution $\bar{\boldsymbol{x}}^*$ is closer to the subspace spanned by $\{U_1, \ldots, U_k\}$ than the solution $\bar{\boldsymbol{x}}$.

Furthermore, each solution of the problem $(\mathscr{P})$ can be approximated, if the perturbation parameter $\boldsymbol{\alpha}$ is properly chosen. When the multiplicity of λ_1 is equal to one, as stated in Theorem 3, there are exactly two global solutions. In this case, $\boldsymbol{\alpha}$ becomes a scalar and has exactly two possible directions, which are mutual opposite and, respectively, lead to the two global solutions (see Example 1). For general cases, there may be infinite number of global solutions for the problem $(\mathscr{P})$, and we will show that there is a one-to-one correspondence between solutions of the problem $(\mathscr{P})$ and directions of $\boldsymbol{\alpha}$. In the problem (4), variables y_i, $i = 1, \ldots, k$ are removed by solving the following minimization problem

$$\min\{-2\boldsymbol{\alpha}^T \boldsymbol{y}_k \ \mid \boldsymbol{y}_k^T \boldsymbol{y}_k = r^2 - \boldsymbol{y}_\ell^T \boldsymbol{y}_\ell, \ \boldsymbol{y}_k \in \mathbb{R}^k\}.$$

Its solution is

$$\boldsymbol{y}_k = h\boldsymbol{\alpha}, \ \ h = \frac{1}{\|\boldsymbol{\alpha}\|}\sqrt{r^2 - \boldsymbol{y}_\ell^T \boldsymbol{y}_\ell},$$

i.e., the point falls on the boundary of the sphere in (4) and has the same direction with $\boldsymbol{\alpha}$. If $\|\boldsymbol{\alpha}\|$ keeps unchanged, the problem (4) always has the same solution and the scalar h also keeps unchanged. Thus, each direction of $\boldsymbol{\alpha}$ is corresponding to a solution $\{y_i\}_{i=1}^{k}$, and all the solutions comprise the surface of a sphere centered at the original in $\mathbb{R}^k$. On the other hand, from the problem (4), we have $\bar{\boldsymbol{y}}_k^T \bar{\boldsymbol{y}}_k = r^2 - \bar{\boldsymbol{y}}_\ell^T \bar{\boldsymbol{y}}_\ell$, which means all global solutions of the problem $(\mathscr{P})$ also comprise the surface of a sphere. Combining Theorem 5, we then conclude that each solution of the problem $(\mathscr{P})$ can be approached as the direction of $\boldsymbol{\alpha}$ is properly chosen and $\|\boldsymbol{\alpha}\|$ approaches zero.

5 Canonical Primal-Dual Algorithm

Based on the results obtained above, a *canonical primal-dual algorithm* is developed, which is matrix inverse free and the essential cost of calculation is only the matrix–vector multiplication.

The main step of this algorithm is to solve the following perturbed canonical dual problem:

$$(\mathscr{P}_\alpha^d) \quad \max\left\{P_\alpha^d(\sigma) = -\boldsymbol{p}^T\boldsymbol{G}(\sigma)^{-1}\boldsymbol{p} - r^2\sigma \mid \sigma \in \mathscr{S}_a^+\right\}$$

Let $\psi(\sigma)$ be its first-order derivative, i.e.,

$$\psi(\sigma) = (P_\alpha^d(\sigma))' = \boldsymbol{p}^T\boldsymbol{G}(\sigma)^{-1}\boldsymbol{G}(\sigma)^{-1}\boldsymbol{p} - r^2.$$

Then the critical point of $P_\alpha^d(\sigma)$ in $\mathscr{S}_a^+$ is corresponding to the solution of the equation $\psi(\sigma) = 0$ in $\mathscr{S}_a^+$. The first- and second-order derivatives of $\psi(\sigma)$ are

$$\psi'(\sigma) = -2\boldsymbol{p}^T\boldsymbol{G}(\sigma)^{-1}\boldsymbol{G}(\sigma)^{-1}\boldsymbol{G}(\sigma)^{-1}\boldsymbol{p},$$
$$\psi''(\sigma) = 6\boldsymbol{p}^T\boldsymbol{G}(\sigma)^{-1}\boldsymbol{G}(\sigma)^{-1}\boldsymbol{G}(\sigma)^{-1}\boldsymbol{G}(\sigma)^{-1}\boldsymbol{p}.$$

It is noticed that $\psi(\sigma)$ is strictly decreasing and strictly convex over $\mathscr{S}_a^+$, $\psi(\sigma)$ will approach $-r^2$ as σ approaches infinity and $\sigma = -\lambda_1$ is a pole of $\psi(\sigma)$.

We use the Lanczos method to compute an approximation for the smallest eigenvalue of $\boldsymbol{Q}$ and a corresponding eigenvector, denoted, respectively, by $\tilde{\lambda}_1$ and $\tilde{U}_1$, where the latter is a unit vector. For choosing an effective perturbation, it is not necessary to calculate all eigenvectors of the smallest eigenvalue, since any one of which will be sufficient to divert the direction of $\boldsymbol{f}$. Here we use $\alpha\tilde{U}_1$ as a perturbation to $\boldsymbol{f}$.

Although the perturbed canonical dual problem $(\mathscr{P}_\alpha^d)$ is strictly concave on $\mathscr{S}_a^+$, its derivative $\psi(\sigma)$ would become ill-conditioned when σ approaches to the pole. Therefore, instead of nonlinear optimization techniques, a bisection method is used to find the root in $(-\lambda_1, +\infty)$ for $\psi(\sigma)$. Each time, as a dual solution $\sigma > -\lambda_1$ is obtained, the value of $\psi(\sigma)$ is calculated and checked to see whether it is equal to zero. For moderate-size problems, it is not hard to calculate $\boldsymbol{G}(\sigma)^{-1}\boldsymbol{p}$ by computing the inverse or decomposition of $\boldsymbol{G}(\sigma)$, but it is not possible for very large-size problems, especially when the memory is very limited. One alternative approach is to solve the following strictly convex minimization problem,

$$\min_{x\in\mathbb{R}^n} x^T\boldsymbol{G}(\sigma)x - 2\boldsymbol{p}^Tx,$$

whose optimal solution is $x = \boldsymbol{G}(\sigma)^{-1}\boldsymbol{p}$. Actually, during iterations, we do not need to calculate $\psi(\sigma)$ every time, especially when σ is on the left side of the root and close to the pole. It is discovered that for a given σ, the value of $\psi(\sigma)$ is equal to the optimal value of the following unconstrained concave maximization problem

$$\max_{z\in\mathbb{R}^n} -\boldsymbol{z}^T\boldsymbol{G}(\sigma)\boldsymbol{G}(\sigma)\boldsymbol{z} + 2\boldsymbol{p}^T\boldsymbol{z} - r^2.$$

By the fact that the value of the target function will increase during the iterations, we can stop solving the problem (5) if the target function is larger than a threshold, and then we claim that σ must be on the left side of the root. Thus, the ill-condition in computing $\psi(\sigma)$ can be prevented as σ approaches to the pole. Since the optimal value is equal to zero when σ is a root of $\psi(\sigma)$, any nonnegative value can be a threshold.

An uncertainty interval should be initialized before the bisection method is applied, and it is used to safeguard that the root is always in intervals of the bisection method. For the right end of the interval, any large enough number can be a candidate. An upper bound can be calculated and then be chosen to be the right end of the uncertainty interval. Let $\bar{\sigma}^* \in (-\lambda_1, +\infty)$ be the root of $\psi(\sigma)$. From the definition of $\psi(\sigma)$, we have

$$\frac{1}{(\lambda_1 + \bar{\sigma}^*)^2} \hat{\boldsymbol{p}}^T \hat{\boldsymbol{p}} - r^2 \geq 0.$$

Hence, $\sqrt{\hat{\boldsymbol{p}}^T \hat{\boldsymbol{p}}}/r = \|\boldsymbol{p}\|/r$ is an upper bound for the root $\bar{\sigma}^*$. However, the bound $\|\boldsymbol{p}\|/r$ may be not tight. A practical way is to let $\sigma = -\lambda_1$ as a starting point and then to update σ recursively by moving a certain step to its right each step. If the first σ that makes the value of $\psi(\sigma)$ be negative is smaller than the upper bound $\|\boldsymbol{p}\|/r$, it is a tighter right end for the uncertainty interval.

Algorithm 1 (Initialization)
Input: Coefficients $\boldsymbol{Q}$, $\boldsymbol{f}$ and r, and an error tolerance ε.
The smallest eigenvalue: Use Lanczos method to obtain $\tilde{\lambda}_1$ and $\tilde{U}_1$.
Perturbation: If existence conditions do not hold, a perturbation is introduced and let

$$\boldsymbol{p} = \boldsymbol{f} + \alpha \tilde{U}_1;$$

otherwise, let $\boldsymbol{p} = \boldsymbol{f}$.
Uncertainty interval: set a step size s_t and a threshold ε_t; let $\sigma = \sigma_\ell = -\tilde{\lambda}_1$.

step 1: Solve the problem (5). If the value of the target function is larger than the threshold ε_t, stop the iteration, let $\sigma = \sigma + s_t$ and go to step 1; otherwise, go to step 2.
step 2: Calculate the value of $\psi(\sigma)$. If $\psi(\sigma) > 0$, set $\sigma_\ell = \sigma$, $\sigma = \sigma + s_t$ and go to step 2; otherwise, let $\sigma_u = \sigma$ and stop.

As the uncertainty interval $[\sigma_\ell, \sigma_u]$ is obtained, the bisection method is applied to find the next iterate for σ, by setting σ be the middle point of the uncertainty interval. The main part of the algorithm is given as follows:

Algorithm 2 (Main)
Do

set $\sigma = (\sigma_\ell + \sigma_u)/2$ and calculate the value of $\psi(\sigma)$;
If $|\psi(\sigma)| < \varepsilon$, then STOP and return σ and x;
Else if $\psi(\sigma) > 0$, update $\sigma_\ell = \sigma$;
Else update $\sigma_u = \sigma$;
End if

End do

6 Numerical Experiments

First, three small-size examples are used to illustrate the application of the canonical duality theory. Then, randomly generated examples for $n \in [500, 5000]$ are presented to demonstrate the efficiency of our method.

6.1 Small-Size Examples

Example 1 The given coefficients are

$$\boldsymbol{Q} = \begin{pmatrix} [r] - 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} [r]0 \\ -1.8 \end{pmatrix}, \text{ and } r = 1.$$

The existence conditions do not hold true for this example. There are two global solutions, $\bar{\boldsymbol{x}}_1 = (0.437, -0.9)^T$ and $\bar{\boldsymbol{x}}_2 = (-0.437, -0.9)^T$, which are red points shown in Fig. 2. In order to show how the perturbation method works, a big perturbation is firstly introduced to the linear coefficient $\boldsymbol{f}$ and let

$$\boldsymbol{p} = (0.5, -1.8)^T.$$

A critical point appears in the interior of $\mathscr{S}_a^+$, which is $\bar{\sigma} = 1.676$ (see Fig. 2b). The corresponding optimal solution for the perturbed problem is $\bar{\boldsymbol{x}}_1^* = (0.74, -0.673)^T$, which is shown as a green point in Fig. 2a. As the perturbation becomes smaller, the solution of the perturbed problem should approach to that of the original problem. We then let

$$\boldsymbol{p} = (0.01, -1.8)^T.$$

The critial point now is $\bar{\sigma} = 1.022$ and the corresponding solution is $\bar{\boldsymbol{x}}_1^* = (0.456, -0.89)^T$ (see Fig. 2d and c).

As pointed out above, the other global solution, $\bar{\boldsymbol{x}}_2$, can also be approximated by just choosing a perturbation with the opposite direction.

Let $\boldsymbol{p} = (-0.5, -1.8)^T$ and $\boldsymbol{p} = (-0.01, -1.8)^T$. The critical point will be the same as that for $\bar{\boldsymbol{x}}_1^*$, $\bar{\sigma} = 1.676$ and $\bar{\sigma} = 1.022$, and their corresponding primal solutions are $\bar{\boldsymbol{x}}_2^* = (-0.74, -0.673)^T$ and $\bar{\boldsymbol{x}}_2^* = (-0.456, -0.89)^T$.

In Fig. 2b, we can see that there is no critical point between $-\lambda_2 = -1$ and $-\lambda_1 = 1$, which suggests that there will no local-nonglobal solution. While there is a critical point between $-\lambda_2 = -1$ and $-\lambda_1 = 1$ in Fig. 2d, by Theorem 4 there must be a local-nonglobal solution and it should locate near one of the global solutions, depending on the perturbation.

Example 2 The matrix $\boldsymbol{Q}$ and radius r are the same as that in Example 1 and $\boldsymbol{f}$ is changed to

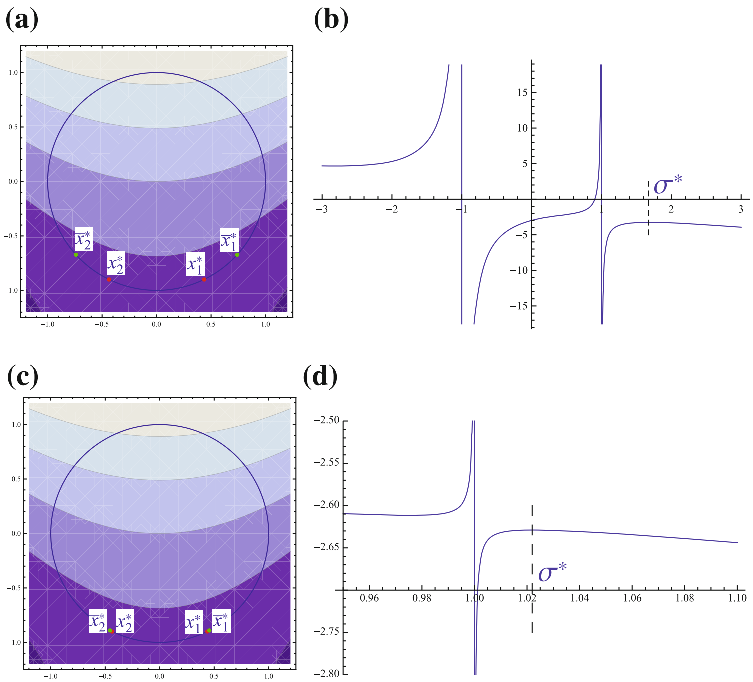


Fig. 2 Example 1: **a** and **c** are contours of the primal function and the boundary of the sphere; **b** and **d** are the graphs of the dual function

$$\boldsymbol{f} = \begin{pmatrix} 0 \\ -3 \end{pmatrix},$$

which is in the same direction of that in Example 1 but has a larger length. We notice that though $\sum_{i=1}^{k} \hat{f}_i^2 \neq 0$ is violated, the condition $\sum_{i=k+1}^{n} \frac{\hat{f}_i^2}{(\lambda_i - \lambda_1)^2} > r^2$ holds true. Thus, the problem is not in the hard case. There is a critical point in the interior of $\mathscr{S}_a^+$, which is shown in Fig. 3b, and it is corresponding to the unique global solution of the primal problem, which is the green point in Fig. 3a.

Example 3 We consider a four-dimensional problem with $\boldsymbol{Q}$, $\boldsymbol{f}$ and r being

$$\boldsymbol{Q} = \begin{pmatrix} [r]-10 & 0 & 2 & -2 \\ 0 & -3 & -4 & 2 \\ 2 & -4 & 7 & -4 \\ -2 & 2 & -4 & 1 \end{pmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} [r]-10 \\ 6 \\ 10 \\ 9 \end{pmatrix}, \text{ and } r = 5.$$

As shown in Fig. 1, the canonical dual function $P^d(\sigma)$ has six critical points

$$\bar{\sigma}_6 = -11.1 < \bar{\sigma}_5 = -10.49 < \bar{\sigma}_4 = -1.84 < \bar{\sigma}_3 = 6.08 < \bar{\sigma}_2 = 8.23 < \bar{\sigma}_1 = 12.58.$$

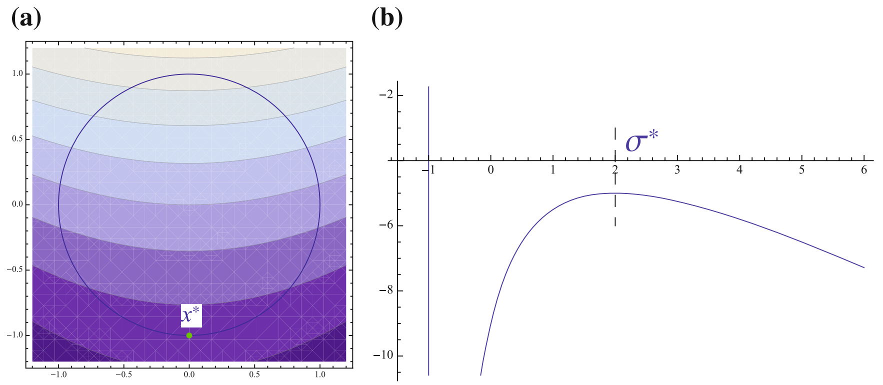


Fig. 3 Example 2: **a** is the contour of the primal function and boundary of the sphere; **b** is the graph of the dual function

It can be verified that $\bar{\sigma}_1$ belongs to $\mathscr{S}_a^+$, i.e., $\boldsymbol{G}(\bar{\sigma}_1) \succ 0$, which can also be observed from Fig. 1 where all the vertical lines represent eigenvalues of matrix $\boldsymbol{Q}$. Thus the corresponding solution

$$\bar{\boldsymbol{x}}_1 = (-4.71, 1.11, 1.25, 0.18)^T$$

is the global solution of the primal problem. While $\bar{\sigma}_2 = 8.23$ is a local minimizer of $P^d(\sigma)$ in $(-\lambda_2, -\lambda_1)$ and thus the corresponding solution

$$\bar{\boldsymbol{x}}_2 = (4.33, 1.05, 0.91, 2.08)^T$$

is the local-nonglobal minimizer.

6.2 Large-Size Examples

Examples with dimensions of 500, 1000, 2000, 3000, and 5000 are randomly generated, including both general and hard cases. For each given dimension, both cases are tested by ten examples, respectively. Thus, there are totally one hundred examples. All elements of the coefficients, $\boldsymbol{Q}$, $\boldsymbol{f}$, and r, are integer numbers in $[-100, 100]$. For each example of the hard case, in order to make $\boldsymbol{f}$ be easily chosen, we use a matrix $\boldsymbol{Q}$ of whom the multiplicity of the smallest eigenvalue is equal to one. The vector $\boldsymbol{f}$ is constructed such that it is perpendicular to the eigenvector of the smallest eigenvalue, and then a proper radius r is selected such that the existence conditions are violated.

Two approaches are used to calculate the value of $\psi(\sigma)$, one using decomposition methods to calculate $\boldsymbol{G}(\sigma)^{-1}\boldsymbol{p}$, for which we use the 'left division' in Matlab, and the other solving the problem (5), for which we use the function 'quadprog' in Matlab.

The tolerance parameter 'TolFun' of 'quadprog' is set to 1e-12. The Lanczos method is implemented by the function 'eigs' of Matlab. The Matlab is of version 7.13 and runned in the platform with Linux 64-bit system and quad CPUs.

The step size s_t, the threshold ε_t and the termination tolerance ε are set to $\|\boldsymbol{p}\|/(200r)$, 0, and 1e-8, respectively. For the hard case, a perturbation αU_1 is added to the vector $\boldsymbol{f}$, and two values of α, 1e-3, and 1e-4, are tried.

Results are shown in Tables 1, 2, 3, and 4, and they contain the number of examples which are successfully solved (Succ.Solv.), the distance of the optimal solution to the boundary of the sphere (Dist.Boun.), the number of iterations in Algorithm 2 (Main) (Numb.Iter.), and the running time (in second) of the algorithm (Runn.Time). The values in the columns of Dist.Boun., Numb.Iter., and Runn.Time are averages of the examples successfully solved. We compare the results of the algorithm adopting

Table 1 General case and $\alpha = 1e - 3$

Dim	Succ. Solv.		Dist. Boun.		Numb. Iter.		Runn. Time.	
	LD	QP	LD	QP	LD	QP	LD	QP
500	10	10	4.716e-09	5.245e-09	28.9	28.6	0.53	1.29
1000	10	10	4.261e-09	3.974e-09	27.1	27.5	1.67	6.25
2000	10	10	3.211e-09	3.822e-09	28.2	27.8	6.52	15.23
3000	10	10	5.674e-09	5.221e-09	26.1	26.4	20.90	72.43
5000	10	10	5.422e-09	3.873e-09	28.6	28.5	71.68	170.34

Table 2 General case and $\alpha = 1e - 4$

Dim	Succ. Solv.		Dist. Boun.		Numb. Iter.		Runn. Time.	
	LD	QP	LD	QP	LD	QP	LD	QP
500	10	10	4.532e-09	4.464e-09	28.9	28.9	0.43	1.16
1000	10	10	3.849e-09	5.931e-09	27.4	27.1	1.47	6.08
2000	10	10	2.648e-09	2.872e-09	27.9	28.5	6.26	15.82
3000	10	10	5.299e-09	5.137e-09	26.2	26.2	20.15	73.60
5000	10	10	3.188e-09	4.005e-09	28.7	28.5	65.71	171.92

Table 3 Hard case and $\alpha = 1e - 3$

Dim	Succ.Solv.		Dist.Boun.		Numb.Iter.		Runn.Time.	
	LD	QP	LD	QP	LD	QP	LD	QP
500	10	10	4.340e-09	6.297e-09	36.0	34.9	0.48	1.11
1000	10	10	4.253e-09	4.904e-09	34.6	34.9	1.54	3.54
2000	10	10	2.808e-09	4.255e-09	35.9	35.8	7.15	15.11
3000	9	10	5.479e-09	4.466e-09	34.0	35.0	19.41	36.01
5000	10	10	3.755e-09	4.705e-09	35.2	35.5	74.79	121.41

Table 4 Hard case and $\alpha = 1e - 4$

Dim	Succ.Solv.		Dist.Boun.		Numb.Iter.		Runn.Time.	
	LD	QP	LD	QP	LD	QP	LD	QP
500	7	9	2.503e-09	4.488e-09	39.6	40.6	0.51	1.36
1000	9	9	3.148e-09	4.482e-09	37.4	38.3	1.56	3.81
2000	5	9	8.668e-09	5.785e-09	38.6	42.6	7.36	17.95
3000	5	10	6.003e-09	3.997e-09	38.4	40.6	20.43	41.06
5000	8	10	4.748e-09	2.814e-09	37.8	38.8	72.72	131.51

'left division' and that of the algorithm adopting 'quadprog' in the same table, where LD denotes 'left division' and QP denotes 'quadprog'.

We can see that the examples are solved very accurately with error allowance being less than 1e-09. The failure in solving some examples is due to 'left division' and 'quadprog' being unable to handle very nearly singular matrices. For general cases, all the examples can be solved within no more than 30 iterations, while for hard cases, the number of iterations is around 40. From the running time, we notice that our method is capable to handle very large problems in reasonable time. The algorithms using 'left division' and 'quadprog' have similar performances in the accuracy and the number of iterations. Whereas the one using 'left division' needs much less time than that of the one using 'quadprog'. However, the one using 'quadprog' is able to solve more examples successfully.

7 Conclusion Remarks

We have presented a detailed study on the quadratic minimization problem with a sphere constraint. By the canonical duality, this nonconvex optimization is equivalent to a unified concave maximization dual problem over a convex domain $\mathscr{S}_a^+$, which is true also for many other global optimization problems under certain conditions (see [26, 42–47]). Based on this canonical dual problem, sufficient and necessary conditions are obtained for both general and hard cases. In order to solve hard-case problems, a perturbation method and the associated polynomial algorithm are proposed. Numerical results demonstrate that the proposed approach is able to solve large-size problems deterministically and efficiently. Combining with the trust region method, the theory and method presented in this paper can be used to solve general global optimizations.

Acknowledgements This research is supported by US Air Force Office of Scientific Research under the grants AFOSR FA2386-16-1-4082 and FA9550-17-1-0151, as well as by a grant from the Australian Government under the Collaborative Research Networks (CRN) program. The main results of this paper have been announced at the 3rd World Congress of Global Optimization, July 9–11, 2013, the Yellow Mountains, China.

References

1. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods, and Applications. Springer, New York (2000)
2. Gao, D.Y.: Panpenalty finite element programming for limit analysis. Comput. Struct. **28**(6), 749–755 (1988)
3. Gao, D.Y.: On the complementary bounding theorems for limit analysis. Int. J. Solids Struct. **24**(6), 545–556 (1988)
4. Conn, A.R., Gould, N.I.M., Toint, P.L.: Trust-Region Methods. SIAM, Philadelphia, PA (2000)
5. Powell, M.J.D.: On trust region methods for unconstrained minimization without derivatives. Math. Program. **97**(3), 605–623 (2003)
6. Jin, Q., Fang, S.C., Xing, W.X.: On the global optimality of generalized trust region subproblems. Optimization **59**(8), 1139–1151 (2010)
7. Xing, WX, Fang, SC, Gao, DY, Sheu, RL, Zhang, L.: Canonical dual solutions to the quadratic programming over a quadratic constraint. In: ICOTA7 (2007)
8. Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
9. Stern, R.J., Wolkowicz, H.: Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. SIAM J. Optim. **5**(2), 286–313 (1995)
10. Ben-Tal, A., Teboulle, M.: Hidden convexity in some nonconvex quadratically constrained quadratic programming. Math. Program. **72**(1), 51–63 (1996)
11. Sorensen, D.C.: Newton's method with a model trust region modification. SIAM J. Numer. Anal. **19**(2), 409–426 (1982)
12. Sorensen, D.C.: Minimization of a large-scale quadratic functionsubject to a spherical constraint. SIAM J. Optim. **7**(1), 141–161 (1997)
13. Rendl, F., Wolkowicz, H.: A semidefinite framework for trust region subproblems with applications to large scale minimization. Math. Program. **77**, 273–299 (1997)
14. Jorge, N., Wright, S.J.: Numerical Optimization, vol. 2. Springer, New York (1999)
15. Rojas, M., Santos, S.A., Sorensen, D.C.: A new matrix-free algorithm for the large-scale trust-region subproblem. SIAM J. Optim. **11**(3), 611–646 (2001)
16. Fortin, C., Wolkowicz, H.: The trust region subproblem and semidefinite programming. Optim. Method Softw. **19**(1), 41–67 (2004)
17. Moré, J.J., Sorensen, D.C.: Computing a trust region step. SIAM J. Sci. Stat. Comput. **4**(3), 553–572 (1983)
18. Hager, W.W.: Minimizing a quadratic over a sphere. SIAM J. Optim. **12**(1), 188–208 (2001)
19. Gay, D.M.: Computing optimal locally constrained steps. SIAM J. Sci. Stat. Comput. **2**(2), 186–197 (1981)
20. Gould, N.I.M., Lucidi, S., Roma, M., Toint, P.L.: Solving the trust-region subproblem using the lanczos method. SIAM J. Optim. **9**(2), 504–525 (1999)
21. Tao, P.D., An, L.T.H.: A dc optimization algorithm for solving the trust-region subproblem. SIAM J. Optim. **8**(2), 476–505 (1998)
22. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solution for global optimization problems. Comput. Chem. Eng. **33**(12), 1964–1972 (2009)
23. Gao, D.Y., Ruan, N., Sherali, H.D.: Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and lagrangian duality. J. Global Optim. **45**(3), 473–497 (2009)
24. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Quart. Appl. Math. **47**, 487–504 (1989)
25. Ciarlet, P.G.: Linear and Nonlinear Functional Analysis with Applications, vol. 130. SIAM (2013)
26. Gao, D.Y.: Perfect duality theory and complete solutions to a class of global optimization problems. Optimization **52**(4–5), 467–493 (2003)

27. Gao, D.Y., Sherali, H.D.: Canonical duality theory: connection between nonconvex mechanics and global optimization. In: Gao, D.Y., Sherali, H.D. (eds.) Adv. Appl. Math. Glob. Optim. Springer, New York (2009)
28. Gao, D.Y.: Complementary finite element method for finite deformation nonsmooth mechanics. J. Eng. Math. **30**(3), 339–353 (1996)
29. Gao, D.Y., Wu, C.: On the triality theory for a quartic polynomial optimisation problem. J. Ind. Manag. Optim. **8**(1), 229–242 (2012)
30. Voisei, M., Zalinescu, C.: Some remarks concerning gao-strang's complementary gap function. Appl. Anal. **90**(6), 1111–1121 (2010)
31. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
32. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland (1976)
33. Gao, D.Y.: Canonical duality theory and solutions to constrained nonconvex quadratic programming. J. Glob. Optim. **29**(4), 377–399 (2004)
34. Morales Silva, D.M., Gao, D.Y.: Canonical Duality Theory and Triality for Solving General Nonconstrained Global Optimization Problems, arXiv preprint arXiv:1210.0180, 2012
35. Morales Silva, D.M., Gao, D.Y.: Complete solutions and triality theory to a nonconvex optimization problem with double-well potential in R^n. Numer. Algebra Control Optim. **3**(2), 271–282 (2013)
36. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. **47**(3), 463–484 (2010)
37. Martínez, J.M.: Local minimizers of quadratic functions on euclidean balls and spheres. SIAM J. Optim. **4**(1), 159–176 (1994)
38. 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)
39. Ruan, N., Gao, D.Y.: Canonical duality approach for non-linear dynamical systems. IMA J. Appl. Math. **79**, 313–325 (2014)
40. Ruan, N., Gao, D.Y.: Global optimal solutions to general sensor network localization problem. Perform. Eval. **75–76**, 1–16 (2014)
41. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Canonical dual approach to solving the maximum cut problem. J. Glob. Optim. **54**(2), 341–351 (2012)
42. Gao, D.Y.: Sufficient conditions and perfect duality in nonconvex minimization with inequality constraints. J. Ind. Manag. Optim. **1**(1), 53–63 (2005)
43. Gao, D.Y.: Complete solutions and extremality criteria to polynomial optimization problems. J. Glob. Optim. **35**(1), 131–143 (2006)
44. Gao, D.Y.: Solutions and optimality criteria to box constrained nonconvex minimization problems. J. Ind. Manag. Optim. **3**(2), 293–304 (2007)
45. Gao, D.Y., Ruan, N., Sherali, H.D.: Canonical dual solutions for fixed cost quadratic programs. In: Chinchuluun, A., Pardalos, P.M., Enkhbat, R., Tseveendorj, I. (eds.) Optimization and Optimal Control, vol. 39, pp. 139–156. Springer, New York (2010)
46. 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, New York (2012)
47. Gao, D.Y., Watson, L.T., Easterling, D.R., Thacker, W.I., Billups, S.C.: Solving the canonical dual of box- and integer-constrained nonconvex quadratic programs via a deterministic direct search algorithm. Optim. Method Softw. **26**(1), 1–14 (2011)

Global Optimal Solution to Quadratic Discrete Programming Problem with Inequality Constraints

Ning Ruan and David Yang Gao

Abstract This paper presents a canonical dual method for solving a quadratic discrete value selection problem subjected to inequality constraints. By using a linear transformation, the problem is first reformulated as a standard quadratic 0–1 integer programming problem. Then, by the canonical duality theory, this challenging problem is converted to a concave maximization over a convex feasible set in continuous space. It is proved that if this canonical dual problem has a solution in its feasible space, the corresponding global solution to the primal problem can be obtained directly by a general analytical form. Otherwise, the problem could be NP-hard. In this case, a quadratic perturbation method and an associated canonical primal-dual algorithm are proposed. Numerical examples are illustrated to demonstrate the efficiency of the proposed method and algorithm.

1 Introduction

Many decision-making problems, such as portfolio selection, capital budgeting, production planning, resource allocation, and computer networks, etc., can often be formulated as quadratic programming problems with discrete variables. See for examples, [4, 5, 9, 24]. In engineering applications, the decision variables can not have arbitrary values. Instead, either some or all of the variables must be selected from a list of integer or discrete values for practical reasons. For examples, structural members may have to be selected from selections available in standard sizes, member thicknesses may have to be selected from the commercially available ones, the number of bolts for a connection must be an integer, the number of reinforcing bars in a concrete member must be an integer, etc. [23]. However, these integer

N. Ruan (✉) · D.Y. Gao
Faculty of Science and Technology, Federation University Australia, Ballarat, VIC 3353, Australia
e-mail: n.ruan@federation.edu.au

D.Y. Gao
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_16

programming problems are computationally highly demanding. Nevertheless, some numerical methods are now available.

Several review articles on nonlinear optimization problems with discrete variables are available [1, 4, 28, 33, 37, 38], and some popular methods have been discussed, including branch and bound methods, a hybrid method that combines a branch and bound method with a dynamic programming technique [29], sequential linear programming, rounding-off techniques, cutting plane techniques [2], heuristic techniques, penalty function approaches, simulated annealing [25], and genetic algorithms, etc. The relaxation methods have also been proposed recently, leading to second order cone programming (SOC) [21] and improved linearization strategy [35].

Branch and bound is perhaps the most widely known and used deterministic method for discrete optimization problems. When applied to linear problems, this method can be implemented in a way to yield a global minimum point. However for nonlinear problems there is no such guarantee, unless the problem is convex. The branch and bound method has been used successfully to deal with problems with discrete design variables. However for the problem with a large number of discrete design variables, the number of subproblems (nodes) becomes large, making the method inefficient.

Simulated annealing (SA) and genetic algorithms (GA) belong to the category of stochastic search methods [22] which based on an element of random choice. Because of this, one has to sacrifice the possibility of an absolute guarantee of success within a finite amount of computation.

Canonical duality theory provides a new and potentially useful methodology for solving a large class of nonconvex/nonsmooth/discrete problems (see the review articles [13, 19]). It was shown in [8, 12] that the Boolean integer programming problems are actually equivalent to certain canonical dual problems in continuous space without duality gap, which can be solved deterministically under certain conditions. This theory has been generalized for solving multi-integer programming [39] and the well-known max cut problems [40]. It is also shown in [13, 16] that by the canonical duality theory, the NP-hard quadratic integer programming problem is identical to a continuous unconstrained Lipschitzian global optimization problem, which can be solved via deterministic methods (but not in polynomial times) (see [20]). The canonical duality theory has been used successfully for solving a large class of challenging problems not only in global optimization, but also in nonconvex analysis and continuum mechanics [17].

In this paper, our goal is to solve a general quadratic programming problem with its decision variables taking values from discrete sets. The elements from these discrete sets are not required to be binary or uniformly distributed. An effective numerical method is developed based on the canonical duality theory [10]. The rest of the paper is organized as follows. Section 2 presents a mathematical statement of the general discrete value quadratic programming problem and how it can be transformed into a standard 0–1 programming problem in higher dimensional space. Section 3 presents a brief review on the canonical duality theory. Detailed canonical dual transformation procedure is presented in Sect. 4 to show how the integer programming problem can be converted to a concave maximization in a convex space. A perturbed computational method is developed in Sect. 5. Some numerical examples are illustrated in Sect. 6 to demonstrate the effectiveness and efficiency of the proposed method. The paper is ended with some concluding remarks.

2 Primal Problem and Equivalent Transformation

The discrete programming problem to be addressed is given below:

$$(\mathscr{P}_a) \quad \min \quad P(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} - \mathbf{c}^T\mathbf{x} \tag{1}$$

$$\text{s.t.} \quad \mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b} \le 0, \tag{2}$$

$$\mathbf{x} = [x_1, x_2, \cdots, x_n]^T, \; x_i \in U_i, \; i = 1, \cdots, n,$$

where $Q = \{q_{ij}\} \in \mathbb{R}^{n\times n}$ is a symmetric matrix, $\mathbf{A} = \{a_{ij}\} \in \mathbb{R}^{m\times n}$ is a matrix with $rank(\mathbf{A}) = m < n$, $\mathbf{c} = [c_1, \cdots, c_n]^T \in \mathbb{R}^n$ and $\mathbf{b} = [b_1, \cdots, b_m]^T \in \mathbb{R}^m$ are given vectors. Here, for each $i = 1, \cdots, n$,

$$U_i = \{u_{i,1}, \cdots, u_{i,K_i}\},$$

where $u_{i,j}$, $j = 1, \cdots, K_i$, are given real numbers. In this paper, we let $K = \sum_{i=1}^n K_i$.

Problem $(\mathscr{P}_a)$ arises in many real-world applications, say, the pipe network optimization problems in water distribution systems, where the choices of pipelines are discrete values. Such problems have been studied extensively by traditional direct approaches (see [41]). Due to the constraint of discrete values, this problem is considered to be NP-hard and the traditional methods can only provide upper bound results. In this paper, we will show that the canonical duality theory will provide either a lower bound approach to this challenging problem, or the global optimal solution under certain conditions.

In order to convert the discrete value problem $(\mathscr{P}_a)$ to the standard 0–1 programming problem, we introduce the following transformation,

$$x_i = \sum_{j=1}^{K_i} u_{i,j} y_{i,j}, \; i = 1, \cdots, n, \tag{3}$$

where, for each $i = 1, \cdots, n, u_{i,j} \in U_i, \; j = 1, \cdots, K_i$. Then, the discrete programming problem $(\mathscr{P}_a)$ can be written as the following 0–1 programming problem:

$$(\mathscr{P}_b) \quad \min \quad P(\mathbf{y}) = \frac{1}{2}\mathbf{y}^T B\mathbf{y} - \mathbf{h}^T\mathbf{y} \tag{4}$$

$$\text{s.t.} \quad \mathbf{g}(\mathbf{y}) = D\mathbf{y} - \mathbf{b} \le 0, \tag{5}$$

$$\sum_{j=1}^{K_i} y_{i,j} - 1 = 0, \ i = 1, \cdots, n, \tag{6}$$

$$y_{i,j} \in \{0, 1\}, \ i = 1, \ldots, n; \ j = 1, \cdots, K_i, \tag{7}$$

where

$$\mathbf{y} = [y_{1,1}, \cdots, y_{1,K_1}, \cdots, y_{n,1}, \cdots, y_{n,K_n}]^T \in \mathbb{R}^K,$$

$$\mathbf{h} = [c_1 u_{1,1}, \cdots, c_1 u_{1,K_1}, \cdots, c_n u_{n,1}, \cdots, c_n u_{n,K_n}]^T \in \mathbb{R}^K,$$

$$B = \begin{bmatrix} q_{1,1}u_{1,1}^2 & \cdots & q_{1,1}u_{1,1}u_{1,K_1} & \cdots & q_{1,n}u_{1,1}u_{n,K_n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ q_{1,1}u_{1,K_1}u_{1,1} & \cdots & q_{1,1}u_{1,K_1}^2 & \cdots & \cdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ q_{n,1}u_{n,K_n}u_{1,1} & \cdots & \cdots & \cdots & q_{n,n}u_{n,K_n}^2 \end{bmatrix} \in \mathbb{R}^{K \times K},$$

$$D = \begin{bmatrix} a_{1,1}u_{1,1} & \cdots & a_{1,1}u_{1,K_1} & \cdots & a_{1,n}u_{n,K_n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m,1}u_{1,1} & \cdots & a_{m,1}u_{1,K_1} & \cdots & a_{m,n}u_{n,K_n} \end{bmatrix} \in \mathbb{R}^{m \times K}.$$

Theorem 1 *Problem* $(\mathscr{P}_b)$ *is equivalent to Problem* $(\mathscr{P}_a)$.

Proof For any $i = 1, 2, \cdots, n$, it is clear that constraints (6) and (7) are equivalent to the existence of only one $j \in \{1, \cdots, K_i\}$, such that $y_{i,j} = 1$ while $y_{i,j} = 0$ for all other j. Thus, from the definition of $\mathbf{y}$, the conclusion follows readily. □

Problem $(\mathscr{P}_b)$ is a standard 0–1 quadratic programming problem with both equality and inequality constraints. Let

$$H = \begin{bmatrix} 1 \cdots 1 & 0 \cdots 0 & \cdots & 0 \cdots 0 \\ 0 \cdots 0 & 1 \cdots 1 & \cdots & 0 \cdots 0 \\ \vdots \ddots \vdots & \vdots \ddots \vdots & \ddots & \vdots \ddots \vdots \\ 0 \cdots 0 & 0 \cdots 0 & \cdots & 1 \cdots 1 \end{bmatrix} \in \mathbb{R}^{n \times K}$$

and, for a given integer K, let

$$\mathbf{e}_K = [1, \cdots, 1, \cdots, 1, \cdots, 1]^T \in \mathbb{R}^K.$$

Thus, on the feasible space

$$\mathscr{Y} = \{\mathbf{y} \in \mathbb{R}^K : D\mathbf{y} \le \mathbf{b}, \quad H\mathbf{y} = \mathbf{e}_n, \quad \mathbf{y} \in \{0, 1\}^K\}, \tag{8}$$

the integer constrained problem $(\mathscr{P}_b)$ can be reformulated as a standard constrained 0–1 programming problem:

$$(\mathscr{P}_c): \quad \min \left\{ P(\mathbf{y}) = \frac{1}{2}\mathbf{y}^T B\mathbf{y} - \mathbf{h}^T\mathbf{y} | \ \ \mathbf{y} \in \mathscr{Y} \right\}. \tag{9}$$

3 Canonical Duality Theory: A Brief Review

The basic idea of the canonical duality theory can be demonstrated by solving the following general nonconvex problem (the primal problem $(\mathscr{P})$ in short)

$$(\mathscr{P}): \ \min_{\mathbf{x}\in\mathscr{X}_a} \left\{ P(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, \mathbf{A}\mathbf{x}\rangle - \langle \mathbf{x}, \mathbf{f}\rangle + W(\mathbf{x}) \right\}, \tag{10}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a given symmetric indefinite matrix, $\mathbf{f} \in \mathbb{R}^n$ is a given vector (input), $\langle \mathbf{x}, \mathbf{x}^*\rangle$ denotes the bilinear form between $\mathbf{x}$ and its dual variable $\mathbf{x}^*$, $\mathscr{X}_a \subset \mathbb{R}^n$ is a given feasible space, and $W : \mathscr{X}_a \to \mathbb{R} \cup \{\infty\}$ is a general nonconvex objective function.

It must be emphasized that, different from the objective function extensively used in mathematical optimization, a real-valued function $W(\mathbf{x})$ is called to be *objective* in continuum physics and the canonical duality theory only if (see [10] Chap. 6, p. 288)

$$W(\mathbf{x}) = W(\mathbf{Q}\mathbf{x}) \quad \forall\, \mathbf{x} \in \mathscr{X}_a, \quad \forall \mathbf{Q} \in \mathscr{Q},$$

where $\mathscr{Q} = \{\mathbf{Q} \in \mathbb{R}^{n\times n} | \ \ \mathbf{Q}^{-1} = \mathbf{Q}^T \quad \det \mathbf{Q} = 1\}$ is a special rotation group.

Geometrically speaking, an objective function does not depend on the rotation, but only on certain measure of its variable. In Euclidean space $\mathbb{R}^n$, the simplest objective function is the ℓ_2-norm $\|\mathbf{x}\|$ in $\mathbb{R}^n$ since $\|\mathbf{Q}\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{Q}^T\mathbf{Q}\mathbf{x} = \|\mathbf{x}\|^2 \ \ \forall \mathbf{Q} \in \mathscr{Q}$. By Cholesky factorization, any positive definite matrix has a unique decomposition $C = D^*D$. Thus, any convex quadratic function is objective. Physically, an objective function does not depend on observers [7], which is essential for any real-world mathematical modeling.

The **key step** in the canonical duality theory is to choose a nonlinear operator

$$\boldsymbol{\xi} = \Lambda(\mathbf{x}) : \mathscr{X}_a \to \mathscr{E}_a \subset \mathbb{R}^p \tag{11}$$

and a *canonical function* $V : \mathscr{E}_a \to \mathbb{R}$ such that the nonconvex objective function $W(\mathbf{x})$ can be recast by adopting a canonical form $W(\mathbf{x}) = V(\Lambda(\mathbf{x}))$. Thus, the primal problem $(\mathscr{P})$ can be written in the following canonical form:

$$(\mathscr{P}): \min_{\mathbf{x}\in\mathscr{X}_a} \{P(\mathbf{x}) = V(\Lambda(\mathbf{x})) - U(\mathbf{x})\}, \tag{12}$$

where $U(\mathbf{x}) = \langle \mathbf{x}, \mathbf{f}\rangle - \frac{1}{2}\langle \mathbf{x}, \mathbf{A}\mathbf{x}\rangle$. By the definition introduced in [10], a differentiable function $V(\boldsymbol{\xi})$ is said to be a *canonical function* on its domain $\mathscr{E}_a$ if the duality mapping $\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi})$ from $\mathscr{E}_a$ to its range $\mathscr{S}_a \subset \mathbb{R}^p$ is invertible. Let $\langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle$ denote the bilinear form on $\mathscr{E}_a \times \mathscr{S}_a$. Thus, for the given canonical function $V(\boldsymbol{\xi})$, its Legendre conjugate $V^*(\boldsymbol{\varsigma})$ can be defined uniquely by the Legendre transformation (cf. Gao [10])

$$V^*(\boldsymbol{\varsigma}) = \text{sta}\{\langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle - V(\boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \mathscr{E}_a\}, \tag{13}$$

where the notation sta$\{g(\boldsymbol{\xi}) | \boldsymbol{\xi} \in \mathscr{E}_a\}$ stands for finding stationary point of $g(\boldsymbol{\xi})$ on $\mathscr{E}_a$. It is easy to prove that the following canonical duality relations hold on $\mathscr{E}_a \times \mathscr{S}_a$:

$$\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) \;\Leftrightarrow\; \boldsymbol{\xi} = \nabla V^*(\boldsymbol{\varsigma}) \;\Leftrightarrow\; V(\boldsymbol{\xi}) + V^*(\boldsymbol{\varsigma}) = \langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle. \tag{14}$$

By this one-to-one canonical duality, the nonconvex term $W(\mathbf{x}) = V(\Lambda(\mathbf{x}))$ in the problem $(\mathscr{P})$ can be replaced by $\langle \Lambda(\mathbf{x}); \boldsymbol{\varsigma}\rangle - V^*(\boldsymbol{\varsigma})$ such that the nonconvex function $P(\mathbf{x})$ is reformulated as the Gao-Strang total complementary function [10]:

$$\Xi(\mathbf{x}, \boldsymbol{\varsigma}) = \langle \Lambda(\mathbf{x}); \boldsymbol{\varsigma}\rangle - V^*(\boldsymbol{\varsigma}) - U(\mathbf{x}): \quad \mathscr{X}_a \times \mathscr{S}_a \to \mathbb{R}. \tag{15}$$

By using this total complementary function, the canonical dual function $P^d(\boldsymbol{\varsigma})$ can be obtained as

$$\begin{aligned} P^d(\boldsymbol{\varsigma}) &= \text{sta}\{\Xi(\mathbf{x}, \boldsymbol{\varsigma}) \mid \mathbf{x} \in \mathscr{X}_a\} \\ &= U^\Lambda(\boldsymbol{\varsigma}) - V^*(\boldsymbol{\varsigma}), \end{aligned} \tag{16}$$

where $U^\Lambda(\mathbf{x})$ is defined by

$$U^\Lambda(\boldsymbol{\varsigma}) = \text{sta}\{\langle \Lambda(\mathbf{x}); \boldsymbol{\varsigma}\rangle - U(\mathbf{x}) \mid \mathbf{x} \in \mathscr{X}_a\}. \tag{17}$$

In many applications, the geometrically nonlinear operator $\Lambda(\mathbf{x})$ is usually a quadratic function [3, 34]

$$\Lambda(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, D_k\mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{b}_k\rangle, \tag{18}$$

where $D_k \in \mathbb{R}^{n\times n}$ and $\mathbf{b}_k \in \mathbb{R}^n (k = 1, \cdots, p)$. Let $\boldsymbol{\varsigma} = [\varsigma_1, \cdots, \varsigma_p]^T$. In this case, the canonical dual function can be written in the following form:

$$P^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\langle \mathbf{F}(\boldsymbol{\varsigma}), \mathbf{G}^{-1}(\boldsymbol{\varsigma})\mathbf{F}(\boldsymbol{\varsigma})\rangle - V^*(\boldsymbol{\varsigma}), \tag{19}$$

where

$$\mathbf{G}(\boldsymbol{\varsigma}) = \mathbf{A} + \sum_{k=1}^{p} \varsigma_k D_k, \quad \mathbf{F}(\boldsymbol{\varsigma}) = \mathbf{f} - \sum_{k=1}^{p} \varsigma_k \mathbf{b}_k.$$

Let

$$\mathscr{S}_a^+ = \{\boldsymbol{\varsigma} \in \mathbb{R}^p | \ G(\boldsymbol{\varsigma}) \succ 0\}.$$

It is easy to prove that $\mathscr{S}_a^+$ is convex. Moreover, $\mathscr{S}_a^+$ is nonempty as long as there exists one $D_k \succ 0$.

Therefore, the canonical dual problem can be proposed as

$$(\mathscr{P}^d): \quad \max\{P^d(\boldsymbol{\varsigma}) | \ \boldsymbol{\varsigma} \in \mathscr{S}_a^+\}. \tag{20}$$

which is a concave maximization problem over a convex set $\mathscr{S}_a^+ \subset \mathbb{R}^p$.

Theorem 2 **([10]).** *Problem $(\mathscr{P}^d)$ is canonically dual to $(\mathscr{P})$ in the sense that if $\bar{\boldsymbol{\varsigma}}$ is a critical point of $P^d(\boldsymbol{\varsigma})$, then*

$$\bar{\mathbf{x}} = \mathbf{G}^{-1}(\bar{\boldsymbol{\varsigma}})\mathbf{F}(\bar{\boldsymbol{\varsigma}}) \tag{21}$$

is a critical point of $\Pi(\mathbf{x})$ and

$$P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}) = P^d(\bar{\boldsymbol{\varsigma}}). \tag{22}$$

If $\bar{\boldsymbol{\varsigma}}$ is a solution to $(\mathscr{P}^d)$, then $\bar{\mathbf{x}}$ is a global minimizer of $(\mathscr{P})$ and

$$\min_{\mathbf{x} \in \mathscr{X}_a} P(\mathbf{x}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}) = \max_{\boldsymbol{\varsigma} \in \mathscr{S}_a^+} P^d(\boldsymbol{\varsigma}). \tag{23}$$

Conversely, if $\bar{\mathbf{x}}$ is a solution to $(\mathscr{P})$, it must be in the form of (21) for critical solution $\bar{\boldsymbol{\varsigma}}$ of $P^d(\boldsymbol{\varsigma})$.

To help explaining the theory, we consider a simple nonconvex optimization in $\mathbb{R}^n$:

$$\min P(\mathbf{x}) = \frac{1}{2}\alpha(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda)^2 - \mathbf{x}^T\mathbf{f}, \ \forall \mathbf{x} \in \mathbb{R}^n, \tag{24}$$

where $\alpha, \lambda > 0$ are given parameters. The criticality condition $\nabla P(\mathbf{x}) = 0$ leads to a nonlinear algebraic equation system in $\mathbb{R}^n$

$$\alpha(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda)\mathbf{x} = \mathbf{f}. \tag{25}$$

Clearly, to solve this n-dimensional nonlinear algebraic equation directly is difficult. Also traditional convex optimization theory can not be used to identify global minimizer. However, by the canonical dual transformation, this problem can be solved. To do so, we let $\xi = \Lambda(u) = \frac{1}{2}\|\mathbf{x}\|^2 - \lambda \in \mathbb{R}$. Then, the nonconvex function $W(\mathbf{x}) = \frac{1}{2}\alpha(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda)^2$ can be written in canonical form $V(\xi) = \frac{1}{2}\alpha\xi^2$. Its Legendre conjugate is given by $V^*(\varsigma) = \frac{1}{2}\alpha^{-1}\varsigma^2$, which is strictly convex. Thus, the total complementary function for this nonconvex optimization problem is

$$\Xi(\mathbf{x}, \varsigma) = (\frac{1}{2}\|\mathbf{x}\|^2 - \lambda)\varsigma - \frac{1}{2}\alpha^{-1}\varsigma^2 - \mathbf{x}^T\mathbf{f}. \tag{26}$$

For a fixed $\varsigma \in \mathbb{R}$, the criticality condition $\nabla_{\mathbf{x}}\Xi(\mathbf{x}, \varsigma) = 0$ leads to

$$\varsigma\mathbf{x} - \mathbf{f} = 0. \tag{27}$$

For each $\varsigma \neq 0$, the Eq. (27) gives $\mathbf{x} = \mathbf{f}/\varsigma$ in vector form. Substituting this into the total complementary function Ξ, the canonical dual function can be easily obtained as

$$\begin{aligned} P^d(\varsigma) &= \{\Xi(\mathbf{x}, \varsigma) | \nabla_{\mathbf{x}}\Xi(\mathbf{x}, \varsigma) = 0\} \\ &= -\frac{\|\mathbf{f}\|^2}{2\varsigma} - \frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma, \quad \forall \varsigma \neq 0. \end{aligned} \tag{28}$$

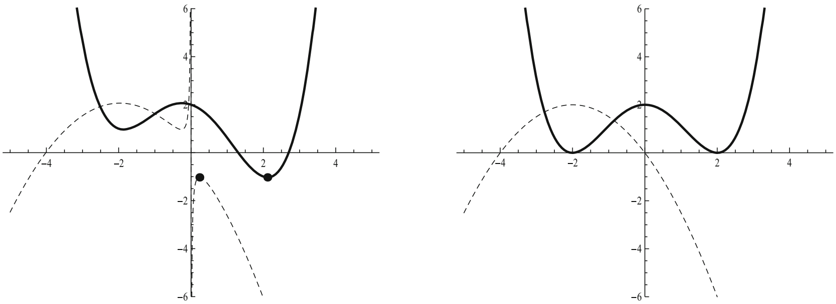
The critical point of this canonical function is obtained by solving the following dual algebraic equation

$$(\alpha^{-1}\varsigma + \lambda)\varsigma^2 = \frac{1}{2}\|\mathbf{f}\|^2. \tag{29}$$

For any given parameters α, λ and the vector $\mathbf{f} \in \mathbb{R}^n$, this cubic algebraic equation has at most three roots satisfying $\varsigma_1 \geq 0 \geq \varsigma_2 \geq \varsigma_3$, and each of these roots leads to a critical point of the nonconvex function $P(\mathbf{x})$, i.e., $\mathbf{x}_i = \mathbf{f}/\varsigma_i$, $i = 1, 2, 3$. By the fact that $\varsigma_1 \in \mathscr{S}_a^+ = \{\varsigma \in \mathbb{R} \mid \varsigma > 0\}$, then Theorem 1 tells us that $\mathbf{x}_1$ is a global minimizer of $P(\mathbf{x})$.

Consider one dimension problem with $\alpha = 1$, $\lambda = 2$, $f = \frac{1}{2}$, the primal function and canonical dual function are shown in Fig. 1, where, $x_1 = 2.11491$ is a global minimizer of $P(x)$, $\varsigma_1 = 0.236417$ is a global maximizer of $P^d(\varsigma)$, and $P(x_1) = -1.02951 = P^d(\varsigma_1)$ (See the two black dots).

If we let $\mathbf{f} = 0$, the graph of $P(\mathbf{x})$ is symmetric (i.e., the so-called double-well potential or the Mexican hat for $n = 2$ [11]) with infinite number of global minimizers satisfying $\|\mathbf{x}\|^2 = 2\lambda$. In this case, the canonical dual $P^d(\varsigma) = -\frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma$ is strictly concave with only one critical point (local maximizer) $\varsigma_3 = -\alpha\lambda < 0$ (for $\alpha, \lambda > 0$). The corresponding solution $\mathbf{x}_3 = \mathbf{f}/\varsigma_3 = 0$ is a local maximizer. By the canonical dual equation (29) we have $\varsigma_1 = \varsigma_2 = 0$ located on the boundary of $\mathscr{S}_a^+$, which corresponding to the two global minimizers $x_{1,2} = \pm\sqrt{2\lambda}$ for $n = 1$, see Fig. 1b.



(a) $f = 0.5$

(b) $f = 0$

Fig. 1 Graphs of $P(\mathbf{x})$ (solid) and $P^d(\varsigma)$ (dashed)

This simple example shows a fundamental issue in global optimization, i.e., the optimal solutions of a nonconvex problem depends sensitively on the linear term (input) $\mathbf{f}$. Geometrically speaking, the objective function $W(\mathbf{x})$ in $P(\mathbf{x})$ possesses certain symmetry. If there is no linear term, i.e., the *subjective function* in $P(\mathbf{x})$, the nonconvex problem usually has more than one global minimizer due to the symmetry. Traditional direct approaches and the popular SDP method are usually failed to deal with this situation. By the canonical duality theory, we understand that in this case the canonical dual function has no critical point in its open set $\mathscr{S}_a^+$. Therefore, by adding a linear perturbation $\mathbf{f}$ to break this symmetry, the canonical duality theory can be used to solve the nonconvex problems to obtain one of global optimal solutions. This idea was originally from Gao's work (1996) on post-buckling analysis of large deformed beam. The potential energy of this beam model is a double-well function, similar to this example, without the force $\mathbf{f}$, the beam could have two buckling states (corresponding to two minimizers) and one un-buckled state (local maximizer). Later on (2008) in the Gao and Ogden work on analytical solutions in phase transformation [14], they further discovered that the nonconvex system has no phase transition unless the force distribution $f(x)$ vanished at certain points. They also discovered that if force field $f(x)$ changes dramatically, all the Newton type direct approaches failed even to find any local minimizer. This discovery is fundamentally important for understanding NP-hard problems in global optimization and chaos in nonconvex dynamical systems. The linear perturbation method has been used successfully for solving global optimization problems [16, 18, 32, 40]. Comprehensive reviews of the canonical duality theory and its applications in nonconvex analysis and global optimization can be found in [11, 13, 15].

4 Canonical Dual Problem

Now we are ready to apply the canonical duality theory for solving the integer programming problem $(\mathscr{P}_c)$ presented in Sect. 2. As indicated in [12, 13], the key step for solving this NP-hard problem is to use a so-called canonical measure $\boldsymbol{\rho} = \{y_i(y_i - 1)\} \in \mathbb{R}^K$ such that the integer constraint $y_i \in \{0, 1\}$ can be equivalently written in the canonical form

$$\boldsymbol{\rho} = \mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K) = \{y_i(y_i - 1)\} = 0 \in \mathbb{R}^K$$

where the notation $\mathbf{s} \circ \mathbf{t} := [s_1 t_1, s_2 t_2, \ldots, s_K t_K]^T$ denotes the Hadamard product for any two vectors $\mathbf{s}, \mathbf{t} \in \mathbb{R}^K$. Thus, the so-called *geometrically admissible measure* Λ can be defined as

$$\begin{aligned} \boldsymbol{\xi} &= \Lambda(\mathbf{y}) = \{D\mathbf{y} - \mathbf{b},\ \ H\mathbf{y} - \mathbf{e}_n,\ \ \mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K)\} \\ &= \{\boldsymbol{\varepsilon},\ \ \boldsymbol{\delta},\ \ \boldsymbol{\rho}\} \in \mathscr{E} = \mathbb{R}^{m+n+K}. \end{aligned}$$

Let

$$U(\mathbf{y}) = -P(\mathbf{y}) = \mathbf{h}^T \mathbf{y} - \frac{1}{2}\mathbf{y}^T B \mathbf{y},$$

and define

$$V(\boldsymbol{\xi}) = \begin{cases} 0 & \text{if } \boldsymbol{\varepsilon} \le 0, \boldsymbol{\delta} = 0, \boldsymbol{\rho} = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Clearly, the constraints in $\mathscr{Y}$ can be replaced by the canonical transformation $V(\Lambda(\mathbf{y}))$ and the primal problem $(\mathscr{P}_c)$ can be equivalently written in the standard *canonical form* [13]

$$(\mathscr{P}):\quad \min\left\{\Pi(\mathbf{y}) = V(\Lambda(\mathbf{y})) - U(\mathbf{y}):\quad \mathbf{y} \in \mathbb{R}^K\right\}. \tag{30}$$

By the fact that $V(\boldsymbol{\xi})$ is convex, lower, semi-continuous on $\mathscr{E}$, its sub-differential leads to the canonical dual variable $\boldsymbol{\varsigma} = (\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}) \in \partial V(\boldsymbol{\xi}) \in \mathscr{E}^* = \mathbb{R}^{m+n+K}$, and its Fenchel super-conjugate (cf. Rockafellar [30])

$$\begin{aligned} V^\sharp(\boldsymbol{\varsigma}) &= \sup\{\langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle - V(\boldsymbol{\xi}):\ \boldsymbol{\xi} \in \mathscr{E}\} \\ &= \begin{cases} 0 & \text{if } \boldsymbol{\sigma} \ge 0,\ \ \boldsymbol{\tau} \ne 0,\ \ \boldsymbol{\mu} \ne 0 \\ +\infty & \text{otherwise} \end{cases} \end{aligned} \tag{31}$$

is also convex, l.s.c. on $\mathscr{E}^*$. By convex analysis, the following generalized canonical duality relations

$$\boldsymbol{\varsigma} \in \partial V(\boldsymbol{\xi}) = \mathscr{E}_a^* \quad \Leftrightarrow \quad \boldsymbol{\xi} \in \partial V^\sharp(\boldsymbol{\varsigma}) = \mathscr{E}_a \quad \Leftrightarrow \quad V(\boldsymbol{\xi}) + V^\sharp(\boldsymbol{\varsigma}) = \langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle \tag{32}$$

hold on $\mathscr{E} \times \mathscr{E}^*$, where

$$\mathscr{E}_a = \{\boldsymbol{\xi} = \{\boldsymbol{\varepsilon}, \boldsymbol{\delta}, \boldsymbol{\rho}\} \in \mathscr{E} |\ \boldsymbol{\varepsilon} \leq 0,\ \ \boldsymbol{\delta} = 0,\ \ \boldsymbol{\rho} = 0\},$$

$$\mathscr{E}_a^* = \{\boldsymbol{\varsigma} = \{\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}\} \in \mathscr{E}^* |\ \ \boldsymbol{\sigma} \geq 0,\ \ \ \boldsymbol{\tau} \neq 0,\ \ \ \boldsymbol{\mu} \neq 0\}$$

are effective domains of V and $V^\sharp$, respectively. The last equality in (32) is equivalent to the following KKT complementarity conditions:

$$\boldsymbol{\varepsilon}^T \boldsymbol{\sigma} = 0,\ \ \boldsymbol{\delta}^T \boldsymbol{\tau} = 0,\ \ \boldsymbol{\rho}^T \boldsymbol{\mu} = 0. \tag{33}$$

Clearly, the condition $\boldsymbol{\mu} \neq 0$ leads to the integer condition $\boldsymbol{\rho} = \{y_i(y_i - 1)\} = 0 \in \mathbb{R}^K$. Let

$$\mathbf{F}(\boldsymbol{\varsigma}) = \mathbf{h} - D^T \boldsymbol{\sigma} - H^T \boldsymbol{\tau} + \boldsymbol{\mu}, \tag{34}$$

$$\mathbf{G}(\boldsymbol{\mu}) = B + 2\text{Diag}\ (\boldsymbol{\mu}). \tag{35}$$

Thus, on $\mathbb{R}^K \times \mathscr{E}_a^*$, the total complementary function $\varXi$ associated with $\varPi(\mathbf{y})$ can be written as

$$\begin{aligned}\varXi(\mathbf{y}, \boldsymbol{\varsigma}) &= \langle \varLambda(\mathbf{y}); \boldsymbol{\varsigma} \rangle - V^\sharp(\boldsymbol{\varsigma}) - U(\mathbf{y}) \\ &= \frac{1}{2} \mathbf{y}^T \mathbf{G}(\boldsymbol{\mu}) \mathbf{y} - \mathbf{F}^T(\boldsymbol{\varsigma}) \mathbf{y} - \boldsymbol{\sigma}^T \mathbf{b} - \boldsymbol{\tau}^T \mathbf{e}_n.\end{aligned}$$

The criticality condition $\nabla_{\mathbf{y}} \varXi(\mathbf{y}, \boldsymbol{\varsigma}) = 0$ leads to the canonical equilibrium equation

$$\mathbf{G}(\boldsymbol{\mu})\mathbf{y} - \mathbf{F}(\boldsymbol{\varsigma}) = 0. \tag{36}$$

Let $\mathscr{S}_a \subset \mathscr{E}_a^*$ be a canonical dual space:

$$\mathscr{S}_a = \{\boldsymbol{\varsigma} = (\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathscr{E}_a^* :\ \ \det \mathbf{G}(\boldsymbol{\varsigma}) \neq 0\ \}. \tag{37}$$

Then on $\mathscr{S}_a$, the canonical dual function can be finally formulated as

$$\begin{aligned}\varPi^d(\boldsymbol{\varsigma}) &= \text{sta}\{\varXi(\mathbf{y}, \boldsymbol{\varsigma})\ \ :\ \ \mathbf{y} \in \mathbb{R}^K\} \\ &= -\frac{1}{2} \mathbf{F}^T(\boldsymbol{\varsigma}) \mathbf{G}^{-1}(\boldsymbol{\mu}) \mathbf{F}(\boldsymbol{\varsigma}) - \boldsymbol{\sigma}^T \mathbf{b} - \boldsymbol{\tau}^T \mathbf{e}_n.\end{aligned} \tag{38}$$

Theorem 3 (**Complementary-Dual Principle**). *If* $\bar{\boldsymbol{\varsigma}} = (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}})$ *is a KKT point of* $\varPi^d(\boldsymbol{\varsigma})$ *on* $\mathscr{S}_a$*, then the vector*

$$\bar{\mathbf{y}}(\bar{\boldsymbol{\varsigma}}) = \mathbf{G}^{-1}(\bar{\boldsymbol{\mu}}) \mathbf{F}(\bar{\boldsymbol{\varsigma}}) \tag{39}$$

is a KKT point of Problem ($\mathscr{P}$) *and*

$$\Pi(\bar{\mathbf{y}}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{40}$$

Proof By introducing the Lagrange multiplier vectors $\boldsymbol{\xi} = \{\boldsymbol{\varepsilon}, \boldsymbol{\delta}, \boldsymbol{\rho}\} \in \mathscr{E}_a$ to relax the inequality constraints[1] in $\mathscr{E}_a^*$, the Lagrangian function associated with the dual function $\Pi^d(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu})$ becomes

$$L(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}, \boldsymbol{\varepsilon}, \boldsymbol{\rho}) = \Pi^d(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}) - \boldsymbol{\varepsilon}^T\boldsymbol{\sigma} - \boldsymbol{\delta}^T\boldsymbol{\tau} - \boldsymbol{\rho}^T\boldsymbol{\mu}.$$

Then, in terms of $\mathbf{y} = G^{-1}(\boldsymbol{\mu})\mathbf{F}(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu})$, the criticality condition $\nabla_{\boldsymbol{\varsigma}} L(\boldsymbol{\varsigma}, \boldsymbol{\xi}) = 0$ leads to

$$\frac{\partial L(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}, \boldsymbol{\varepsilon}, \boldsymbol{\delta}, \boldsymbol{\rho})}{\partial \boldsymbol{\sigma}} = D\mathbf{y} - \mathbf{b} - \boldsymbol{\varepsilon} = 0,$$
$$\frac{\partial L(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}, \boldsymbol{\varepsilon}, \boldsymbol{\delta}, \boldsymbol{\rho})}{\partial \boldsymbol{\tau}} = H\mathbf{y} - \mathbf{e}_n - \boldsymbol{\delta} = 0,$$
$$\frac{\partial L(\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}, \boldsymbol{\varepsilon}, \boldsymbol{\delta}, \boldsymbol{\rho})}{\partial \boldsymbol{\mu}} = \mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K) - \boldsymbol{\rho} = 0,$$

as well as the KKT conditions

$$\boldsymbol{\sigma} \geq 0, \quad \boldsymbol{\varepsilon} \leq 0, \quad \boldsymbol{\sigma}^T\boldsymbol{\varepsilon} = 0, \tag{41}$$
$$\boldsymbol{\tau} \neq 0, \quad \boldsymbol{\delta} = 0, \quad \boldsymbol{\delta}^T\boldsymbol{\tau} = 0. \tag{42}$$
$$\boldsymbol{\mu} \neq 0, \quad \boldsymbol{\rho} = 0, \quad \boldsymbol{\rho}^T\boldsymbol{\mu} = 0. \tag{43}$$

They can be written as:

$$D\mathbf{y} - \mathbf{b} \leq 0, \tag{44}$$
$$H\mathbf{y} - \mathbf{e}_n = 0, \tag{45}$$
$$\mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K) = 0, \tag{46}$$

This proves that if $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}})$ is a KKT point of $\Pi^d(\boldsymbol{\varsigma})$, then the vector

$$\bar{\mathbf{y}}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}}) = \mathbf{G}^{-1}(\bar{\boldsymbol{\mu}})\mathbf{F}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}})$$

is a KKT point of Problem ($\mathscr{P}$).

[1] The inequality $\det \mathbf{G}(\boldsymbol{\varsigma}) \neq 0$ is not a constraint since the Lagrange multiplier for this inequality is identical zero.

Again, by the complementary conditions (41)–(43) and (39), we have

$$\begin{aligned}
\Pi^d(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}}) &= -\frac{1}{2}\mathbf{F}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}})^T \mathbf{G}(\bar{\boldsymbol{\mu}})^{-1}\mathbf{F}(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}}) - \bar{\boldsymbol{\sigma}}^T \mathbf{b} - \bar{\boldsymbol{\tau}}^T \mathbf{e}_n \\
&= \frac{1}{2}\bar{\mathbf{y}}^T B\bar{\mathbf{y}} - \mathbf{h}^T\bar{\mathbf{y}} + \bar{\boldsymbol{\sigma}}^T(D\bar{\mathbf{y}} - \mathbf{b}) + \bar{\boldsymbol{\tau}}^T(H\bar{\mathbf{y}} - \mathbf{e}_n) + \bar{\boldsymbol{\mu}}^T(\bar{\mathbf{y}} \circ (\bar{\mathbf{y}} - \mathbf{e}_K)) \\
&= \frac{1}{2}\bar{\mathbf{y}}^T B\bar{\mathbf{y}} - \mathbf{h}^T\bar{\mathbf{y}} = \Pi(\bar{\mathbf{y}}).
\end{aligned}$$

Therefore, the theorem is proved. □

Theorem 3 shows that the strong duality (40) holds for all KKT points of the primal and dual problems. In continuum mechanics, this theorem solved a 50-year-old problem and is known as the Gao principle [27]. In nonconvex analysis, this theorem can be used for solving a large class of fully nonlinear partial differential equations.

Remark 1. As we have demonstrated that by the generalized canonical duality (32), all KKT conditions can be recovered for both equality and inequality constraints. Generally speaking, the nonzero Lagrange multiplier condition for the linear equality constraint is usually ignored in optimization textbooks. But it can not be ignored for nonlinear constraints. It is proved recently [26] that the popular augmented Lagrange multiplier method can be used mainly for linear constrained problems. Since the inequality constraint $\boldsymbol{\mu} \neq 0$ produces a nonconvex feasible set $\mathscr{E}_a^*$, this constraint can be replaced by either $\boldsymbol{\mu} < 0$ or $\boldsymbol{\mu} > 0$. But the condition $\boldsymbol{\mu} < 0$ is corresponding to $\mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K) \geq 0$, this leads to a nonconvex open feasible set for the primal problem. By the fact that the integer constraints $y_i(y_i - 1) = 0$ are actually a special case (boundary) of the boxed constraints $0 \leq y_i \leq 1$, which is corresponding to $\mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K) \leq 0$, we should have $\boldsymbol{\mu} > 0$ (see [8] and [12, 16]). In this case, the KKT condition (43) should be replaced by

$$\boldsymbol{\mu} > 0, \quad \mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K) \leq 0, \quad \boldsymbol{\mu}^T[\mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K)] = 0. \tag{47}$$

Therefore, as long as $\boldsymbol{\mu} \neq 0$ is satisfied, the complementarity condition in (47) leads to the integer condition $\mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K) = 0$. Similarly, the inequality $\boldsymbol{\tau} \neq 0$ can be replaced by $\boldsymbol{\tau} > 0$.

By this remark, we can introduce a convex subset of the dual feasible space $\mathscr{S}_a$:

$$\mathscr{S}_a^+ = \{\boldsymbol{\varsigma} = (\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathscr{E}^* : \quad \boldsymbol{\sigma} \geq 0, \quad \boldsymbol{\tau} > 0, \quad \boldsymbol{\mu} > 0, \quad \mathbf{G}(\boldsymbol{\mu}) \succ 0\}. \tag{48}$$

Then the canonical dual problem can be eventually proposed as the following

$$(\mathscr{P}^d) \quad \max \left\{ \Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\mathbf{F}^T(\boldsymbol{\varsigma})\mathbf{G}^{-1}(\boldsymbol{\mu})\mathbf{F}(\boldsymbol{\varsigma}) - \boldsymbol{\sigma}^T\mathbf{b} - \boldsymbol{\tau}^T\mathbf{e}_n | \ \ \boldsymbol{\varsigma} \in \mathscr{S}_a^+ \right\}. \tag{49}$$

It is easy to check that $\Pi^d(\boldsymbol{\varsigma})$ is concave on the convex open set $\mathscr{S}_a^+$. Therefore, if $\mathscr{S}_a^+$ is not empty, this canonical dual problem can be solved easily by convex minimization techniques.

Theorem 4 *Assume that $\bar{\boldsymbol{\varsigma}} = (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}})$ is a KKT point of $\Pi^d(\boldsymbol{\varsigma})$ and $\bar{\mathbf{y}} = \mathbf{G}^{-1}(\bar{\boldsymbol{\mu}})\mathbf{F}(\bar{\boldsymbol{\varsigma}})$. If $\bar{\boldsymbol{\varsigma}} \in \mathscr{S}_a^+$, then $\bar{\mathbf{y}}$ is a global minimizer of $\Pi(\mathbf{y})$ and $\bar{\boldsymbol{\varsigma}}$ is a global maximizer of $\Pi^d(\boldsymbol{\varsigma})$ with*

$$\Pi(\bar{\mathbf{y}}) = \min_{\mathbf{y}\in\mathbb{R}^K} \Pi(\mathbf{y}) = \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}) \tag{50}$$

Proof It is easy to check that the total complementary function $\Xi(\mathbf{y}, \boldsymbol{\varsigma})$ is a saddle function on the open set $\mathbb{R}^K \times \mathscr{S}_a^+$, i.e., convex (quadratic) in $\mathbf{y} \in \mathbb{R}^K$ and concave (linear) in $\boldsymbol{\varsigma} \in \mathscr{S}_a^+$. Therefore, if $(\bar{\mathbf{y}}, \bar{\boldsymbol{\varsigma}})$ is a critical point of $\Xi(\mathbf{y}, \boldsymbol{\varsigma})$, we must have

$$\begin{aligned}
\Pi^d(\bar{\boldsymbol{\varsigma}}) &= \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} P^d(\boldsymbol{\varsigma}) = \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \min_{\mathbf{y}\in\mathbb{R}^K} \Xi(\mathbf{y}, \boldsymbol{\varsigma}) = \min_{\mathbf{y}\in\mathbb{R}^K} \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \Xi(\mathbf{y}, \boldsymbol{\varsigma}) \\
&= \min_{\mathbf{y}\in\mathbb{R}^K} \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \left\{ \frac{1}{2}\mathbf{y}^T \mathbf{G}(\boldsymbol{\mu})\mathbf{y} - (\mathbf{h} - D^T\boldsymbol{\sigma} - H^T\boldsymbol{\tau} + \boldsymbol{\mu})^T \mathbf{y} - \boldsymbol{\sigma}^T \mathbf{b} - \boldsymbol{\tau}^T \mathbf{e}_n \right\} \\
&= \min_{\mathbf{y}\in\mathbb{R}^K} \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \left\{ \frac{1}{2}\mathbf{y}^T B\mathbf{y} - \mathbf{h}^T\mathbf{y} + \boldsymbol{\sigma}^T(D\mathbf{y} - \mathbf{b}) + \boldsymbol{\tau}^T(H\mathbf{y} - \mathbf{e}_n) + \boldsymbol{\mu}^T[\mathbf{y} \circ (\mathbf{y} - \mathbf{e}_K)] \right\} \\
&= \min_{\mathbf{y}\in\mathbb{R}^K} \max_{\boldsymbol{\varsigma}\in\mathscr{S}_a^+} \{ \frac{1}{2}\mathbf{y}^T B\mathbf{y} - \mathbf{h}^T\mathbf{y} + \langle \Lambda(\mathbf{y}); \boldsymbol{\varsigma} \rangle \}
\end{aligned} \tag{51}$$

Note that

$$\min_{\boldsymbol{\varsigma}\in\mathscr{E}^*} \{V^\sharp(\boldsymbol{\varsigma})\} = V^\sharp(\bar{\boldsymbol{\varsigma}}) = 0, \quad \min_{\boldsymbol{\xi}\in\mathscr{E}} \{V(\boldsymbol{\xi})\} = V(\bar{\boldsymbol{\xi}}) = 0.$$

Thus, it follows from (51) that

$$\begin{aligned}
\Pi^d(\bar{\boldsymbol{\varsigma}}) &= \min_{\mathbf{y}\in\mathbb{R}^K} \max_{\boldsymbol{\varsigma}\in\mathscr{E}^*} \{ \frac{1}{2}\mathbf{y}^T B\mathbf{y} - \mathbf{h}^T\mathbf{y} + \langle \Lambda(\mathbf{y}); \boldsymbol{\varsigma} \rangle - V^\sharp(\boldsymbol{\varsigma}) \} \\
&= \min_{\mathbf{y}\in\mathbb{R}^K} \{ \frac{1}{2}\mathbf{y}^T B\mathbf{y} - \mathbf{h}^T\mathbf{y} \} + \max_{\boldsymbol{\varsigma}\in\mathscr{E}^*} \{ \langle \Lambda(\mathbf{y}); \boldsymbol{\varsigma} \rangle - V^\sharp(\boldsymbol{\varsigma}) \} \\
&= \min_{\mathbf{y}\in\mathbb{R}^K} \{ \frac{1}{2}\mathbf{y}^T B\mathbf{y} - \mathbf{h}^T\mathbf{y} + V(\Lambda(\mathbf{y})) \} \\
&= \min_{\mathbf{y}\in\mathbb{R}^K} \Pi(\mathbf{y}) = \min_{\mathbf{y}\in\mathscr{Y}} P(\mathbf{y}).
\end{aligned}$$

This completes the proof. □

Remark 2. By the fact that $\mathscr{S}_a^+$ is an open convex set, if the problem $(\mathscr{P})$ has multiple global minimizers, then its canonical dual solutions could be located on the boundary of $\mathscr{S}_a^+$ as illustrated in Sect. 3 and in [12, 31]. In order to solve this problem, we let

$$\mathcal{S}_c^+ = \{\boldsymbol{\varsigma} = (\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathcal{S}_a^+ : \quad \boldsymbol{\mu} \geq 0, \quad \boldsymbol{\tau} \geq 0, \quad \mathbf{G}(\boldsymbol{\mu}) \succeq 0\}.$$

Then on this closed convex domain, the relaxed concave maximization problem

$$(\mathcal{P}^\sharp) \quad \max\{\Pi^d(\boldsymbol{\varsigma}) : \quad \boldsymbol{\varsigma} \in \mathcal{S}_c^+\} \tag{52}$$

has at least one solution $\bar{\boldsymbol{\varsigma}} = (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}})$. If the corresponding $\bar{\mathbf{y}} = \mathbf{G}^{-1}(\bar{\boldsymbol{\mu}})\mathbf{F}(\bar{\boldsymbol{\varsigma}})$ is feasible, then $\bar{\mathbf{y}}$ is a global minimizer of the primal problem $(\mathcal{P})$. If $\mathbf{G}(\bar{\boldsymbol{\mu}})$ is singular, than $\mathbf{G}^{-1}(\bar{\boldsymbol{\mu}})$ can be replaced by the Moore–Penrose generalized inverse $\mathbf{G}^\dagger$ (see [31]). Otherwise, the relaxed canonical dual $(\mathcal{P}^\sharp)$ provides a lower bound approach to the primal problem $(\mathcal{P})$, i.e.,

$$\min_{\mathbf{y} \in \mathcal{Y}} P(\mathbf{y}) \geq \max_{\boldsymbol{\varsigma} \in \mathcal{S}_c^+} \Pi^d(\boldsymbol{\varsigma}).$$

This is one of the main advantages of the canonical duality theory.

5 Canonical Perturbation Method

In fact, Problem $(\mathcal{P}^d)$ can be rewritten as a convex minimization problem:

$$\begin{aligned} \min \quad & \frac{1}{2}\mathbf{F}^T(\boldsymbol{\varsigma})\mathbf{G}^{-1}(\boldsymbol{\mu})\mathbf{F}(\boldsymbol{\varsigma}) + \boldsymbol{\sigma}^T\mathbf{b} + \boldsymbol{\tau}^T\mathbf{e}_n, \\ \text{s.t.} \quad & \boldsymbol{\varsigma} \in \mathcal{S}_a^+. \end{aligned}$$

If the primal problem has a unique global minimal solution, this canonical dual problem may have a unique critical point in $\mathcal{S}_a^+$ which can be obtained easily by well-developed nonlinear minimization techniques. Otherwise, the canonical dual function $\Pi^d(\boldsymbol{\varsigma})$ may have critical point $\bar{\boldsymbol{\varsigma}}$ located on the boundary of $\mathcal{S}_a^+$, where the matrix $\mathbf{G}(\boldsymbol{\mu})$ is singular. In order to handle this issue, $(\mathcal{P}^d)$ can be relaxed to a semi-definite programming problem:

$$\begin{aligned} \min \quad & g + \boldsymbol{\sigma}^T\mathbf{b} + \boldsymbol{\tau}^T\mathbf{e}_n, \\ \text{s.t.} \quad & g \geq \frac{1}{2}\mathbf{F}^T(\boldsymbol{\varsigma})\mathbf{G}^\dagger(\boldsymbol{\mu})\mathbf{F}(\boldsymbol{\varsigma}), && (53) \\ & \mathbf{G}(\boldsymbol{\mu}) \succeq 0, && (54) \\ & \boldsymbol{\varsigma} \in \mathcal{E}^*, \quad \boldsymbol{\sigma} \geq 0, \quad \boldsymbol{\mu} > 0, && (55) \end{aligned}$$

where the parameter g is actually the Gao-Strang pure complementary gap function [19], and $\mathbf{G}^\dagger$ represents the Moore–Penrose generalized inverse of $\mathbf{G}$. Since $\boldsymbol{\tau}$ is a Lagrange multiplier for the linear equality $H\mathbf{y} = \mathbf{e}_n$, the condition $\boldsymbol{\tau} \neq 0$ can be ignored in this section as long as the final solution $\mathbf{y}$ is feasible.

Lemma 1 (**Schur Complementary Lemma**). *Let*

$$A = \begin{bmatrix} B & C^T \\ C & D \end{bmatrix},$$

If $B \succ 0$, then A is positive (semi) definite if and only if the matrix $D - CB^{-1}C^T$ is positive (semi) definite. If $B \succeq 0$, then, A is positive semi-definite if and only if the matrix $D - CB^{-1}C^T$ is positive semi-definite and $(I - BB^{-1})C = 0$.

According to Lemma 1, (53) is equivalent to

$$\begin{bmatrix} \mathbf{G}(\boldsymbol{\mu}) & \mathbf{F}(\boldsymbol{\varsigma}) \\ \mathbf{F}^T(\boldsymbol{\varsigma}) & 2g \end{bmatrix} \succeq 0.$$

Thus, the canonical dual problem $(\mathscr{P}^d)$ can be further relaxed to the following stardard semi-definite problem (SDP):

$$\begin{aligned} \min \quad & g + \boldsymbol{\sigma}^T \mathbf{b} + \boldsymbol{\tau}^T \mathbf{e}_n, \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{G}(\boldsymbol{\mu}) & \mathbf{F}(\boldsymbol{\varsigma}) \\ \mathbf{F}^T(\boldsymbol{\varsigma}) & 2g \end{bmatrix} \succeq 0, \quad \mathbf{G}(\boldsymbol{\mu}) \succeq 0, \\ & \boldsymbol{\varsigma} \in \mathscr{E}^*, \quad \boldsymbol{\sigma} \geq 0, \quad \boldsymbol{\mu} > 0. \end{aligned}$$

Although the SDP relaxation can be used theoretically to solve the canonical dual problem for the case that Π^d has critical points on the boundary $\partial \mathscr{S}_a^+$, in practice, the matrix $\mathbf{G}(\boldsymbol{\mu})$ will be ill-conditioning when the dual solution approaches to $\partial \mathscr{S}_a^+$. In order to solve this type of challenging problems, a canonical perturbation method has been suggested [16, 32]. Let

$$\begin{aligned} \Xi_{\delta_k}(\mathbf{y}, \boldsymbol{\varsigma}) &= \Xi(\mathbf{y}, \boldsymbol{\varsigma}) + \frac{\delta_k}{2} \|\mathbf{y} - \mathbf{y}_k\|^2 \\ &= \frac{1}{2} \mathbf{y}^T \mathbf{G}_{\delta_k}(\boldsymbol{\mu}) \mathbf{y} - \mathbf{F}_{\delta_k}^T(\boldsymbol{\varsigma}) \mathbf{y} - \boldsymbol{\sigma}^T \mathbf{b} - \boldsymbol{\tau}^T \mathbf{e}_n + \frac{\delta_k}{2} \mathbf{y}_k^T \mathbf{y}_k, \end{aligned}$$

where, $\{\delta_k\}$ is a bounded sequence of positive real numbers, $\{\mathbf{y}_k\} \in \mathbb{R}^K$ is a set of given vectors, $\mathbf{G}_{\delta_k}(\boldsymbol{\mu}) = \mathbf{G}(\boldsymbol{\mu}) + \delta_k I$, $\mathbf{F}_{\delta_k}^T(\boldsymbol{\varsigma}) = \mathbf{F}^T(\boldsymbol{\varsigma}) + \delta_k \mathbf{y}_k$. Let

$$\mathscr{S}_{\delta_k}^+ = \{\boldsymbol{\varsigma} \in \mathscr{S}_a : \ \mathbf{G}_{\delta_k}(\boldsymbol{\mu}) \geq 0\}.$$

Clearly, we have $\mathscr{S}_a^+ \subset \mathscr{S}_{\delta_k}^+$. Therefore, the perturbed canonical dual problem can be expressed as

$$\begin{aligned} (\mathscr{P}_{\delta_k}^d) \quad \max \quad & \Pi_{\delta_k}^d(\boldsymbol{\varsigma}) = -\frac{1}{2} \mathbf{F}_{\delta_k}^T(\boldsymbol{\varsigma}) \mathbf{G}_{\delta_k}^{\dagger}(\boldsymbol{\mu}) \mathbf{F}_{\delta_k}(\boldsymbol{\varsigma}) - \boldsymbol{\sigma}^T \mathbf{b} - \boldsymbol{\tau}^T \mathbf{e}_n, \\ \text{s.t.} \quad & \boldsymbol{\varsigma} \in \mathscr{S}_{\delta_k}^+. \end{aligned}$$

Based on this perturbed problem, the following canonical primal-dual algorithm can be proposed for solving the nonconvex problem $(\mathscr{P})$.

Algorithm 1 (**Canonical Primal-Dual Algorithm**)
Given initial data $\delta_0 > 0$, $\mathbf{y}_0 \in \mathbb{R}^K$, and error allowance $\varepsilon > 0$, let $k = 0$.

1. Solve the perturbed canonical dual problem $(\mathscr{P}^d_{\delta_k})$ to obtain $\boldsymbol{\varsigma}_k \in \mathcal{S}^+_{\delta_k}$.

2. Compute $\tilde{\mathbf{y}}_{k+1} = [\mathbf{G}_{\delta_k}(\boldsymbol{\varsigma}_k)]^\dagger \mathbf{F}_{\delta_k}(\boldsymbol{\varsigma}_k)$ and let
$\mathbf{y}_{k+1} = \mathbf{y}_k + \beta_k(\tilde{\mathbf{y}}_{k+1} - \mathbf{y}_k), \quad \beta_k \in [0, 1]$.

3. If $|P(\mathbf{y}_{k+1} - P(\mathbf{y}_k)| \leq \varepsilon$, then stop, $\mathbf{y}_{k+1}$ is the optimal solution. Otherwise, let $k = k + 1$, go back to step 1.

In this algorithm, $\{\beta_k\} \in [0, 1]$ are given parameters, which change the search directions. Clearly, if $\beta_k = 1$, we have $\mathbf{y}_{k+1} = \tilde{\mathbf{y}}_{k+1}$.

The key step in this algorithm is to solve the perturbed canonical dual problem $(\mathscr{P}^d_{\delta_k})$, which is equivalent to

$$\begin{aligned} \min \quad & g + \boldsymbol{\sigma}^T \mathbf{b} + \boldsymbol{\tau}^T \mathbf{e}_n, \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{G}(\boldsymbol{\mu}) + \delta_k I & \mathbf{F}(\boldsymbol{\varsigma}) + \delta_k \mathbf{y}_k \\ \mathbf{F}^T(\boldsymbol{\varsigma}) + \delta_k \mathbf{y}_k & 2g \end{bmatrix} \succeq 0, \\ & \mathbf{G}(\boldsymbol{\mu}) \succeq 0, \\ & \boldsymbol{\varsigma} \in \mathscr{S}, \quad \boldsymbol{\sigma} \geq 0, \quad \boldsymbol{\mu} > 0. \end{aligned}$$

This problem can be solved by a well-known software package named SeDuMi [36].

6 Numerical Experience

All data and computational results presented in this section are produced by Matlab. In order to save space and fit the matrix in the paper, we round our these results up to two decimals.

Example 1. 5-dimensional problem.
Consider Problem $(\mathscr{P}_a)$ with $\mathbf{x}=[x_1, \cdots, x_5]^T$, while $x_i \in \{2, 3, 5\}, i=1, \cdots, 5$,

$$Q = \begin{bmatrix} 3.43 & 0.60 & 0.39 & 0.10 & 0.60 \\ 0.60 & 2.76 & 0.32 & 0.65 & 0.49 \\ 0.39 & 0.32 & 2.07 & 0.59 & 0.39 \\ 0.10 & 0.65 & 0.59 & 2.62 & 0.30 \\ 0.60 & 0.49 & 0.39 & 0.30 & 3.34 \end{bmatrix},$$

$$\mathbf{c} = [38.97, -24.17, 40.39, -9.65, 13.20]^T,$$

$$\mathbf{A} = \begin{bmatrix} 0.94 & 0.23 & 0.04 & 0.65 & 0.74 \\ 0.96 & 0.35 & 0.17 & 0.45 & 0.19 \\ 0.58 & 0.82 & 0.65 & 0.55 & 0.69 \\ 0.06 & 0.02 & 0.73 & 0.30 & 0.18 \end{bmatrix},$$

$$\mathbf{b} = [11.49, 9.32, 14.43, 5.66]^T.$$

Under the transformation (3), this problem is transformed into the 0–1 programming Problem ($\mathscr{P}$), where

$$\mathbf{y} = [y_{1,1}, y_{1,2}, y_{1,3}, \cdots, y_{5,1}, y_{5,1}, y_{5,3}]^T \in \mathbb{R}^{15},$$

$$B = \begin{bmatrix} 13.71 & 20.56 & 34.27 & 2.40 & 3.61 & 6.01 & 1.58 & 2.37 & 3.95 & 0.39 & 0.58 & 0.97 & 2.38 & 3.57 & 5.95 \\ 20.56 & 30.84 & 51.41 & 3.61 & 5.41 & 9.01 & 2.37 & 3.55 & 5.92 & 0.58 & 0.88 & 1.46 & 3.57 & 5.36 & 8.93 \\ 34.27 & 51.41 & 85.68 & 6.01 & 9.01 & 15.02 & 3.95 & 5.92 & 9.87 & 0.97 & 1.46 & 2.43 & 5,95 & 8.93 & 14.88 \\ 2.40 & 3.61 & 6.01 & 11.05 & 16.57 & 27.61 & 1.27 & 1.91 & 3.18 & 2.61 & 3.91 & 6.52 & 1.95 & 2.93 & 4.88 \\ 3.61 & 5.41 & 9.01 & 16.57 & 24.85 & 41.42 & 1.91 & 2.86 & 4.77 & 3.91 & 5.87 & 9.78 & 2.93 & 4.39 & 7.32 \\ 6.01 & 9.01 & 15.02 & 27.61 & 41.42 & 69.03 & 3.18 & 4.77 & 7.96 & 6.52 & 9.78 & 16.31 & 4.88 & 7.32 & 12.20 \\ 1.58 & 2.37 & 3.95 & 1.27 & 1.91 & 3.18 & 8.27 & 12.40 & 20.67 & 2.37 & 3.55 & 5.92 & 1.57 & 2.36 & 3.93 \\ 2.37 & 3.55 & 5.92 & 1.91 & 2.86 & 4.77 & 12.40 & 18.60 & 31.00 & 3.55 & 5.33 & 8.89 & 2.36 & 3.53 & 5.90 \\ 3.95 & 5.92 & 9.87 & 3.18 & 4.77 & 7.96 & 20.67 & 31.00 & 51.67 & 5.92 & 8.86 & 14.81 & 3.93 & 5.90 & 9.83 \\ 0.39 & 5.58 & 0.97 & 2.61 & 3.91 & 6.52 & 2.37 & 3.55 & 5.92 & 10.50 & 15.74 & 26.24 & 1.20 & 1.80 & 3.00 \\ 0.58 & 0.88 & 1.46 & 3.91 & 5.87 & 9.78 & 3.55 & 5.33 & 8.89 & 15.74 & 23.62 & 39.36 & 1.80 & 2.70 & 4.50 \\ 0.97 & 1.46 & 2.43 & 6.52 & 9.78 & 16.31 & 5.92 & 8.89 & 14.81 & 26.24 & 39.36 & 65.60 & 3.00 & 4.50 & 7.51 \\ 2.38 & 3.57 & 5.95 & 1.95 & 2.93 & 4.88 & 1.57 & 2.36 & 3.93 & 1.20 & 1.80 & 3.00 & 13.35 & 20.02 & 33.37 \\ 3.57 & 5.36 & 8.93 & 2.93 & 4.39 & 7.32 & 2.36 & 3.54 & 5.90 & 1.80 & 2.70 & 4.50 & 20.02 & 30.04 & 50.06 \\ 5.95 & 8.93 & 14.88 & 4.88 & 7.32 & 12.20 & 3.93 & 5.90 & 9.83 & 3.00 & 4.50 & 7.51 & 33.37 & 50.06 & 83.43 \end{bmatrix},$$

$$\begin{aligned} \mathbf{h} = [&77.95, 116.92, 194.87, -48.34, -72.51, -120.85, 80.78, 121.17 \\ &201.96, -19.29, -28.94, -48.23, 26.39, 39.59, 65.99]^T, \end{aligned}$$

$$D = \begin{bmatrix} 1.88 & 2.83 & 4.71 & 0.47 & 0.70 & 1.17 & 0.09 & 0.12 & 0.22 & 1.30 & 1.94 & 3.24 & 1.49 & 2.23 & 3.72 \\ 1.91 & 2.87 & 4.78 & 0.71 & 1.06 & 1.77 & 0.34 & 0.51 & 0.85 & 0.90 & 1.35 & 2.25 & 0.38 & 0.57 & 0.94 \\ 1.15 & 1.72 & 2.88 & 1.64 & 2.46 & 4.11 & 1.30 & 1.95 & 3.25 & 1.09 & 1.64 & 2.74 & 1.37 & 2.06 & 3.43 \\ 0.12 & 0.18 & 0.30 & 0.03 & 0.05 & 0.08 & 1.46 & 2.20 & 3.66 & 0.59 & 0.89 & 1.48 & 0.37 & 0.55 & 0.92 \end{bmatrix},$$

$$H = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{5\times 15}.$$

The canonical dual problem can be stated as follows:

$$\begin{aligned} (\mathscr{P}^d)\text{Maximize } \ & \Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\mathbf{F}(\boldsymbol{\varsigma})^T\mathbf{G}^{\dagger}(\boldsymbol{\mu})\mathbf{F}(\boldsymbol{\varsigma}) - \boldsymbol{\sigma}^T\mathbf{b} - \boldsymbol{\tau}^T\mathbf{e}_5 \\ \text{subject to } \ & \boldsymbol{\varsigma} = (\boldsymbol{\sigma}, \boldsymbol{\tau}, \boldsymbol{\mu}) \in \mathbb{R}^{4+5+15}, \ \ \boldsymbol{\sigma} \geq 0, \boldsymbol{\mu} > 0. \end{aligned}$$

By solving this dual problem with the sequential quadratic programming method in the optimization Toolbox within the Matlab environment, we obtain

$$\bar{\boldsymbol{\sigma}} = [0, 0, 0, 0]^T,$$

$$\bar{\boldsymbol{\tau}} = [73.90, -106.70, 111.95, -59.27, -0.01]^T,$$

and

$$\bar{\boldsymbol{\mu}} = [39.34, 22.07, 12.49, 33.56, 3.01, 76.14, 61.00, 35.52 \\ 18.78, 1.47, 41.96, 0.001, 0.001, 0.006]^T.$$

It is clear that $\bar{\boldsymbol{\varsigma}} = (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}}) \in \mathcal{S}_a^+$. Thus, from Theorem 4,

$$\begin{aligned}\bar{\mathbf{y}} &= (B + 2\text{Diag}\,(\bar{\boldsymbol{\mu}}))^{\dagger}(\mathbf{h} - D^T\bar{\boldsymbol{\sigma}} - H^T\bar{\boldsymbol{\tau}} + \bar{\boldsymbol{\mu}}) \\ &= [0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0]^T\end{aligned}$$

is the global minimizer of Problem ($\mathscr{P}$) with $\Pi^d(\bar{\boldsymbol{\varsigma}}) = -227.87 = \Pi(\bar{\mathbf{y}})$. The solution to the original primal problem can be calculated by using the transformation

$$\bar{x}_i = \sum_{j=1}^{K_i} u_{i,j}\bar{y}_{i,j}, \;\; i = 1, 2, 3, 4, 5,$$

to give

$$\bar{\mathbf{x}} = [5, 2, 5, 2, 2]^T$$

with $P(\bar{\mathbf{x}}) = -227.87$.

Example 2. 10-dimensional problem. Consider Problem ($\mathscr{P}_a$), with $\mathbf{x} = [x_1, \cdots, x_{10}]^T$, while $x_i \in \{1, 2, 4, 7, 9\}, \;\; i = 1, \cdots, 10$,

$$Q = \begin{bmatrix} 6.17 & 0.62 & 0.46 & 0.37 & 0.56 & 0.66 & 0.67 & 0.85 & 0.57 & 0.44 \\ 0.62 & 5.63 & 0.29 & 0.56 & 0.79 & 0.29 & 0.43 & 0.69 & 0.49 & 0.39 \\ 0.46 & 0.29 & 5.81 & 0.55 & 0.22 & 0.55 & 0.36 & 0.27 & 0.51 & 0.91 \\ 0.37 & 0.56 & 0.55 & 6.10 & 0.28 & 0.42 & 0.44 & 0.34 & 0.75 & 0.44 \\ 0.56 & 0.79 & 0.22 & 0.28 & 4.75 & 0.40 & 0.55 & 0.42 & 0.49 & 0.44 \\ 0.66 & 0.29 & 0.55 & 0.42 & 0.40 & 5.71 & 0.32 & 0.57 & 0.65 & 0.70 \\ 0.67 & 0.43 & 0.36 & 0.44 & 0.55 & 0.32 & 5.27 & 0.56 & 0.37 & 0.85 \\ 0.85 & 0.69 & 0.27 & 0.34 & 0.42 & 0.57 & 0.56 & 5.91 & 0.15 & 0.62 \\ 0.57 & 0.49 & 0.51 & 0.75 & 0.49 & 0.65 & 0.37 & 0.15 & 4.51 & 0.46 \\ 0.44 & 0.39 & 0.91 & 0.44 & 0.44 & 0.70 & 0.85 & 0.62 & 0.46 & 5.73 \end{bmatrix},$$

$$\mathbf{f} = [0.89, 0.03, 0.49, 0.17, 0.98, 0.71, 0.50, 0.47, 0.06, 0.68]^T,$$

$$\mathbf{A} = \begin{bmatrix} 0.04 & 0.82 & 0.97 & 0.83 & 0.83 & 0.42 & 0.02 & 0.20 & 0.05 & 0.94 \\ 0.07 & 0.72 & 0.65 & 0.08 & 0.80 & 0.66 & 0.98 & 0.49 & 0.74 & 0.42 \\ 0.52 & 0.15 & 0.80 & 0.13 & 0.06 & 0.63 & 0.17 & 0.34 & 0.27 & 0.98 \\ 0.10 & 0.66 & 0.45 & 0.17 & 0.40 & 0.29 & 0.11 & 0.95 & 0.42 & 0.30 \\ 0.82 & 0.52 & 0.43 & 0.39 & 0.53 & 0.43 & 0.37 & 0.92 & 0.55 & 0.70 \end{bmatrix},$$

$$\mathbf{b} = [33.76, 37.07, 26.75, 25.46, 37.36]^T.$$

By solving the canonical dual problem of Problem ($\mathscr{P}_a$), we obtain

$$\bar{\boldsymbol{\sigma}} = [0, 0, 0, 0, 0]^T,$$

$$\bar{\boldsymbol{\tau}} = [-19.99, -20.12, -18.13, -18.37, -14.32, \\ -17.13, -18.46, -19.73, -17.65, -16.55]^T,$$

and

$$\bar{\boldsymbol{\mu}} = [9.51, 0.97, 21.93, 53.36, 74.34, 9.95, 0.21, 20.53, 51.01, 71.35 \\ 8.68, 0.77, 19.68, 48.03, 66.94, 8.30, 1.77, 21.91, 52.13, 72.27 \\ 6.40, 1.54, 17.39, 41.19, 57.04, 7.57, 1.98, 21.10, 49.77, 68.90 \\ 9.15, 0.16, 18.79, 46.72, 65.34, 9.82, 0.09, 19.90, 49.63, 69.45 \\ 8.76, 0.13, 17.92, 44.60, 62.39, 6.26, 4.03, 24.60, 55.48, 76.04]^T,$$

It is clear that $\bar{\boldsymbol{\varsigma}} = (\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\mu}}) \in \mathcal{S}_a^+$. Therefore,

$$\bar{\mathbf{y}} = [1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, \\ 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0]^T$$

is the global minimizer of the problem ($\mathscr{P}$) with $\Pi^d(\bar{\boldsymbol{\varsigma}}) = 45.54 = \Pi(\bar{\mathbf{y}})$. The solution to the original primal problem is

$$\bar{\mathbf{x}} = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1]^T$$

with $P(\bar{\mathbf{x}}) = 45.54$.

Example 3. Relatively large size problems.

Consider Problem ($\mathscr{P}_a$) with $n = 20, 50, 100, 200$, and 300. Let these five problems be referred to as Problem (1), $\cdots$, Problem (5), respectively. Their coefficients are generated randomly with uniform distribution. For each problem, $q_{ij} \in (0, 1)$,

Table 1 Numerical results for large scale integer programming problems

n	m	CPU Time (s)
20	5	1.77
50	5	6.23
100	5	26.05
200	5	136.29
300	5	408.59

$a_{ij} \in (0, 1)$, for $i = 1, \cdots, n$; $j = 1, \cdots, n$, and $c_i \in (0, 1)$, $x_i \in \{1, 2, 3, 4, 5\}$, for $i = 1, \cdots n$. Without loss of generality, we ensure that the constructed Q is a symmetric matrix. Otherwise, we let $Q = \frac{Q+Q^T}{2}$. Furthermore, let Q be diagonally dominated. For each x_i, its lower bound is $l_i = 1$, and its upper bound is $u_i = 5$. Let $l = [l_1, \cdots, l_n]^T$ and $u = [u_1, \cdots, u_n]^T$. The right-hand sides of the linear constraints are chosen such that the feasibility of the test problem is satisfied. More specifically, we set $\mathbf{b} = \sum_j a_{ij} l_j + 0.5 \cdot (\sum_j a_{ij} u_j - \sum_j a_{ij} l_j)$.

We then construct the canonical problem of each of the five problems. It is solved by using the sequential quadratic programming method with active set strategy from the Optimization Toolbox within the Matlab environment. The specifications of the personal notebook computer used are: Window 7 Enterprise, Intel(R), Core(TM)(2.50 GHZ). Table 1 presents the numerical results, where m is number of linear constraints in Problem $I(\mathscr{P}_a)$.

From Table 1, we see that the algorithm based on the canonical dual method can solve large scale problems with reasonable computational time. Furthermore, for each of the five problems, the solution obtained is a global optimal solution. For the case of $n = 300$, the equivalent problem in the form of Problem $(\mathscr{P}_b)$ has 1500 variables. For such a problem, there are 2^{1500} possible combinations.

7 Conclusion

We have presented a canonical duality approach for solving a general quadratic discrete value selection problem with linear constraints. Our results show that this NP-hard problem can be converted to a continuous concave dual maximization problem over a convex space without duality gap. For certain given data, if this canonical dual has a KKT point in the dual feasible space $\mathscr{S}_a^+$, the problem can be solved easily by well-developed convex optimization methods. Otherwise, a canonical perturbation method is proposed, which can be used to deal with challenging cases when the primal problem has multiple global minimizers. Several examples, including some relatively large scale ones, were solved effectively by using the method proposed.

Remanning open problems include how to solve the canonical dual problem $(\mathscr{P}^d)$ more efficiently instead of using the SDP approximation. Also, for the given data Q, $\mathbf{c}$, $\mathbf{A}$, $\mathbf{b}$, the existence condition for the canonical dual problem having KKT point in $\mathscr{S}_a^+$ is fundamentally important for understanding NP-hard problems. If the canonical dual $(\mathscr{P}^d)$ has no KKT point in the closed set $\mathscr{S}_c^+ = \mathscr{S}_a^+ \cup \partial\mathscr{S}_a^+$, the primal problem is equivalent to the following canonical dual problem (see Eq. (67) in [16])

$$\min \text{sta}\{\Pi^d(\boldsymbol{\varsigma}) |\ \boldsymbol{\varsigma} \in \mathscr{S}_a\}, \tag{56}$$

i.e., to find the minimal stationary value of Π^d on $\mathscr{S}_a$. Since the feasible set $\mathscr{S}_a$ is nonconvex, to solve this canonical dual problem is very difficult. Therefore, it is a conjecture that the primal problem $(\mathscr{P})$ could be NP-hard if its canonical dual $(\mathscr{P}^d)$ has no KKT point in the closed set $\mathcal{S}_a^+$ [12]. In this case, one alternative approach for solving $(\mathscr{P})$ is the canonical dual relaxation $(\mathscr{P}^\sharp)$. Although the relaxed problem $(\mathscr{P}^\sharp)$ is convex, by Remark 2 we know that there exists a duality gap between the primal problem $(\mathscr{P})$. It turns out that the associated SDP method provides only a lower bound approach for solving the primal problem. Further researches are needed to know how big is this duality gap, how much does this relaxation lose, and how to solve the nonconvex canonical dual problem (56).

Acknowledgements The research is supported by US Air Force Office of Scientific Research under grants FA2386-16-1-4082 and FA9550-17-1-0151.

References

1. Arora, J.S., Huang, M.W., Hsieh, C.C.: Methods for optimization of nonlinear problems with discrete variables: a review. Struct. Optim. **8**, 69–85 (1994)
2. Balas, E., Ceria, S., Cornuéjols, G.: A lift-and-project cutting plane algorithm for mixed 0–1 programs. Math. Program. **58**, 295–324 (1993)
3. Belmiloudi, A.: Stabilization, Optimal and Robust Control: Theory and Applications in Biological and Physical Science. Springer, London (2008)
4. Bertsimas, D., Shioda, R.: Classification and regression via integer optimization. Oper. Res. **55**(2), 252–271 (2007)
5. Chen, D.S., Batson, R.G., Dang, Y.: Applied Integer Programming: Modeling and Solution. Wiley, New Jersey (2010)
6. Ciarlet, P.G.: Linear and Nonlinear Functional Analysis with Applications. SIAM, Philadelphia (2013)
7. Marsden, J.E., Hughes, T.J.R.: Mathematical Foundations of Elasticity. Prentice-Hall, Upper Saddle River (1983)
8. Fang, S.C., Gao, D.Y., Sheu, R.L., Wu, S.Y.: Canonical dual approach to solving 0–1 quadratic programming problems. J. Ind. Manag. Optim. **4**(4), 125–142 (2008)
9. Floudas, C.A.: Nonlinear and Mixed-Integer Optimization: Theory Methods and Applications. Kluwer Academic Publishers, Dordrecht (2000)
10. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications, p. 454. Kluwer Academic Publishers, Dordrecht (2000)
11. Gao, D.Y.: Nonconvex semi-linear problems and canonical duality solutions, In: Advances in Mechanics and Mathematics, vol. II, pp. 261–312. Kluwer Academic Publishers, Dordrecht (2003)
12. Gao, D.Y.: Solutions and optimality criteria to box constrained nonconvex minimization problem. J. Ind. Manag. Optim. **3**(2), 293–304 (2007)

13. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solution for global optimization problems. Comput. Chem. Eng. **33**, 1964–1972 (2009)
14. Gao, D.Y., Ogden, R.W.: Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation. Q. J. Mech. Appl. Math **61**(4), 497–522 (2008)
15. Gao, D.Y., Ruan, N.: Solutions and optimality criteria for nonconvex quadratic-exponential minimization problem. Math. Method. Oper. Res. **67**, 479–496 (2008)
16. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. **47**, 463–484 (2010)
17. 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: Boginski, V.L., Commander, C.W., Pardalos, P.M., Ye, Y.Y. (eds.) Sensors: Theory, Algorithms and Applications, vol. 61, pp. 37–54. Springer, Berlin (2012)
18. Gao, D.Y., Ruan, N., Sherali, H.D.: Solutions and optimality criteria for nonconvex constrained global optimization problems. J. Glob. Optim. **45**, 473–497 (2009)
19. Gao, D.Y., Sherali, H.D.: Canonical duality: connection between nonconvex mechanics and global optimization. In: Advances in Application Mathematics and Global Optimization, pp. 249–316. Springer, Berlin (2009)
20. Gao, D.Y., Watson, L.T., Easterling, D.R., Thacker, W.I., Billups, S.C.: Solving the canonical dual of box- and integer-constrained nonconvex quadratic programs via a deterministic direct search algorithm. Optim. Methods Softw. **28**(2), 313–326 (2013)
21. Ghaddar, B., Vera, J.C., Anjos, M.: Second-order cone relaxations for binary quadratic polynomial programs. SIAM J. Optim. **21**, 391–414 (2011)
22. Holland, J.H.: Adaptation in Natural and Artificial System. The University of Michigan Press, Ann, Arbor, MI (1975)
23. Huang, M.W., Arora, J.S.: Optimal design with discrete variables: Some numerical experiments. INT. J. Numer. Meth. Eng. **40**, 165–188 (1997)
24. Karlof, J.K.: Integer Programming: Theory and Practice. CRC Press, Boca Raton (2006)
25. Kincaid, R.K., Padula, S.L.: Minimizing distortion and internal forces in truss structures by simulated annealing. In: Proceedings of the 31st AIAA SDM Conference, Long Beach, CA, pp. 327–333 (1990)
26. Latorrie, V., Gao, D.Y.: Canonical duality for solving general nonconvex constrained problems. Optim. Lett. **10**(8), 1–17 (2015)
27. Li, S.F., Gupta, A.: On dual configuration forces. J. Elast. **84**, 13–31 (2006)
28. Loh, H.T., Papalambros, P.Y.: Computational implementation and tests of a sequential linearization approach for solving mixed-discrete nonlinear design optimization. J. Mech. Des. ASME **113**, 335–345 (1991)
29. Marsten, R.E., Morin, T.L.: A hybrid approach to discrete mathematical programming. Math. Program. **14**, 21–40 (1978)
30. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
31. Ruan, N., Gao, D.Y.: Global optimal solutions to a general sensor network localization problem. Perform. Eval. **75–76**, 1–16 (2014)
32. Ruan, N., Gao, D.Y., Jiao, Y.: Canonical dual least square method for solving general nonlinear systems of equations. Comput. Optim. Appl. **47**, 335–347 (2010)
33. Sandgren, E.: Nonlinear integer and discreter programming in mechanical design optimizatin. J. Mech. Des. ASME. **112**, 223–229 (1990)
34. Santos, H.A.F.A., Moitinho De Almeida, J.P.: Dual extremum principles for geometrically exact finite strain beams. Int/ J. Nonlinear Mech. Elsevier **46**(1), 151 (2010)
35. Sherali, H.D., Smith, J.C.: An improved linearization strategy for zero-one quadratic programming problems. Optim. Lett. **1**(1), 33–47 (2007)
36. Sturn, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cone. Optim. Meth. Softw. **11**, 625–653 (1999)
37. Thanedar, P.B., Vanderplaats, G.N.: A survey of discrete variable optimization for structural design. J. Struct. Eng. ASCE **121**, 301–306 (1994)

38. Wang, S., Teo, K.L., Lee, H.W.J.: A new approach to nonlinear mixed discrete programming problems. Eng. Optim. **30**(3), 249–262 (1998)
39. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Global extremal conditions for multi-integer quadratic programming. J. Ind. Manag. Optim. **4**(2), 213–225 (2008)
40. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Canonical dual approach to solving the maximum cut problem. J. Glob. Optim. **54**(2), 341–351 (2012)
41. Zecchin, A.C., Simpson, A.R., Maier, H.R., Nixon, J.B.: Parametric study for an ant algorithm applied to water distribution system optimization. IEEE Trans. Evol. Comput. **9**(2), 175–191 (2005)

Global Optimization Solutions to a Class of Nonconvex Quadratic Minimization Problems with Quadratic Constraints

Yu Bo Yuan

Abstract This paper studies the nonconvex quadratic minimization problems with quadratic constraints (called it as $\mathscr{P}_{qq}$). These problems are from computational science, machine learning, data mining, pattern recognition, computational mechanics, and so on. When the quadratical matrix in the objective function is non-definition, it is very difficult to get the global optimization solutions. There is a very powerful method proposed by David Gao and it is called as *canonical duality*. It can help to convert $\mathscr{P}_{qq}$ into a concave maximization dual problem over a convex set. In this work, we employ it to deal with a special class of $\mathscr{P}_{qq}$. The canonical duality problems are formulated and the equation between optimization solution of $\mathscr{P}_{qq}$ and canonical duality problem is presented in Theorem 1. Two conditions are given in Theorem 2. Under these conditions, we can prove that the canonical duality problem has a unique nonzero solution in the dual space. An algorithm is proposed to find out the global optimization solutions. Several examples are illustrated to show that the conditions are active and the proposed method is effective.

1 Introduction

In recent years, nonconvex quadratic minimization problems with quadratic constraints have attracted more and more attentions. The problems are arising from applications in diverse fields such as computational science, machine learning, data mining, pattern recognition, computational mechanics, and so on.

In 2012, Feng, Lin, Sheu, and Xia [7] had studied the (nonconvex) quadratic minimization problem with one quadratic constraint(QP1QC). They showed that under given assumption, the nonconvex (QP1QC) problem could be solved through a dual approach with no duality gap. In 2014, Fabian and Gabriel [6] had considered quadratic minimization problems with finitely many linear equality and a single (nonconvex) quadratic inequality constraints. They characterized the strong duality,

Y.B. Yuan (✉)
Department of Computer Science and Engineering, East China University of Science and Technology, Shanghai 200237, China
e-mail: ybyuan@ecust.edu.cn

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_17

necessary and sufficient optimality conditions with or without the Slater assumption geometrically.

In 2013, Tuy and Tuan [28] had studied new strong duality conditions for multiple constrained quadratic optimization based on the topological minimax theorem. Their results showed that many quadratic programs to be solved by solving one or just a few semidefinite programs. In the last work of Tuy and Hoai-Phuong [27], they had proposed novel approach to get more appropriate approximate optimal solutions of the problems. In 2007, Jeyakumar et al. [21] had studied necessary global optimality conditions for special classes of quadratic optimization problems such as weighted least squares with ellipsoidal constraints, quadratic minimization with binary constraints, and so on.

In 2013, Misener and Floudas [23] had introduced the global mixed-integer quadratic optimizer(GloMIQO). The problems can be considered as the special cases of $\mathscr{P}_{qq}$. They proposed a novel algorithm to solve the problems based on branch-and-bound method. In 2013, Peter et al. [22] had studied the spatial branch-and-bound method [26]. They proposed a novel method to perturb infeasible iterates along Mangasarian–Fromovitz directions to feasible points. Their numerical results showed that their proposed algorithm could perform well even for optimization problems where the standard branch-and-bound method did not converge to the correct optimal value.

In 2012, Yuan, Fang, and Gao [33] had considered a class of quadrinomial minimization problems with one quadratic constraint. In that work, the objective function is fourth order polynomial. Before this work, the canonical duality was employed to solve the altering support vector machine [32] and the corresponding problems with linear inequality constraints had been studied [31].

The nonconvex quadratic minimization problems with quadratic constraints can be formulated as follows (($\mathscr{P}_{qq}$) in short)

$$(\mathscr{P}_{qq}): \quad \min \left\{ P(x) = \frac{1}{2} x^T A x - f^T x \;\;:\;\; x \in \mathscr{X}_a \right\}, \tag{1}$$

where $A = A^T \in \mathbb{R}^{n \times n}$ is an indefinite matrix, the feasible space $\mathscr{X}_a$ is defined by

$$\mathscr{X}_a \triangleq \left\{ x \in \mathbb{R}^n | \; \frac{1}{2} x^T Q_i x + b_i^T x \leqslant c_i, i = 1, 2, \cdots, m \right\}, \tag{2}$$

in which, ${Q_i}^T = Q_i \in \mathbb{R}^{n \times n}$ $(i = 1, 2, \cdots, m)$ are given nonsingular matrices, $b_i \in \mathbb{R}^n$, $(i = 1, 2, \cdots, m)$ are given vectors which control the geometric centers. $c_i (i = 1, 2, \cdots, m) \in \mathbb{R}$ are given input constants.

In order to make sure that the feasible space $\mathscr{X}_a$ is nonempty, the quadratic constraints must satisfy the *Slater regularity condition*, i.e., there exists one point x_0 such that $\frac{1}{2} x_0^T Q_i x_0 + b_i^T x_0 \leqslant c_i, i = 1, 2, \cdots, m$.

In this work, one hard restriction is given that $f \neq 0 \in \mathbb{R}^n$. The restriction is very important to guarantee the uniqueness of global optimization solution of $(\mathscr{P}_{qq})$. In physics, $P(x) = \frac{1}{2}x^T Ax - f^T x$ means energy function. The first part $\frac{1}{2}x^T Ax$ means kinetic energy or elastic energy or other one. The second part $f^T x$ means work under an input force f. If force $f = 0$(object is on the stable state), the problems may have infinite global optimization solutions. For example, we consider the following problem

$$\min_{(x,y)\in\mathbb{R}^2} \left\{P(x, y) = -x^2 - y^2\right\}$$
$$s.t. \quad x^2 + y^2 \leq 4.$$

This problem has infinite solutions (x, y) in $\mathbb{R}^2$ and $x^2 + y^2 = 4$. In another word, the boundary points of feasible space are the global optimization solutions. If force $f = (2, 2)^T$, the problem is formulated as follows

$$\min_{(x,y)\in\mathbb{R}^2} \left\{P(x, y) = -x^2 - y^2 - 2x - 2y\right\}$$
$$s.t. \quad x^2 + y^2 \leq 4.$$

This problem has unique global optimization solution $(x^*, y^*) = (\sqrt{2}, \sqrt{2})$ (Fig. 1).

It is known that linear mixed 0–1, fractional, polynomial, bilevel, generalized linear complementarity problems, can be reformulated as special cases of $(\mathscr{P}_{qq})$. Such problems have attracted the attention of many researchers in recent years. The problem of minimizing nonconvex quadratic function with one convex quadratic constraint arises from applying the trust region method in solving unconstrained optimization. It was first proposed by Celis, Dennis, and Tapia (see in [2] and developed by Powell and Yuan in 1990 and 1991(see in [25, 30]. The subproblem of trust region method is described as follows

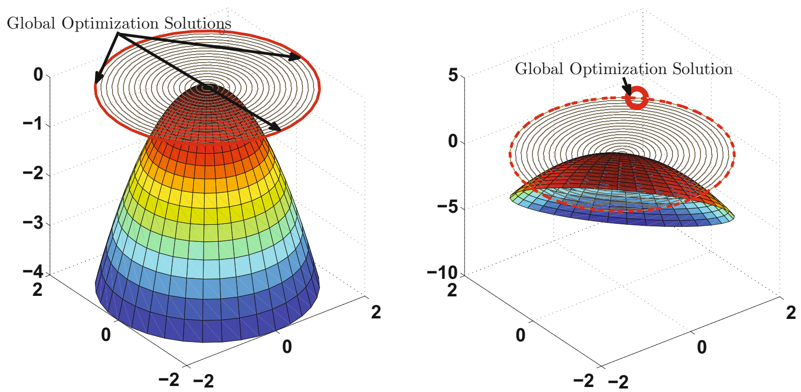


Fig. 1 Flowchart to show the difference between $f = 0$ and $f \neq 0$. The *color* surface is the figure of the objective function on the feasible space (the disk). The *left* one is $f = 0$ and it shows that the problem has infinite global optimization solutions on the feasible space's boundary. The *right* one is $f = (2, 2)^T$ and it shows that the problem has unique global optimization solution on the point $(\sqrt{2}, \sqrt{2})$

$$(\mathscr{S}\mathscr{T}\mathscr{R}):\begin{array}{l}\min_{\delta\in\mathbb{R}^n}\ \{P_k(\delta)=f(x^{(k)})+g^{(k)^T}\delta+\frac{1}{2}\delta^T B^{(k)}\delta\}\\ s.t.\ \ \frac{1}{2}\|\delta\|^2\leqslant\rho_k.\end{array}\tag{3}$$

In which, δ is the objective vector(after solving the model (3), we can construct the next iteration points with ($x^{(k+1)}=x^{(k)}+\delta$)), $g^{(k)}=\nabla f(x^{(k)})$ is the gradient vector, $B^{(k)}$ is the Hessian matrix or approximate matrix of Hessian and ρ_k is the trust region parameter. If the objective function is nonconvex, this problem is NP-hard [24].

With two (general) convex quadratic constraints, recently, the problem is termed as the extended trust region subproblem(see in [2, 25, 29, 30]. In general, it is proved to be NP-hard(see in [1, 24]. Actually, the extended trust region subproblem is a special case of our presented problem ($\mathscr{P}_{qq}$). It can be formulated as follows [29]

$$(\mathscr{S}\mathscr{T}\mathscr{T}\mathscr{R}):\begin{array}{l}\min_{x\in\mathbb{R}^n}\ \{P(x)=\frac{1}{2}x^T Ax-f^T x\}\\ s.t.\ \ \frac{1}{2}x^T Q_1 x+b_1^T x\leqslant c_1,\\ \qquad\ \frac{1}{2}x^T Q_2 x+b_2^T x\leqslant c_2.\end{array}\tag{4}$$

Motivated by the difficulty of solving these problems, we are looking for some good and powerful method to check out the global optimization solution. There is a very powerful method proposed by Gao David (see in [12, 18]) and it is called as *canonical duality*. The idea is from Legendre duality(presented and explored by Ekeland (readers can refer to [3–5]). It is proved that it has some advantages in global optimization and nonlinear mechanics (see in [8–20]). In this work, we employ it to deal with a special class of $\mathscr{P}_{qq}$ and use it to convert $\mathscr{P}_{qq}$ into a concave maximization dual problem over a convex set.

The paper is organized as follows. In Sect. 2, one novel definition is introduced and stated as *complementary positive definite matrix group*. The basic procedure is presented to convert $\mathscr{P}_{qq}$ into a concave maximization dual problem. Two theorems are presented to support us to find out the global optimization solution. Main result in Theorem 1 is the equation between optimization solution of $\mathscr{P}_{qq}$ and canonical duality problem. Main result in Theorem 2 is to give conditions to make sure that the canonical duality problem has a unique optimization solution. In Sect. 3, we present the basic framework of the proposed algorithm. In Sect. 4, several examples are illustrated to show the correctness of given conditions and effectiveness of presented theorems. Finally, we make a conclusion.

2 Canonical Duality Problem

2.1 Complementary Positive Definite Matrix

In order to study the existence of the problem $\mathscr{P}_{qq}$, we introduce a definition.

Definition 1 For a given matrix $A \in \mathbb{R}^{n\times n}$, $\mathscr{G}_+(A) \subset \mathbb{R}^{n\times n}$ is called as *complementary positive definite matrix group* of A, if for any $B \in \mathscr{G}_+(A)$, $A + B$ is positive definite. Mathematically,

$$\mathscr{G}_+(A) \triangleq \left\{B \in \mathbb{R}^{n\times n} | A + B \succ 0\right\}. \tag{5}$$

Especially, if $A + B = I$, B is called *identity complementary matrix* of A, where I is the identity matrix of order n by n.

With the same idea, a new definition on *complementary negative definite matrix group* can be given.

2.2 Canonical Duality Problem of $\mathscr{P}_{qq}$

Following the standard procedure and ideas proposed by David Gao [15–18], we construct the geometrical mapping as follows

$$\boldsymbol{\varepsilon}(x) = \{\varepsilon_i(x)\} = \left\{\frac{1}{2}x^T Q_i x + b_i^T x - c_i, i = 1, 2, \cdots, m\right\} : \mathbb{R}^n \to \mathbb{R}^m. \tag{6}$$

The indicator is defined by

$$\mathscr{I}(\boldsymbol{\varepsilon}) = \begin{cases} 0, & \text{if } \boldsymbol{\varepsilon} \le 0 \in \mathbb{R}^m, \\ +\infty, & \text{otherwise.} \end{cases}$$

With the indicator, the quadratic constraints in ($\mathscr{P}_{qq}$) can be relaxed and ($\mathscr{P}_{qq}$) takes the unconstrained form as following

$$(\mathscr{P}): \quad \min\left\{P(x) = \mathscr{I}(\boldsymbol{\varepsilon}(x)) + \frac{1}{2}x^T Ax - f^T x \ : \ x \in \mathbb{R}^n\right\}. \tag{7}$$

Because $\mathscr{I}(\boldsymbol{\varepsilon})$ is convex and lower semi-continuous on $\mathbb{R}^m$, their canonical dual variable $\boldsymbol{\sigma}$ satisfies the following duality relation

$$\boldsymbol{\sigma} \in \partial^- \mathscr{I}(\boldsymbol{\varepsilon}) \Leftrightarrow \boldsymbol{\varepsilon} \in \partial^- \mathscr{I}^*(\boldsymbol{\sigma}) \Leftrightarrow \mathscr{I}(\boldsymbol{\varepsilon}) + \mathscr{I}^*(\boldsymbol{\sigma}) = \boldsymbol{\varepsilon}^T \boldsymbol{\sigma}, \tag{8}$$

where ∂^- is called the sub-differential of $\mathscr{I}$ in convex analysis. $\mathscr{I}^*(\boldsymbol{\sigma})$ is *Fenchel sup-conjugate* of $\mathscr{I}$ by

$$\mathscr{I}^*(\boldsymbol{\sigma}) = \sup_{\boldsymbol{\varepsilon}\in\mathbb{R}^m} \{\boldsymbol{\varepsilon}^T \boldsymbol{\sigma} - \mathscr{I}(\boldsymbol{\varepsilon})\} = \begin{cases} 0, & if\ \boldsymbol{\sigma} \geqslant 0, \\ +\infty, & otherwise. \end{cases} \tag{9}$$

The canonical dual function of $P(x)$ is defined by the following equation (referred to [8])

$$P^d(\boldsymbol{\sigma}) = Q^{\Lambda}(\boldsymbol{\sigma}) - \mathscr{I}^*(\boldsymbol{\sigma}), \tag{10}$$

where

$$\begin{aligned} Q^{\Lambda}(\boldsymbol{\sigma}) &= sta\ \left\{\boldsymbol{\varepsilon}^T\boldsymbol{\sigma} + \tfrac{1}{2}x^T Ax - f^T x\right\} \\ &= -\tfrac{1}{2}F(\boldsymbol{\sigma})^T G(\boldsymbol{\sigma})^{-1}F(\boldsymbol{\sigma}) - c^T\boldsymbol{\sigma}, \end{aligned} \tag{11}$$

in which the notation $sta\{* : x \in \mathbb{R}^n\}$ is the operator to find out the stationary point in the space $\mathbb{R}^n$, $G(\boldsymbol{\sigma})$, $F(\boldsymbol{\sigma})$ and c are defined by

$$G(\boldsymbol{\sigma}) = (A + \sum_{i=1}^{m} Q_i\sigma_i), \quad F(\boldsymbol{\sigma}) = (f - \sum_{i=1}^{m} b_i\sigma_i), \quad c = (c_1, c_2, \cdots, c_m)^T, \tag{12}$$

where σ_i is the i-th element of $\boldsymbol{\sigma}$.

The dual feasible space is defined by

$$\mathscr{S} \triangleq \left\{\boldsymbol{\sigma} \in \mathbb{R}^m | \boldsymbol{\sigma} \geqslant 0 \in \mathbb{R}^m, \quad \det(G(\boldsymbol{\sigma})) \neq 0\right\}. \tag{13}$$

The canonical dual problem ($\mathscr{P}^d$ in short) associated with ($\mathscr{P}_{qq}$) can be eventually formulated as follows

$$(\mathscr{P}^d) : \max_{\boldsymbol{\sigma}\in\mathscr{S}} \left\{P^d(\boldsymbol{\sigma})\right\}. \tag{14}$$

2.3 Two Important Theorems

In order to show that there is no duality gap, the following theorem is presented.

Theorem 1 If A, Q_i, b_i, f_i, c_i, $i = 1, 2, \cdots, m$, are given with definitions in ($\mathscr{P}_{qq}$) such that the dual feasible space

$$\mathscr{Y} \triangleq \left\{\boldsymbol{\sigma} \in \mathscr{S} \mid G(\boldsymbol{\sigma})^{-1}F(\boldsymbol{\sigma}) \in \mathscr{X}\right\} \tag{15}$$

is not empty, the problem

$$(\mathscr{P}^d) : \max_{\boldsymbol{\sigma}\in\mathscr{Y}} \left\{P^d(\boldsymbol{\sigma})\right\}, \tag{16}$$

is canonically (perfectly) dual to ($\mathscr{P}_{qq}$). In another words, if $\bar{\boldsymbol{\sigma}}$ is a solution of the dual problem ($\mathscr{P}^d$),

$$\bar{x} = G(\bar{\boldsymbol{\sigma}})^{-1}F(\bar{\boldsymbol{\sigma}}) \tag{17}$$

is a solution of ($\mathscr{P}_{qq}$) and

$$P(\bar{x}) = P^d(\bar{\boldsymbol{\sigma}}). \tag{18}$$

Proof. If $\bar{\boldsymbol{\sigma}}$ is a solution of the dual problem $(\mathscr{P}^d)$ such that (17) holds, it must satisfy the KKT conditions. Then, according to the complementarity conditions, we have

$$\bar{\boldsymbol{\sigma}} \perp \nabla P^d(\bar{\boldsymbol{\sigma}}) \quad \text{and} \quad P^d(\bar{\boldsymbol{\sigma}}) = 0. \tag{19}$$

Let us pay attention to the (13) and $\bar{x}$ must satisfy the constraints, we have

$$\begin{array}{l} \frac{1}{2}\bar{x}^T Q_i \bar{x} + b_i^T \bar{x} - c_i \leqslant 0, \\ \bar{\sigma}_i \perp \frac{1}{2}\bar{x}^T Q_i \bar{x} + b_i^T \bar{x} - c_i, \\ \bar{\sigma}_i \geqslant 0, i = 1, 2, \cdots, m. \end{array} \tag{20}$$

This result shows that $\bar{x} = G(\bar{\boldsymbol{\sigma}})^{-1} F(\bar{\boldsymbol{\sigma}})$ is a KKT point of $(\mathscr{P}_{qq})$.

Next, we show the equivalence between the primal problem and canonical duality one. According to complementarity conditions (20), we have

$$c_i \bar{\sigma}_i = \frac{1}{2}\bar{x}^T Q_i \bar{\sigma}_i \bar{x} + b_i^T \bar{\sigma}_i \bar{x}, i = 1, 2, \cdots, m. \tag{21}$$

Thus, in terms of $\bar{x} = G(\bar{\boldsymbol{\sigma}})^{-1} F(\bar{\boldsymbol{\sigma}}) = (A + \sum_{i=1}^{m} Q_i \bar{\sigma}_i)^{-1}(f - \sum_{i=1}^{m} b_i \bar{\sigma}_i)$, we have

$$\sum_{i=1}^{m} Q_i \bar{\sigma}_i \bar{x} + \sum_{i=1}^{m} b_i \bar{\sigma}_i = f - A\bar{x}, \tag{22}$$

then

$$\begin{aligned} P^d(\bar{\boldsymbol{\sigma}}) &= -\tfrac{1}{2} F(\bar{\boldsymbol{\sigma}})^T G(\bar{\boldsymbol{\sigma}})^{-1} F(\bar{\boldsymbol{\sigma}}) - c^T \bar{\boldsymbol{\sigma}}, \\ &= -\tfrac{1}{2}\bar{x}^T G(\bar{\boldsymbol{\sigma}})\bar{x} - \sum_{i=1}^{m} (\tfrac{1}{2}\bar{x}^T Q_i \bar{\sigma}_i \bar{x} + b_i^T \bar{\sigma}_i \bar{x}), \\ &= -\tfrac{1}{2}\bar{x}^T A\bar{x} - (f - A\bar{x})^T \bar{x}, \\ &= \tfrac{1}{2}\bar{x}^T A\bar{x} - f^T \bar{x}, \\ &= P(\bar{x}), \end{aligned}$$

which shows that there is no duality gap between $(\mathscr{P}_{qq})$ and $(\mathscr{P}^d)$. The proof of the theorem is concluded.

In order to get the optimization solution of $(\mathscr{P}_{qq})$, we introduce the following subset

$$\mathscr{S}_+ = \{\boldsymbol{\sigma} \in \mathscr{S} \;\mid\; G(\boldsymbol{\sigma}) \text{ is positive definite}\}. \tag{23}$$

In order to hold on the uniqueness of optimal duality solution, the following existence theorem is presented.

Theorem 2 For any given symmetrical matrixes $A, Q_i, \in \mathbb{R}^{n\times n}$, $\mathscr{G}_+(A)$ (defined by (5)) is the complementary positive definite matrix group of A, $f, b_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}$, $i = 1, 2, \cdots, m$, if the following two conditions are satisfied

$C_1 : \sum_{i=1}^{m} Q_i \in \mathscr{G}_+(A)$;

C_2 : there must exist one $k(1 \leq k \leq m)$ such that Q_k is positive definite and $Q_k \in \mathscr{G}_+(A)$, moreover,

$$\|D_k A^{-1} f\| > \|b_k^T D_k^{-1}\| + \sqrt{\|b_k^T D_k^{-1}\|^2 + 2|c_k|}, \tag{24}$$

where $Q_k = D_k^T D_k$ and $\| * \|$ is some vector norm.

Then, the canonical duality problem (16) has a unique nonzero solution $\bar{\boldsymbol{\sigma}}$ in the space $\mathscr{S}_+$.

Proof. If the condition C_1 is satisfied, the dual feasible space defined by (23) is nonempty.

If C_2 is also satisfied, we can get two results, the first one is that there is one positive definite matrix D_k such that

$$Q_k = D_k^T D_k.$$

The second one is that the stationary point of quadratic objective function is out of the convex constraint defined by $\frac{1}{2}x^T Q_k x + b_k^T x \leq c_k$. The first one is easy to be proved because Q_k is symmetrical positive definite. Next, we will show how to get the the second result. Because D_k from the first result is also positive definite, we have

$$\begin{aligned}
&\tfrac{1}{2} f^T (A^{-1})^T Q_k A^{-1} f + b_k^T A^{-1} f \\
&= \tfrac{1}{2} f^T (A^{-1})^T D_k^T D_k A^{-1} f + b_k^T D_k^{-1} D_k A^{-1} f \\
&\geq |\tfrac{1}{2} f^T (A^{-1})^T D_k^T D_k A^{-1} f| - |b_k^T D_k^{-1} D_k A^{-1} f| \\
&\geq \tfrac{1}{2} \|D_k A^{-1} f\|^2 - \|b_k^T D_k^{-1}\| \|D_k A^{-1} f\| \\
&= \tfrac{1}{2} (\|D_k A^{-1} f\| - \|b_k^T D_k^{-1}\|)^2 - \tfrac{1}{2} \|b_k^T D_k^{-1}\|^2,
\end{aligned}$$

if we pay attention to (24), from the above inequalities, the following inequalities are easy to obtain,

$$\begin{aligned}
&\tfrac{1}{2} f^T (A^{-1})^T Q_k A^{-1} f + b_k^T A^{-1} f \\
&\geq |c_k| \geq c_k.
\end{aligned}$$

So, $A^{-1} f$ is out of the constraint. According to complementary theory,

$$\bar{\sigma}_k \neq 0.$$

Then, there is nonzero solution for the canonical duality problem in the space $\mathscr{S}_+$. Because the objective function is concave and differentiable in the space $\mathscr{S}_+$, the canonical duality solution is unique. The proof of theorem is concluded.

3 Algorithm

In this section, an algorithm is proposed to solve the problem ($\mathscr{P}_{qq}$). The basic procedures are listed in Algorithm 1.

The algorithm has two important parts. The first one is to judge the conditions. The other one is to get the duality optimization solution.

In the first part, we need to complete two important steps, they are from the computation of eigenvalues of $A + \sum_{k=1}^{m} Q_k$ and $A + Q_i$. If we recall the parameters, n is the dimensional number of input variable x and m is the number of constrains. The complexity of the first part is

$$T(m,n) = O(m \times n^2). \tag{25}$$

In the other part, the time complexity comes from the method to solve the canonical duality problem. The complexity is

$$T(m,n) = O(n \times m^2). \tag{26}$$

The complexity of the final time complexity of our proposed algorithm is

$$T(m,n) = O(m \times n^2) + O(n \times m^2). \tag{27}$$

4 Applications

In this section, several examples are illustrated to show how to use the presented theory to solve the problems. We employ the Quasi-Newton method to solve the canonical duality problems.

Example 1. First of all, let us consider two-dimensional quadratic minimization problem with one quadratic constraint. If we take $A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$, $Q = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$, $f = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$, $c = 3$, the following minimization problem is obtained,

$$\min_{x \in \mathbb{R}^2} \left\{ P(x) = x_1^2 - 0.5 * x_2^2 - 3x_1 - 3x_2 \right\} \tag{28}$$

such that

$$2x_1^2 + (x_2 - 1)^2 \le 4. \tag{29}$$

Algorithm 1 CDQN-QPQS algorithm

1: Input: the matrices A ,Q_i, $i = 1, 2, \cdots, m$, the load item f and linear items b_i, $i = 1, 2, \cdots, m$, the const items of constraints c_i, $i = 1, 2, \cdots, m$, algorithm stop cutoff ϵ;;
2: Initialization: $\boldsymbol{\sigma}^0 = 0$, $H^0 = \mathbf{I}$, $\alpha^0 = 1$, $i := 0$ and $ps := 0$;
3: Computing: $F^i = P^d(\boldsymbol{\sigma}^i)$ and $g^i = \nabla P^d(\boldsymbol{\sigma}^i)$;
$mes = min(eig(A + \sum_{k=1}^{m} Q_k))$;
4: **if** $mes > 0$ **then**
5: **for** $k = 1 : m$ **do**
6: $v_k = min(eig(Q_k))$;
7: **if** $v_k > 0$ **then**
8: $ps = ps + 1$;
9: $Av_k = min(eig(A + Q_k))$;
10: **else**
11: $ps = ps$;
12: $Av_k = -1$;
13: **end if**
14: **end for**
15: Compute $mv = max(Av)$;
16: **if** $ps < 0$ or $mv < 0$ **then**
17: $ss = 0$;
18: **else**
19: Select Q_i such that $v_i > 0$ and $Av_i > 0$, let $PQ = Q_i, k = i$;
computing orthogonal decomposition of $PQ = D^T D$;
$lp = \|DA^{-1}f\|; rp = \|b_k^T D^{-1}\| + \sqrt{\|b_k^T D^{-1}\|^2 + 2|c_k|}$;
20: **if** $lp < rp$ **then**
21: $ss = 0$;
22: **else**
23: Quasi-Newton method is employed to solve the canonical duality problem
ss=arg$\{\max_{\boldsymbol{\sigma} \in \mathscr{Y}} \left\{P^d(\boldsymbol{\sigma})\right\}\}$;
24: **end if**
25: **end if**
26: **else**
27: $ss = 0$;
28: **end if**
29: Computing the global optimization solution: $\bar{\boldsymbol{\sigma}} = ss$,

$$G(\bar{\boldsymbol{\sigma}}) = (A + \sum_{i=1}^{m} Q_i \bar{\boldsymbol{\sigma}}), \quad F(\bar{\boldsymbol{\sigma}}) = (f - \sum_{i=1}^{m} b_i \bar{\boldsymbol{\sigma}}),$$

and

$$\bar{x} = G(\bar{\boldsymbol{\sigma}})^{-1} F(\bar{\boldsymbol{\sigma}});$$

30: Output: $\bar{x}$; $P(\bar{x})$; $\bar{\boldsymbol{\sigma}}$; $P^d(\bar{\boldsymbol{\sigma}})$;

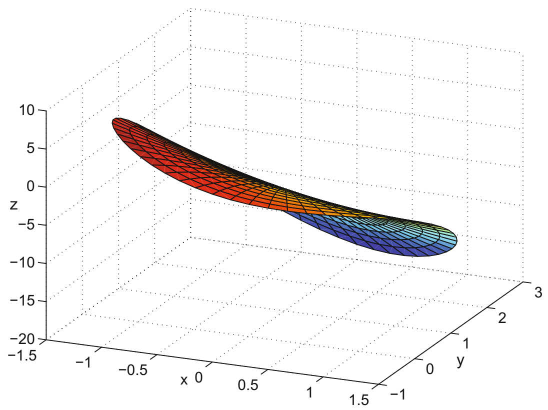


Fig. 2 The graph of $p(x)$ on the quadratic constraint

This problem is to search the global minimize value of $P(x)$ in the inner part of elliptic sphere whose boundary is determined by $2x_1^2 + (x_2 - 1)^2 = 4$(can be seen in Fig. 2).

We can easily verify that condition C_1 in Theorem 2 is satisfied because the eigenvalues of matrix $A + Q$ are 6 and 1. C_2 is also satisfied because $\|DA^{-1}f\| = 5.1962$ and

$$\|b^T D^{-1}\| + \sqrt{\|b^T D^{-1}\|^2 + 2|c|} = 4.2426.$$

where $Q = D^T D$.

The corresponding dual problem is

$$\max_{\sigma\in\mathbb{R}} \left\{ P^d(\sigma) = -\frac{1}{2}\left(\frac{9}{4\sigma+2} + \frac{(3+2\sigma)^2}{2\sigma-1}\right) - 3\sigma \right\} \tag{30}$$

such that $\sigma \geqslant 0$.

Then we can present the solution of this problem. This dual problem has a unique solution:

$$\bar{\sigma} = 1.5358.$$

The canonical duality global maximize value is

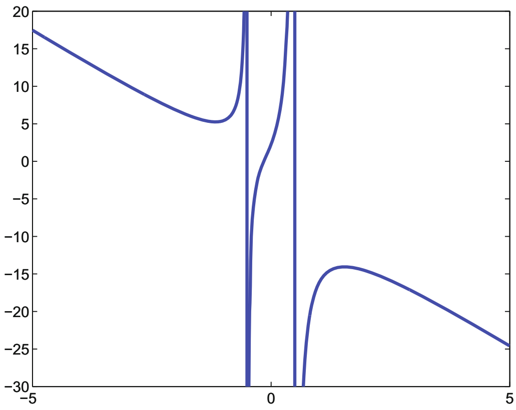
$$P^d(\bar{\sigma}) = -14.0576.$$

The graph of the canonical duality problem $P^d(\sigma)$ on the interval $[-5, 5]$ is shown in Fig. 3. In this figure, we easily see that $\bar{\sigma} = 1.5358$ is the global maximizer and $\max P^d(\sigma) = P^d(1.5) = -14.0576$.

The optimal solution of primal problem can be obtained by

$$\bar{x} = (A + \bar{\sigma}Q)^{-1}(f - b\bar{\sigma}) = \begin{pmatrix} 0.3684 \\ 2.9309 \end{pmatrix}.$$

Fig. 3 The graph of $P^d(\sigma)$ on the interval $[-5, 5]$



It is very easy to verify that

$$P(\bar{x}) = -14.0576 = P^d(\bar{\sigma}).$$

Let us pay attention to the solution, $\bar{\sigma} = 1.5358$ shows that the solution $\bar{x}$ is on the boundary of the feasible space, in fact, we can understand this from Fig. 2. We can easily check that $\bar{x} = \begin{pmatrix} 0.3684 \\ 2.9309 \end{pmatrix}$ satisfies $2x_1^2 + (x_2 - 1)^2 = 4$.

Example 2. We now consider three-dimensional quadratic minimization problem with two quadratic constraints. If we take $A = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$, $Q_1 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, $Q_2 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, $f = \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix}$, $b_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix}$, $c_1 = 2, c_2 = 2$, the following minimization problem is obtained,

$$\min_{x \in \mathbb{R}^3} \left\{ P(x) = -x_1^2 + x_2^2 - x_3^2 - 4x_1 - 2x_2 - 4x_3 \right\} \tag{31}$$

such that

$$2x_1^2 + 2x_2^2 + 2x_3^2 \leqslant 2, \tag{32}$$

and

$$2x_1^2 - 0.5x_2^2 + 2x_3^2 - 3x_1 \leqslant 2. \tag{33}$$

This problem is to look for the global minimize value of $P(x)$ in the communal inner part of one parabolic and one sphere which boundary is determined by $2x_1^2 + 2x_2^2 + 2x_3^2 = 2$ and $2x_1^2 - 0.5x_2^2 + 2x_3^2 - 3x_1 = 2$ (can be seen in Fig. 4).

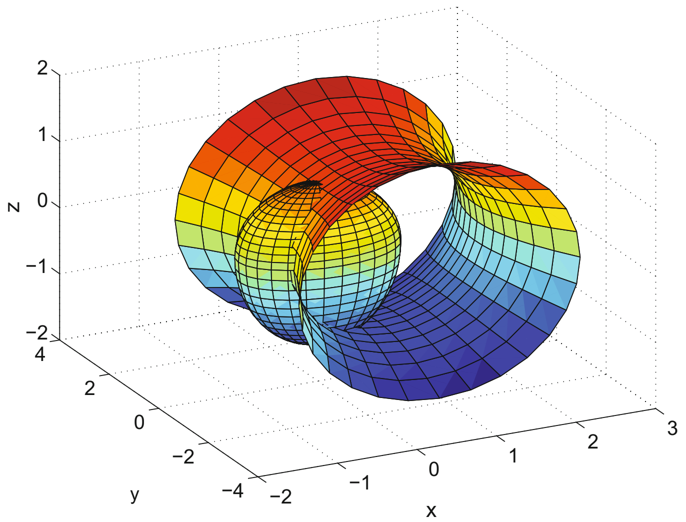


Fig. 4 Two quadratic constrains figure bounded by $2x_1^2 + 2x_2^2 + 2x_3^2 \leqslant 2$ and $2x_1^2 - 0.5x_2^2 + 2x_3^2 - 3x_1 \leqslant 2$

Also, we can easily verify that condition C_1 in Theorem 2 is satisfied because the eigenvalues of $A + Q_1 + Q_2$ are 5, 6 and 6, the eigenvalues of $A + Q_1$ are 2, 2 and 6. C_2 is satisfied because $\|D_1 A^{-1} f\| = 6$ and

$$\|b_1^T D_1^{-1}\| + \sqrt{\|b_1^T D_1^{-1}\|^2 + 2|c_1|} = 2.$$

where $Q_1 = D_1^T D_1$.

According to canonical duality theory, the canonical dual problem of (31) is as follows:

$$\max_{(\sigma_1,\sigma_2)\in\mathbb{R}^2} \{P^d(\sigma_1,\sigma_2) = -\tfrac{1}{2}(\tfrac{(4+3\sigma_2)^2}{4(\sigma_1+\sigma_2)-2} + \tfrac{4}{4\sigma_1-\sigma_2+2} + \tfrac{16}{4(\sigma_1+\sigma_2)-2}) - 2\sigma_1 - 2\sigma_2\} \tag{34}$$

such that $\sigma_1 \geqslant 0, \sigma_2 \geqslant 0$.

Then we can get the solution of this problem as following:

$$\bar{\sigma}_1 = 1.9447, \bar{\sigma}_2 = 0.$$

The canonical duality global maximized value is

$$P^d(\bar{\sigma}_1, \bar{\sigma}_2) = -6.8627.$$

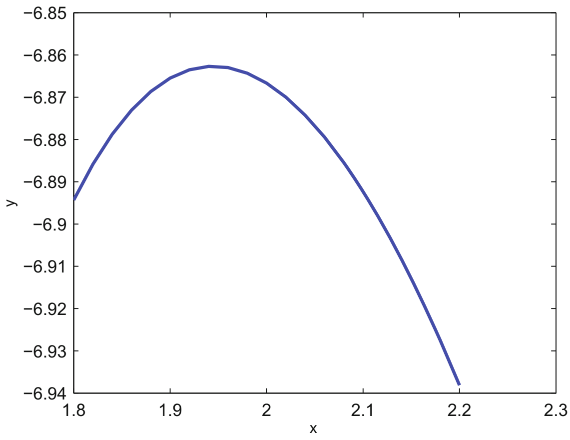


Fig. 5 The graph of $P^d(\sigma_1, 0)$ on the interval [1.8, 2.2]

The graph of canonical duality problem objective function $P^d(\bar{\sigma}_1, 0)$ on the interval [1.8, 2.2] is shown in Fig. 5. In its figure, we easily guarantee that $(\bar{\sigma}_1 = 1.9447, \bar{\sigma}_2 = 0)$ is the global maximize point and $\max P^d(\sigma_1, 0) = P^d(1.9447, 0) = -6.8627$.

The optimal solution of primal problem can be obtained by

$$\bar{x} = (A + \bar{\sigma}_1 Q_1 + \bar{\sigma}_2 Q_2)^{-1}(f - b_1\bar{\sigma}_1 - b_2\bar{\sigma}_2) = \begin{pmatrix} 0.6922 \\ 0.2045 \\ 0.6922 \end{pmatrix}.$$

It is very easy to verify that

$$P(1.9447, 0) = -6.8627 = P^d(0.6922, 0.2045, 0.6922).$$

Example 3. We now consider four-dimensional quadratic minimization problem with three quadratic inequalities. constraint. If we take

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -20 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}, Q_1 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 24 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, Q_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$

$$Q_3 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, f = \begin{pmatrix} 4 \\ 12 \\ -2 \\ 2 \end{pmatrix}, b_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, b_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$b_3 = \begin{pmatrix} 0 & 0 & -1 & 0 \end{pmatrix}^T$, $c_1 = 9$, $c_2 = c_3 = 8.5$, the following minimization problem is obtained,

$$\min_{x\in\mathbb{R}^4} \left\{P(x) = (x_1-1)^2 + x_2^2 - 10x_3^2 - 4x_4^2 - 12x_2 + 2x_3 - 2x_4\right\} \tag{35}$$

such that

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 \leqslant 9, \tag{36}$$

and

$$\frac{1}{2}(x_1-1)^2 + x_2^2 + 12x_3^2 + x_4^2 \leqslant 9, \tag{37}$$

and

$$x_1^2 - x_2^2 + \frac{1}{2}(x_3-1)^2 + x_4^2 \leqslant 9. \tag{38}$$

The equality of the first constraint is sphere, the second one is ellipsoid and the last one is hyperboloid.

Also, we can easily verify that condition C_1 in Theorem 2 is satisfied because the eigenvalues of $A+Q_1+Q_2+Q_3$ are 4, 7, 7 and 10, the eigenvalues of $A+Q_1$ are 4, 4, 4 and 6. C_2 is satisfied because $\|D_1A^{-1}f\| = 8.9855$ and

$$\|b_1^TD_1^{-1}\| + \sqrt{\|b_1^TD_1^{-1}\|^2 + 2|c_1|} = 4.2426.$$

where $Q_1 = D_1^TD_1$.

According to canonical duality theory, the canonical dual problem of (35) is as follows

$$\max_{(\sigma_1,\sigma_2,\sigma_3)\in\mathbb{R}^3} \Big\{P^d(\sigma_1,\sigma_2,\sigma_3) = -\tfrac{1}{2}\big(\tfrac{(4+\sigma_2)^2}{2\sigma_1+\sigma_2+2\sigma_3+2} + \tfrac{144}{2(\sigma_1+\sigma_2-\sigma_3)+2} + \tfrac{(\sigma_3-2)^2}{24\sigma_1+2\sigma_2+\sigma_3-20} + \tfrac{4}{2(\sigma_1+\sigma_2+\sigma_3)+4}\big) - 9\sigma_1 - 8.5\sigma_2 - 8.5\sigma_3\Big\} \tag{39}$$

such that $\sigma_i \geqslant 0, i = 1, 2, 3$.

Then we can get the solution of this problem as following:

$$\bar{\sigma}_1 = 1.1983, \bar{\sigma}_2 = 0, \bar{\sigma}_3 = 0.$$

The canonical duality global maximized value is

$$P^d(\bar{\sigma}_1,\bar{\sigma}_2,\bar{\sigma}_3) = -29.5216.$$

The optimal solution of primal problem can be obtained by

$$\bar{x} = (A+\bar{\sigma}_1Q_1+\bar{\sigma}_2Q_2+\bar{\sigma}_3Q_3)^{-1}(f - b_1\bar{\sigma}_1 - b_2\bar{\sigma}_2 - b_3\bar{\sigma}_3) = \begin{pmatrix} 0.9098 \\ 2.7293 \\ -0.2283 \\ 0.3127 \end{pmatrix}.$$

It is very easy to verify that

$$P(1.1983, 0, 0) = -29.5216 = P^d(0.9098, 2.7293, -0.2283, 0.3127).$$

Example 4. Here, we present a ten-dimensional nonconvex quadratic programming with two quadratic constraints. If we let

$$A = \begin{pmatrix} -2 & 0.5 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & -13 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 1 \\ 1 & 1 & -14 & 0.5 & 0 & 1 & 1 & 1 & 0.5 & 0.5 \\ 1 & 1 & 0.5 & -2 & 0.5 & 1 & 1 & 0.5 & 0.5 & 0.5 \\ 1 & 1 & 0 & 0.5 & 3 & 0.5 & 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 & 1 & 0.5 & -6 & 0.5 & 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 1 & 1 & 0.5 & 0.5 & -1 & 0.5 & 1 & 0 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & -13 & 0.5 & 1 \\ 0.5 & 0.5 & 0.5 & 0.5 & 1 & 0.5 & 1 & 0.5 & -14 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0 & 0 & 1 & 0.5 & -5 \end{pmatrix},$$

$$Q_1 = \begin{pmatrix} 10 & 0.5 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 25 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 1 \\ 1 & 1 & 24 & 0.5 & 0 & 1 & 1 & 1 & 0.5 & 0.5 \\ 1 & 1 & 0.5 & 12 & 0.5 & 1 & 1 & 0.5 & 0.5 & 0.5 \\ 1 & 1 & 0 & 0.5 & 9 & 0.5 & 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 & 1 & 0.5 & 12 & 0.5 & 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 1 & 1 & 0.5 & 0.5 & 9 & 0.5 & 1 & 0 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 25 & 0.5 & 1 \\ 0.5 & 0.5 & 0.5 & 0.5 & 1 & 0.5 & 1 & 0.5 & 31 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0 & 0 & 1 & 0.5 & 9 \end{pmatrix},$$

$$Q_2 = \begin{pmatrix} 5 & 0.5 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 9 & 1 & 1 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 1 \\ 1 & 1 & 8 & 0.5 & 0 & 1 & 1 & 1 & 0.5 & 0.5 \\ 1 & 1 & 0.5 & 6 & 0.5 & 1 & 1 & 0.5 & 0.5 & 0.5 \\ 1 & 1 & 0 & 0.5 & 4 & 0.5 & 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 & 1 & 0.5 & 6 & 0.5 & 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 1 & 1 & 0.5 & 0.5 & 4 & 0.5 & 1 & 0 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 13 & 0.5 & 1 \\ 0.5 & 0.5 & 0.5 & 0.5 & 1 & 0.5 & 1 & 0.5 & 13 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0 & 0 & 1 & 0.5 & 4 \end{pmatrix},$$

the other corresponding coefficients are listed as follows

$$f = (-27\ 1\ 8\ 6\ -18\ 6\ 30\ 1\ 13\ 4)\,,$$
$$b_1 = (10\ 8\ 6\ 10\ 14\ 9\ 9\ 9\ 11\ 8),$$
$$b_2 = (16\ 9\ 14\ 15\ 9\ 7\ 9\ 13\ 6\ 11),$$

and $c_1 = c_2 = 20$.

Similar with the other three examples, we can easily verify that condition C_1 in Theorem 2 is satisfied because the eigenvalues of $A + Q_1 + Q_2$ are listed as follows {6.3635, 8.7851, 10.4295, 11.0365, 14.1675, 16.3225, 17.7796, 22.3263, 26.8958, 36.8937}, the eigenvalues of $A + Q_1$ are listed as follows {2.8501, 4.6238, 5.8963, 6.9364, 8.6136, 9.8704, 10.7933, 11.8424, 15.0207, 22.5531}. C_2 is satisfied because $\|D_1 A^{-1} f\| = 615.4753$ and

$$\|b_1^T D_1^{-1}\| + \sqrt{\|b_1^T D_1^{-1}\|^2 + 2|c_1|} = 11.3815.$$

where $Q_1 = D_1^T D_1$.

The canonical duality solution is

$$\bar{\sigma}_1 = 1.6894, \bar{\sigma}_2 = 0.$$

The canonical duality global maximize value is

$$P^d(\bar{\sigma}_1, \bar{\sigma}_2) = -149.6523.$$

The optimal solution of primal problem can be obtained by

$$\bar{x} = (-2.6417, -0.1485, 0.1268, -0.2010, -1.9168, -0.3058,$$
$$1.5580, -0.3112, 0.0087, -0.2020).$$

It is very easy to verify that

$$P(1.6984, 0) = -149.6523 = P^d(\bar{x}).$$

5 Conclusions

Nonconvex quadratic minimization problems with quadratic constraints are well known because they are very difficult to find out the global optimization solutions. In this paper, we have employed the canonical duality to convert them into a concave maximization dual problem over a convex set. With the presented conditions in Theorem 2, we have proved that the canonical duality problem (16) has a unique nonzero solution $\bar{\boldsymbol{\sigma}}$ in the space $\mathscr{S}_+$. With Theorem 1, we can find out the global optimization solutions $\bar{x}$ of the class of nonconvex quadratic minimization problems with quadratic constraints by (17). Several numerical examples can show that the given conditions and results in Theorems 1 and 2 are correct.

Acknowledgements The author would like to offer sincere thanks to reviewers. Their comments and suggestions are very important to improve the presentation and technical sounds. The author would like to offer his sincere thanks to Professor David Yang Gao at the Federation University Australia. Six years ago, he taught me some good ideas to deal with some difficult mathematical problems. Especially, he told me that the canonical duality theory is a very powerful principle to discover the secrets between mathematics and mechanics. This research has been supported by the National Natural Science Foundation under Grant(No. 61001200).

References

1. Anstreicher, K., Wolkowicz, H.: On Lagrangian relaxation of quadratic matrix constraints. SIAM J. Matrix Anal. Appl. **22**(1), 41–55 (2000)
2. Celis, M.R., Dennis, J.E., Tapia, R.A.: A trust region strategy for nonlinear equality constrained optimization. In: Boggs, P.T., Byrd, R.H., Schnabel, R.B. (eds.) Numerical Optimization 1994 (Proceedings of the SIAM Conference on Numerical Optimization. Boulder, CO), SIAM, Philadelphia, pp. 71–82 (1985)
3. Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland, (1976)
4. Ekeland, I.: Convexity Methods in Hamiltonian Mechanics, pp. 1–247. Springer, Berlin (1990)
5. Ekeland, I.: Legendre duality in nonconvex optimization and calculus of variations. SIAM J. Control Optim. **15**(4), 905–934 (1977)
6. Fabian, F.B., Gabriel, C.: A geometric characterization of strong duality in nonconvex quadratic programming with linear and nonconvex quadratic constraints. Math. Program. **145**(1–2), 263–290 (2014)
7. Feng, J.M., Lin, G.X., Sheu, R.L., Yong, X.: Duality and solutions for quadratic programming over single non-homogeneous quadratic constraint. J. Glob. Optim. **54**(2), 275–293 (2012)
8. Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. **17**(1/4), 127–160 (2000)
9. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications, pp. xviii+454. Kluwer Academic Publishers, Dordrecht (2000)
10. Gao, D.Y.: Duality-Mathematics. Wiley Encycl. Electr. Electr. Eng. **6**(1), 68–77 (1999)
11. Gao, D.Y: Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications. IMA J. Appl. Math. **61**(1), 199–235 (1998)
12. Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory. Appl. Mech. Rev. **50**(11), 64–71 (1997)
13. Gao, D.Y.: General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics. Meccanica **34**(1), 169–198 (1999)
14. Gao, D.Y.: Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications. Nonlinear Anal. **42**(7), 1161–1193 (2000)
15. Gao, D.Y.: Perfect duality theory and complete set of solutions to a class of global optimization. Optimization **52**(4–5), 467–493 (2003)
16. Gao, D.Y.: Complete solutions to constrained quadratic optimization problems. J. Glob. Optim. **29**(2), 377–399 (2004)
17. Gao, D.Y.: Sufficient conditions and canonical duality in nonconvex minimization with inequality constraints. J. Ind. Manag. Optim. **1**(1), 53–63 (2005)
18. Gao, D.Y.: Complete solutions and extremality criteria to polynomial optimization problems. J. Glob. Optim. **35**(1), 131–143 (2006)

19. Gao, D.Y., Strang, G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Quart. Appl. Math. **47**(3), 487–504 (1989)
20. Gao, D.Y., Strang, G.: Dual extremum principles in finite deformation elastoplastic analysis. Acta Appl. Math. **17**(1), 257–267 (1989)
21. Jeyakumar, V., Rubinov, A.M., Wu, Z.Y.: Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions. Math. Program. Ser. A **110**(3), 521–541 (2007)
22. Kirst, P., Stein, O., Steuermann, P.: An Enhanced Spatial Branch-and-Bound Method in Global Optimization with Nonconvex Constraints. IOR-Preprint, Karlsruhe Institute of Technology (KIT). 2/2013, March 2013
23. Misener, R., Floudas, C.A.: GloMIQO: global mixed-integer quadratic optimizer. J. Glob. Optim. **57**(1), 3–50 (2013)
24. Pardalos, P.M., Schnitger, G.: Checking local optimality in constrained quadratic programming is NP-hard. Oper. Res. Lett. **7**(1), 33–35 (1988)
25. Powell, M.J.D., Yuan, Y.: A trust region algorithm for equality constrained optimization. Math. Program. **49**(1), 189–211 (1990)
26. Smith, E.M.B., Pantelides, C.C.: A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. Comput. Chem. Eng. **23**(4–5), 457–478 (1999)
27. Tuy, H., Hoaiphuong, N.T.: A robust algorithm for quadratic optimization under quadratic constraints. J. Glob. Optim. **37**(4), 557–569 (2007)
28. Tuy, H., Tuan, H.D.: Generalized S-Lemma and strong duality in nonconvex quadratic programming. J. Glob. Optim. **56**(3), 1045–1072 (2013)
29. Ye, Y., Zhang, S.: New results on quadratic minimization. SIAM J. Optim. **14**(1), 245–267 (2003)
30. Yuan, Y.: On a subproblem of trust region algorithms for constrained optimization. Math. Program. **47**(1), 53–63 (1990)
31. Yuan, Y.B.: Optimal solutions to a class of nonconvex minimization problems with linear inequalities constraints. Appl. Math. Comput. **203**(1), 142–152 (2008)
32. Yuan, Y.B.: Canonical duality solution for alternating support vector machine. J. Ind. Manag. Optim. **8**(3), 611–621 (2012)
33. Yuan, Y.B., Fang, S.C., Gao, D.Y.: Perfect Duality Theory and Optimal Solutions to Nonconvex Quadratic Minimization Problems with Quadratic Constraints. J. Glob. Optim. **52**(1), 195–209 (2012)

On Minimal Distance Between Two Surfaces

Daniel Morales-Silva and David Yang Gao

Abstract This article corrects the results presented in [7] (D.Y. Gao and Wei-Chi, Yang, Optimization, 57(5), 705–714, 2008) which were challenged in [13] (M.D. Voisei, C. Zalinescu, Optimization, 60(5), 593–602, 2011). We aim to use the points of view presented in [13] (M.D. Voisei, C. Zalinescu, Optimization, 60(5), 593–602, 2011) to modify the original results and highlight that the consideration of the Gao–Strang total complementary function and the canonical duality theory are indeed quite useful for solving a class of real-world global optimization problems with nonconvex constraints. Additionally, we demonstrate how a perturbed canonical dual method can be used to solve the counter example presented in [13] (M.D. Voisei, C. Zalinescu, Optimization, 60(5), 593–602, 2011) which has multiple global minimum solutions.

1 Introduction and Primal Problem

Minimal distance problems between two surfaces arise naturally from many applications, which have been recently studied by both engineers and scientists (see [10, 11]). In this article, the problem presents a quadratic minimization problem with equality constraints: we let $\mathbf{x} := (\mathbf{y}, \mathbf{z})$ and

$$(\mathscr{P}) : \ \min \left\{ \Pi(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{z}\|^2 : \ h(\mathbf{y}) = 0, \ g(\mathbf{z}) = 0 \right\}, \tag{1}$$

where $h : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are defined by

$$h(\mathbf{y}) := \frac{1}{2}\left(\mathbf{y}^t \mathbf{A}\mathbf{y} - r^2\right), \tag{2}$$

D. Morales-Silva (✉) · D.Y. Gao
Faculty of Science and Technology, Federation University Australia,
Ballarat, VIC 3353, Australia
e-mail: d.moralessilva@federation.edu.au

D.Y. Gao
e-mail: d.gao@federation.edu.au

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3_18

$$g(\mathbf{z}) := \frac{1}{2}\alpha\left(\frac{1}{2}\|\mathbf{z}-\mathbf{c}\|^2 - \eta\right)^2 - \mathbf{f}^t(\mathbf{z}-\mathbf{c}), \tag{3}$$

in which, $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a positive definite matrix, α, r and η are positive numbers, and $\mathbf{f}, \mathbf{c} \in \mathbb{R}^n$ are properly chosen so that these two surfaces

$$\mathscr{Y}_c := \{\mathbf{y} \in \mathbb{R}^n : \; h(\mathbf{y}) = 0\}$$

and

$$\mathscr{Z}_c := \{\mathbf{z} \in \mathbb{R}^n : \; g(\mathbf{z}) = 0\}$$

are disjoint such that if $\mathbf{z} \in \mathscr{Z}_c$ then $h(\mathbf{z}) > 0$. For example, it can be proved that if $\mathbf{c} = 0$, $r > 0$, $\eta > 0.5r^2$, $\|\mathbf{f}\| < 0.5(0.5r^2 - \eta)^2/r$, $\alpha = 1$ and $\mathbf{A} = \mathbf{I}$ ($\mathbf{I}$ stands for the identity matrix of size n) then, $\mathscr{Y}_c \cap \mathscr{Z}_c = \emptyset$ and if $\mathbf{z} \in \mathscr{Z}_c$ then $h(\mathbf{z}) > 0$. Notice that the feasible set $\mathscr{X}_c = \mathscr{Y}_c \times \mathscr{Z}_c \subset \mathbb{R}^n \times \mathbb{R}^n$, defined by

$$\mathscr{X}_c = \{\mathbf{x} \in \mathbb{R}^n \times \mathbb{R}^n : \; h(\mathbf{y}) = 0, \; g(\mathbf{z}) = 0\},$$

is, in general, nonconvex.

By introducing Lagrange multipliers $\lambda, \mu \in \mathbb{R}$ to relax the two equality constraints in $\mathscr{X}_c$, the classical Lagrangian associated with the constrained problem ($\mathscr{P}$) is

$$L(\mathbf{x}, \lambda, \mu) = \frac{1}{2}\|\mathbf{y}-\mathbf{z}\|^2 + \lambda h(\mathbf{y}) + \mu g(\mathbf{z}). \tag{4}$$

By nonconvexity of the constraint $g(\mathbf{z})$, this Lagrangian may have multiple local minima. The identification of the global minimizer has been a fundamentally difficult task in global optimization. It is well-known in the field of optimization that the classical Lagrangian duality theory cannot be used alone for solving general nonconvex problems due to a so-called duality gap produced by the Lagrange multiplier method.

Canonical duality theory is a newly developed, potentially useful methodology, which is composed mainly of (i) a *canonical dual transformation*, (ii) a *complementary-dual principle,* and (iii) an associated *triality theory*. The canonical dual transformation can be used to formulate dual problems without duality gap; the complementary-dual principle shows that the canonical dual problem is equivalent to the primal problem in the sense that they have the same set of KKT points; while the triality theory can be used to identify both global and local extrema. The canonical duality theory has been successfully used for solving many global optimization problems with box/integer constraints [4] and nonconvex polynomial constraints [3, 5, 6].

The minimal distance between two nonconvex surfaces was first studied via the canonical duality theory in [7]. However, instead of the vector-valued variable $\mathbf{x} \in \mathbb{R}^n \times \mathbb{R}^n$, the two sub-vectors $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ were considered as independent variables (it is known that a bi-convex function may not be convex in the whole space), it turns out that certain global optimality condition was missing. The main goal of this paper is to re-solve this challenging problem via the canonical duality theory. In the next section, we will show how to correctly use the canonical dual transformation to convert the nonconvex constrained problem into a canonical dual problem. The missing global optimality condition in the original paper [7] is naturally obtained in Sect. 2 such that Theorems 1 and 2 proposed in [7] are represented in a correct canonical dual feasible space. Applications are illustrated in Sect. 3. Results show that in order to solve this problem by the canonical duality theory, the primal problem must have a unique global minimizer. Otherwise, a perturbation method can be used to find one of the global optimal solutions.

2 Canonical Dual Problem and Global Optimal Solution

In order to use the canonical dual transformation method, the key step is to introduce a so-called *geometrical operator* $\xi = \Lambda(\mathbf{z})$ and a *canonical function* $V(\xi)$ such that the nonconvex function

$$W(\mathbf{z}) = \frac{1}{2}\alpha \left(\frac{1}{2}\|\mathbf{z} - \mathbf{c}\|^2 - \eta \right)^2 \tag{5}$$

in $g(\mathbf{z})$ can be written in the so-called canonical form $W(\mathbf{z}) = V(\Lambda(\mathbf{z}))$. By the definition introduced in [1] (see Chap. 6 or Definition 8.1 in [6]), a differentiable function $V : \mathscr{V}_a \subset \mathbb{R} \to \mathbb{R}$ is called a *canonical function* if the duality relation $\varsigma = DV(\xi) : \mathscr{V}_a \to \mathscr{V}_a^* \subset \mathbb{R}$ is invertible. Thus, for the nonconvex function defined by (5), we let

$$\xi = \Lambda(\mathbf{z}) = \frac{1}{2}\|\mathbf{z} - \mathbf{c}\|^2,$$

then the quadratic function $V(\xi) := \frac{1}{2}\alpha(\xi - \eta)^2$ is a canonical function on the domain $\mathscr{V}_a = \{\xi \in \mathbb{R} : \ \xi \geq 0\}$ since the duality relation

$$\varsigma = DV(\xi) = \alpha(\xi - \eta) : \mathscr{V}_a \to \mathscr{V}_a^* = \{\varsigma \in \mathbb{R} : \ \varsigma \geq -\alpha\eta\}$$

is invertible. By the Legendre transformation, the conjugate function of V can be uniquely defined on $\mathscr{V}_a^*$ by

$$V^*(\varsigma) = \xi(\varsigma)\varsigma - V(\xi(\varsigma)) = \frac{1}{2\alpha}\varsigma^2 + \eta\varsigma, \tag{6}$$

where $\xi(\varsigma)$ is such that $\varsigma = DV(\xi(\varsigma))$.

It is easy to prove that the following canonical relations

$$\xi = DV^*(\varsigma) \Leftrightarrow \varsigma = DV(\xi) \Leftrightarrow V(\xi) + V^*(\varsigma) = \xi\varsigma \tag{7}$$

hold for every $(\xi, \varsigma) \in \mathscr{V}_a \times \mathscr{V}_a^*$. Thus, replacing $W(\mathbf{z})$ in the nonconvex function $g(\mathbf{z})$ by $V(\Lambda(\mathbf{z})) = \Lambda(\mathbf{z})\varsigma - V^*(\varsigma)$, the nonconvex Lagrangian $L(\mathbf{x}, \lambda, \mu)$ can be written in the Gao–Strang *total complementary function* form

$$\Xi(\mathbf{x}, \lambda, \mu, \varsigma) = \frac{1}{2}\|\mathbf{y} - \mathbf{z}\|^2 + \lambda h(\mathbf{y}) + \mu(\Lambda(\mathbf{z})\varsigma - V^*(\varsigma) - \mathbf{f}^t(\mathbf{z} - \mathbf{c})), \tag{8}$$

with $\Xi : (\mathbb{R}^n \times \mathbb{R}^n) \times \mathbb{R} \times \mathbb{R} \times \mathscr{V}_a^* \to \mathbb{R}$. Through this total complementary function, the canonical dual function can be defined by

$$\Pi^d(\lambda, \mu, \varsigma) = \Xi(\mathbf{x}, \lambda, \mu, \varsigma) \text{ where } \mathbf{x} \in \mathbb{R}^n \times \mathbb{R}^n \text{ satisfies } \nabla_{\mathbf{x}}\Xi(\mathbf{x}, \lambda, \mu, \varsigma) = 0. \tag{9}$$

In order to have the explicit form of Π^d, we need to calculate

$$\nabla_{\mathbf{x}}\Xi(\mathbf{x}, \lambda, \mu, \varsigma) = \begin{bmatrix} \mathbf{y} - \mathbf{z} + \lambda\mathbf{A}\mathbf{y} \\ \mathbf{z} - \mathbf{y} + \mu\varsigma(\mathbf{z} - \mathbf{c}) - \mu\mathbf{f} \end{bmatrix}.$$

Let the dual feasible space $\mathscr{S}_a$ be defined by

$$\mathscr{S}_a := \{(\lambda, \mu, \varsigma) \in \mathbb{R} \times \mathbb{R} \times \mathscr{V}_a^* : (1 + \mu\varsigma)(\mathbf{I} + \lambda\mathbf{A}) - \mathbf{I} \text{ is invertible}\}, \tag{10}$$

where $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the identity matrix. Clearly, if $(\lambda, \mu, \varsigma) \in \mathscr{S}_a$ we have that $\nabla_{\mathbf{x}}\Xi(\mathbf{x}, \lambda, \mu, \varsigma) = 0$ if and only if

$$\mathbf{x}(\lambda, \mu, \varsigma) = \begin{bmatrix} \mu((1 + \mu\varsigma)(\mathbf{I} + \lambda\mathbf{A}) - \mathbf{I})^{-1}(\mathbf{f} + \varsigma\mathbf{c}) \\ \mu(\mathbf{I} + \lambda\mathbf{A})((1 + \mu\varsigma)(\mathbf{I} + \lambda\mathbf{A}) - \mathbf{I})^{-1}(\mathbf{f} + \varsigma\mathbf{c}) \end{bmatrix}. \tag{11}$$

Then on $\mathscr{S}_a$, the canonical dual function Π^d can be well defined by

$$\Pi^d(\lambda, \mu, \varsigma) = \Xi(\mathbf{x}(\lambda, \mu, \varsigma), \lambda, \mu, \varsigma),$$

where $\mathbf{x}(\lambda, \mu, \varsigma)$ is given by (11).

By the fact that the stationary points of the function Ξ play a key role in the canonical duality theory, let us put in evidence what conditions the stationary points of Ξ must satisfy:

$$\nabla_{\mathbf{x}}\Xi(\mathbf{x}, \lambda, \mu, \varsigma) = \begin{bmatrix} \mathbf{y} - \mathbf{z} + \lambda\mathbf{A}\mathbf{y} \\ \mathbf{z} - \mathbf{y} + \mu\varsigma(\mathbf{z} - \mathbf{c}) - \mu\mathbf{f} \end{bmatrix} = 0, \tag{12}$$

$$\frac{\partial\Xi}{\partial\lambda}(\mathbf{x}, \lambda, \mu, \varsigma) = h(\mathbf{y}) = 0, \tag{13}$$

$$\frac{\partial\Xi}{\partial\mu}(\mathbf{x}, \lambda, \mu, \varsigma) = \Lambda(\mathbf{z})\varsigma - V^*(\varsigma) - \mathbf{f}^t(\mathbf{z} - \mathbf{c}) = 0, \tag{14}$$

$$\frac{\partial\Xi}{\partial\varsigma}(\mathbf{x}, \lambda, \mu, \varsigma) = \mu(\Lambda(\mathbf{z}) - DV^*(\varsigma)) = 0. \tag{15}$$

The following Lemma can be found in [13]. Its proof is presented for completeness.

Lemma 1 *Consider* $(\mathbf{x}, \lambda, \mu, \varsigma)$ *a stationary point of* Ξ *then the following are equivalent:*

$$\mu = 0 \;\; \Leftrightarrow \;\; \lambda = 0 \;\; \Leftrightarrow \;\; \mathbf{x} \notin \mathscr{X}_c. \tag{16}$$

Proof. Consider the contrapositive form of this statement, namely, if $\mu = 0$, then from (12) we have $\mathbf{y} = \mathbf{z}$. This implies that $\lambda \mathbf{A}\mathbf{y} = 0$ but $\mathbf{y} \neq 0$ since $\|\mathbf{y}\| = r$ by (12) and $\mathbf{A}$ is invertible, therefore $\lambda = 0$. If $\lambda = 0$, then from (12), $\mathbf{y} = \mathbf{z}$ and so $(\mathbf{y}, \mathbf{z}) \notin \mathscr{X}_c$ because $\mathscr{Y}_c \cap \mathscr{Z}_c = \emptyset$. Dually, if $\mu \neq 0$ then from (14), $\Lambda(\mathbf{z}) = DV^*(\varsigma)$ which combined together with (7) and (13) provides $\mathbf{z} \in \mathscr{Z}_c$. Since $\mathbf{y} \in \mathscr{Y}_c$, from (12), it has been proven that $\mathbf{x} \in \mathscr{X}_c$.

Now we are ready to re-introduce Theorems 1 & 2 of Gao and Yang ([7]).

Theorem 1. (Complementary-dual principle): *If* $(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma})$ *is a stationary point of* Ξ *such that* $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \in \mathscr{S}_a$ *then* $\bar{\mathbf{x}}$ *is a critical point of* $(\mathscr{P})$ *with* $\bar{\lambda}$ *and* $\bar{\mu}$ *its Lagrange multipliers,* $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma})$ *is a stationary point of* Π^d *and*

$$\Pi(\bar{\mathbf{x}}) = L(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}) = \Xi(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = \Pi^d(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}). \tag{17}$$

Proof. From Lemma 1, we must have that $\bar{\lambda}$ and $\bar{\mu}$ are different than zero, otherwise they both will be zero and $(0, 0, \varsigma) \notin \mathscr{S}_a$ for any $\varsigma \in \mathbb{R}$ which contradicts the assumption that $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \in \mathscr{S}_a$. Furthermore $\bar{\mathbf{x}} \in \mathscr{X}_c$, clearly $\bar{\mathbf{x}}$ is a critical point of $(\mathscr{P})$ with $\bar{\lambda}$ and $\bar{\mu}$ its Lagrange multipliers and

$$\Pi(\bar{\mathbf{x}}) = L(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}) = \Xi(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma}).$$

On the other hand, since $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \in \mathscr{S}_a$, Eqs. (11) and (12) are equivalent, therefore it is easily proven that

$$\frac{\partial \Xi}{\partial t}(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = \frac{\partial \Pi^d}{\partial t}(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = 0,$$

where t is either λ, μ or ς. This implies that $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma})$ is a stationary point of Π^d and

$$\Xi(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = \Pi^d(\bar{\lambda}, \bar{\mu}, \bar{\varsigma})$$

The proof is complete.

Following the canonical duality theory, in order to identify the global minimizer of $(\mathscr{P})$, we first need to look at the Hessian of Ξ:

$$\nabla^2_{\mathbf{x}} \Xi(\mathbf{x}, \lambda, \mu, \varsigma) = \begin{bmatrix} \mathbf{I} + \lambda \mathbf{A} & -\mathbf{I} \\ -\mathbf{I} & (1 + \mu\varsigma)\mathbf{I} \end{bmatrix}. \tag{18}$$

This matrix is positive definite if and only if $\mathbf{I} + \lambda\mathbf{A}$ and $(1 + \mu\varsigma)(\mathbf{I} + \lambda\mathbf{A}) - \mathbf{I}$ are positive definite (see Theorem 7.7.6 in [9]). With this, we define $\mathscr{S}_a^+ \subset \mathscr{S}_a$ as follows:

$$\mathscr{S}_a^+ := \{(\lambda, \mu, \varsigma) \in \mathscr{S}_a : \ \mathbf{I} + \lambda\mathbf{A} \succ 0 \text{ and } (1 + \mu\varsigma)(\mathbf{I} + \lambda\mathbf{A}) - \mathbf{I} \succ 0\}. \tag{19}$$

Remark 1. Comparing this canonical dual feasible set $\mathscr{S}_a^+$ with the one in the original paper by Gao & Yang (where it was denoted as $\mathscr{S}_c$ by Equation (16) in [7]), we can see clearly that the two sets are different. In [7], the total complementary function Ξ is convex in each $\mathbf{y}$ and $\mathbf{z}$ on $\mathscr{S}_c$, but it may not be convex in $\mathbf{x} = (\mathbf{y}, \mathbf{z})$.

Therefore, on the refined canonical dual feasible space $\mathscr{S}_a^+$, Theorem 2 of Gao and Yang ([7]) can be represented as the following.

Theorem 2. *Suppose that $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \in \mathscr{S}_a^+$ is a stationary point of Π^d with $\bar{\mu} \geq 0$. Then $\bar{\mathbf{x}}$ defined by (11) is the only global minimizer of Π on $\mathscr{X}_c$, and*

$$\Pi(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathscr{X}_c} \Pi(\mathbf{x}) = \max_{(\lambda, \mu, \varsigma) \in \mathscr{S}_a^+} \Pi^d(\lambda, \mu, \varsigma) = \Pi^d(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}). \tag{20}$$

Proof. Since $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \in \mathscr{S}_a^+$, it is clear from Lemma 1 that $\bar{\mu} > 0$, $\bar{\lambda} \neq 0$ and $\bar{\mathbf{x}} \in \mathscr{X}_c$, otherwise, $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = (0, 0, \bar{\varsigma}) \notin \mathscr{S}_a$ and this contradicts the assumption that $(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \in \mathscr{S}_a^+ \subset \mathscr{S}_a$. From Eq. (18) it is clear that $\Xi(\cdot, \bar{\lambda}, \bar{\mu}, \bar{\varsigma})$ is a strictly convex function, on the other hand $\bar{\mathbf{x}} \in \mathscr{X}_c$ is a stationary point of $\Xi(\cdot, \bar{\lambda}, \bar{\mu}, \bar{\varsigma})$, therefore $\bar{\mathbf{x}}$ is the only global minimizer of $\Xi(\cdot, \bar{\lambda}, \bar{\mu}, \bar{\varsigma})$. From (6), notice that V is a strictly convex function, therefore $V^*(\varsigma) = \sup\{\xi\varsigma - V(\xi) : \xi \geq 0\}$ and since $\bar{\mu} \geq 0$ by assumption we have that

$$\Xi(\mathbf{x}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \leq L(\mathbf{x}, \bar{\lambda}, \bar{\mu}), \ \forall \mathbf{x} \in \mathbb{R}^{n \times n}, \tag{21}$$

in particular, $\Xi(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = L(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu})$. Suppose now that there exists $\mathbf{x}' \in \mathscr{X}_c \setminus \{\bar{\mathbf{x}}\}$ such that

$$\Pi(\mathbf{x}') \leq \Pi(\bar{\mathbf{x}}),$$

we would have the following:

$$L(\mathbf{x}', \bar{\lambda}, \bar{\mu}) = \Pi(\mathbf{x}') \leq \Pi(\bar{\mathbf{x}}) = L(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}),$$

but because of (21) this is equivalent to

$$\Xi(\mathbf{x}', \bar{\lambda}, \bar{\mu}, \bar{\varsigma}) \leq L(\mathbf{x}', \bar{\lambda}, \bar{\mu}) \leq L(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}) = \Xi(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}, \bar{\varsigma}).$$

This contradicts the fact that $\bar{\mathbf{x}}$ is the only global minimizer of $\Xi(\cdot, \bar{\lambda}, \bar{\mu}, \bar{\varsigma})$, therefore, we must have that

$$\Pi(\bar{\mathbf{x}}) < \Pi(\mathbf{x}), \ \forall \mathbf{x} \in \mathscr{X}_c \setminus \{\bar{\mathbf{x}}\}.$$

Remark 2. Theorem 2 ensures that a stationary point in $\mathscr{S}_a^+$ will give us a unique global optimal solution of $(\mathscr{P})$. Dually, the existence of a unique global optimal solution of $(\mathscr{P})$ is necessary in order to find a stationary point of Π^d in $\mathscr{S}_a^+$. Therefore, it is clear to us that the examples provided in [13] do not contradict any of the present results.

In the next section, we will show that even if the primal problem $(\mathscr{P})$ has more than one global optimal solution, a perturbed canonical duality theory can be still used to find one of the optimal solutions.

3 Numerical Results

The graphs in this section were obtained using WINPLOT,[1] while the numerical computations using Maxima.[2]

3.1 Distance Between a Sphere and a Nonconvex Surface Defined by a Fourth-Degree Polynomial Equation

Let $n = 3$, $\eta = 2$, $\alpha = 1$, $\mathbf{f} = (2, 1, 1)$, $\mathbf{c} = (4, 5, 0)$, $r = 2\sqrt{2}$ and $\mathbf{A} = \mathbf{I}$ (Fig. 1). In this case, the sets $\mathscr{S}_a$ and $\mathscr{S}_a^+$ are given by:

$$\mathscr{S}_a = \{(\lambda, \mu, \varsigma) \in \mathbb{R}^3 : (1 + \mu\varsigma)(1 + \lambda) \neq 1\}, \tag{22}$$

$$\mathscr{S}_a^+ = \{(\lambda, \mu, \varsigma) \in \mathbb{R}^3 : 1 + \lambda > 0,\ (1 + \mu\varsigma)(1 + \lambda) > 1\}. \tag{23}$$

Using Maxima, we can find the following stationary point of Π^d in $\mathscr{S}_a^+$:

$$(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = (0.9502828628898, 1.06207786194864, 0.30646555192966).$$

Then the global minimizer of $(\mathscr{P})$ is given by Eq. (11):

$$\bar{\mathbf{y}} = \begin{pmatrix} 2.161477484004744 \\ 1.696777196962463 \\ 0.67004643869564 \end{pmatrix}, \quad \bar{\mathbf{z}} = \begin{pmatrix} 4.215492495576614 \\ 3.309195489378083 \\ 1.306780086728456 \end{pmatrix}.$$

[1]R. Parris: Peanut Software Homepage: http://math.exeter.edu/rparris/, Version 1.54 (2012).

[2]See Maxima.sourceforge.net. Maxima, a Computer Algebra System. Version 5.22.1 (2010). http://maxima.sourceforge.net/.

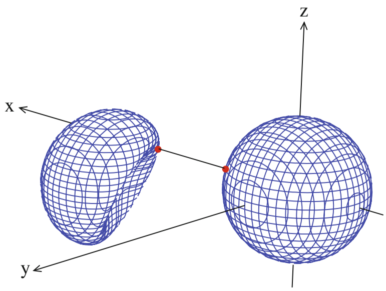


Fig. 1 Distance between a sphere and a nonconvex surface defined by a polynomial equation

3.2 Distance Between an Ellipsoid and a Nonconvex Surface Defined by a Fourth-Degree Polynomial Equation

Let $n = 3$, $\eta = 2$, $\alpha = 1$, $\mathbf{f} = (-2, -2, 1)$, $\mathbf{c} = (-4, -5, 0)$, $r = 2\sqrt{2}$ (Fig. 2) and

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{bmatrix}.$$

Using Maxima, we can find the following stationary point of Π^d in $\mathscr{S}_a^+$:

$$(\bar{\lambda}, \bar{\mu}, \bar{\varsigma}) = (0.84101802234162, 1.493808342458642, 0.12912817444352).$$

To put in evidence that this stationary point is in fact in $\mathscr{S}_a^+$, notice that the eigenvalues of $\mathbf{A}$ are given by:

$$\begin{aligned} \beta_1 &= \tfrac{4}{\sqrt{3}} \cos\left(\tfrac{4\pi}{3} + \tfrac{\theta}{3}\right) + 4 \approx 3.460811127 \\ \beta_2 &= \tfrac{4}{\sqrt{3}} \cos\left(\tfrac{2\pi}{3} + \tfrac{\theta}{3}\right) + 4 \approx 2.324869129 \\ \beta_3 &= \tfrac{4}{\sqrt{3}} \cos\left(\tfrac{\theta}{3}\right) + 4 \qquad\quad \approx 6.214319743, \end{aligned}$$

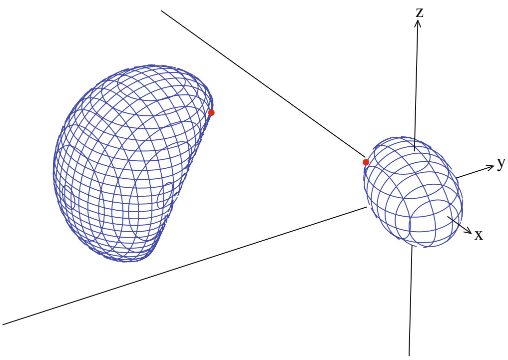


Fig. 2 Distance between an ellipsoid and a nonconvex surface defined by a polynomial equation

with $\theta = \cos^{-1}\left(\frac{3\sqrt{3}}{8}\right)$. Then, the matrices $\mathbf{I} + \bar{\lambda}\mathbf{A}$ and $(1 + \bar{\mu}\bar{\varsigma})(\mathbf{I} + \bar{\lambda}\mathbf{A}) - \mathbf{I}$ are similar to

$$\begin{bmatrix} 3.910604529727413 & 0 & 0 \\ 0 & 2.955256837074665 & 0 \\ 0 & 0 & 6.226354900456345 \end{bmatrix}$$

and

$$\begin{bmatrix} 3.664931769065526 & 0 & 0 \\ 0 & 2.525304438283014 & 0 \\ 0 & 0 & 6.42737358375643 \end{bmatrix}$$

respectively. Finally, the global minimizer of $(\mathscr{P})$ is given by Eq. (11):

$$\bar{\mathbf{y}} = \begin{pmatrix} -1.121270493506938 \\ -0.83025443673537 \\ 0.66262025515374 \end{pmatrix}, \quad \bar{\mathbf{z}} = \begin{pmatrix} -4.091279940255224 \\ -4.009023330835817 \\ 1.807730500535487 \end{pmatrix}.$$

3.3 Example Given in [13]

Let $n = 2, \alpha = \eta = 1, \mathbf{c} = (1, 0), \mathbf{f} = \left(\frac{\sqrt{6}}{96}, 0\right), r = 1$ and $\mathbf{A} = \mathbf{I}$. As it was pointed out in [13], there are no stationary points in $\mathscr{S}_a^+$. Under the present conditions of Theorem 2, this is expected since the problem has more than one solution (see Fig. 3). In [13], the following was found to be one of the global minimizers of $(\mathscr{P})$:

$$\bar{\mathbf{y}} = \begin{pmatrix} 0.5872184947 \\ 0.8094284647 \end{pmatrix}, \quad \bar{\mathbf{z}} = \begin{pmatrix} 1.012757759 \\ 1.395996491 \end{pmatrix}.$$

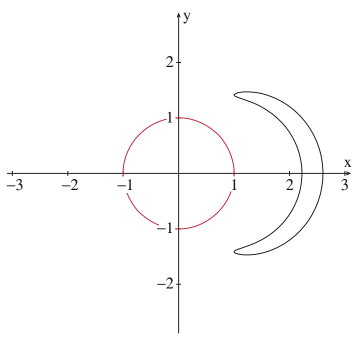


Fig. 3 Example given in [13]

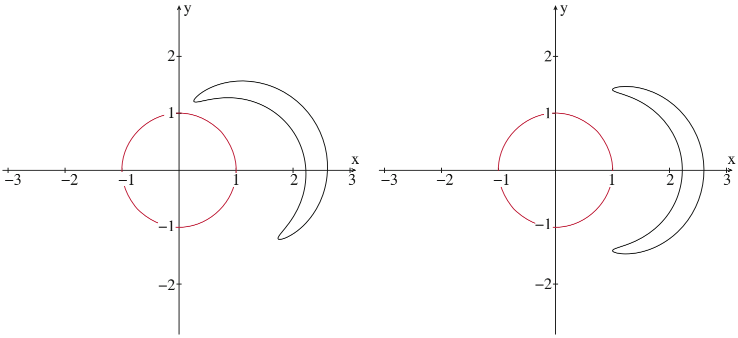


Fig. 4 Perturbations of Example given in [13] with $k = 64$ (*left*) and $k = 10^5$ (*right*)

Notice that $\mathscr{S}_a$ and $\mathscr{S}_a^+$ are defined as in Eqs. (22) and (23).
In order to solve this problem, we will introduce a perturbation. Instead of the given $\mathbf{f}$, we will consider $\mathbf{f}_k = \left(\frac{\sqrt{6}}{96}, \frac{1}{k}\right)$ for $k > 50$ (Fig. 4).
The following table summarizes the results for different values of k.

k	$(\bar{\lambda}_k, \bar{\mu}_k, \bar{\varsigma}_k) \in \mathscr{S}_a^+$	$\bar{\mathbf{x}}_k = (\bar{\mathbf{y}}_k, \bar{\mathbf{z}}_k)$
64	(0.2284381,5.319007,−0.0219068)	$\bar{\mathbf{y}} = \begin{pmatrix} 0.2250312 \\ 0.9743515 \end{pmatrix}, \bar{\mathbf{z}} = \begin{pmatrix} 0.2764370 \\ 1.1969306 \end{pmatrix}$
1000	(0.6926569,16.01863,−0.0248297)	$\bar{\mathbf{y}} = \begin{pmatrix} 0.5656039 \\ 0.8246770 \end{pmatrix}, \bar{\mathbf{z}} = \begin{pmatrix} 0.9573734 \\ 1.3958953 \end{pmatrix}$
10000	(0.7214940,16.42599,−0.0254434)	$\bar{\mathbf{y}} = \begin{pmatrix} 0.5850814 \\ 0.8109745 \end{pmatrix}, \bar{\mathbf{z}} = \begin{pmatrix} 1.0072142 \\ 1.3960878 \end{pmatrix}$
100000	(0.7243521,16.46345,−0.0255083)	$\bar{\mathbf{y}} = \begin{pmatrix} 0.5870050 \\ 0.8095833 \end{pmatrix}, \bar{\mathbf{z}} = \begin{pmatrix} 1.0122034 \\ 1.3960066 \end{pmatrix}$

Remark 3. The key idea of the perturbation method is to destroy certain symmetry in the original problem such that its canonical dual problem has a unique solution in $\mathscr{S}_a^+$. The combination of the linear perturbation method and canonical duality theory for solving nonconvex optimization problems was first proposed in [12] with successful applications in solving some NP-complete problems [14]. High-order perturbation methods for solving integer programming problems were discussed in [4].

4 Concluding Remarks and Future Research

We have demonstrated the correct application of the canonical duality theory for solving a nonconvex constrained global optimization problem.

Actually, the Gao–Strang total complementary function Ξ and the canonical dual form Π^d presented in this paper are special cases of the *sequential canonical dual* *transformation* developed in [1] (Chap. 4). To see this fact, let us introduce the first-level geometrical operator $\boldsymbol{\xi}_0 = \Lambda_0(\mathbf{x}) = \{h(\mathbf{y}), g(\mathbf{z})\} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^2$ and the associated canonical function

$$V_0(\boldsymbol{\xi}_0) = \begin{cases} 0 & \text{if } \boldsymbol{\xi}_0 = \{0, 0\} \\ +\infty & \text{otherwise} \end{cases}$$

Its Fenchel conjugate can be defined easily by

$$V_0^\sharp(\boldsymbol{\xi}_0^*) = \sup\{\boldsymbol{\xi}_0^t \boldsymbol{\xi}_0^* - V(\boldsymbol{\xi}_0) : \ \boldsymbol{\xi}_0 \in \mathbb{R}^2\} = 0 \ \ \forall \boldsymbol{\xi}_0^* \in \mathbb{R}^2.$$

Let the first-level canonical dual variable $\boldsymbol{\xi}_0^* = \{\lambda, \mu\}$, the associated total complementary function $\Xi_0(\mathbf{x}, \boldsymbol{\xi}_0)$ is

$$\Xi_0(\mathbf{x}, \boldsymbol{\xi}_0^*) = (\Lambda_0(\mathbf{x}))^t \boldsymbol{\xi}_0^* - V_0^\sharp(\boldsymbol{\xi}_0^*) + \Pi(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{z}\|^2 + \lambda h(\mathbf{y}) + \mu g(\mathbf{z}), \quad (24)$$

which is exactly the classical Lagrangian $L(\mathbf{x}, \lambda, \mu)$ defined by (4). Since the first-level geometrical operator $\Lambda_0(\mathbf{x})$ is nonconvex, a second-level (partial) geometrical operator $\xi = \Lambda(\mathbf{z})$ and the associated canonical function $V(\xi)$ are introduced in Sect. 2 such that the vector-valued nonconvex geometrical operator $\Lambda_0(\mathbf{x}) = \{h(\mathbf{y}), V(\Lambda(\mathbf{z}))\}$ is represented in a canonical form.

For inequality constraints, say $\boldsymbol{\xi}_0 \leq 0$, the canonical function

$$V_0(\boldsymbol{\xi}_0) = \begin{cases} 0 & \text{if } \boldsymbol{\xi}_0 \leq 0 \\ +\infty & \text{otherwise} \end{cases} \quad (25)$$

is convex. Its Fenchel conjugate

$$V_0^\sharp(\boldsymbol{\xi}_0^*) = \begin{cases} 0 & \text{if } \boldsymbol{\xi}_0^* \geq 0 \\ +\infty & \text{otherwise} \end{cases} \quad (26)$$

is also convex. In this case, the generalized canonical duality relations (see [6])

$$\boldsymbol{\xi}_0^* \in \partial V_0(\boldsymbol{\xi}_0) \quad \Leftrightarrow \quad \boldsymbol{\xi}_0 \in \partial V^\sharp(\boldsymbol{\xi}_0^*) \quad \Leftrightarrow \quad \boldsymbol{\xi}_0^t \boldsymbol{\xi}_0^* = V_0(\boldsymbol{\xi}) + V^\sharp(\boldsymbol{\xi}_0^*) \quad (27)$$

are equivalent to the well-known KKT conditions:

$$\boldsymbol{\xi}_0 \leq 0, \quad \boldsymbol{\xi}_0^* \geq 0, \quad \boldsymbol{\xi}_0^t \boldsymbol{\xi}_0^* = 0. \quad (28)$$

Actually, it can be proved that for equality constraints, Lemma 1 is also a special case of the generalized canonical duality conditions. Therefore, as a unified methodology, the canonical duality theory covers both the Lagrange multiplier method and KKT theory as two special cases.

The canonical min–max duality statement in the triality theory, i.e., Theorem 2 in this paper, can be used to identify the global minimizer; while the double-min and double-max duality statements of the triality theory can be used to identify both the biggest local min and local max, respectively. Recently, an open question left on the double-min duality theory has been solved (see [8]). The sequential canonical duality theory can be used not only for solving high-order polynomial optimization and nonconvex variational problems (see [1] Chap. 4 and recent review article [3]), but also for modeling multiscale complex systems. Indeed, the so-called von Kármán paradox was discovered by the canonical duality theory, which leads to several large deformed beam models in engineering mechanics ([1], Chap. 7). By the fact that the canonical duality theory was developed from nonconvex mechanics, where the *objective function* has its own physical meaning (see Definition 6.1.2 in [1], page 288), to correctly understand this theory with real-world applications may need some necessary background in continuum physics and systems theory. Therefore, it is important to refine the canonical duality theory in order to bridge the existing gap between mathematical optimization and engineering sciences, which will be a future work.

To summarize, we have the following conclusions:

- Theorem 1 (the complementary-dual principle) holds correctly on $\mathscr{S}_a$.
- Theorem 2 (the first statement of the triality theory) holds on the refined canonical dual feasible space $\mathscr{S}_a^+$.
- The examples presented in [13] do not contradict the canonical duality theory presented in this work.
- The combination of the perturbation and the canonical duality theory is an important method for solving nonconvex optimization problems which have more than one global optimal solution (see also [15]).
- If Π^d has a stationary point in $\mathscr{S}_a^+$ (with $\mu \geq 0$), then Π has only one global minimizer. The reverse of this statement remains an open question, as well as the condition that ensures the existence of the stationary point in $\mathscr{S}_a^+$.
- The total complementary function Ξ is indeed useful for global optimization problems in real-world applications, at least for the problem $(\mathscr{P})$ studied in this work. It is worth to continue studying this topic both theoretically and numerically in order to develop efficient algorithms for solving challenging problems with general nonconvex constraints.

Acknowledgements This research is supported by US Air Force Office of Scientific Research under the grants FA2386-16-1-4082 and FA9550-17-1-0151. Comments and suggestions from several anonymous referees are sincerely acknowledged.

References

1. Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory. Methods and Applications. Springer, New York (2000)
2. Gao, D.Y.: Solutions and optimality criteria to box constrained non-convex minimization problem. Ind. Manage. Optim. **3**(3), 293–304 (2007)

3. Gao, D.Y.: Canonical duality theory: unified understanding and generalized solution for global optimization problems. Comput. Chem. Eng. **33**, 1964–1972 (2009)
4. Gao, D.Y., Ruan, N.: Solutions to quadratic minimization problems with box and integer constraints. J. Glob. Optim. **47**, 463–484 (2010)
5. Gao, D.Y., Ruan, N., Sherali, H.: Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality. J. Glob. Optim. **45**, 473–497 (2009)
6. Gao, D.Y., Sherali, H.D.: Canonical duality: connection between nonconvex mechanics and global optimization. In: Advances in Application Mathematics and Global Optimization, pp. 249-316. Springer, Berlin (2009)
7. Gao, D.Y.: Yang, Wei-Chi: Minimal distance between two non-convex surfaces. Optimization **57**(5), 705–714 (2008)
8. Gao, D.Y., Wu, C.: On the triality theory for a quartic polynomial optimization problem. J. Ind. Manag. Optim. **8**(1), 229–242 (2012)
9. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1985)
10. Johnson, D.E., Cohen, E.: A framework for efficient minimum distance computations. In: IEEE Proceedings International Conference on Robotics and Automation, Leuven, Belgium, pp. 3678–3684 (1998)
11. Patoglu, V.; Gillespie, R.B.: Extremal distance maintenance for parametric curves and surfaces. In: Proceedings International Conference on Robotics and Automation, Washington, DC, pp. 2817–2823 (2002)
12. 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)
13. Voisei, M.D., Zalinescu, C.: A counter-example to "Minimal distance between two non-convex surfaces". Optimization **60**(5), 593–602 (2011)
14. Wang, Z.B., Fang, S.C., Gao, D.Y., Xing, W.X.: Canonical dual approach to solving the maximum cut problem. J. Glob. Optim. **54**, 341–352 (2012)
15. Wu, C., Gao, D.Y.: Canonical primal-dual method for solving nonconvex minimization problems. In: Gao, D.Y., Latorre, V., Ruan, N. (eds.) Advances in Canonical Duality Theory. Springer, Berlin

Index

D.Y. Gao et al. (eds.), *Canonical Duality Theory*, Advances in Mechanics and Mathematics 37, DOI 10.1007/978-3-319-58017-3

T

V