THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

This thesis in electronic version is provided to the Library by the author. In the case where its contents is different from the printed version, the printed version shall prevail.

POLYNOMIAL OPTIMIZATION AND ITS APPLICATIONS

ZHANG XINZHEN

Ph.D

The Hong Kong

Polytechnic University

2010

The Hong Kong Polytechnic University
Department of Applied Mathematics

Polynomial Optimization and its Applications

by Zhang Xinzhen

A thesis submitted in partial fulfilment of
the requirements for the degree of Doctor of Philosophy

May 2010

CERTIFICATE OF ORIGINALITY

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of a university or other institute of higher learning, except where due acknowledgment is made in the text.

XIN ZHEN ZHANG

Abstract

The main purposes of this thesis are to solve some polynomial optimization problems and to find their applications. The polynomial optimization problems involved in this thesis include the cubic spherical optimization problems and bi-quadratic optimization problems.

The main contributions of this thesis are as follows:

In this thesis, we first consider a new model, the truncated generalized diffusion tensor imagine (GDTI) model in medical engineering, which overcomes the drawback that water movement in biological tissues often shows non-Gaussian diffusion behavior. In the GDTI model, polynomial associated with even order tensors reflects the magnitude of the signal, while polynomial associated with odd order tensors reflects the phase of the signal. Moreover, we use the apparent skewness coefficient (ASC) value to measure the phase of non-Gaussian signals. We present some properties of related tensors and propose a direct computation method for calculating the ASC value.

We discuss the general cubic spherical optimization problems, which include the cubic one-spherical/two-spherical/three-spherical optimization problems. For these three problems, we present their NP-hardnesses and discuss the complexity results of some special cases. For the NP-hardness cases, some approximation solution methods for them are established.

Then we study the bi-quadratic optimization problem over two unit spheres. At first, the problem is equivalently transformed into computing the largest M-eigenvalue of related tensor. Based on the reformulation, power method for computing the largest eigenvalue of a matrix is modified to compute the largest M-eigenvalue of a tensor. To get a good approximation of the largest M-eigenvalue of a tensor, we introduce

an initialization technique. The given numerical experiments show that the modified method performs well.

Finally, we discuss the bi-quadratic optimization problems with quadratic constraints. First, we generalize the SDP relaxation scheme for approximately solving NP-hard quadratic optimization to solve bi-quadratic optimization problems. Then we show that each r-bound approximation solution of the relaxed bi-linear SDP problems can be used to generate in randomized polynomial time an $\mathcal{O}(r)$-approximation solution for bi-quadratic optimization problems. Furthermore, we show that when the number of constraints is not larger than two, bi-quadratic optimization problems are equivalent to their corresponding SDP relaxation problems, which generalizes the result in [33]. Then, we present some approximation solutions with some quality bounds for the bi-quadratic maximization model with some assumptions. For bi-quadratic optimization problems with two constraints, some approximation solutions are established. Finally, we extend the results from real cases to complex cases.

Underlying papers

This thesis is based on the following four papers written by the author during the period of stay in Department of Applied Mathematics, The Hong Kong Polytechnic University as a graduate student.

1. Y.J. Wang, L.Q. Qi and X.Z. Zhang, "A practical method for computing the largest M-eigenvalue of a fourth order partially symmetric tensor," *Numerical Linear Algebra with Applications,* 16 (2009) 589-601.

2. X.Z. Zhang, C. Ling and L.Q. Qi, "Semidefinite relaxation bounds for bi-quadratic optimization problems with quadratic constraints, " *to appear in: Journal of global optimization.*

3. X.Z. Zhang, C. Ling, L.Q. Qi, Ed.X. Wu, "The measure of diffusion skewness and kurtosis in magnetic resonance imaging," *Pacific Journal of Optimization,* 6 (2010) 391-404.

4. X.Z. Zhang, L.Q. Qi and Y.Y. Ye, "The cubic spherical optimization problems," *2009, Technical Report.*

In addition, some other papers are listed which are written by the author during the period of her Ph.D study.

1. I.M. Bomze, C. Ling, L.Q. Qi and X.Z. Zhang, "Standard bi-quadratic optimization problems and unconstrained polynomial reformulations," *Submitted.*

2. C. Ling, X.Z. Zhang and L.Q. Qi, "Semidefinite relaxation approximation for multivariate bi-quadratic optimization with quadratic constraints," *Submitted.*

3. L.Q. Qi, Y. Xu, Y.X. Yuan and X.Z. Zhang, "A cone constrained convex program: structure and algorithms. "*Submitted.*

Acknowledgements

First and foremost, I wish to express my deep gratitude to my supervisor Professor Liqun Qi for introducing me to the research topics of this thesis. During the past three years, Prof. Qi, as a brilliant mathematician, has never ceased to inspire; as a kind teacher, has never lacked encouraging words; as a dedicated scholar, has never yielded to challenges. He has been an exemplary role model for my work and study in future. Especially I wish to thank Professor Liqun Qi for his careful reading and providing useful comments and suggestions to all the papers that have been written over the past three years. It has been an invaluable experience for me to pursue my PhD degree under the supervision of Prof. Qi in the Department of Applied Mathematics at The Hong Kong Polytechnic University.

I would like to express my sincere appreciation to Professor Chen Ling (Hangzhou Dianzi University, China) for his comments and suggestions, constant encouragement and help. I had many fruitful and stimulating discussions with him. I also wish to thank Professor Yiju Wang (Qufu Normal University, China) for his help, encouragement and guidance. It has been a precious experience for me to study for my PhD degree under their help.

I am greatly indebted to Professor Yinyu Ye (Stanford University, USA), Professor Ed Xuekui Wu (The University of Hong Kong, Hong Kong), Professor I.M. Bomze (University of Vienna, Austria) and Professor Yaxiang Yuan (Chinese Academy of Sciences, China), who gave me much help and guidance. I also wish to thank Professor Soonyi Wu (National Cheng-Kung University, Taiwan) for his help and encouragement. I wish to thank Dr Xun Li (The Hong Kong Polytechnic University, Hong Kong) for his help. In addition, I would like to thank a number of editors and anonymous referees from some journals, who have given me valuable comments and suggestions which have

contributed to a better presentation of the papers. This thesis is based on the results obtained in these papers.

I have fond memory of many colleagues and my friends who have helped me in one way or the others during my stay at The Hong Kong Polytechnic University. In particular, to Dr. Deren Han, Dr. Liping Zhang, Dr. Gaohang Yu, Mr. Yi Xu, I am grateful to them for their friendship. Thanks are also due to the Research Committee of The Hong Kong Polytechnic University for offering me financial support over the nearly three years. I shall always treasure the stimulating research environment and the excellent research facilities of the Department of Applied Mathematics of The Hong Kong Polytechnic University.

Last but not least, I would like to thank my parents. They gave me the whole life and have always encouraged me to keep on studying. Many thanks to my husband, Yuanhua Ni, and my son, Zhiheng Ni. I thank them for their love, encouragement, support, and tolerance. I thank them for always being by my side, even during those moments when I was astray and down.

Contents

Chapter 1

Introduction

Polynomial optimization problems have wide applications such as in independent component analysis [11], blind channel equalization in digital communication [40], sensor localization [8], strong ellipticity condition in solid mechanics [29, 30, 61, 63, 69], entanglement problem in quantum physics [17] and so on. The theory and applications of (constrained) polynomial optimization problems have attracted more and more interests. This chapter gives some preliminarys and motivations of the polynomial optimization problems we will consider in the thesis.

1.1 The Cubic Spherical Optimization Problems

1.1.1 The Three-dimensional Cubic Spherical Optimization Problems

In medical engineering, a popular magnetic resonance imaging (MRI) model is said to be diffusion tensor imaging (DTI) model. DTI model has wide applications in biological and clinical applications [1]. For example, it may be used to study the properties of water molecule diffusion in the brain, particularly for white matter fibers. Such properties can be used to detect abnormalities and diseases in such tissues [2, 3].

In DTI model, a diffusion tensor D involved is a three dimensional symmetric ma-

trix, that is, $D \in S^{3\times 3}$. It is well known that D has six independent elements, which is obtained by MRI techniques. Based upon the obtained D, the medical engineering researchers can calculate some characteristic quantities, such as three eigenvalues, the mean diffusivity (MD) and the fractional anisotropy (FA), and so on. These quantities are rotationally invariant, that is, independent from the choice of the laboratory coordinate system.

However, in DTI model, one needs assume a perfect Gaussian distribution for the water molecule movement [1]. In fact, water in biological structures often shows non-Gaussian diffusion behavior, which affects the use of the DTI model. To overcome this drawback, some new MRI models were introduced by medical engineering researchers.

One model is said to be diffusion kurtosis imaging (DKI) model, established in [27, 36]. For a particular direction specified by a unit vector x, based on the denotations that a gradient strength g, a pulse duration δ and a time interval Δ between the centers of the diffusion sensitizing gradient pulse, DKI model has the following form

$$\ln(\frac{S(b)}{S(0)}) = -b \sum_{i_1,i_2=1}^{3} D^{app}_{i_1 i_2} x_{i_1} x_{i_2} + \frac{1}{6} b^2 (\frac{1}{3} \sum_{i_1=1}^{3} D^{app}_{i_1 i_1})^2 \sum_{i_1,i_2,i_3,i_4=1}^{3} W^{app}_{i_1 i_2 i_3 i_4} x_{i_1} x_{i_2} x_{i_3} x_{i_4},$$

where matrix D^{app} is the apparent diffusion coefficient, the fourth order three dimensional symmetric tensor W^{app} is the apparent diffusional kurtosis (DK) tensor. The parameter b is given by the usual expression $b = (\gamma\delta g)^2(\Delta - \frac{\delta}{3})$, where γ is the gyromagnetic ratio. For this model, D-eigenvalues for a DK tensor was introduced in [58]. They showed that the largest, the smallest and the average D-eigenvalues of a DK tensor correspond with the largest, the smallest and the average apparent kurtosis coefficients (AKC) of a water molecule in the space, respectively. Some computational methods for related anisotropy value of AKC were presented there. Later, Han *et al.* studied the properties of the extreme values and directions associated to DK tensor in [22, 55].

Another model was introduced by Liu *et al.* in [34]. They introduced the so-called generalized diffusion tensors imaging (GDTI) model, stated as following, to characterize

the non-Gaussian diffusion of the water molecules in tissues

$$\begin{aligned}\ln\left(\frac{S(b)}{S(0)}\right) &= -\sum_{i_1,i_2=1}^{3} D^{(2)}_{i_1i_2} b^{(2)}_{i_1i_2} + \sum_{i_1,\cdots,i_4=1}^{3} D^{(4)}_{i_1\cdots i_4} b^{(4)}_{i_1\cdots i_4} + \cdots \\ &\quad +(-1)^n \sum_{i_1,\cdots,i_{2n}=1}^{3} D^{(2n)}_{i_1\cdots i_{2n}} b^{(2n)}_{i_1\cdots i_{2n}} + \cdots \\ &\quad +j\left(-\sum_{i_1,i_2,i_3=1}^{3} D^{(3)}_{i_1i_2i_3} b^{(3)}_{i_1i_2i_3} + \sum_{i_1,\cdots,i_5=1}^{3} D^{(5)}_{i_1\cdots i_5} b^{(5)}_{i_1\cdots i_5} + \cdots \right. \\ &\quad \left. +(-1)^n \sum_{i_1,\cdots,i_{2n+1}=1}^{3} D^{(2n+1)}_{i_1\cdots i_{2n+1}} b^{(2n+1)}_{i_1\cdots i_{2n+1}} + \cdots \right).\end{aligned} \tag{1.1.1}$$

Here j is the square root of -1, and $D^{(n)}$ $(n \geq 2)$ are n-th order coefficient tensors which can be determined by using some common statistical methods such as the least square estimate method and Monte-Carlo simulations. It is not difficult to see that the tensors $b^{(n)}$ $(n \geq 2)$ in (1.1.1) are functions of the direction, the magnitude and the timing of the diffusion-encoding gradients. More precisely, if the magnetic field gradient is a constant vector over the considered time, by [34], element $b^{(n)}_{i_1i_2\cdots i_n}$ of tensor $b^{(n)}$ can be written as

$$b^{(n)}_{i_1i_2\cdots i_n} = (\gamma g \delta)^n \left(\Delta - \frac{n-1}{n+1}\delta\right) x_{i_1}x_{i_2}\cdots x_{i_n}, \quad i_1, i_2, \cdots, i_n = 1, 2, 3. \tag{1.1.2}$$

From (1.1.1), it is obvious that in the case of Gaussian diffusion, all the tensors $D^{(n)}$ of orders higher than two are zero. For non-Gaussian diffusion, however, those higher order tensors become significant and it is important to recognize that the higher order terms in (1.1.1) have to be considered in such situations.

Furthermore, from (1.1.1), we can see that the real part of the logarithmic signal is solely determined by the even order tensors and only affects the magnitude of the signal, while the imaginary part is completely governed by odd order tensors and only affects the phase of the signal. This shows that the DTI model may fail to identify the underlying structure [35] with the diffusion behavior of the non-Gaussian signal with the asymmetry. This point is even clearer for the one modeled by Phantom 4 in [34]. We refer readers to [22, 26, 55, 56] and references therein for non-Gaussian diffusion with the symmetry.

1.1.2 The General Cubic Spherical Optimization Problems

In the previous subsection, variables of polynomial involved in DKI and GDTI models are three dimensional. But in some other applications, such as signal processing and independent component analysis, the cubic spherical optimization problems with higher dimensional variables will be involved. Motivated by this, we are ready to consider the general cubic spherical optimization problems, which include the cubic one-spherical/two-spherical/three-spherical optimization problems.

It is well known that homogeneous multivariate polynomials have simple expressions via tensors. Hence, we give some definitions which are involved in cubic spherical optimization problems.

A tensor $\mathcal{A} \in \Re^{n\times n\times n}$ is said to be symmetric, if its element $\mathcal{A}_{ijk}$ is invariant under any permutation of indices (i, j, k), [31, 32, 51]. Tensor $\mathcal{B} \in \Re^{n\times n\times q}$ is called partially symmetric with respect to the first two indices in the sense that $\mathcal{B}_{ijk} = \mathcal{B}_{jik}$, for all $i, j = 1, 2, \cdots, n$, and $k = 1, 2, \cdots, q$.

Based on these conceptions, the cubic one-spherical/two-spherical/three-spherical optimization problems can be written as follows, respectively

$$\begin{array}{ll} \min\limits_{x\in\Re^n} & g_1(x) := \mathcal{A}x^3 = \sum\limits_{i,j,k=1}^{n} \mathcal{A}_{ijk}x_ix_jx_k \\ \text{s.t.} & \|x\| = 1, \end{array} \tag{1.1.3}$$

$$\begin{array}{ll} \min\limits_{x\in\Re^n, z\in\Re^q} & g_2(x, z) := \mathcal{B}x^2z = \sum\limits_{i,j=1}^{n}\sum\limits_{k=1}^{q} \mathcal{B}_{ijk}x_ix_jz_k \\ \text{s.t.} & \|x\| = 1, \quad \|z\| = 1 \end{array} \tag{1.1.4}$$

and

$$\begin{array}{ll} \min\limits_{x\in\Re^n, y\in\Re^p, z\in\Re^q} & g_3(x, y, z) := \mathcal{C}xyz = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{p}\sum\limits_{k=1}^{q} \mathcal{C}_{ijk}x_iy_jz_k \\ \text{s.t.} & \|x\| = 1, \quad \|y\| = 1, \quad \|z\| = 1, \end{array} \tag{1.1.5}$$

where $\mathcal{C} \in \Re^{n\times p\times q}$ is a third order tensor.

These three problems arise from the best rank-one approximation to the third order symmetric tensor $\mathcal{A}$, the third order partially symmetric tensor $\mathcal{B}$, and the third order tensor $\mathcal{C}$, respectively. The best rank-one approximation problem has many applications in signal and image processing, wireless communication systems, and independent component analysis, see [14, 12, 31, 32, 37, 51, 74] for details.

In [57], Qi *et al.* presented a Z-eigenvalue method by direct computation for solving (1.1.3) with $n = 3$. In [74], for solving (1.1.5), they proposed some methods, including the generalized Rayleigh-Newton iteration method, alternating least squares method and Jacobi Gauss-Newton iteration method. These methods approximately solved (1.1.5), but did not guarantee the convergence even locally and quality bound.

Furthermore, Nesterov in [43] showed that (1.1.3) is NP-hard to solve. Thus, it is not expected that there exist efficient exact methods to find global minimizers of the NP-hard problems. It is important to find approximation algorithms for solving NP-hard problems. To characterize the approximation algorithms, some quality measures of approximation ratio are introduced as follows.

Definition 1.1.1 *Let $0 \leq \epsilon < 1$ and $\mathfrak{A}$ be an approximation algorithm for the minimization problem $\boldsymbol{P}$. We say $\mathfrak{A}$ is a $(1-\epsilon)$-approximation algorithm if for any instance of $\boldsymbol{P}$ the algorithm $\mathfrak{A}$ returns a feasible solution with its corresponding objective value p such that*

$$p - p_{min} \leq \epsilon(p_{max} - p_{min}),$$

where $p_{max}(resp.\ p_{min})$ is the maximum (resp. minimum) value of the objective of $\boldsymbol{P}$.

Furthermore, we say that the problem $\boldsymbol{P}$ has a polynomial time approximation scheme (PTAS) if for every $\epsilon > 0$, there exists a $(1-\epsilon)$-approximation algorithm.

Definition 1.1.2 *The problem has an r-bound approximation solution for the given minimization model $\boldsymbol{P}$, if there is an algorithm $\mathfrak{A}$ whose complexity is polynomial such that when applied to $\boldsymbol{P}$, it returns a feasible solution with objective value p such that*

$$\begin{cases} rp \leq p_{\min} \leq p, & \text{if } p_{\min} \geq 0, \\ p_{\min} \leq p \leq rp_{\min}, & \text{if } p_{\min} < 0, \end{cases}$$

where $p_{\min}$ is the minimum value of the problem and $0 < r \leq 1$. The feasible solution is said to be an r-bound approximation solution of the minimization model. The algorithm $\mathfrak{A}$ is said to be an r-bound approximation algorithm.

For convenience of notations, we call feasible solution appeared in Definition 1.1.1 as $(1-\epsilon)-$approximation solution.

In a similar way, the definitions of approximation ratio can be presented for maximization problem and are omitted here. Based on these definitions, the approximation algorithms for considered NP-hard problems will be studied.

1.2 Bi-quadratic Optimization Problems

Bi-quadratic optimization problems studied in the thesis include the following problems: bi-quadratic optimization problem over two unit spheres and bi-quadratic optimization problems with quadratic constraints.

1.2.1 Bi-quadratic Optimization Problem over Unit Spheres

Bi-quadratic optimization problem over two unit spheres has the following form:

$$\begin{array}{rl} \max & f(x,y) = \mathcal{F}xxyy \\ \text{s.t.} & x^\top x = 1,\ y^\top y = 1 \\ & x \in \Re^m, y \in \Re^n \end{array} \tag{1.2.6}$$

where $\mathcal{F}$ is a fourth order $(m \times m \times n \times n)$-dimensional partially symmetric tensor. Here, a tensor $\mathcal{F}$ is said to be fourth order partially symmetric if

$$\mathcal{F}_{ijkl} = \mathcal{F}_{jikl} = \mathcal{F}_{ijlk}, \quad \text{for all} \ \ i,j = 1,\cdots,m, \ \ \text{and} \ \ k,l = 1,\cdots,n.$$

The problem (1.2.6) arises from the nonlinear elastic materials analysis and the entanglement problem in quantum physics.

In the nonlinear elastic materials analysis, both the ellipticity and strong ellipticity play important roles, especially when a material is required to satisfy a number of important statical and dynamical properties [13, 10]. By use of tensor expression of the elastic material, the strong ellipticity condition and the ordinary ellipticity conditions can be characterized by the positiveness and the nonnegativity of the minimization model of (1.2.6) with $m = n = 3$, respectively. In this sense, tensor $\mathcal{F}$ is said to be positive definite or positive semi-definite on $\Re^3 \times \Re^3$, respectively. Recently, Qi *et al.* [54, 21] established a necessary and sufficient condition for this by introducing the concept of M-eigenvalues of tensor.

In quantum physics, the entanglement describes a certain type of correlations between subsystems of the full quantum system, and the standard mathematical formulation of a composite quantum system is stated in terms of density matrix [18]. It is shown that to identify whether a state, i.e., a density matrix, is entangled or not for a general quantum state is considered to be a hard problem [20] and many attempts have been made for this problem there. One attempt to solve the problem is to consider the identification of the separability by use of the natural geometrical structure of the problem, see [68]. For the case of a quantum system with two subsystems, it can be formulated as the problem of finding the closest separable state to any given state [17]. That is, for given density matrix $A \in \Re^{pq\times pq}$, find a separable density matrix X which minimizes the distance $\|A - X\|_{\mathrm{F}}$. A matrix X is said to be a separable density matrix iff for some positive integer N, there exist some matrices $X^i \in \mathcal{S}_+^{p\times p}$, $Y^i \in \mathcal{S}_+^{q\times q}$ and $\rho_i > 0$ such that

$$X = \sum_{i=1}^{N} \rho_i X^i \otimes Y^i$$

with $\sum\limits_{i=1}^{N} \rho_i = 1$. To solve the problem, Dahl *et al.* [17] applied the Frank-Wolfe minimizing method [7], involved the minimization model of (1.2.6). Furthermore, they used an alternating eigenvalue maximization method to obtain its solution which may be inefficient in computation.

More recently, minimization model of (1.2.6) was studied in [33]. The authors showed that the problem is NP-hard to solve and presented some approximation solutions based on SOS and SDP relaxation scheme.

1.2.2 Bi-quadratic Optimization Problems with Quadratic Constraints

The bi-quadratic optimization problems with quadratic constraints include the following two forms:

$$\begin{array}{ll} \min & f(x,y) := \mathcal{F}xxyy \\ \text{s.t.} & x^\top A_p x \geq 1,\ p = 1, \cdots, m_1, \\ & y^\top B_q y \geq 1,\ q = 1, \cdots, n_1, \end{array} \tag{1.2.7}$$

and

$$\begin{aligned} \max \quad & f(x,y) = \mathcal{F}xxyy \\ \text{s.t.} \quad & x^\top A_p x \le 1,\ p = 0, 1, \cdots, m_1, \\ & y^\top B_q y \le 1,\ q = 1, \cdots, n_1, \end{aligned} \tag{1.2.8}$$

where the matrices $A_p \in \Re^{m\times m}$ $(p = 1, 2, \cdots, m_1)$ and $B_q \in \Re^{n\times n}$ $(q = 1, 2, \cdots, n_1)$ are symmetric positive semidefinite, whereas $A_0 \in \Re^{m\times m}$ is symmetric indefinite matrix.

The bi-quadratic optimization problems (1.2.7) and (1.2.8) are natural generalizations of bi-quadratic optimization over unit spheres problem (1.2.6). Furthermore, these two problems can be regarded as the generalizations of general quadratic optimization problems. For example, if there exist matrices $C \in \Re^{m\times m}$ and $D \in \Re^{n\times n}$ such that $\mathcal{F} = C \otimes D$ where $\otimes$ denotes the standard Kronecker product, then the minimization model (1.2.7) will be equivalent to solving the following two quadratic optimization problems:

$$\begin{aligned} \min \quad & x^\top C x \\ \text{s.t.} \quad & x^\top A_p x \ge 1,\ p = 1, \cdots, m_1 \end{aligned} \tag{1.2.9}$$

and

$$\begin{aligned} \min \quad & y^\top D y \\ \text{s.t.} \quad & y^\top B_q y \ge 1,\ q = 1, \cdots, n_1, \end{aligned} \tag{1.2.10}$$

which were shown to be NP-hard even when C and D are positive definite due to [38]. In fact, the general quadratic maximization problem is also NP-hard from [42].

Therefore, it is reasonable to recall how to solve the NP-hard quadratic optimization problems. A popular approach to approximately solving the considered problem is to use their SDP relaxation problems [24, 48, 62]. Before proceeding, we present the SDP relaxation scheme, illustrated by problem (1.2.9).

It is easy to see that (1.2.9) can be rewritten as

$$\begin{aligned} \min \quad & C \bullet (xx^\top) \\ \text{s.t.} \quad & A_p \bullet (xx^\top) \ge 1,\ p = 1, \cdots, m_1. \end{aligned} \tag{1.2.11}$$

So that, (1.2.9) can be written as the following matrix form

$$\begin{aligned} \min \quad & C \bullet X \\ \text{s.t.} \quad & A_p \bullet X \ge 1,\ p = 1, \cdots, m_1, \\ & X \succeq 0, \\ & \text{rank}(X) = 1. \end{aligned} \tag{1.2.12}$$

By eliminating the rank-one requirement, the SDP relaxation problem of (1.2.9) is obtained as

$$\begin{array}{ll} \min & C \bullet X \\ \text{s.t.} & A_p \bullet X \geq 1,\ p = 1, \cdots, m_1, \\ & X \succeq 0. \end{array} \tag{1.2.13}$$

The obtained SDP relaxation problem is used for solving NP-hard quadratic optimization problem. Notice the following quadratic minimization form, studied in [38]

$$\begin{array}{ll} \min & \|x\| \\ \text{s.}t. & \sum\limits_{l \in I_i} |h_l^H x| \geq 1, \quad i = 1, 2, \cdots, m, \\ & x \in \mathbf{F}^n, \end{array} \tag{1.2.14}$$

where $\mathbf{F}$ is either $\Re$ or $\mathbf{C}$. It was shown that the SDP relaxation for (1.2.14) could provid an $\mathcal{O}(m^2)$ approximation solution in the real case and an $\mathcal{O}(m)$ approximation solution in the complex case.

For general quadratic maximization form

$$\begin{array}{ll} \max & x^\top A x \\ \text{s.}t. & x^\top A_i x \leq 1, \quad i = 1, 2, \cdots, m \end{array} \tag{1.2.15}$$

where A_i, for all $i = 1, 2, \cdots, m$ are symmetric positive matrices with positive semidefinite sum and A is an arbitrary matrix, the relative accuracy between (1.2.15) and its corresponding SDP relaxation problem is shown to be $\frac{1}{2\ln(2m^2)}$ in [42], which improves bound established in [44] for the case when all A_i are of rank 1. Sturm and Zhang in [65] presented a matrix decomposition method to get an approximation solution for quadratic problem over the intersection of an ellipsoid and a half-plane by solving its corresponding SDP relaxation. Later, the decomposition method was extended to the case that the constraints are two quadratic inequalities in [72]. Furthermore, the results were strengthened for the complex Hermitian matrices cases in [25].

Later, He *et al.* studied the more general cases that

$$\begin{array}{ll} \min & x^\top A x \\ \text{s.}t. & x^\top A_i x \geq 1, \quad i = 0, 1, \cdots, m, \\ & x \in \mathbf{F}^n \end{array} \tag{1.2.16}$$

and

$$\begin{array}{ll} \max & x^\top Ax \\ \text{s.}t. & x^\top A_i x \le 1, \quad \forall i = 0, 1, \cdots, m \\ & x \in \mathbf{F}^n \end{array} \tag{1.2.17}$$

where $\mathbf{F}$ is either the real field $\Re$ or the complex field $\mathbf{C}$. They proved that for (1.2.16), provided that matrix A and all but one of A_k are positive semidefinite, the ratio between the optimal value of (1.2.16) and its SDP relaxation would be upper bounded by $\mathcal{O}(m^2)$ as $\mathbf{F} = \Re$, and by $\mathcal{O}(m)$ as $\mathbf{F} = \mathbf{C}$. For the maximization model (1.2.17), they proved that the ratio is bounded from below by $\mathcal{O}(\frac{1}{\log m})$ for both the real and complex case and at most one of A_k are indefinite.

More recently, higher order polynomial optimizations received much attention. Problem (1.2.6) was studied in [33]. They showed that the problem is NP-hard to solve and there is no polynomial time algorithm returning bounds with finite relative quality bound. Based upon the complexity analysis, some approximation methods by SDP relaxation were presented in their paper. Latest, Quartic polynomial optimization with quadratic constraints, of the following form, were considered in [39]

$$\begin{array}{ll} \max & \mathcal{G}xxxx \\ s.t. & x^\top A_i x \le 1, \quad i = 1, 2, \cdots, m, \end{array}$$

and

$$\begin{array}{ll} \min & \mathcal{G}xxxx \\ s.t. & x^\top A_i x \ge 1, \quad i = 1, 2, \cdots, m, \end{array}$$

where $A_i \in \Re^{n\times n}$ for $i = 1, \cdots, m$ are positive semidefinite matrices and $\mathcal{G}$ is a fourth order $(n\times n\times n\times n)$-dimensional symmetric tensor. It was proved that each α-approximate solution of the relaxed SDP can be used to generate in randomized polynomial time an $\mathcal{O}(\alpha)$-approximate solution for the original optimization, where $\mathcal{O}(\cdot)$ depends on the dimension of variables and the number of constraints.

1.3 Notation

To conclude this chapter, we present some notations that will be used throughout the thesis. Tensor $\mathcal{A}$ denotes a third order $(n \times n \times n)$-dimensional real symmetric tensor, $\mathcal{B}$ denotes a third order $(n \times n \times q)$-dimensional real partially symmetric tensor and

$\mathcal{C}$ denotees a third order $(n \times p \times q)$-dimensional real tensor. $\mathcal{F}$ is a fourth order $(m \times m \times n \times n)$-dimensional real partially symmetric tensor. $\mathcal{G}$ is a fourth order $(n \times n \times n \times n)$-dimensional symmetric tensor. We assume that $m, n, p, q \geq 2$ without specification.

Let $\Re$ denote the real number field and $\mathbf{C}$ denote the complex number field. The spaces of n-dimensional real and complex vectors are denoted by $\Re^n$ and $\mathbf{C}^n$, respectively. The spaces of $n \times n$ real symmetric and complex Hermitian matrices are denoted by $\mathcal{S}^n$ and $\mathcal{H}^n$, respectively. Matrix $Z \in \mathcal{H}^n$ means that $\mathrm{Re}(Z)$ is symmetric and $\mathrm{Im}(Z)$ is skew-symmetric, where $\mathrm{Re}(Z)$ and $\mathrm{Im}(Z)$ stand for the real and imaginary part of Z, respectively. For two real matrices A and B with the same dimension, $A \bullet B$ stands for usual matrix inner product, i.e., $A \bullet B = \mathrm{tr}(A^\top B)$, where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. In addition, $\|A\|_F$ denotes the Frobenius norm of A, i.e., $\|A\|_F = (A \bullet A)^{1/2}$, and I_n denotes the $n \times n$ identity matrix. For two complex matrices A and B, their inner product

$$A \bullet B = \mathrm{Re}(\mathrm{tr}(A^H B)) = \mathrm{tr}\left(\mathrm{Re}(A)^\top \mathrm{Re}(B) + \mathrm{Im}(A)^\top \mathrm{Im}(B)\right),$$

where A^H denotes the conjugate transpose of matrix A. The notation $A \succeq 0$ ($\succ 0$) means that A is positive semidefinite (positive definite).

Chapter 2

The Measure of Diffusion Skewness and Kurtosis in Magnetic Resonance Imaging

2.1 Introduction

In this chapter, we consider the following lower-order approximation of (1.1.1)

$$\ln\left(\frac{S(b)}{S(0)}\right) = -\sum_{i_1,i_2=1}^{3} D^{(2)}_{i_1 i_2} b^{(2)}_{i_1 i_2} + \sum_{i_1,i_2,i_3,i_4=1}^{3} D^{(4)}_{i_1 i_2 i_3 i_4} b^{(4)}_{i_1 i_2 i_3 i_4} - j \sum_{i_1,i_2,i_3=1}^{3} D^{(3)}_{i_1 i_2 i_3} b^{(3)}_{i_1 i_2 i_3}, \tag{2.1.1}$$

which is obtained by truncating (1.1.1) to the fourth order tensor and contains useful information of the signal. Moreover, the first two terms of (2.1.1) are related to the magnitude of the signal and the last term of (2.1.1) is related to the phase of the signal. The second order tensor $D^{(2)}$ is the diffusion tensor. For convenience of notation, we call the third order tensor $D^{(3)}$ and the fourth order tensor $D^{(4)}$ in (2.1.1) the diffusion skewness (DS) tensor and the diffusion kurtosis (DK) tensor, respectively. On the other hand, it is important to note that the values $D^{(2)}_{i_1 i_2}, D^{(3)}_{i_1 i_2 i_3}$ and $D^{(4)}_{i_1 i_2 i_3 i_4}$ in (2.1.1) are not independent of the coordinate system. That is, these values will be changed when the coordinate system is rotated. However, to understand the biological and clinical meaning of the corresponding tensors in (2.1.1), the quantities and parameters which

are independent from coordinate system choices, denoted by invariants, are needed. Therefore, it is important to find, measure and calculate the invariants involved in the model (2.1.1).

Recall that the main invariants of the diffusion tensor $D^{(2)}$ are its eigenvalues, which have already been widely used in the DTI technique [1]. Recently, some important invariants, based on the definition of D-eigenvalue, related to $D^{(4)}$ in the DKI model were presented by Qi *et al.* in [58]. Moreover, a method for calculating D-eigenvalues was presented there. Motivated by these, in this chapter, we discuss the quantities and parameters associated with the DS tensor $D^{(3)}$ in (2.1.1), which include the largest and the smallest apparent skewness coefficients (ASC) values. Then we study their computation formulas and relationships.

This chapter is organized as follows. In Section 2.2, we discuss some further properties of the invariants of $D^{(4)}$. In Section 2.3, based on the concept of Z-eigenvalues of tensors [51], we show that the largest and the smallest ASC values are invariant under coordinate rotations and may have important biological and clinical meanings. In Section 2.4, we propose numerical methods to calculate the largest and the smallest ASC values and the AKC values. In Section 2.5, we provide some numerical examples for calculating ASC values. Some final conclusions are made in Section 2.6.

2.2 The AKC Values

In this section, we first summarize the concept and properties of AKC values, then further discuss some properties of the D-eigenvalues and Kelvin eigenvalues of $D^{(4)}$. To this end, let us write

$$D = (\gamma g \delta)^2 \left(\Delta - \frac{1}{3}\delta \right) D^{(2)} \tag{2.2.2}$$

and

$$\mathcal{W} = (\gamma g \delta)^4 \left(\Delta - \frac{3}{5}\delta \right) D^{(4)}. \tag{2.2.3}$$

In practice, D is positive definite. Then the apparent diffusion coefficient (ADC) [1] is

$$D_{app} = Dx^2 \equiv \sum_{i,j=1}^{3} D_{ij} x_i x_j.$$

Let the eigenvalues of D be $\alpha_1 \geq \alpha_2 \geq \alpha_3 > 0$, then the mean diffusivity [1] can be calculated by

$$M_D = \frac{\alpha_1 + \alpha_2 + \alpha_3}{3}.$$

As [51, 58], we denote Dx and $\mathcal{W}x^3$ as two vectors in $\Re^3$ with their ith component as

$$(Dx)_i = \sum_{j=1}^{3} D_{ij}x_j$$

and

$$(\mathcal{W}x^3)_i = \sum_{j,k,l=1}^{3} \mathcal{W}_{ijkl}x_j x_k x_l,$$

respectively, for $i = 1, 2, 3$. Based on these notations, Qi *et al.* in [58] introduced the following concepts of D-eigenvalues and D-eigenvectors of $\mathcal{W}$, which is a generalization of Z-eigenvalues and Z-eigenvectors presented in [51].

Definition 2.2.1 *A real number λ is said to be a Z-eigenvalue of the mth order n-dimensional symmetric tensor $\bar{\mathcal{W}}$, if there exists a real vector $x \in \Re^n$ satisfying the following system*

$$\begin{cases} \bar{\mathcal{W}}x^{m-1} = \lambda x, \\ \|x\| = 1. \end{cases}$$

Vector x is called the Z-eigenvector associated with Z-eigenvalue λ.

Definition 2.2.2 *A real number λ is said to be a D-eigenvalue of $\mathcal{W}$, if there exists a real vector x such that*

$$\begin{cases} \mathcal{W}x^3 = \lambda Dx, \\ Dx^2 = 1. \end{cases} \tag{2.2.4}$$

The real vector x is called the D-eigenvector of $\mathcal{W}$ associated with the D-eigenvalue λ.

For the fourth order three dimensional symmetric tensor $\mathcal{W}$, it is easy to see that a D-eigenvalue reduces to a Z-eigenvalue when D is an identity matrix. Furthermore, from the definition of D-eigenvalues, a key formula for the tensor $\mathcal{W}$ is as follows:

$$K_{app}(x) = \frac{M_D^2}{D_{app}^2}\mathcal{W}x^4, \tag{2.2.5}$$

where $K_{app}(x)$ is the AKC value at the direction x, and

$$\mathcal{W}x^4 \equiv \sum_{i,j,k,l=1}^{3} \mathcal{W}_{ijkl}x_ix_jx_kx_l.$$

Denote the largest and the smallest AKC values as $K_{\max}$ and $K_{\min}$, respectively. Then we have the following results which were proved in [58].

Theorem 2.2.1 *D-eigenvalues of $\mathcal{W}$ are real numbers and always exist. If x is a D-eigenvector associated with a D-eigenvalue λ, then*

$$\lambda = \mathcal{W}x^4.$$

Denote the largest and the smallest D-eigenvalues of $\mathcal{W}$ as $\lambda^D_{\max}$ and $\lambda^D_{\min}$ respectively. Then the largest AKC value is

$$K_{\max} = M_D^2\lambda^D_{\max} \tag{2.2.6}$$

and the smallest AKC value is

$$K_{\min} = M_D^2\lambda^D_{\min}. \tag{2.2.7}$$

Theorem 2.2.2 *The D-eigenvalues of $\mathcal{W}$ are invariant under rotations of coordinate systems.*

From these two theorems, we know that $K_{\max}$ and $K_{\min}$ are also invariants of $\mathcal{W}$. In the rest of this section, we discuss some further properties of D-eigenvalues of $\mathcal{W}$.

Before proceeding, we recall the following reformulated matrix $W \in \Re^{6\times 6}$, denoted by the Kelvin matrix, of tensor $\mathcal{W}$

$$W = \begin{pmatrix} \mathcal{W}_{1111} & \mathcal{W}_{1122} & \mathcal{W}_{1133} & \sqrt{2}\mathcal{W}_{1112} & \sqrt{2}\mathcal{W}_{1113} & \sqrt{2}\mathcal{W}_{1123} \\ \mathcal{W}_{1122} & \mathcal{W}_{2222} & \mathcal{W}_{2233} & \sqrt{2}\mathcal{W}_{2212} & \sqrt{2}\mathcal{W}_{2213} & \sqrt{2}\mathcal{W}_{2223} \\ \mathcal{W}_{1133} & \mathcal{W}_{2233} & \mathcal{W}_{3333} & \sqrt{2}\mathcal{W}_{3312} & \sqrt{2}\mathcal{W}_{3313} & \sqrt{2}\mathcal{W}_{3323} \\ \sqrt{2}\mathcal{W}_{1112} & \sqrt{2}\mathcal{W}_{2212} & \sqrt{2}\mathcal{W}_{3312} & 2\mathcal{W}_{1212} & 2\mathcal{W}_{1213} & 2\mathcal{W}_{1223} \\ \sqrt{2}\mathcal{W}_{1113} & \sqrt{2}\mathcal{W}_{2213} & \sqrt{2}\mathcal{W}_{3313} & 2\mathcal{W}_{1213} & 2\mathcal{W}_{1313} & 2\mathcal{W}_{1323} \\ \sqrt{2}\mathcal{W}_{1123} & \sqrt{2}\mathcal{W}_{2223} & \sqrt{2}\mathcal{W}_{3323} & 2\mathcal{W}_{1223} & 2\mathcal{W}_{1323} & 2\mathcal{W}_{2323} \end{pmatrix}.$$

Obviously, matrix W has six eigenvalues and six eigenvectors. Suppose that σ is an eigenvalue of W with its corresponding eigenvector $v \in \Re^6$. For vector v, there exists a matrix $V \in \Re^{3\times 3}$ such that

$$V = \begin{pmatrix} v_1 & \frac{1}{\sqrt{2}}v_4 & \frac{1}{\sqrt{2}}v_5 \\ \frac{1}{\sqrt{2}}v_4 & v_2 & \frac{1}{\sqrt{2}}v_6 \\ \frac{1}{\sqrt{2}}v_5 & \frac{1}{\sqrt{2}}v_6 & v_3 \end{pmatrix}.$$

For simplicity, the eigenvalue σ with its eigenvector v of the matrix W is said to be a Kelvin eigenvalue of $\mathcal{W}$, and its corresponding matrix V is said to be a Kelvin eigentensor of $\mathcal{W}$.

Based on these definitions, we can assert that the Kelvin eigenvalues of tensor $\mathcal{W}$ are also invariants. In fact, the theory about Kelvin eigenvalues of $\mathcal{W}$ can be traced back to Kelvin 150 years ago [66] and has been discussed in [4]. We wonder how about the relations between D-eigenvalues and Kelvin eigenvalues. By the definitions of D-eigenvalues and Kelvin eigenvalues, the following propositions hold.

Proposition 2.2.1 *Let $\mathcal{W}$ be a fourth order three dimensional fully symmetric tensor, and let σ be a Kelvin eigenvalue of $\mathcal{W}$, associated with a Kelvin eigentensor V. If there exists a vector $x \in \Re^3$ such that $V = xx^T$, then σ is a D-eigenvalue of $\mathcal{W}$.*

Proposition 2.2.2 *Let σ_1, σ_2, $\cdots$, σ_6 be 6 Kelvin eigenvalues of $\mathcal{W}$. Suppose $D = I$. Then we have*

$$-\sum_{m=1}^{6}(-\sigma_m)_+ \le \lambda^D_{\min} \le \lambda^D_{\max} \le \sum_{m=1}^{6}(\sigma_m)_+,$$

where $(a)_+ = \max\{a, 0\}$.

Proof. It is easy to verify that we have the spectral decomposition of $\mathcal{W}$ as follows

$$\mathcal{W} = \sum_{m=1}^{6} \sigma_m E^m \otimes E^m, \tag{2.2.8}$$

where $\otimes$ denotes the outer tensor product,

$$E^m = \begin{pmatrix} \varepsilon^m_{11} & \frac{1}{\sqrt{2}}\varepsilon^m_{12} & \frac{1}{\sqrt{2}}\varepsilon^m_{13} \\ \frac{1}{\sqrt{2}}\varepsilon^m_{12} & \varepsilon^m_{22} & \frac{1}{\sqrt{2}}\varepsilon^m_{23} \\ \frac{1}{\sqrt{2}}\varepsilon^m_{13} & \frac{1}{\sqrt{2}}\varepsilon^m_{23} & \varepsilon^m_{33} \end{pmatrix},$$

and $\varepsilon^m = (\varepsilon^m_{11}, \varepsilon^m_{22}, \varepsilon^m_{33}, \varepsilon^m_{12}, \varepsilon^m_{13}, \varepsilon^m_{23})^T$ is the mth normalized eigenvector of W [4]. It is clear that for each m, E^m is a symmetric matrix satisfying trace $((E^m)^2) = 1$, which implies that $\mu^2_{m1} + \mu^2_{m2} + \mu^2_{m3} = 1$, where $\mu_{m1} \le \mu_{m2} \le \mu_{m3}$ are three eigenvalues of E^m. By (2.2.8), we have that for any $x = (x_1, x_2, x_3)^T$,

$$\begin{aligned} \mathcal{W}x^4 &= \sum_{i,j,l,k=1}^{3} \mathcal{W}_{ijkl} x_i x_j x_k x_l \\ &= \sum_{m=1}^{6} \sigma_m (x^T E^m x)^2. \end{aligned} \tag{2.2.9}$$

It is well known that $\mu_{m1} \le x^T E^m x \le \mu_{m3}$ for any $m = 1, \cdots, 6$. This implies that $0 \le (x^T E^m x)^2 \le \max\{\mu^2_{m1}, \mu^2_{m3}\} \le 1$. Therefore, by (2.2.9), we obtain the desired result and complete the proof. □

Now we discuss the independence of eigenvalues of fourth order three dimensional tensor. We first give the following definition.

Definition 2.2.3 *A set S consisted of the functions*

$$y_i = f_i(x_1, x_2, \cdots, x_n), \quad i = 1, 2, \cdots, m, \tag{2.2.10}$$

which are defined on the region Ω in $\Re^n$, is said to be functionally dependent on Ω, if there exist an index i_0 and a function φ defined on an appropriate region in $\Re^{m-1}$, such that

$$\begin{aligned} y_{i_0} \equiv\ & \varphi(f_1(x_1, x_2, \cdots, x_n), f_2(x_1, x_2, \cdots, x_n), \cdots, \\ & f_{i_0-1}(x_1, x_2, \cdots, x_n), f_{i_0+1}(x_1, x_2, \cdots, x_n), \cdots, f_m(x_1, x_2, \cdots, x_n)) \end{aligned}$$

holds for any $(x_1, x_2, \cdots, x_n) \in \Omega$. If for any sub-region Ω' of Ω, there are no i_0 and such function φ that

$$y_{i_0} \equiv \varphi(y_1, y_2, \cdots, y_{i_0-1}, y_{i_0+1}, \cdots, y_m)$$

holds on Ω', then the function set S is said to be functionally independent on Ω.

For the functional independence, we have the following theorem.

Theorem 2.2.3 *Suppose that $m \le n$ and there exists an mth order determinant $|A|$ in the Jacobian matrix of the functions set (2.2.10) such that $|A| \ne 0$ holds on Ω. Then the functions set S is functionally independent on Ω.*

It is important to note that the trace Π_K of matrix W corresponding with tensor $\mathcal{W}$ is an important invariant, which characterizes the average AKC value on a spherical surface and has physics significance. In addition, the largest D-eigenvalue $\lambda^D_{\max}$ and the smallest D-eigenvalue $\lambda^D_{\min}$ of $\mathcal{W}$ play an important role in the diffusion analysis of the water molecule in biological tissue. From Proposition 2.2.2, we see that the largest D-eigenvalue and the smallest D-eigenvalue of $\mathcal{W}$ can be estimated with an interval determined by the Kelvin eigenvalues of $\mathcal{W}$. However, this result does not mean that there must be some functional dependence between the largest (smallest) D-eigenvalues and Kelvin eigenvalues of $\mathcal{W}$. In fact, the following example shows that both $\{\lambda^D_{\max}, \lambda^D_{\min}, \Pi_K\}$ and $\{\lambda^D_{\max}, \lambda^D_{\min}, \sigma_{\max}, \sigma_{\min}\}$ are functionally independent on a considered region, where $\sigma_{\max}$ and $\sigma_{\min}$ denote the largest and smallest Kelvin eigenvalues of $\mathcal{W}$, respectively.

Example 2.2.1 Let $\mathcal{W}$ be a fourth order three dimensional fully symmetric tensor with $\mathcal{W}_{1111} = t_1$, $\mathcal{W}_{2222} = t_2$, $\mathcal{W}_{3333} = t_3$, $\mathcal{W}_{1122} = t_4$ and its other elements are zero, and let $D = I$. Consider the case where $0 < t_1 < t_3 < 3t_4 < t_2$, $t_1 < t_4$ and $t_1 t_2 < t_4^2$. By Definition 2.2.2, it is easy to obtain that the D-eigenvalues of $\mathcal{W}$ are as follows

$$\lambda_1 = t_1, \quad \lambda_2 = t_2, \quad \lambda_3 = t_3, \quad \lambda_4 = \frac{t_1 t_3}{t_1 + t_3}, \quad \lambda_5 = \frac{t_2 t_3}{t_2 + t_3}.$$

Under the given conditions, it is easy to see that the largest and smallest D-eigenvalues of W are

$$\lambda^D_{\max} = F_1(t_1, t_2, t_3, t_4) := t_2 \quad \text{and} \quad \lambda^D_{\min} = F_2(t_1, t_2, t_3, t_4) := \frac{t_1 t_3}{t_1 + t_3},$$

respectively.

On the other hand, it is clear that the trace Π_K in sense of Kelvin

$$\Pi_K = F_3(t_1, t_2, t_3, t_4) := t_1 + t_2 + t_3 + 2t_4.$$

Moreover, by direct computation, we obtain that the set consisted of all Kelvin eigenvalues of $\mathcal{W}$ is

$$\left\{ \frac{t_1 + t_2 + \sqrt{(t_1 - t_2)^2 + 4t_4^2}}{2}, \frac{t_1 + t_2 - \sqrt{(t_1 - t_2)^2 + 4t_4^2}}{2}, t_3, 2t_4, 0, 0 \right\},$$

which implies that the largest and smallest Kelvin eigenvalues of $\mathcal{W}$ are

$$\sigma_{\max} = F_4(t_1, t_2, t_3, t_4) := \frac{t_1 + t_2 + \sqrt{(t_1 - t_2)^2 + 4t_4^2}}{2}$$

and

$$\sigma_{\min} = F_5(t_1, t_2, t_3, t_4) := \frac{t_1 + t_2 - \sqrt{(t_1 - t_2)^2 + 4t_4^2}}{2},$$

respectively. Based on these above, it is easy to verify that the Jacobian matrices of $\hat{F} := (F_1, F_2, F_3)$ and $\tilde{F} := (F_1, F_2, F_4, F_5)$ are

$$\nabla \hat{F}(t_1, t_2, t_3, t_4) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \dfrac{t_3^2}{(t_1+t_3)^2} & 0 & \dfrac{t_1^2}{(t_1+t_3)^2} & 0 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

and

$$\nabla \tilde{F}(t_1, t_2, t_3, t_4) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \dfrac{t_3^2}{(t_1+t_3)^2} & 0 & \dfrac{t_1^2}{(t_1+t_3)^2} & 0 \\ \dfrac{1}{2}\left(1 + \dfrac{t_1 - t_2}{\sqrt{(t_1-t_2)^2 + 4t_4^2}}\right) & \dfrac{1}{2}\left(1 - \dfrac{t_1 - t_2}{\sqrt{(t_1-t_2)^2 + 4t_4^2}}\right) & 0 & \dfrac{2t_4}{\sqrt{(t_1-t_2)^2 + 4t_4^2}} \\ \dfrac{1}{2}\left(1 - \dfrac{t_1 - t_2}{\sqrt{(t_1-t_2)^2 + 4t_4^2}}\right) & \dfrac{1}{2}\left(1 + \dfrac{t_1 - t_2}{\sqrt{(t_1-t_2)^2 + 4t_4^2}}\right) & 0 & -\dfrac{2t_4}{\sqrt{(t_1-t_2)^2 + 4t_4^2}} \end{pmatrix},$$

respectively. It is easy to see that for $\hat{F}$ and $\tilde{F}$, the conditions required in Theorem 2.2.3 are satisfied. Hence, we know that both $\{\lambda_{\max}^D, \lambda_{\min}^D, \Pi_K\}$ and $\{\lambda_{\max}^D, \lambda_{\min}^D, \sigma_{\max}, \sigma_{\min}\}$ are functionally independent on

$$\Omega := \{(t_1, t_2, t_3, t_4) \mid 0 < t_1 < t_3 < 3t_4 < t_2,\ t_1 < t_4 \text{ and } t_1 t_2 < t_4^2\}.$$

2.3 The ASC Values

As mentioned, we may use ASC values to characterize the phase of the magnetic resonance signal in biological tissues. Let us write

$$\mathcal{A} = (\gamma g \delta)^3 \left(\Delta - \frac{2}{4}\delta\right) D^{(3)}, \tag{2.3.11}$$

which is a third order three dimensional fully symmetric tensor. It is easy to see that $\mathcal{A}$ has ten independent elements because of symmetry. For those elements of $\mathcal{A}$ which are equal to each other, we use the element $\mathcal{A}_{ijk}$ with $i \le j \le k$ to represent them. That is, if we say that $\mathcal{A}_{122} = 4$, this automatically implies that $\mathcal{A}_{212} = \mathcal{A}_{221} = 4$. Then, the ten

independent elements of $\mathcal{A}$ are $\mathcal{A}_{111}; \mathcal{A}_{222}; \mathcal{A}_{333}; \mathcal{A}_{112}; \mathcal{A}_{113}; \mathcal{A}_{223}; \mathcal{A}_{122}; \mathcal{A}_{133}; \mathcal{A}_{233}; \mathcal{A}_{123}$. We call $\mathcal{A}_{111}; \mathcal{A}_{222}; \mathcal{A}_{333}$ the diagonal elements of $\mathcal{A}$. We denote $S_{app}(x)$ the apparent skewness coefficient at direction x as follows

$$S_{app}(x) = \frac{\mathcal{A}x^3}{\|x\|^3}, \tag{2.3.12}$$

where

$$\mathcal{A}x^3 \equiv \sum_{i,j,k=1}^{3} \mathcal{A}_{ijk} x_i x_j x_k.$$

We denote $\mathcal{A}x^2$ as a vector in $\Re^3$ with its ith component as

$$(\mathcal{A}x^2)_i = \sum_{j,k=1}^{3} \mathcal{A}_{ijk} x_j x_k,$$

for $i = 1, 2, 3$. Denote the largest and the smallest ASC values as $S_{\max}$ and $S_{\min}$ respectively. Then

$$\begin{aligned} S_{\max} = \quad & \max \quad \mathcal{A}x^3 \\ & \text{s.t} \quad \|x\|^2 = 1, \end{aligned} \tag{2.3.13}$$

and

$$\begin{aligned} S_{\min} = \quad & \min \quad \mathcal{A}x^3 \\ & \text{s.t} \quad \|x\|^2 = 1. \end{aligned} \tag{2.3.14}$$

The critical points of (2.3.13) and (2.3.14) satisfy the following system for some $\lambda \in \Re$ and $x \in \Re^3$:

$$\begin{cases} \mathcal{A}x^2 = \lambda x, \\ \|x\|^2 = 1. \end{cases} \tag{2.3.15}$$

From Definition 2.2.1, we have the following two theorems which can be proved by a similar way to that in [58].

Theorem 2.3.1 *For tensor $\mathcal{A}$, Z-eigenvalues always exist. If x is a Z-eigenvector associated with a Z-eigenvalue λ, then*

$$\lambda = \mathcal{A}x^3.$$

Denote the largest and the smallest Z-eigenvalues of $\mathcal{A}$ as $\lambda^Z_{\max}$ and $\lambda^Z_{\min}$ respectively. Then the largest ASC value is

$$S_{\max} = \lambda^Z_{\max}, \tag{2.3.16}$$

and the smallest ASC value is

$$S_{\min} = \lambda^Z_{\min}. \tag{2.3.17}$$

Theorem 2.3.2 *[53] The Z-eigenvalues of $\mathcal{A}$ are invariant under rotations of coordinate systems.*

Remark 2.3.1 *By these two theorems, $S_{\max}$ and $S_{\min}$ are also invariants of $\mathcal{A}$, and can be calculated by a similar method to that given in [58], which will be presented in Section 2.4. On the other hand, from definitions of Z-eigenvalues and Z-eigenvectors, we know that λ is a Z-eigenvalue of $\mathcal{A}$ with its corresponding eigenvector x if and only if $-\lambda$ is a Z-eigenvalue of $\mathcal{A}$ with the associated eigenvector $-x$. Hence, $\lambda^Z_{\min} = -\lambda^Z_{\max}$.*

Denote the unit sphere as

$$\Xi := \{x \in \Re^3 \ : \ x_1^2 + x_2^2 + x_3^2 = 1\}.$$

Then the average ASC value over the Ξ is defied as

$$M_\Xi = \frac{1}{\Xi} \int \int_\Xi S_{\text{app}}(x) dA = \frac{1}{4\pi} \int \int_\Xi \frac{\mathcal{A}x^3}{\|x\|^3} dA, \tag{2.3.18}$$

where the denominator $\Xi = 4\pi$ is the area of the surface Ξ. Here, we slightly abuse the symbol Ξ for both the surface and its area.

Noting the fact that $\mathcal{A}$ is an odd order full symmetric tensor, it is obvious that for any closed surface Λ with symmetry about the origin, the average ASC value over Λ is equal to zero. Specially, it holds that $M_\Xi = 0$.

2.4 Computation of the ASC and AKC Values

Now we are ready to describe direct methods to obtaining all Z-eigenvalues of $\mathcal{A}$ and D-eigenvalues of $\mathcal{W}$, respectively. Then $S_{\max}$, $S_{\min}$, $K_{\max}$ and $K_{\min}$ can be calculated.

The first method is used to find all the Z-eigenvalues of $\mathcal{A}$. The key idea here is to reduce the three variable system (2.3.15) to a system of two variables. Here, we regard λ as a parameter instead of a variable. Then, we may use the Sylvester formula of the resultant of a two variable system [16] to solve this system.

Based on the consideration above, we state the following theorem which generalizes Theorem 3 in [57] and can be proved in a similar way to that used in [58] .

Theorem 2.4.1 *(a) If $\mathcal{A}_{211} = \mathcal{A}_{311} = 0$, then $x = (\pm 1, 0, 0)^T$ are two Z-eigenvector of $\mathcal{A}$ associated with the Z-eigenvalue $\lambda = \pm\mathcal{A}_{111}$, respectively.*

(b) For any real root t of the following equations:

$$\begin{cases} \mathcal{A}_{211}t^3 + (2\mathcal{A}_{212} - \mathcal{A}_{111})t^2 + (\mathcal{A}_{222} - 2\mathcal{A}_{112})t - \mathcal{A}_{122} = 0, \\ \mathcal{A}_{311}t^2 + 2\mathcal{A}_{312}t + \mathcal{A}_{322} = 0, \end{cases} \tag{2.4.19}$$

$$x = \pm\frac{1}{\sqrt{t^2+1}}(t, 1, 0)^T \tag{2.4.20}$$

is a Z-eigenvector of $\mathcal{A}$ with the Z-eigenvalue $\lambda = \mathcal{A}x^3$.

(c) $\lambda = \mathcal{A}x^3$ and

$$x = \frac{\pm(u, v, 1)^T}{\sqrt{u^2+v^2+1}} \tag{2.4.21}$$

constitute a Z-eigenpair of $\mathcal{A}$, where u and v are a real solution pair of the following polynomial equations:

$$\begin{cases} -\mathcal{A}_{311}u^3 - 2\mathcal{A}_{312}u^2v - \mathcal{A}_{322}uv^2 + (\mathcal{A}_{111} - 2\mathcal{A}_{313})u^2 + 2(\mathcal{A}_{112} - \mathcal{A}_{323})uv \\ \qquad + \mathcal{A}_{122}v^2 + (2\mathcal{A}_{113} - \mathcal{A}_{333})u + 2\mathcal{A}_{123}v + \mathcal{A}_{133} = 0, \\ \mathcal{A}_{211}u^2 - \mathcal{A}_{311}u^2v - 2\mathcal{A}_{312}uv^2 + 2(\mathcal{A}_{212} - \mathcal{A}_{313})uv + 2\mathcal{A}_{213}u - \mathcal{A}_{322}v^3 \\ \qquad + (\mathcal{A}_{222} - 2\mathcal{A}_{323})v^2 + 2(\mathcal{A}_{223} - \mathcal{A}_{333})v + \mathcal{A}_{233} = 0. \end{cases} \tag{2.4.22}$$

All the Z-eigenpairs of tensor $\mathcal{A}$ are given by (a) if $\mathcal{A}_{211} = \mathcal{A}_{311} = 0$, (b) and (c) otherwise.

We regard the polynomial equation system (2.4.22) as equations of u. We may write it as

$$\begin{cases} \alpha_0u^3 + \alpha_1u^2 + \alpha_2u + \alpha_3 = 0, \\ \beta_0u^2 + \beta_1u + \beta_2 = 0, \end{cases}$$

where $\alpha_0, \cdots, \alpha_3, \beta_0, \beta_1, \beta_2$ are polynomials of v, which can be calculated by (2.4.22). It has complex solutions if and only if its resultant vanishes [16]. By the Sylvester theorem [16], its resultant is equal to the determinant of the following 5×5 matrix:

$$\begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & 0 \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_0 & \beta_1 & \beta_2 & 0 & 0 \\ 0 & \beta_0 & \beta_1 & \beta_2 & 0 \\ 0 & 0 & \beta_0 & \beta_1 & \beta_2 \end{pmatrix},$$

which is a one-dimensional polynomial of v.

To find approximate solutions of all the real roots of a one dimensional polynomial, we can use the following Sturm Theorem in [41].

Theorem 2.4.2 *Let ψ be a nonconstant polynomial of degree l, with real coefficients and let c_1 and c_2 be two real numbers such that $c_1 < c_2$ and $\psi(c_1)\psi(c_2) \neq 0$. The sequence $\psi_0, \psi_1, \cdots, \psi_l$ defined by*

$$\psi_0 = \psi, \quad \psi_1 = \psi', \quad \psi_{i+1} = -\psi_{i-1} \mod \psi_i, \quad i = 1, 2, \cdots, l-1$$

and $\psi_{l+1} \equiv 0$ is called a sequence of Sturm. Denote by $v(x)$ the number of changes of signs in the sequence $\psi_0(x), \psi_1(x), \cdots, \psi_l(x))$. Then the number of distinct real roots of ψ on the interval (c_1, c_2) is equal to $v(c_1) - v(c_2)$.

We may find the approximate solutions of all the real roots of this one-dimensional polynomial such that their differences with the exact solutions are within a given error bound. We then substitute them to (2.4.22) to find the corresponding approximate real solutions of u. Correspondingly, approximate values of all the Z-eigenvalues and Z-eigenvectors can be obtained. Based on this, we can obtain the largest and smallest ASC values.

Now we are ready to present the method used to find all the D-eigenvalues of $\mathcal{W}$, which is similar as above and is based on the following theorem given in [58].

Theorem 2.4.3 *Let $\bar{W}$ be a fourth order three dimensional tensor such that its entries satisfy $\bar{\mathcal{W}}_{ijkl} = \sum_{h=1}^{3} \bar{d}_{ih}\mathcal{W}_{ijkl}$ for $i, j, k, l = 1, 2, 3$, where $\bar{d}_{ih}$ is the ih row hth column element in the inverse D^{-1} of D. Then we have*

(a) If $\bar{\mathcal{W}}_{2111} = \bar{\mathcal{W}}_{3111} = 0$, then $\lambda = \dfrac{\bar{\mathcal{W}}_{111}}{D_{11}}$ is a D-eigenvalue of $\mathcal{W}$ with a D-eigenvector $x = (\pm\sqrt{\dfrac{1}{D_{11}}}, 0, 0)^T$.

(b) For any real root t of the following equations:

$$\begin{cases} -\bar{\mathcal{W}}_{2111}t^4 + (\bar{\mathcal{W}}_{1111} - 3\bar{\mathcal{W}}_{2112})t^3 + 3(\bar{\mathcal{W}}_{1112} - \bar{\mathcal{W}}_{2122})t^2 \\ +(3\bar{\mathcal{W}}_{1122} - \bar{\mathcal{W}}_{2222})t + \bar{\mathcal{W}}_{1222} = 0, \\ \bar{\mathcal{W}}_{3111}t^3 + 3\bar{\mathcal{W}}_{3112}t^2 + 3\bar{\mathcal{W}}_{3122}t + \bar{\mathcal{W}}_{3222} = 0, \end{cases} \tag{2.4.23}$$

$$x = \pm \frac{1}{\sqrt{D_{11}t^2 + 2D_{12}t + D_{22}}}(t, 1, 0)^T \tag{2.4.24}$$

is a D-eigenvector of $\mathcal{W}$ *with the* $D-$*eigenvalue* $\lambda = \mathcal{W}x^4$.

(c) $\lambda = Wx^4$ *and*

$$x = \frac{\pm(u, v, 1)^T}{\sqrt{D_{11}u^2 + 2D_{12}uv + 2D_{13}u + D_{22}v^2 + 2D_{23}v + D_{33}}} \tag{2.4.25}$$

constitute a D-eigenpair of $\mathcal{W}$*, where* u *and* v *are a real solution of the following polynomial equations:*

$$\begin{cases} -\bar{\mathcal{W}}_{3111}u^4 - 3\bar{\mathcal{W}}_{3112}u^3v + (\bar{\mathcal{W}}_{1111} - 3\bar{\mathcal{W}}_{3113})u^3 - 3\bar{\mathcal{W}}_{3122}u^2v^2 \\ +(3\bar{\mathcal{W}}_{1112} - 6\bar{\mathcal{W}}_{3123})u^2v + +(3\bar{\mathcal{W}}_{1113} - 3\bar{\mathcal{W}}_{3133})u^2 \\ -3\bar{\mathcal{W}}_{3223}uv^2 - \bar{\mathcal{W}}_{3222}uv^3 + 3\bar{\mathcal{W}}_{1122}uv^2 + (6\bar{\mathcal{W}}_{1123} - 3\bar{\mathcal{W}}_{3233})uv \\ +(3\bar{\mathcal{W}}_{1133} - \bar{\mathcal{W}}_{3333})u + \bar{\mathcal{W}}_{1222}v^3 + 3\bar{\mathcal{W}}_{1223}v^2 + 3\bar{\mathcal{W}}_{1233}v + \bar{\mathcal{W}}_{1333} = 0, \\ -\bar{\mathcal{W}}_{3111}u^3v + \bar{\mathcal{W}}_{2111}u^3 - 3\bar{\mathcal{W}}_{3112}u^2v^2 + (3\bar{\mathcal{W}}_{2112} - 3\bar{\mathcal{W}}_{3113})u^2v \\ +3\bar{\mathcal{W}}_{2113}u^2 - 3\bar{\mathcal{W}}_{3122}uv^3 + (3\bar{\mathcal{W}}_{2122} - 6\bar{\mathcal{W}}3123)uv^2 \\ +(6\bar{\mathcal{W}}_{2123} - 3\bar{\mathcal{W}}_{3133})uv + 3\bar{\mathcal{W}}_{2133}u + 3\bar{\mathcal{W}}_{2223}v^2 - \bar{\mathcal{W}}_{3222}v^4 \\ +(\bar{\mathcal{W}}_{2222} - 3\bar{\mathcal{W}}_{3223})v^3 - 3\bar{\mathcal{W}}_{3233}v^2 + (3\bar{\mathcal{W}}_{2233} - \bar{\mathcal{W}}_{3333})v + \bar{\mathcal{W}}_{2333} = 0. \end{cases} \tag{2.4.26}$$

All the D-eigenpairs of tensor $\mathcal{W}$ *are given by (a) if* $\bar{\mathcal{W}}_{2111} = \bar{\mathcal{W}}_{3111} = 0$*, by (b) and (c) otherwise.*

2.5 Numerical Examples

In this section, we present preliminary numerical experiments for the DS tensor with the method presented in Section 2.4. The computation was done on a personal computer (Pentium IV, 2.8GHz) by running MatlabR2006a. A numerical example for DK tensor can be found in [58]. That example is derived from data of MRI experiments on the white matter of rat spinal cord specimen fixed in formalin. The MRI experiments were conducted on a 7 Tesla MRI scanner at Laboratory of Biomedical Imaging and Signal Processing at The University of Hong Kong.

For the test examples below, we choose the parameters in (1.1.2) as follows

$$\triangle = 1, \quad \delta = 0.5, \quad g = 1, \quad \gamma = 1.$$

Then the tensor $\mathcal{A}$ in (2.3.11) becomes $\mathcal{A} = \frac{3}{32}D^{(3)}$.

By Theorem 2.4.1, we can obtain all the Z-eigenvalues of $\mathcal{A}$, and the associated eigenvectors. As mentioned in Remark 2.3.1, $-\lambda$ must be another Z-eigenvalue of it when λ is a Z-eigenvalue of $\mathcal{A}$. Throughout this section, we present only the nonnegative Z-eigenvalues and the corresponding Z-eigenvectors of $\mathcal{A}$ in the following tables.

Example 2.5.1 This example was taken from [34], conducted by Monte-Carlo simulations using computer-synthesized phantoms with a Y-shape tube. The Y-shape tube is asymmetric and the DTI technique fails to identify this structure.

For this example, the ten independent elements of $D^{(3)}$ are $D^{(3)}_{111} = -2.36, D^{(3)}_{112} = 47.9, D^{(3)}_{113} = 0, D^{(3)}_{122} = -0.773, D^{(3)}_{123} = -0.575, D^{(3)}_{133} = 0.282, D^{(3)}_{222} = -28.7, D^{(3)}_{223} = 0, D^{(3)}_{233} = 3.61, D^{(3)}_{333} = 0.488$ in unit of $10^{-8}mm^3/s$.

The numerical results for Example 2.5.1 are listed in the Table 2.5.1.

number	x_1	x_2	x_3	$\lambda \times 10^7$
(1)	0	-1.0000	0	0.2691
(2)	-0.0062	-1.0000	-0.0002	0.2691
(3)	-0.8514	0.5244	0.0097	0.4922
(4)	0.8480	0.5299	-0.0108	0.4548
(5)	-0.0431	0.0557	0.9975	0.0044
(6)	0.0494	-0.0684	0.9964	0.0049

Table 2.5.1: Z-eigenvalues and eigenvectors of $\mathcal{A}$ in Example 2.5.1

From Table 2.5.1, we can see that there are 12 Z -eigenvalues and corresponding Z-eigenvectors for $\mathcal{A}$, and the largest and smallest Z-eigenvalues of $\mathcal{A}$ are 0.4922×10^{-7} and -0.4922×10^{-7}, which attained at $(-0.8514, 0.5244, 0.0097)^T$ and $(0.8514, -0.5244, -0.0097)^T$, respectively. This implies that $S_{\max} = 0.4922 \times 10^{-7}$ and $S_{\min} = -0.4922 \times 10^{-7}$.

In order to illustrate the efficiency of our method, we also calculate the Z-eigenvalues and corresponding Z-eigenvectors of ten third order three dimensional full symmetric tensors which are constructed randomly in the following example.

Example 2.5.2 The elements of tensor $\mathcal{A}$ are drawn by a normal distribution with mean zero and standard deviation one.

Using the method provided in Section 2.4, we compute all the Z-eigenvalues of $\mathcal{A}$, and the associated eigenvectors. In Table 2.5.2, the largest Z-eigenvalue and the corresponding Z-eigenvectors are listed for ten tensors. Moreover, in Table 2.5.3, all the nonnegative Z-eigenvalues with corresponding Z-eigenvectors are presented for Tensor 1 in ten tensors.

Tensor	x_1	x_2	x_3	$\lambda^Z_{\max}$
1	-0.5784	0.7896	0.2050	2.1161
2	-0.8364	-0.0495	0.5459	3.2879
3	-0.6272	-0.2393	-0.7411	2.6702
4	-0.0836	-0.8832	-0.5467	2.9957
5	0.7021	-0.6410	0.3100	2.5146
6	-0.7327	0.6778	0.0612	4.1874
7	0.1531	0.5353	0.8307	3.5715
8	0.7981	-0.5944	0.0991	4.2279
9	-0.6308	-0.6893	-0.3563	3.3815
10	-0.2657	0.7381	-0.6201	3.4800

Table 2.5.2: The Largest Z-eigenvalues with Z-eigenvectors for ten tensors

number	x_1	x_2	x_3	λ
(1)	-0.3518	-0.9140	0.2020	0.9434
(2)	-0.5784	0.7896	0.2050	2.1161
(3)	-0.4346	-0.6970	-0.5704	1.6851
(4)	0.9455	0.1980	-0.2585	1.4644
(5)	0.0836	-0.5452	0.8341	1.5940
(6)	0.8322	-0.1726	0.5269	0.5171
(7)	0.3823	-0.1797	-0.9064	0.0165

Table 2.5.3: Nonnegative Z-eigenvalues and Z-eigenvectors of Tensor 1

2.6 Final Conclusion

In this chapter, in order to overcome the drawback of DTI model, we consider the truncated GDTI model. Based on the model, we introduce the concept of diffusion skewness in magnetic resonance imaging and discuss the measure of the diffusion skewness and kurtosis. The diffusion skewness and kurtosis provide two dimensionless values for characterizing the phase of the signal in tissues and the degree of non-Gaussian of the diffusion displacement probability distribution, respectively. For the water molecule with Gaussian distribution in biological structures, the skewness and kurtosis are zero. But, for non-Gaussian signal with asymmetry about the origin, the skewness and the kurtosis have significant values. Based on the Z-eigenvalues and D-eigenvalues of tensor, the methods for calculating the largest (smallest) ASC values and largest (smallest) AKC values are presented. These ASC and AKC values are the principal invariants under rotations of coordinate systems and can be calculated in any Cartesian coordinate system. We hope that these quantities and properties can be used in practice.

Chapter 3

The Cubic Spherical Optimization Problems

3.1 Introduction

In this chapter, we study the cubic spherical optimization problems, which include the cubic one-spherical/two-spherical/three-spherical optimization problems.

Motivated by the NP-hardness of the cubic one-spherical optimizaiton problem, we first analyze the complexity of problems (1.1.4) and (1.1.5). In Section 3.2, we will show that the other two cubic spherical optimization problems are both NP-hard. Thus, it is not expected that there exist efficient exact methods to find global minimizers of these cubic spherical optimization problems when n, p, q are large. Therefore, we are interested in finding approximation algorithms for solving cubic spherical optimization problems.

This chapter is organized as follows. In Section 3.2, we establish the NP-hardness of (1.1.4) and (1.1.5). For problem (1.1.5), we discuss some polynomial time solvable cases in Section 3.3 and study some cases for which polynomial time approximation schemes (PTAS) exist. In Section 3.5, we present polynomial time approximation solution methods, which return a $\frac{1}{\sqrt{\min\{n,p,q\}}}$ bound for (1.1.5). In Section 3.6, we discuss some approximation algorithms for solving one-spherical optimization problem. In the

last section, a practical method for the three-spherical optimization problem is presented and some related numerical results are reported.

3.2 NP-Hardness of the Two-spherical/Three-spherical Optimization Problems

It was proved in [43] that optimization problem (1.1.3) is NP-hard. But the question remains: would (1.1.4) or (1.1.5) be easy to solve globally? In this section, we show that they are both NP-hard to solve.

To establish the NP-hardness of (1.1.4), we first show that problem (1.1.4) is equivalent to a quartic maximization problem.

Theorem 3.2.1 *For each k, let B^k be a symmetric $n\times n$ matrix with its (i,j)th element as $\mathcal{B}_{ijk}$. Then the cubic two-spherical optimization problem (1.1.4) is equivalent to the following quartic maximization problem:*

$$\begin{array}{ll}\max\limits_{x\in\Re^n} & \sum\limits_{k=1}^{q}\langle B^k x, x\rangle^2 \\ \text{s.t.} & \|x\| = 1.\end{array} \tag{3.2.1}$$

Proof. Denote $\alpha(x)$ as a vector in $\Re^q$, with its kth component as

$$\alpha(x)_k = \sum_{i,j=1}^{n} \mathcal{B}_{ijk} x_i x_j.$$

Then we may rewrite (1.1.4) as

$$\min_{\|x\|=1}\min_{\|z\|=1}\langle \alpha(x), z\rangle = \min_{\|x\|=1} -\|\alpha(x)\| = -\max_{\|x\|=1}\|\alpha(x)\|.$$

This is equivalent to

$$\max_{\|x\|=1}\|\alpha(x)\|^2 = \max_{\|x\|=1}\sum_{k=1}^{q}\langle B^k x, x\rangle^2,$$

which is problem (3.2.1). □

According to Theorem 4 of [43], problem (3.2.1) is NP-hard. Hence, we have the following conclusion.

Corollary 3.2.1 *The cubic two-spherical optimization problem (1.1.4) is NP-hard.*

Now let us go to the hardness of (1.1.5). To obtain the NP-hardness of (1.1.5), we need the following results, which shows that if $n = p$ and $\mathcal{C}$ is partially symmetric, then the cubic three-spherical optimization problem (1.1.5) always has a global minimizer (x^*, y^*, z^*) with $x^* = y^*$.

Theorem 3.2.2 *Suppose that tensor $\mathcal{C}$ in problem (1.1.5) is a third order $(n \times n \times q)$-dimensional partially symmetric tensor with respect to the first two indices. Then there exist $x^* \in \Re^n$ and $z^* \in \Re^q$ such that (x^*, x^*, z^*) is a global minimizer of (1.1.5).*

Proof. Let $(\bar{x}, \bar{y}, \bar{z})$ be an optimal solution of (1.1.5) with optimal value $\lambda \leq 0$. As the constraints of (1.1.5) satisfy the linear independence constraint qualification, according to the optimality conditions, $(\bar{x}, \bar{y}, \bar{z})$ satisfies the following equations for Lagrangian multipliers α, β and γ:

$$\begin{cases} \sum_{j=1}^{n}\sum_{k=1}^{q} \mathcal{C}_{ijk}\bar{y}_j\bar{z}_k = \alpha\bar{x}_i, \quad i = 1, \cdots, n, \\ \sum_{i=1}^{n}\sum_{k=1}^{q} \mathcal{C}_{ijk}\bar{x}_i\bar{z}_k = \beta\bar{y}_j, \quad j = 1, \cdots, n, \\ \sum_{i=1}^{n}\sum_{j=1}^{n} \mathcal{C}_{ijk}\bar{x}_i\bar{y}_j = \gamma\bar{z}_k, \quad k = 1, \cdots, q, \\ \|\bar{x}\| = 1, \\ \|\bar{y}\| = 1, \\ \|\bar{z}\| = 1. \end{cases}$$

Then,

$$\alpha = \alpha\sum_{i=1}^{n}\bar{x}_i^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{q}\mathcal{C}_{ijk}\bar{x}_i\bar{y}_j\bar{z}_k = \lambda.$$

Similarly, we have $\beta = \lambda$ and $\gamma = \lambda$. Thus, we have

$$\begin{cases} \sum_{j=1}^{n}\sum_{k=1}^{q} \mathcal{C}_{ijk}\bar{y}_j\bar{z}_k = \lambda\bar{x}_i, \quad i = 1, \cdots, n, \\ \sum_{i=1}^{n}\sum_{k=1}^{q} \mathcal{C}_{ijk}\bar{x}_i\bar{z}_k = \lambda\bar{y}_j, \quad j = 1, \cdots, n, \\ \sum_{i=1}^{n}\sum_{j=1}^{n} \mathcal{C}_{ijk}\bar{x}_i\bar{y}_j = \lambda\bar{z}_k, \quad k = 1, \cdots, q, \\ \|\bar{x}\| = 1, \\ \|\bar{y}\| = 1, \\ \|\bar{z}\| = 1. \end{cases} \tag{3.2.2}$$

We denote $\bar{\mathcal{C}}(\bar{z})$ as a real $n \times n$ symmetric matrix with its (i,j)-th entry being $\bar{\mathcal{C}}_{ij}(\bar{z}) = \sum_{k=1}^{q} \mathcal{C}_{ijk}\bar{z}_k$. Then, (3.2.2) leads to

$$\begin{cases} \bar{\mathcal{C}}(\bar{z})(-\bar{y}) &= -\lambda\bar{x}, \\ \bar{\mathcal{C}}(\bar{z})\bar{x} &= -\lambda(-\bar{y}), \\ \|\bar{x}\| = 1, & \|-\bar{y}\| = 1, \end{cases}$$

which indicates that $\bar{x}, -\bar{y}$ are left-singular and right-singular vectors for $-\lambda$ of $\bar{\mathcal{C}}(\bar{z})$, respectively. From the symmetric property of $\bar{\mathcal{C}}(\bar{z})$, we know that either λ or $-\lambda$ is the eigenvalue of $\bar{\mathcal{C}}(\bar{z})$. Now we consider two cases.

If λ is an eigenvalue of $\bar{\mathcal{C}}(\bar{z})$ associated with an eigenvector x^*, we denote $z^* = \bar{z}$. Then, $\mathcal{C}x^*x^*z^* = \lambda$.

If $-\lambda$ is an eigenvalue of $\bar{\mathcal{C}}(\bar{z})$ associated with an eigenvector x^*, we denote $z^* = -\bar{z}$. Then, $\mathcal{C}x^*x^*z^* = \lambda$.

Therefore, (x^*, x^*, z^*) is a global minimizer of optimization problem (1.1.5). □

With this result in hand, we can assert that optimization problem (1.1.4) can be regarded as a special case of (1.1.5). Thus the NP-hardness of (1.1.5) is also established.

Corollary 3.2.2 *The cubic three-spherical optimization problem (1.1.5) is NP-hard.*

As mentioned earlier, (1.1.4) is a special case of (1.1.5). Therefore, from now on we focus on problems (1.1.3) and (1.1.5).

3.3 Polynomial Time Solvable Cases

In spite of NP-hardness of (1.1.5), there exist some special cases which can be solved in polynomial time. We present a couple of such cases in this section.

For the tensor $\mathcal{C}$ in (1.1.5), it is easy to see that there exist matrices $A^i \in \Re^{p\times q}$, $B^j \in \Re^{q\times n}$, and $C^k \in \Re^{n\times p}$ for $i = 1, 2, \cdots, n, j = 1, 2, \cdots, p$ and $k = 1, 2, \cdots, q$ such

that $\mathcal{C}xyz$ can be rewritten as

$$\mathcal{C}xyz = \sum_{i=1}^{n}(y^\top A^i z)x_i = \sum_{j=1}^{p}(z^\top B^j x)y_j = \sum_{k=1}^{q}(x^\top C^k y)z_k. \tag{3.3.3}$$

Based on this observation, we have the following theorem.

Theorem 3.3.1 *Suppose that there exist orthogonal matrices $P \in \Re^{n\times n}$ and $Q \in \Re^{p\times p}$ such that $P^\top C^k Q$ for all $k = 1, 2, \cdots, q$ are $n \times p$ diagonal matrices. Then (1.1.5) is polynomial time solvable.*

Proof. Without loss of generality, after some orthogonal transformations, we may assume that $n \geq p$ and C^k $(k = 1, 2, \cdots, q)$ are $n \times p$ diagonal matrices. Then

$$\mathcal{C}xyz = \sum_{i,j,k=1}^{n,p,q} C^k_{ij}x_i y_j z_k = \sum_{j=1}^{p} x_j y_j \sum_{k=1}^{q} C^k_{jj} z_k.$$

Therefore, we have that

$$\begin{aligned}
& \max_{\|x\|=1}\max_{\|y\|=1}\max_{\|z\|=1} \mathcal{C}xyz \leq \max_{\|x\|=1}\max_{\|y\|=1}\max_{\|z\|=1} \sum_{j=1}^{p} \tfrac{x_j^2+y_j^2}{2} \left|\sum_{k=1}^{q} C^k_{jj} z_k\right| \\
= & \max_{\|x\|=\|y\|=1} \sum_{j=1}^{p} \tfrac{x_j^2+y_j^2}{2} \max_{\|z\|=1} \sum_{k=1}^{q} C^k_{jj} z_k \leq \max_{\|x\|=\|y\|=1} \sum_{j=1}^{p} \tfrac{x_j^2+y_j^2}{2} \sum_{k=1}^{q} C^k_{j^*j^*} z^*_k \\
= & \sum_{k=1}^{q} C^k_{j^*j^*} z^*_k,
\end{aligned}$$

where j^* and z^* satisfy $\sum_{k=1}^{q} C^k_{j^*j^*} z^*_k = \max_{1\leq j\leq p}\max_{\|z\|=1} \sum_{k=1}^{q} C^k_{jj} z_k$. Note that j^* and z^* can be found in polynomial time. Let x^* and y^* be the unit vectors such that their j^*-th entries are 1 and the rest of entries are 0, then we have that $\max_{\|x\|=\|y\|=\|z\|=1} \mathcal{C}xyz = \mathcal{C}x^*y^*z^*$.

Thus we can assert that $(x^*, y^*, -z^*)$ is an optimal solution of (1.1.5), which complete the proof. □

Remark 3.3.1 *It is necessary to show that $\max_{\|x\|=\|y\|=\|z\|=1} \mathcal{C}xyz = \max_{\|x\|=1}\max_{\|y\|=1}\max_{\|z\|=1} \mathcal{C}xyz$. Suppose that (x^*, y^*, z^*) is an optimal solution of the cubic three-spherical optimization problem. Then for any optimal solution of the multilinear optimization $(\bar{x}, \bar{y}, \bar{z})$, $\mathcal{C}x^*y^*z^* \geq \mathcal{C}\bar{x}\bar{y}\bar{z}$. On the other hand, there holds*

$$\mathcal{C}\bar{x}\bar{y}\bar{z} = \max_{\|y\|=1}\max_{\|z\|=1} \mathcal{C}\bar{x}yz \geq \max_{\|y\|=1}\max_{\|z\|=1} \mathcal{C}x^*yz = \mathcal{C}x^*y^*z^*,$$

where the last equality comes from the fact that (y^*, z^*) is singular vector pair associated with the largest singular value of $\mathcal{C}x^*$. So we have the result.

Similarly, if A^i for $i = 1, 2, \cdots, n$ or B^j for $j = 1, 2, \cdots, p$ can be simultaneously diagonalized by orthogonal matrices, then (1.1.5) is polynomial time solvable.

To get another polynomial time solvable case, we recall (1.2.6), the bi-quadratic optimization problem over unit spheres, considered in [33].

Suppose that λ is the optimal value of (1.1.5) with an optimal solution $(\bar{x}, \bar{y}, \bar{z})$. Without loss of generality, we assume that $\lambda < 0$.

It is easy to show that the cubic three-spherical optimization problem (1.1.5) can be reformulated as the following problem

$$\begin{array}{ll} \max\limits_{x\in\Re^n, y\in\Re^p,} & h_1(x,y) := \|\mathcal{C}xy\|^2 \\ \text{s.t.} & x^\top x = 1, \quad y^\top y = 1. \end{array}$$

In fact,

$$h_1(x,y) = \sum_{k=1}^{q}(\sum_{i=1}^{n}\sum_{j=1}^{p}\mathcal{C}_{ijk}x_iy_j)^2 = \sum_{i,s=1}^{n}\sum_{j,t=1}^{p}(\sum_{k=1}^{q}\mathcal{C}_{ijk}\mathcal{C}_{stk})x_iy_jx_sy_t.$$

Thus, (1.1.5) can be reformulated as the following optimization problem

$$\begin{array}{ll} \min\limits_{x\in\Re^n, y\in\Re^p} & \sum\limits_{i,s=1}^{n}\sum\limits_{j,t=1}^{p}(-\sum\limits_{k=1}^{q}\mathcal{C}_{ijk}\mathcal{C}_{stk})x_iy_jx_sy_t \\ \text{s.t.} & x^\top x = 1, \quad y^\top y = 1. \end{array} \tag{3.3.4}$$

It is easy to see that $(\bar{x}, \bar{y})$ is an optimal solution of optimization problem (3.3.4) with the optimal value $-\lambda^2$ when $(\bar{x}, \bar{y}, \bar{z})$ is an optimal solution of (1.1.5).

Furthermore, for any $x \in \Re^n$ and $y \in \Re^p$, there hold

$$\begin{aligned}
& \sum_{i,s=1}^{n}\sum_{j,t=1}^{p}\sum_{k=1}^{q}\mathcal{C}_{ijk}\mathcal{C}_{stk}x_iy_jx_sy_t \\
= & \sum_{i,s=1}^{n}\tfrac{1}{2}\sum_{j,t=1}^{p}(\sum_{k=1}^{q}(\mathcal{C}_{ijk}\mathcal{C}_{stk}y_jy_t+\mathcal{C}_{sjk}\mathcal{C}_{itk}y_jy_t))x_ix_s \\
= & \sum_{i,s=1}^{n}\sum_{j,t=1}^{p}\tfrac{1}{2}\sum_{k=1}^{q}(\mathcal{C}_{ijk}\mathcal{C}_{stk}+\mathcal{C}_{sjk}\mathcal{C}_{itk})x_iy_jx_sy_t \\
= & \sum_{j,t=1}^{p}(\sum_{i,s=1}^{n}\tfrac{1}{2}\sum_{k=1}^{q}(\mathcal{C}_{ijk}\mathcal{C}_{stk}+\mathcal{C}_{sjk}\mathcal{C}_{itk})x_ix_s)y_jy_t \\
= & \sum_{j,t=1}^{p}\tfrac{1}{2}\sum_{i,s=1}^{n}(\tfrac{1}{2}\sum_{k=1}^{q}(\mathcal{C}_{ijk}\mathcal{C}_{stk}+\mathcal{C}_{sjk}\mathcal{C}_{itk})x_ix_s+\tfrac{1}{2}(\mathcal{C}_{itk}\mathcal{C}_{sjk}+\mathcal{C}_{stk}\mathcal{C}_{ijk})x_ix_s)y_jy_t \\
= & \sum_{i,s=1}^{n}\sum_{j,t=1}^{p}\tfrac{1}{4}\sum_{k=1}^{q}(\mathcal{C}_{ijk}\mathcal{C}_{stk}+\mathcal{C}_{sjk}\mathcal{C}_{itk}+\mathcal{C}_{itk}\mathcal{C}_{sjk}+\mathcal{C}_{stk}\mathcal{C}_{ijk})x_ix_sy_jy_t \\
= & \sum_{i,s=1}^{n}\sum_{j,t=1}^{p}\tfrac{1}{4}\sum_{k=1}^{q}(\mathcal{C}_{ijk}\mathcal{C}_{stk}+\mathcal{C}_{sjk}\mathcal{C}_{itk}+\mathcal{C}_{itk}\mathcal{C}_{sjk}+\mathcal{C}_{stk}\mathcal{C}_{ijk})x_iy_jx_sy_t.
\end{aligned}$$

Therefore, by letting tensor $\tilde{\mathcal{C}}$ have elements

$$\tilde{\mathcal{C}}_{ijst}=\frac{1}{4}\sum_{k=1}^{q}(\mathcal{C}_{ijk}\mathcal{C}_{stk}+\mathcal{C}_{sjk}C_{itk}+\mathcal{C}_{itk}\mathcal{C}_{sjk}+\mathcal{C}_{stk}\mathcal{C}_{ijk}),$$

the optimization problem (3.3.4) will be equivalent to the following optimization problem

$$\begin{aligned}
h_{2min} := \min_{x\in\Re^n, y\in\Re^p} \quad & h_2(x,y)=\sum_{i,s=1}^{n}\sum_{j,t=1}^{p}-\tilde{\mathcal{C}}_{ijst}x_iy_jx_sy_t \\
\text{s.t.} \quad & x^\top x=1, \quad y^\top y=1,
\end{aligned} \tag{3.3.5}$$

where $\tilde{\mathcal{C}}$ satisfy the symmetric property: $\tilde{\mathcal{C}}_{ijst}=\tilde{\mathcal{C}}_{sjit}=\tilde{\mathcal{C}}_{itsj}$ for all $i,s=1,\cdots,n$ and $j,t=1,2,\cdots,p$.

Therefore, the cubic optimization problem (1.1.5) may be regarded as subclasses of the bi-quadratic optimization problem (1.2.6). According to Corollary 3.7 and Theorem 4.1 in [33], we have the following result.

Theorem 3.3.2 *If $min\{n,p,q\}=2$, then the cubic three-spherical optimization problem (1.1.5) is polynomial time solvable.*

3.4 Polynomial Time Approximation Scheme Cases of Three-spherical Optimization Problem

Now we are ready to present some special cases of (1.1.5) which have PTAS. Before proceeding, we consider the cubic polynomial optimization of the special form

$$\begin{array}{ll} \min & \sum\limits_{i,j,k=1,}^{n,p,q} \mathcal{C}_{ijk} x_i y_j z_k \\ \text{s.t.} & \|x\| = 1,\ \|y\| = 1,\ \|z\| = 1, \\ & z \geq 0. \end{array} \tag{3.4.6}$$

In this case, one can choose $z \in \Re^q_+$ to be from grid points $\{0, \sqrt{\frac{1}{d}}, \cdots, \sqrt{\frac{d-1}{d}}, 1\}$ such that $z_1^2 + \cdots + z_q^2 = 1$, for some given integer d. They represent uniform grid points on the partial sphere $\{z \in \Re^q_+ : \|z\| = 1\}$. The total number of such feasible grid points is $\binom{q+d-1}{d}$ which is a polynomial in q for any fixed integer $d \geq 1$.

For each feasible grid point $\bar{z}$, one can solve the maximum singular value of the matrix $\mathcal{C}(\bar{z})$ with its (i,j)-th entry being $\mathcal{C}(\bar{z})_{ij} = \mathcal{C}_{ijk}\bar{z}_k$. It's well known that the singular value of matrix can be solved in polynomial time. Let $\bar{z}$ be the one among these grid points whose objective value is the greatest and $\bar{x}, \bar{y}$ be the corresponding left-singular and right-singular vectors. Then, by Bomze and de Klerk [9], the feasible solution pair $(-\bar{x}, \bar{y}, \bar{z})$ returns a $(1-\frac{1}{d})$-approximation solution to (3.4.6). In conclusion, the following result is established.

Theorem 3.4.1 *There is a PTAS for solving problem (3.4.6).*

Note that the variables x, y and z are equal "in rights", and so Theorem 3.4.1 is valid when we consider the case that $z \geq 0$ is replaced by $x \geq 0$ or $y \geq 0$. In fact, for the original optimization, if we know in advance the sign of optimal vector x^* or y^* or z^*, then the PTAS can be modified slightly. For instance, for A^i, $i = 1, \cdots, n$ as in (3.3.3), if all entries in matrix A^i have the same sign pattern, then we can see that

$$|\mathcal{C}xyz| = |y^\top(\sum_{i=1}^n A^i x_i)z| \leq |y|^\top |\sum_{i=1}^n A^i x_i||z|,$$

where $|\cdot|$ denotes the entries being the absolute value of corresponding entries. Hence the sign of optimal solution (x^*, y^*, z^*) can be known in advance, which leads to a PTAS. In a similar pattern, there exists a PTAS if each matrix B^j or C^k has the same sign.

Furthermore, notice the fact that the number of sign patterns for x, y and z are at most 2^n, 2^p and 2^q, respectively. Thus, if $\min\{n, p, q\}$ is fixed, then a PTAS for (1.1.5) can be obtained by solving subproblem (3.4.6) at most $2^{\min\{n,p,q\}}$ times, which gives the following statement.

Corollary 3.4.1 *If* $\min\{n, p, q\}$ *is fixed, there exists a PTAS based on the grid sampling on simplex for solving (1.1.5).*

On the other hand, if we recall (3.3.5), combining with Corollary 3.7 and Theorem 4.1 in [33], then we have the following result.

Theorem 3.4.2 *If* $min\{n, p, q\}$ *is fixed, then there exists a PTAS based on sum of squares for solving (1.1.5).*

3.5 Relative Quality Bound

In general, Definition 1.1.2 is stronger than Definition 1.1.1. However, we have the following result for odd order spherical optimization problems based on the following theorem.

Theorem 3.5.1 *For an odd order spherical optimization problem, the following results hold.*

If there is a $(1-\epsilon)$*-approximation solution with* $0 \le \epsilon \le \frac{1}{2}$*, then the corresponding value is a* $(1-2\epsilon)$*-bound.*

If there is an r*-bound, then its corresponding approximation solution is a* $\frac{1-r}{2}$*-approximation solution.*

The proof is trivial since the absolute value of the maximum value and the absolute value of the minimum value is the same, so that we omit the details of the proof.

In the subsequent discussion, we will present a quite simple but efficient method for (1.1.5) by finding the singular values of matrices.

Theorem 3.5.2 *There exists a* $\frac{1}{\sqrt{\min\{n,p,q\}}}$*-bound for (1.1.5).*

Proof. To get the relative quality bound, we rewrite (1.1.5) as follows

$$\mathcal{C}xyz = \sum_{k=1}^{q}(x^\top C^k y)z_k,$$

for some suitable matrices C^k as in (3.3.3).

For $k = 1, 2, \cdots, q$, we solve the following problem by finding the singular values of a matrix in polynomial time

$$\begin{aligned} \max \quad & x^\top C^k y \\ \text{s.t.} \quad & \|x\| = 1, \\ & \|y\| = 1. \end{aligned}$$

Let k^* be the one whose objective value is the largest and x^*, y^* be the optimizer and $z^*_{k^*}$ be -1 accordingly and the rest of entries be 0. Then for any feasible solution pair (x, y, z), there holds

$$\begin{aligned} (\mathcal{C}xyz)^2 &= (\sum_{k=1}^{q}(x^\top C^k y)z_k)^2 = \sum_{i,j=1}^{q}(x^\top C^i y)(x^\top C^j y)z_i z_j \\ &\le q\sum_{i=1}^{q}(x^\top C^i y)^2 z_i^2 \le q\sum_{i=1}^{q}(x^{*\top} C^{k^*} y^*)^2 z_i^2 \le q(x^{*\top} C^{k^*} y^*)^2. \end{aligned}$$

Furthermore, there holds

$$(x^{*\top} C^{k^*} y^*)^2 = (\mathcal{C}x^* y^* z^*)^2 \le \max_{\|x\|=\|y\|=\|z\|=1}(\mathcal{C}xyz)^2 \le q(x^{*\top} C^{k^*} y^*)^2,$$

which implies the result. □

3.6 Approximate Algorithms for One-Spherical Optimization Problem

In this section, we turn our attention to (1.1.3). It's easy to solve (1.1.3) when $n \leq 3$ by the direct computation method proposed in the second chapter. Hence, we focus on (1.1.3) for $n \geq 4$.

3.6.1 Approximation Bound Based On Sum of Squares (SOS)

It is easy to see that optimization problem (1.1.3) is equivalent to the following optimization problem

$$\begin{aligned} p_{min} := \max \quad & \gamma \\ \text{s.t.} \quad & \mathcal{A}x^3 - \gamma \geq 0, \ \forall\, x \in \{x \in \Re^n \mid \|x\| = 1\}. \end{aligned}$$

Since the function $\mathcal{A}x^3$ is an odd function, the above optimization problem can be reformulated equivalently to :

$$\begin{aligned} p_{mins} := \max \quad & \beta \\ \text{s.t.} \quad & -(\mathcal{A}x^3)^2 - \beta(x^\top x)^3 \geq 0, \ \forall\, x \in \Re^n. \end{aligned} \tag{3.6.7}$$

Obviously, $p_{mins} = -p_{min}^2$.

Let $N \geq 0$ be an integer. Consider the N-order SOS relaxation:

$$\begin{aligned} p_N := \max \quad & \beta \\ \text{s.t.} \quad & (x^\top x)^N(-(\mathcal{A}x^3)^2 - \beta(x^\top x)^3) \succeq_{SOS} 0, \ x \in \Re^n. \end{aligned} \tag{3.6.8}$$

where the inequality $g \succeq_{SOS} 0$ means that g is SOS, i.e., a sum of squares of other polynomials. It is easy to see that $p_N \geq p_{mins}$.

To obtain the convergence result, we need the following lemma in [60].

Lemma 3.6.1 *Let $p(x)$ be a homogeneous positive definite polynomial of degree m, where m is even and $x \in \Re^n$. Then for any integer N satisfying*

$$N \geq \frac{nm(m-1)}{4\log 2}\varepsilon(p) - \frac{n+m}{2},$$

where $\varepsilon(p) = \frac{\sup\limits_{\|x\|=1} p(x)}{\inf\limits_{\|x\|=1} p(x)}$, $(x^\top x)^N p(x)$ is SOS.

Based upon Lemma 3.6.1, the approximation ratio is presented as follows.

Theorem 3.6.1 *For an even N satisfying $N \geq \frac{15n}{2\log 2} - \frac{n+6}{2}$, we have*

$$0 \leq \frac{p_{mins} - p_N}{-p_{mins}} \leq \frac{15n}{(N + \frac{n+6}{2})2\log 2 - 15n}. \tag{3.6.9}$$

Proof. It is easy to see that 0 is the maximum of $-(\mathcal{A}x^3)^2$ over sphere and

$$p_{mins}(x^\top x)^3 \leq -(\mathcal{A}x^3)^2 \leq 0, \quad \forall x \in \Re^n.$$

Taking $\beta < p_{mins}$, we have

$$(p_{mins} - \beta)(x^\top x)^3 \leq -(\mathcal{A}x^3)^2 - \beta(x^\top x)^3 \leq -\beta(x^\top x)^3, \quad \forall x \in \Re^n.$$

Hence $\varepsilon(-(\mathcal{A}x^3)^2 - \beta(x^\top x)^3) \leq \frac{-\beta}{p_{mins}-\beta}$.
Fix N satisfying $N > \frac{15n}{2\log 2} - \frac{n+6}{2}$, and choose

$$\beta_N = p_{mins} + \frac{15np_{mins}}{(N + \frac{n+6}{2})2\log 2 - 15n} < p_{mins}.$$

Then

$$N = \frac{15n}{2\log 2}\frac{-\beta_N}{p_{mins} - \beta_N} - \frac{n+6}{2}.$$

From Lemma 3.6.1, we can see that $(x^\top x)^N((-\mathcal{A}x^3)^2 - \beta_N(x^\top x)^3)$ is SOS. On the other hand, It is easy to see that $p_N \leq p_{mins}$. Therefore,

$$0 \leq \frac{p_{mins} - p_N}{-p_{mins}} \leq \frac{p_{mins} - \beta_N}{-p_{mins}} = \frac{15n}{(N + \frac{n+6}{2})2\log 2 - 15n}.$$

This completes the proof. □

3.6.2 Approximation Bound Based On Semidefinite Relaxation Scheme

In general, SOS relaxation algorithm might not be quite efficient for practical problems when n is large. Therefore, we discuss approximation bound based on semidefinite relaxation.

It is well known that semidefinite relaxation method is an efficient method for even order polynomial optimization problem. Therefore, we first reformulate (1.1.3) as the following quartic optimization problem

$$\begin{aligned} \min_{\bar{x}\in\Re^{n+1}} \quad & \bar{\mathcal{G}}\bar{x}^4 = \sum_{i,j,k,l=0}^{n} \bar{\mathcal{G}}_{ijkl}\bar{x}_i\bar{x}_j\bar{x}_k\bar{x}_l \\ \text{s.t.} \quad & \bar{x}_0 = 1, \quad \|\bar{x}\| = 2, \end{aligned} \tag{3.6.10}$$

where $\bar{\mathcal{G}}$ is a fourth order $((n+1)\times(n+1)\times(n+1)\times(n+1))$-dimensional real symmetric tensor with

$$\bar{\mathcal{G}}_{ijkl} = \begin{cases} \frac{1}{4\mathcal{A}_{ijk}} & \text{for } l = 1, \text{and } i,j,k \neq 1, \\ 0, & \text{Otherwise.} \end{cases}$$

Consider the following SDP relaxation of (3.6.10),

$$\begin{aligned} \min \quad & (\bar{\mathcal{G}}X) \bullet X \\ \text{s.t.} \quad & X_{11} = 1, \quad I \bullet X = 2, \\ & X \succeq 0. \end{aligned} \tag{3.6.11}$$

Based on the relationship in [6]: $\mathcal{S}_+^{m\times m} \subseteq \{X \in \mathcal{S}^{m\times m} : \|X\|_F \leq I_m \bullet X\}$, we may relax (3.6.11) as

$$\begin{aligned} \min \quad & (\bar{\mathcal{G}}X) \bullet X \\ \text{s.t.} \quad & X_{11} = 1, \quad I \bullet X = 2, \\ & \|X\|_F \leq 2. \end{aligned} \tag{3.6.12}$$

Stack up the entries of symmetric matrix X, denoted by $vec_S(\cdot)$. Then there exists a suitable matrix Q such that the quadratic SDP relaxation can be rewritten in the vector form

$$\begin{aligned} \min \quad & vec_S(X)^\top Q vec_S(X) \\ \text{s.t.} \quad & (vec_S(X))_1 = 1, \ vec_S(I_{n+1})^\top vec_S(X) = 2, \\ & \|vec_S(X)\| \leq 2. \end{aligned}$$

After eliminating two variables by the first two constraints, there exist two suitable quadratic functions $q_0(w)$ and $q_1(w)$ such that the relaxation can be rewritten as

$$\begin{aligned} \min \quad & q_0(w) \\ \text{s.t.} \quad & q_1(w) \leq 4. \end{aligned}$$

Using the matrix decomposition in [65], we can see that the standard SDP relaxation for above quadratic optimization problem is tight. By the obtained optimization w^* of above quadratic optimization, together with the stack relation between vector w and matrix X, we can get a feasible solution of (3.6.12), which returns a lower bound for (3.6.11) and (1.1.3) in polynomial time.

3.7 Practical Method for Solving Problem (1.1.5)

It's easy to see that the approximation solution given in section 3.5 is not a KKT point of problem (1.1.5). Therefore, we will present a modified power method to get a 'good' solution of problem (1.1.5). Without loss of generality, we assume that $n \leq p \leq q$.

Algorithm 3.7.1

- *Initial step: Input tensor* $\mathcal{C}$ *and matrices* A^i*, with entries*

$$(A^i)_{jk} = \mathcal{C}_{ijk}, \quad \forall i = 1, 2, \cdots, n.$$

  Substep 1: Compute the largest singular value λ^i *and its corresponding left-singular and right-singular vectors* y^i, z^i *for matrices* A^i*,* $i = 1, 2, \cdots, n$*.*

  Substep 2: Choose the index i_0 *with the singular value* λ^{i_0} *being the largest among* $\{\lambda^1, \lambda^2, \cdots, \lambda^n\}$*.*

  Substep 3: Let $y^0 = y^{i_0}, z^0 = z^{i_0}$*,* x^0 *be the vector with* $x^0_{i_0} = -1$ *and the rest of entries be* 0*. Let* $l = 0$*.*

- *Iterative step: Compute matrix* $\mathcal{C}x^l \in \Re^{p\times q}$ *with its entries*

$$(\mathcal{C}x^l)_{jk} = \sum_{i=1}^{n} \mathcal{C}_{ijk} x^l_i.$$

  Compute the largest singular value and its corresponding left-singular and right-singular vectors $y^{l+1}, -z^{l+1}$ *for matrix* $\mathcal{C}x^l$*. Let* x^{ln} *be the vector with entries* $x^{ln}_i = \sum_{j=1}^{p}\sum_{k=1}^{q} \mathcal{C}_{ijk} y^{l+1}_j z^{l+1}_k$*. Take* $x^{l+1} = \frac{x^{ln}}{\|x^{ln}\|}$ *and* $l = l + 1$*.*

Obviously, Algorithm 3.7.1 includes two parts: initial step and iterative step. By means of Theorem 3.5.2, the initial step provides a lower bound satisfying

$$\mathcal{C}xyz \geq \sqrt{n}\mathcal{C}x^0y^0z^0, \quad \forall \ \ \|x\| = \|y\| = \|z\| = 1.$$

Furthermore,

$$\mathcal{C}x^{l+1}y^{l+1}z^{l+1} = -\|\mathcal{C}y^{l+1}z^{l+1}\| \leq \mathcal{C}x^ly^{l+1}z^{l+1} \leq \mathcal{C}x^ly^lz^l,$$

which indicates that the sequence $\{\mathcal{C}x^ly^lz^l\}$ generated by Algorithm 3.7.1 is non-increasing. Based upon this observation, we establish the following convergence result.

Theorem 3.7.1 *Suppose that the sequence $\{x^l, y^l, z^l\}$ generated by Algorithm 3.7.1 is infinite. Then each accumulation point of the sequence $\{x^l, y^l, z^l\}$ is a KKT point of the problem (1.1.5).*

Proof. From Iterative step of Algorithm 3.7.1, we have

$$\begin{aligned} \mathcal{C}x^l y^{l+1} &= -\lambda^l z^{l+1}, \\ \mathcal{C}x^l z^{l+1} &= -\lambda^l y^{l+1}, \\ \mathcal{C}y^{l+1} z^{l+1} &= -\|\mathcal{C}y^{l+1} z^{l+1}\| x^{l+1}, \\ \|x^{l+1}\| &= \|y^{l+1}\| = \|z^{l+1}\| = 1. \end{aligned}$$

Note that $\mathcal{C}xyz$ is bounded on the unit sphere. Combining this with the fact that $\{\mathcal{C}x^{l+1}y^{l+1}z^{l+1}\}$ is non-increasing, we know that $\{\mathcal{C}x^{l+1}y^{l+1}z^{l+1}\}$ is convergent and there holds

$$\begin{aligned} & \mathcal{C}x^{l+1}y^{l+1}z^{l+1} - \mathcal{C}x^l y^l z^l \\ = \; & \mathcal{C}x^{l+1}y^{l+1}z^{l+1} - \mathcal{C}x^l y^{l+1} z^{l+1} + \mathcal{C}x^l y^{l+1} z^{l+1} - \mathcal{C}x^l y^l z^l \\ \leq \; & \mathcal{C}x^{l+1}y^{l+1}z^{l+1} - \mathcal{C}x^l y^{l+1} z^{l+1} \\ = \; & \mathcal{C}(x^{l+1} - x^l) y^{l+1} z^{l+1} \leq 0, \end{aligned}$$

where in the last inequality we use the fact that $\mathcal{C}x^{l+1}y^{l+1}z^{l+1} \leq \mathcal{C}xy^{l+1}z^{l+1}$ for any $\|x\| = 1$ from the obtained x^{l+1}.

From the convergence of sequence $\{\mathcal{C}x^l y^l z^l\}$, we know $\mathcal{C}x^{l+1}y^{l+1}z^{l+1} - \mathcal{C}x^l y^l z^l \to 0$. Furthermore, $\|\mathcal{C}y^{l+1}z^{l+1}\| \geq |\mathcal{C}x^0 y^0 z^0| > 0$, which leads to $\|x^{l+1} - x^l\| \to 0$ since $\mathcal{C}y^{l+1}z^{l+1} = -\|\mathcal{C}y^{l+1}z^{l+1}\| x^{l+1}$. Therefore, $\{x^l\}$ is convergent. As a result, $\{\lambda^l\}$ is also convergent.

Suppose that (x^*, y^*, z^*) is an accumulation point of $\{(x^l, y^l, z^l)\}$. Without loss of generality, we assume that

$$\lim_{l\to\infty} x^l = x^*, \quad \lim_{l_k\to\infty} y^{l_k} = y^*, \quad \lim_{l_k\to\infty} z^{l_k} = z^*, \quad \lim_{l\to\infty} \lambda^l = \lambda^*.$$

Then, there holds

$$\begin{aligned} \mathcal{C}x^* y^* &= -\lambda^* z^* \\ \mathcal{C}x^* z^* &= -\lambda^* y^*, \\ \mathcal{C}y^* z^* &= -\|\mathcal{C}y^* z^*\| x^*, \\ \|x^*\| &= \|y^*\| = \|z^*\| = 1. \end{aligned}$$

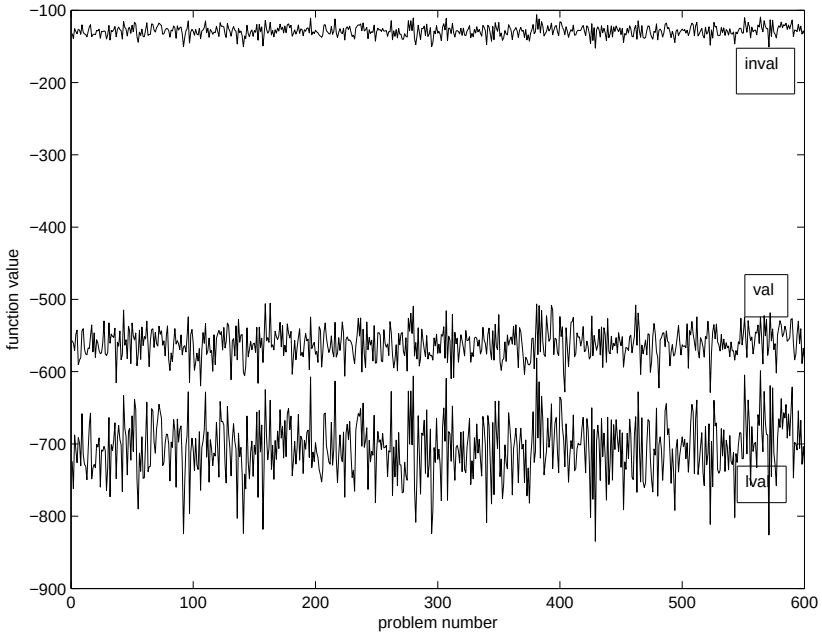

Figure 3.7.1: Numerical Results of Example 3.7.1

As a consequence, we have

$$-\lambda^* = -\|\mathcal{C}y^*z^*\| = \mathcal{C}x^*y^*z^*,$$

which completes the proof. □

To illustrate that the procedure performs well, the following two examples are considered.

Example 3.7.1 *The example is taken from [74]. Tensor* $\mathcal{C}$ *is a random low-rank* $(40 \times 30 \times 40)$*-dimensional tensor generated as the sum of* 20 *rank-one tensors, each rank-one tensor* $x \otimes y \otimes z$ *is generated with components of* x, y *and* z *uniformly distributed in* $(0, 1)$*. Figure 3.7.1 shows the numerical results, where "inval" and "lval" denote the initial value and corresponding lower bound obtained by the initial step of Algorithm 3.7.1, respectively. "val" denotes the value obtained by Algorithm 3.7.1. From Figure 3.7.1, we can claim that the Algorithm performs well.*

Example 3.7.2 *Consider the problem (1.1.5) with* $(n \times p \times q)$*-dimensional tensor* C *whose entries are uniformly distributed in* $(0, 1)$*.*

Table 3.7.2 reports the average value of 100 problems, where "Dim" denotes the dimension of Tensor $\mathcal{C}$, "lval" denotes the average lower bound by the initial step and Theorem 3.5.2, "alval" denotes the average value obtained by Algorithm 3.7.1. "CPU" denotes the average working time of the computer excluding the input/output time.

Dim	upval	alval	CPU
$(3 \times 5 \times 10)$	-6.7970	-6.2303	0.0441
$(5 \times 10 \times 30)$	-20.5247	-19.4263	0.0563
$(10 \times 30 \times 50)$	-63.2307	-61.3158	0.0798
$(30 \times 50 \times 100)$	-197.9735	-193.6653	0.4516
$(50 \times 100 \times 300)$	-618.4687	-612.3631	3.8347
$(100 \times 300 \times 500)$	-1945.3	-1936.4	22.5809

Table 3.7.2: The numerical results of Example 3.7.2

3.8 Some Remarks

In this chapter, we discuss the cubic spherical optimization problems. We present the NP-hardness of the cubic two-spherical/three-spherical optimization problems. Since the cubic two-spherical optimization is a special case of the three-spherical optimization problem, we focus on the cubic three-spherical optimization problem. We discuss some special cases of three-spherical optimization which can be polynomial time solvable or have PTAS. For general cases, we present a quality bound. For one-spherical optimization, some approximation bounds are discussed based on SOS and SDP schemes. Finally, a practical method for solving the cubic three-spherical optimization problem is proposed. However, there is little discussion for the cubic one-spherical optimization problem. So in future study, how to obtain approximation bound will be considered and it is possible to design some practical methods.

Chapter 4

A Practical Method for the Bi-quadratic Optimization Problem Over Unit Spheres

4.1 Introduction

To solve problem (1.2.6), we first consider its optimal condition following the tensor product notation used in [51, 15, 70]. Before proceeding, we denote $\mathcal{F} \cdot xyy$ as a vector with its ith component being $\sum_{j=1}^{m} \sum_{k,l=1}^{n} \mathcal{F}_{ijkl} x_j y_k y_l$, and $\mathcal{F}xxy\cdot$ as a vector with its lth component being $\sum_{i,j=1}^{m} \sum_{k=1}^{n} \mathcal{F}_{ijkl} x_i x_j y_k$. For any minimizer (x, y) of (1.2.6), by the optimality theory [7, 47], there exist $\lambda, \mu \in \Re$ such that

$$\begin{cases} \mathcal{F} \cdot xyy = \lambda x \\ \mathcal{F}xxy\cdot = \mu y \\ x^\top x = 1 \\ y^\top y = 1 \end{cases}$$

Certainly, the optimal condition can further be simplified as

$$\begin{cases} \mathcal{F} \cdot xyy = \lambda x \\ \mathcal{F}xxy\cdot = \lambda y \\ x^\top x = 1 \\ y^\top y = 1 \end{cases} \tag{4.1.1}$$

If λ, x and y are real solutions of (4.1.1), λ is said to be an M-eigenvalue of tensor $\mathcal{F}$, and x and y are said to be a left M-eigenvector and a right M-eigenvector of tensor $\mathcal{F}$, associated with the M-eigenvalue λ, respectively [54]. Here, the letter M is borrowed from word *material*.

By the discussion above, we know that problem (1.2.6) is equivalently transformed into computing the largest M-eigenvalue of tensor $\mathcal{F}$, i.e., solving (4.1.1). For this system, it seems not difficult to solve. However, this is not true since neither equation of system (4.1.1) is linear. Moreover, this problem is also shown to be NP-hard [33]. Motivated by the fact that this problem is a subproblem of the entanglement problem, in practice, we may aim to find a "good" solution of the problem. In the next section, we will propose a practical method to compute the largest M-eigenvalue of tensor $\mathcal{F}$ based on the power method for computing the largest eigenvalue in magnitude of a matrix [19]. Compared with the alternating eigenvalue maximization method for solving (4.1.1) proposed in [17], the computation cost of our method is less. As for the validity of this method, it is guaranteed theoretically for the convex case. To make the conclusion hold generally, we introduce a translation technique into the method. Furthermore, to make the generated sequence converge to a good solution of the problem, we also develop an initialization scheme in Section 4.3. The given numerical experiments in Section 4.4 show that the proposed method could generate a well-approximated point to the global maximizer of our concerned problem.

4.2 Practical Power Method and Its Convergence

It is well known that the power method is an efficient method for computing the largest eigenvalue (in the sense of absolute value) of a matrix [19]. This method was extended to compute the best rank-1 approximations to higher-order tensors, i.e., the largest

Z-eigenvalue (also in the sense of absolute value) of higher-order tensors [32, 31, 54]. Motivated by this, we propose a modified power method to compute the largest M-eigenvalue of a tensor.

Algorithm 4.2.1

Initialization Step. Take initial points $x^0 \in \Re^m, y^0 \in \Re^n$ *with* $\|x^0\| = \|y^0\| = 1$, *and let* $t = 0$;

Iterative Step. Execute the following procedures alternatively until convergence:

$$\bar{x}^{t+1} = \mathcal{F} \cdot x^t y^t y^t, \quad x^{t+1} = \frac{\bar{x}^{t+1}}{\|\bar{x}^{t+1}\|};$$

$$\bar{y}^{t+1} = \mathcal{F} x^{t+1} x^{t+1} y^t \cdot, \quad y^{t+1} = \frac{\bar{y}^{t+1}}{\|\bar{y}^{t+1}\|};$$

$$t = t + 1.$$

In the following, we would give a theoretical analysis to the method. For the objective function $F(x, y)$, from (1.2.6), we know that it is a bi-quadratic function with respect to variables x, y, respectively. That is, the function can be written as

$$F(x, y) = \mathcal{F} x x y y = x^\top B(y) x = y^\top C(x) y,$$

where $B(y)$ and $C(x)$ are respectively symmetric matrices in $\Re^{m\times m}$ and $\Re^{n\times n}$ with components

$$B_{ij}(y) = \sum_{k,l=1}^{n} \mathcal{F}_{ijkl} y_k y_l, \qquad C_{kl}(x) = \sum_{i,j=1}^{m} \mathcal{F}_{ijkl} x_i x_j.$$

Based on this analysis, we have the following conclusion for Algorithm 4.2.1.

Theorem 4.2.1 *Suppose that for any* $x \in \Re^m, y \in \Re^n$, *the matrices* $B(y)$ *and* $C(x)$ *are both positive definite. Then the generated sequence* $\{F(x^t, y^t)\}$ *by Algorithm 4.2.1 is nondecreasing.*

Proof. From the assumption, we know that for any fixed $x \in \Re^m$, the function $F(x, y)$ is strictly convex with respect to $y \in \Re^n$, and similarly, the function $F(x, y)$ is strictly convex with respect to $x \in \Re^m$ for any fixed $y \in \Re^n$. Thus, for any $t \geq 0$,

$$F(x^{t+1}, y^t) - F(x^t, y^t) \geq \langle x^{t+1} - x^t, \nabla_x F(x^t, y^t) \rangle. \tag{4.2.2}$$

To show the conclusion, we are ready to show the nonnegativity of the inner product term in the right-hand side. Since

$$\nabla_x F(x^t, y^t) = 2\mathcal{F} \cdot x^t y^t y^t,$$

from Algorithm 4.2.1, one has

$$x^{t+1} = \frac{\nabla_x F(x^t, y^t)}{\|\nabla_x F(x^t, y^t)\|}.$$

Recalling the Cauchy-Schwartz inequality and the choice of x^{t+1}, we conclude from (4.2.2) that $F(x^{t+1}, y^t) \geq F(x^t, y^t)$ and it holds strictly if $x^{t+1} \neq x^t$.

Similarly, we can conclude that that $F(x^{t+1}, y^{t+1}) \geq F(x^{t+1}, y^t)$ and it holds strictly if $y^{t+1} \neq y^t$.

Combining these two cases, we obtain the desired result. □

Before giving an analysis to Theorem 4.2.1, we first give an explanation of the condition of Theorem 4.2.1 by introducing the following definition.

Definition 4.2.1 *A fourth order partially symmetric tensor $\mathcal{F}$ is said to be positive definite on $\Re^m \times \Re^n$ if for any nonzero vectors $x \in \Re^m$ and $y \in \Re^n$, it holds that*

$$\mathcal{F}xxyy = \sum_{i,j=1}^{m} \sum_{k,l=1}^{n} \mathcal{F}_{ijkl} x_i x_j y_k y_l > 0.$$

From this definition, we know that the matrices $B(y)$ and $C(x)$ are both positive definite for any nonzero vectors $x \in \Re^m$ and $y \in \Re^n$ if and only if tensor $\mathcal{F}$ is positive definite on $\Re^m \times \Re^n$.

From Theorem 4.2.1, we may conclude that the generated sequence $\{(x^t, y^t)\}$ converges to a stationary point of problem (1.2.6) in the "convex" case [7]. Now, one key

problem is posed naturally: how about the algorithm for the general case? That is, for the case that either matrix $B(y)$ or matrix $C(x)$ is not positive definite. To solve this problem, we now make a translation to the corresponding tensor in the objective function by introducing the following "identity" tensor $\mathcal{I} \in \Re^{m\times m\times n\times n}$:

$$\mathcal{I}_{ijkl} = \begin{cases} 1, & \text{if } i = j \text{ and } k = l \\ 0, & \text{otherwise} \end{cases}$$

Take $\tau \in R$ such that $\tau > \max\{|\lambda| \mid \lambda \text{ is an M-eigenvalue of tensor } \mathcal{F}\}$ and set

$$\bar{F}(x, y) = \tau \mathcal{I}xxyy + \mathcal{F}xxyy \triangleq \bar{\mathcal{F}}xxyy.$$

It is easy to see that tensor $\bar{\mathcal{F}}$ is positive definite on $\Re^m \times \Re^n$ and has the same symmetry property as tensor $\mathcal{F}$. Furthermore, if x and y constitute a pair of M-eigenvectors of tensor $\bar{\mathcal{F}}$ associated with M-eigenvalue λ, then they are also a pair of M-eigenvectors of tensor $\mathcal{F}$ associated with M-eigenvalue $(\lambda-\tau)$. Since function $\bar{F}(x, y)$ satisfies the assumptions in Theorem 4.2.1, we can apply Algorithm 2.1 to compute the largest M-eigenvalue of tensor $\bar{\mathcal{F}}$ and hence we can obtain the largest M-eigenvalue of tensor $\mathcal{F}$.

Now, one more question rises accordingly: How to choose a suitable τ? In fact, this can be solved based on the estimation of the largest eigenvalue of the unfolded matrix of tensor $\mathcal{F}$ defined below.

Define the following index mapping from four indices i, j, k, l to two indices s, t:

$$s = n(i - 1) + k, \quad t = m(j - 1) + l.$$

Using this mapping, we may unfold tensor $\mathcal{F}$ into a matrix $A \in \Re^{mn\times mn}$. From the partial symmetry of tensor $\mathcal{F}$, we know that the unfolded matrix A is symmetric. Based on this representation, the objective function $F(x, y)$ can be written as a quadratic form:

$$F(x, y) = (x \otimes y)^\top A(x \otimes y),$$

where $x \otimes y$ denotes the Kronecker product of vectors x and y which is a vector in $\Re^{mn}$. Based on this, we can immediately obtain the following conclusion.

Proposition 4.2.1 *Suppose matrix A is the unfolded matrix of tensor $\mathcal{F}$. Then tensor $\mathcal{F}$ is positive definite on $\Re^m \times \Re^n$ provided that matrix A is positive definite on $\Re^{mn}$.*

Furthermore, all the M-eigenvalues of tensor $\mathcal{F}$ lie in the interval composed by the smallest eigenvalue and the largest eigenvalue of matrix A.

From the Geršgorin Disc Theorem [19], we know that the magnitude of any eigenvalue of matrix A must be less than or equal to

$$\max_{1\le s\le mn}\sum_{t=1}^{mn}|A_{st}|.$$

This can easily be computed. In fact, since matrix A is symmetric, $\tau = \sum_{1\le i\le j\le mn}|A_{ij}|$ is sufficiently large to guarantee that $(\tau\mathcal{I}+\mathcal{F})$ satisfies the assumption in Theorem 4.2.1.

To end this section, we give an example to show that a fourth order partially symmetric tensor $\mathcal{F}$ may be positive definite on $\Re^m \times \Re^n$ but the corresponding unfolded matrix A is not positive definite on $\Re^{mn}$.

Example 4.2.1 *Consider the following fourth order two-dimensional tensor $\mathcal{F}$ with entries*

$$\mathcal{F}_{1111}=12,\quad \mathcal{F}_{1112}=1,\quad \mathcal{F}_{1122}=2,$$
$$\mathcal{F}_{1222}=1,\quad \mathcal{F}_{1212}=5,\quad \mathcal{F}_{1222}=1,\quad \mathcal{F}_{2222}=12.$$

For this symmetric tensor, the unfolded matrix is:

$$A=\begin{pmatrix} 12 & 1 & 1 & 5\\ 1 & 2 & 5 & 1\\ 1 & 5 & 2 & 1\\ 5 & 1 & 1 & 12\end{pmatrix}$$

It is easy to verify that for $x, y \in \Re^2$,

$$\begin{aligned} F(x,y) &= \mathcal{F}xyxy = \sum_{i,j,k,l=1}^{2}\mathcal{F}_{ijkl}x_ix_jy_ky_l\\ &= (x_1y_1+x_1y_2)^2+(x_1y_2+x_2y_2)^2+(x_2y_1+x_1y_1)^2+(x_2y_1+x_2y_2)^2\\ &\quad +10(x_1y_1+x_2y_2)^2, \end{aligned}$$

and it is positive for any nonzero vectors $x, y \in \Re^2$. However, for $w=(0,1,-1,0)^\top$, $w^\top Aw=-6<0$, i.e., the unfolded matrix A is not positive definite on $\Re^4$.

4.3 Initialization Technique

From the discussion in section 4.2, we know that the given power method can generate a stationary point of problem (1.2.6) generally. However, whether an accumulation point of the generated sequence is a "good" maximizer of problem (1.2.6) or not strongly depends on the initialization. In this sense, we claim that Algorithm 4.2.1 only partially solves problem (1.2.6).

To obtain a good maximizer of problem (1.2.6), we will introduce an initialization technique into Algorithm 4.2.1 inspired by the initialization strategy for computing the best rank-1 symmetric approximation to a symmetric tensor in [59]. This initialization technique is based on the basic fact that the unit eigenvector corresponding to the largest eigenvalue of positive definite and symmetric matrix G is a maximizer of the function $g(x) = x^\top Gx$ over the unit sphere.

Suppose the unfolded matrix A of tensor $\mathcal{F}$ is positive definite, $w \in \Re^{mn}$ is an unit eigenvector of matrix A associated with the largest eigenvalue μ, and (x^*, y^*) is a solution of problem (1.2.6). Then $f(x^*, y^*) \leq \mu$ and the equality holds only when $x^* \otimes y^*$ coincide with $\pm w$. Motivated by this, we may take an initial point (x_0, y_0) in Algorithm 4.2.1 that maximizes the inner product $\langle x \otimes y, w\rangle$ over unit spheres.

To solve the subproblem, we need to fold vector w into a matrix in $\Re^{m\times n}$ in the following way: for $k = 1, 2, \cdots, mn$, set $i = \lceil \frac{k}{n} \rceil$, $j = (k-1) \bmod n \ + 1$ and

$$W_{ij} = w_k.$$

Then the inner product $\langle x \otimes y, w\rangle$ can be expressed as a bi-linear function $x^\top Wy$ and the involved subproblem is as follows

$$\begin{aligned} \max \quad & x^\top Wy \\ \text{s.t.} \quad & x^\top x = 1, y^\top y = 1 \end{aligned}$$

which is also equivalent to the following minimization problem [32]:

$$\begin{aligned} \min \quad & \|W - \mu xy^\top\|_{\mathrm{F}}^2 \\ \text{s.t.} \quad & \mu \in R \\ & x^\top x = 1, y^\top y = 1 \end{aligned} \tag{4.3.3}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm of a matrix. This optimization problem can easily be solved via singular eigenvalue decomposition of matrix W [19]. That is, if matrix W has the following singular eigenvalue decomposition

$$W = U^{\top}\Sigma V = \sum_{i=1}^{r} \sigma_i u_i v_i^{\top}$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and r is the rank of the matrix, then u_1 and v_1 constitute the solution of problem (4.3.3). Thus, the initialization subproblem can be solved by letting $x_0 = u_1, y_0 = v_1$.

4.4 Numerical Experiments and Simulations

Combining discussions in Sections 4.2 and 4.3, we can give a complete algorithm to compute a good approximation of the largest M-eigenvalue of a fourth order partially symmetric tensor.

Algorithm 4.4.1

Initial Step: Input $\mathcal{F}$ and unfold it to obtain matrix A.

Substep 1: Take $\tau = \sum_{1\leq s\leq t\leq mn} |A_{st}|$, set $\bar{\mathcal{F}} = \tau\mathcal{I} + \mathcal{F}$ and unfold $\bar{\mathcal{F}}$ to matrix $\bar{A}$.

Substep 2: Compute the eigenvector w of matrix $\bar{A}$ associated with the largest eigenvalue and fold it into the matrix W.

Substep 3: Compute the singular vectors u_1 and v_1 corresponding to the largest singular value of the matrix W.

Substep 4: Take $x_0 = u_1, y_0 = v_1$, and let $t = 0$.

Iterative Step: Execute the following procedures alternatively until certain convergence criterion is satisfied and output x^*, y^*:

$$\bar{x}_{t+1} = \bar{\mathcal{F}} \cdot x^t y^t y^t, \quad x^{t+1} = \frac{\bar{x}^{t+1}}{\|\bar{x}^{t+1}\|};$$

$$\bar{y}^{t+1} = \bar{\mathcal{F}} x^{t+1} x^{t+1} y^t \cdot, \quad y^{t+1} = \frac{\bar{y}^{t+1}}{\|\bar{y}^{t+1}\|};$$

$t = t + 1$.

Final Step: Output the largest M-eigenvalue of tensor $\mathcal{F}$: $\lambda = \bar{\mathcal{F}}x^*x^*y^*y^* - \tau$, and the associated M-eigenvectors: x^*, y^*.

It is easy to see that the algorithm contains two parts: the initial step and the iterative step. In fact, the initial step, i.e., computing the largest eigenvalue and the corresponding eigenvector of a matrix, is also an iterative scheme [19]. For Algorithm 4.4.1, the computation complexity at each iterative step is of order $\mathrm{O}(m^2n+mn^2)$. Thus, if the largest M-eigenvalue of tensor $\mathcal{A}$ can be generated within few steps, this algorithm can be said to be practical. To check this, we first make the numerical experiments of Algorithm 4.4.1 on two fourth order three dimensional partially symmetric tensors, and then compute the global optimal values of the objective functions by the projected gradient method [10], combined with the uniform grid method in high order accuracy.

Example 4.4.1 *Consider the tensor whose entries are uniformly generated in (-1,1):*

$$\mathcal{F}(1,:,1,:) = \begin{pmatrix} -0.9727 & 0.3169 & -0.3437 \\ -0.6332 & -0.7866 & 0.4257 \\ -0.3350 & -0.9896 & -0.4323 \end{pmatrix}$$

$$\mathcal{F}(1,:,2,:) = \begin{pmatrix} 0.3169 & 0.6158 & -0.0184 \\ -0.7866 & 0.0160 & 0.0085 \\ -0.9896 & -0.6663 & 0.2559 \end{pmatrix}$$

$$\mathcal{F}(1,:,3,:) = \begin{pmatrix} -0.3437 & -0.0184 & 0.5649 \\ 0.4257 & 0.0085 & -0.1439 \\ -0.4323 & 0.2559 & 0.6162 \end{pmatrix}$$

$$\mathcal{F}(2:,1,:) = \begin{pmatrix} -0.6332 & -0.7866 & 0.4257 \\ 0.7387 & 0.6873 & -0.3248 \\ -0.7986 & -0.5988 & -0.9485 \end{pmatrix}$$

$$\mathcal{F}(2,:,2,:) = \begin{pmatrix} -0.7866 & 0.0160 & 0.0085 \\ 0.6873 & 0.5160 & -0.0216 \\ -0.5988 & 0.0411 & 0.9857 \end{pmatrix}$$

$$\mathcal{F}(2,:,3,:) = \begin{pmatrix} 0.4257 & 0.0085 & -0.1439 \\ -0.3248 & -0.0216 & -0.0037 \\ -0.9485 & 0.9857 & -0.7734 \end{pmatrix}$$

$$\mathcal{F}(3,:,1,:) = \begin{pmatrix} -0.3350 & -0.9896 & -0.4323 \\ -0.7986 & -0.5988 & -0.9485 \\ 0.5853 & 0.5921 & 0.6301 \end{pmatrix}$$

$$\mathcal{F}(3,:,2,:) = \begin{pmatrix} -0.9896 & -0.6663 & 0.2559 \\ -0.5988 & 0.0411 & 0.9857 \\ 0.5921 & -0.2907 & -0.3881 \end{pmatrix}$$

$$\mathcal{F}(3,:,3,:) = \begin{pmatrix} -0.4323 & 0.2559 & 0.6162 \\ -0.9485 & 0.9857 & -0.7734 \\ 0.6301 & -0.3881 & -0.8526 \end{pmatrix}$$

For this tensor, its largest M-eigenvalue 2.3227 which is marked in Figure 4.4.1 by the horizontal line.

Example 4.4.2 *Consider the tensor whose entries are uniformly generated in (0,5):*

$$\mathcal{F}(1,:,1,:) = \begin{pmatrix} 1.9832 & 1.0023 & 4.2525 \\ 2.6721 & 3.2123 & 2.8761 \\ 4.6384 & 2.9484 & 4.0319 \end{pmatrix}$$

$$\mathcal{F}(1,:,2,:) = \begin{pmatrix} 1.0023 & 4.9748 & 2.3701 \\ 3.2123 & 1.3024 & 3.2064 \\ 2.9484 & 4.9946 & 3.8951 \end{pmatrix}$$

$$\mathcal{F}(1,:,3,:) = \begin{pmatrix} 4.2525 & 2.3701 & 2.4709 \\ 2.8761 & 3.2064 & 3.4492 \\ 4.0319 & 3.8951 & 0.6581 \end{pmatrix}$$

$$\mathcal{F}(2,:,1,:) = \begin{pmatrix} 2.6721 & 3.2123 & 2.8761 \\ 3.0871 & 0.1393 & 4.4704 \\ 1.7450 & 3.0394 & 4.6836 \end{pmatrix}$$

$$\mathcal{F}(2,:,2,:) = \begin{pmatrix} 3.2123 & 1.3024 & 3.2064 \\ 0.1393 & 4.9456 & 2.9980 \\ 3.0394 & 4.3263 & 0.5925 \end{pmatrix}$$

$$\mathcal{F}(2,:,3,:) = \begin{pmatrix} 2.8761 & 3.2064 & 3.4492 \\ 4.4704 & 2.9980 & 0.4337 \\ 4.6836 & 0.5925 & 4.3514 \end{pmatrix}$$

$$\mathcal{F}(3,:,1,:) = \begin{pmatrix} 4.6384 & 2.9484 & 4.0319 \\ 1.7450 & 3.0394 & 4.6836 \\ 0.3741 & 1.6947 & 2.7677 \end{pmatrix}$$

$$\mathcal{F}(3,:,2,:) = \begin{pmatrix} 2.9484 & 4.9946 & 3.8951 \\ 3.0394 & 4.3263 & 0.5925 \\ 1.6947 & 4.2633 & 0.1524 \end{pmatrix}$$

$$\mathcal{F}(3,:,3,:) = \begin{pmatrix} 4.0319 & 3.8951 & 0.6581 \\ 4.6836 & 0.5925 & 4.3514 \\ 2.7677 & 0.1524 & 2.2336 \end{pmatrix}$$

For this tensor, its largest M-eigenvalue is 26.1187 which is marked in Figure 4.4.2 by the horizontal line.

From Figures 4.4.1 and 4.4.2, we can see that the largest M-eigenvalue can be highly approximated within few steps especially for the second example. In fact, we have done many numerical experiments of Algorithm 4.4.1 on tensors whose entries are uniformly generated in $(0, L)$ for some positive number L, and the numerical results show that this algorithm has a particularly good performance for this kind of tensor.

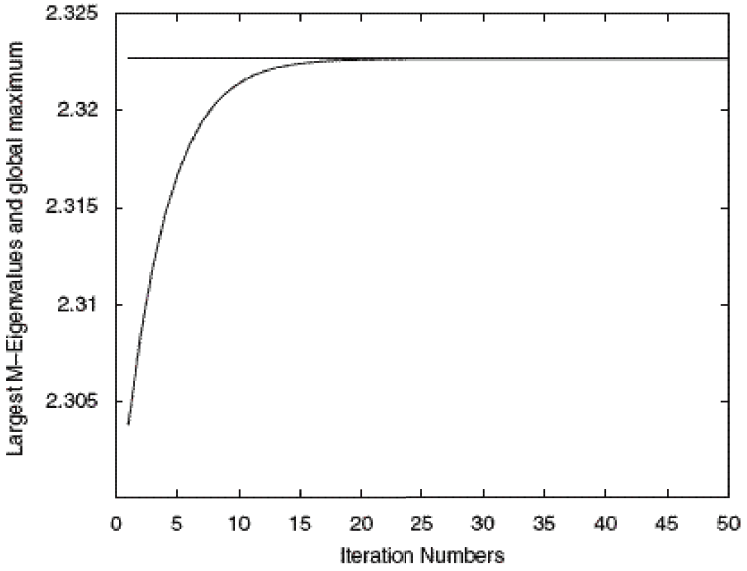

Figure 4.4.1: Numerical Result of Example 4.4.1

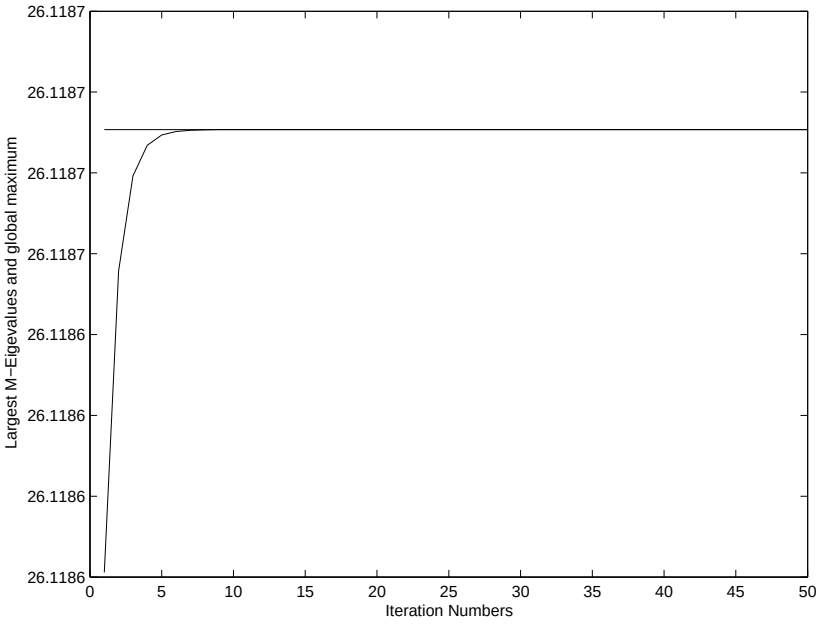

Figure 4.4.2: Numerical Result of Example 4.4.2

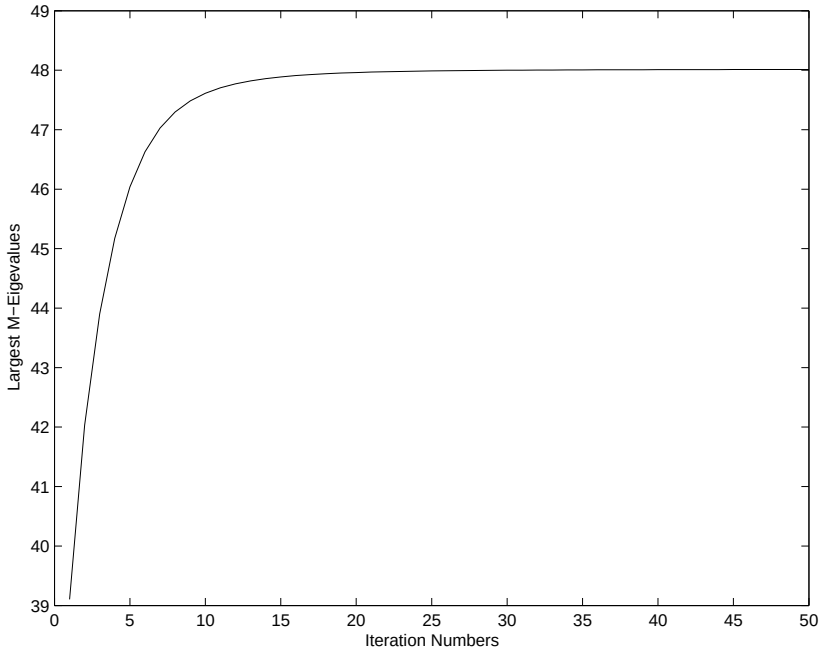

Figure 4.4.3: Numerical Result for a Tensor with $m = 12, n = 18$

We also do two numerical experiments on tensors with high dimensions. Here, we would not compute the largest M-eigenvalue of the tensors using the globalization method since its computation cost is extremely high.

Figure 4.4.3 shows the numerical result of Algorithm 4.4.1 on a tensor whose entries are uniformly generated in (-5,5) with $m = 12, n = 18$, and Figure 4.4.4 shows the numerical result on a tensor whose entries are uniformly generated in (-5,5) with $m = 30, n = 18$.

From the given numerical experiments, we can claim that the numerical result of Algorithm 4.4.1 is good although the iterative procedure is at most linearly convergent in theory [19]. Since the computing cost at the iterative step of the algorithm is very small, the designed algorithm is really efficient in practice especially for the large scale problem.

4.5 Some Comments

Although we can not guarantee that the obtained solution is a global optimization solution, the numerical results indicate that Algorithm 4.4.1 is practical since the solu-

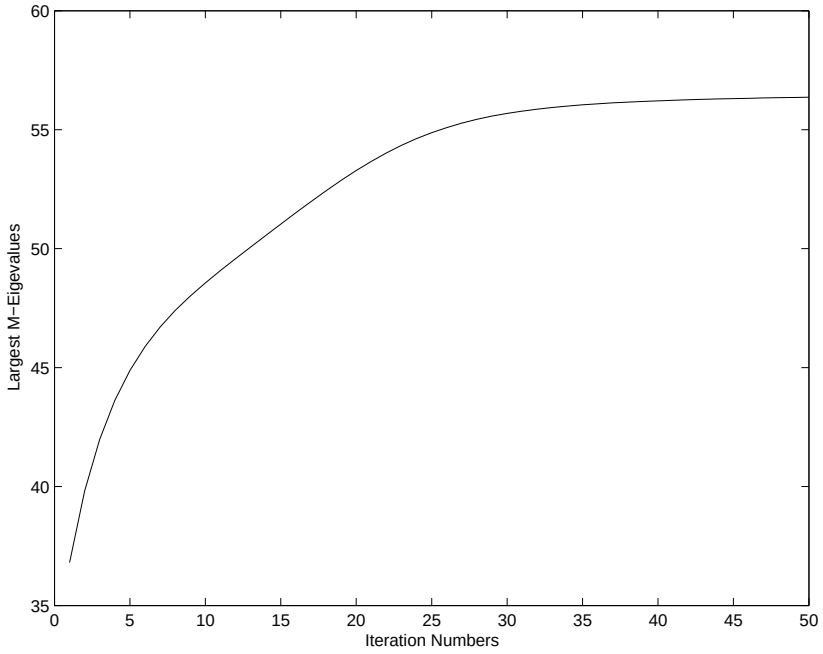

Figure 4.4.4: Numerical Result for a Tensor with $m = 30, n = 18$

tion can be generated within few steps. Hence, we hope that with further research the practical method can be designed with global convergence.

Chapter 5

Bi-quadratic Optimization Problems with Quadratic Constraints

5.1 Introduction

In this chapter, we consider two bi-quadratic polynomial optimization problems (1.2.7) and (1.2.8).

Let $F_{\min}$ and $F_{\max}$ be the optimal values of (1.2.7) and (1.2.8), respectively. Obviously, $F_{\max} \geq 0$. Furthermore, throughout this chapter, we assume that the optimal values $F_{\min}$ and $F_{\max}$ are attainable, which implies that $F_{\min} \geq 0$.

From the fact that (1.2.7) and (1.2.8) can be regarded as the generalization of (1.2.6), which is NP-hard to solve from [33], we can assert that the problems are also NP-hard.

Motivated by the methods for approximately solving (1.2.6) in [33] and the general NP-hard quadratic optimization problem in [42, 38], we consider the following SDP

relaxations of the considered problems

$$\begin{array}{ll} \min & g(X,Y) := (\mathcal{F}X) \bullet Y \\ \text{s.t.} & A_p \bullet X \geq 1,\ p = 1, \cdots, m_1, \\ & B_q \bullet Y \geq 1,\ q = 1, \cdots, n_1, \\ & X \succeq 0,\ Y \succeq 0 \end{array} \tag{5.1.1}$$

and

$$\begin{array}{ll} \max & g(X,Y) = (\mathcal{F}X) \bullet Y \\ \text{s.t.} & A_p \bullet X \leq 1,\ p = 0, 1, \cdots, m_1, \\ & B_q \bullet Y \leq 1,\ q = 1, \cdots, n_1, \\ & X \succeq 0,\ Y \succeq 0, \end{array} \tag{5.1.2}$$

respectively. Denote by $g_{\min}^{\text{sdp}}$ and $g_{\max}^{\text{sdp}}$ the optimal values of (5.1.1) and (5.1.2), respectively. Without loss of generality, we assume that the optimal values are attainable, which implies that $g_{\min}^{\text{sdp}} \geq 0$. It is easy to see that, for the optimization problem (1.2.7) with $m_1 = n_1 = 1$, $A_1 = I_m$ and $B_1 = I_n$, the optimal solution satisfy the equality constraints, that is, the problem can be equivalently reformulated as (1.2.6) studied by Ling *et al.* in [33]. In this case, from [33], we know that its bi-linear SDP relaxation is tight for the problem (1.2.6). However, for a general quadratic/bi-quadratic problem, its SDP relaxation is not tight for the original problem. As stated previous, its SDP relaxation does not always provide a tight approximation in general. However, it does lead to provably approximation solutions for certain type of quadratic optimization problems, see [5, 23, 42], which motivates us to extend the existing methods for quadratic optimization problems to bi-quadratic optimization problems.

This chapter is organized as follows. In Section 5.2, we analyze the approximation ratio of the SDP relaxations for bi-quadratic optimization problems. In Section 5.3, we present a polynomial time approximation algorithm for the bi-quadratic maximization model. In Section 5.4, we extend the approximation bound results obtained in Section 5.2 to the complex cases.

5.2 Bi-linear SDP Relaxation Bounds for Bi-quadratic Optimization Problems

In this section, we study approximation solutions for the bi-quadratic optimization models (1.2.7) and (1.2.8), based upon approximation solutions for their corresponding bi-linear SDP relaxations.

Consider the special case of (1.2.7), in which $m_1 = n_1 = 1$, A_1 and B_1 are positive definite. It is easy to see that the optimal solution pair must satisfy the constraints with equality. In this case, there exists an appropriate tensor $\bar{\mathcal{F}}$ such that (1.2.7) is equivalent to

$$\begin{array}{ll} \min & \bar{\mathcal{F}}xxyy \\ \text{s.t.} & x^\top x = 1, \\ & y^\top y = 1, \end{array}$$

which has no polynomial time algorithm $\mathfrak{A}$ to get a positive bound approximation solution for every instance of (1.2.7), see Theorem 2.2 in [33].

Based on the definition of r-bound, in the following, we argue that there is a finite and data-independent approximation bound between the optimal values of (1.2.7) and its SDP relaxation. To this end, we need some probability estimation results which play important roles in what follows. Lemma 5.2.1 (a) comes from [23], Lemma 5.2.2 comes from [38] and has been used in [39], and Lemma 5.2.3 comes from [64]. In addition, Lemma 5.2.1 (b) can be proved easily by Lemma 5.2.1 (a) and symmetry.

Lemma 5.2.1 *Let A and Z be two real symmetric $n \times n$ matrices with $Z \succeq 0$ and $\mathrm{tr}(AZ) \geq 0$. Let $\xi \sim \mathcal{N}(0, Z)$ be a normal random vector with zero mean and covariance matrix Z. Then the following probability estimation hold.*

(a) For any $0 \leq \gamma \leq 1$ we have

$$\mathrm{Prob}\left\{\xi^\top A\xi < \gamma E[\xi^\top A\xi]\right\} < 1 - \frac{3}{100}.$$

(b) For $\beta \geq 1$, we have

$$\mathrm{Prob}\left\{\xi^\top A\xi > \beta E[\xi^\top A\xi]\right\} < 1 - \frac{3}{100}.$$

Lemma 5.2.2 *Let A and Z be two real symmetric $n \times n$ matrices with $A \succeq 0$ and $Z \succeq 0$. Suppose $\xi \sim \mathcal{N}(0, Z)$ is a normal random vector with zero mean and covariance matrix Z. Then, for any $\gamma > 0$,*

$$\text{Prob}\left\{\xi^\top A\xi < \gamma E[\xi^\top A\xi]\right\} \leq \max\left\{\sqrt{\gamma}, \frac{2(r-1)\gamma}{\pi - 2}\right\},$$

where $r := \min\{\text{rank}(A), \text{rank}(Z)\}$.

Lemma 5.2.3 *Let A and Z be two real symmetric $n \times n$ matrices with $A \succeq 0$ and $Z \succeq 0$. Suppose $\xi \sim \mathcal{N}(0, Z)$ is a normal random vector with zero mean and covariance matrix Z. Then, for any $\gamma > 0$,*

$$\text{Prob}\left\{\xi^\top A\xi > \gamma E[\xi^\top A\xi]\right\} \leq e^{\frac{1}{2}(1-\gamma+\ln\gamma)}.$$

Let $\gamma = \frac{1}{\rho^2}$. It holds that

$$\text{Prob}\left\{\rho^2\xi^\top A\xi > E[\xi^\top A\xi]\right\} \leq e^{\frac{1}{2}(1-\frac{1}{\rho^2}-2\ln\rho)}.$$

Now we are ready to establish the first main result in this section, which characterizes the approximation ratio for the bi-linear SDP relaxation to (1.2.7). Our argumentation is similar to those of [23] and [38].

Theorem 5.2.1 *Suppose that the optimal value of (5.1.1) is nonnegative. Let $(\bar{X}, \bar{Y})$ be an r-bound approximation solution of (5.1.1). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (1.2.7) and the probability that*

$$\frac{r}{10^8 m_1^2 n_1^2} F(\bar{x}, \bar{y}) \leq F_{\min} \leq F(\bar{x}, \bar{y})$$

is at least $\dfrac{1}{2500}$.

Proof. Consider the semidefinite programming of the following form

$$\begin{array}{ll} \min & (\bar{Y}\mathcal{F}) \bullet X \\ \text{s.t.} & A_p \bullet X \geq 1, \ p = 1, \cdots, m_1, \\ & X \succeq 0, \end{array} \tag{5.2.3}$$

where $\bar{Y}\mathcal{F}$ is a symmetric $m \times m$ matrix with

$$(\bar{Y}\mathcal{F})_{kl} = \sum_{k,l=1}^{n} \mathcal{F}_{ijkl}\bar{Y}_{kl}.$$

It is well-known that there exists an optimal solution X^* of (5.2.3) with rank r_{X^*} satisfying $\frac{r_{X^*}(r_{X^*}+1)}{2} \leq m_1$, which can be found in polynomial time; cf. [49] and [25]. Clearly, it holds that

$$(\bar{Y}\mathcal{F}) \bullet X^* \leq (\mathcal{F}\bar{X}) \bullet \bar{Y}.$$

Based upon X^*, we further consider the following standard SDP problem

$$\begin{array}{ll} \min & (\mathcal{F}X^*) \bullet Y \\ \text{s.t.} & B_q \bullet Y \geq 1, \ q = 1, \cdots, n_1, \\ & Y \succeq 0. \end{array} \tag{5.2.4}$$

We can find an optimal solution Y^* of (5.2.4) with rank r_{Y^*} satisfying $\frac{r_{Y^*}(r_{Y^*}+1)}{2} \leq n_1$. Since X^* and Y^* are the optimal solutions of (5.2.3) and (5.2.4) respectively, the matrix pair (X^*, Y^*) satisfies

$$0 \leq (\mathcal{F}X^*) \bullet Y^* \leq (\mathcal{F}\bar{X}) \bullet \bar{Y} \tag{5.2.5}$$

and

$$r_{X^*} \leq \sqrt{2m_1}, \quad r_{Y^*} \leq \sqrt{2n_1}. \tag{5.2.6}$$

Let $\xi \sim \mathcal{N}(0, X^*)$ and $\eta \sim \mathcal{N}(0, Y^*)$ be two independent normal random vectors, whose covariance matrices are X^* and Y^* respectively. From the process of the proof of Theorem 3.3 in [23], it follows by Lemma 5.2.1 (b) and Lemma 5.2.2 that

$$\text{Prob}(\Omega) \geq \frac{3}{100} - m_1 \max\left\{\sqrt{\gamma_1}, \frac{2(r_{X^*}-1)\gamma_1}{\pi - 2}\right\}, \tag{5.2.7}$$

where

$$\Omega = \left\{\min_{1 \leq p \leq m_1} \xi^\top A_p \xi \geq \gamma_1, \ \xi^\top (Y^*\mathcal{F})\xi \leq \mu_1 (Y^*\mathcal{F}) \bullet X^*\right\},$$

$\gamma_1 > 0$ and $\mu_1 \geq 1$. By the assumption that the optimal value of (5.1.1) is nonnegative, we can see that $(\mathcal{F}xx^\top) \bullet Y^* \geq 0$ for any given sample value x of ξ in Ω. Hence, by Lemma 5.2.1 (b), we have

$$\text{Prob}\left\{\eta^\top (\mathcal{F}xx^\top)\eta > \mu_2 (\mathcal{F}xx^\top) \bullet Y^*\right\} < 1 - \frac{3}{100}$$

for every sample value x of ξ in Ω, where $\mu_2 \geq 1$. Note that the above estimation is independent with the sample value x of ξ. Consequently, it is easy to prove that

$$\mathrm{Prob}\left(\left\{\eta^\top(\mathcal{F}\xi\xi^\top)\eta > \mu_2(\mathcal{F}\xi\xi^\top)\bullet Y^*\right\}\cap\Omega\right) \leq \left(1-\frac{3}{100}\right)\mathrm{Prob}(\Omega),$$

which implies that the conditional probability

$$\mathrm{Prob}\left\{\eta^\top(\mathcal{F}\xi\xi^\top)\eta > \mu_2(\mathcal{F}\xi\xi^\top)\bullet Y^* \mid \Omega\right\} \leq 1-\frac{3}{100}. \tag{5.2.8}$$

On the other hand, from the independence of the random variables ξ and η, it follows from Lemma 5.2.2 that for any $\gamma_2 > 0$,

$$\begin{aligned}
\mathrm{Prob}\left\{\min_{1\leq q\leq n_1}\eta^\top B_q\eta < \gamma_2 \mid \Omega\right\} &= \mathrm{Prob}\left\{\min_{1\leq q\leq n_1}\eta^\top B_q\eta < \gamma_2\right\} \\
&\leq \sum_{q=1}^{n_1}\mathrm{Prob}\left\{\eta^\top B_q\eta < \gamma_2 E[\eta^\top B_q\eta]\right\} \\
&\leq n_1\max\left\{\sqrt{\gamma_2}, \frac{2(r_{Y^*}-1)\gamma_2}{\pi-2}\right\}.
\end{aligned}$$

This implies, together with (5.2.8), that

$$\begin{aligned}
&\mathrm{Prob}\left\{\eta^\top(\mathcal{F}\xi\xi^\top)\eta \leq \mu_2(\mathcal{F}\xi\xi^\top)\bullet Y^*, \min_{1\leq q\leq n_1}\eta^\top B_q\eta \geq \gamma_2 \mid \Omega\right\} \\
&\geq 1-\mathrm{Prob}\left\{\eta^\top(\mathcal{F}\xi\xi^\top)\eta > \mu_2(\mathcal{F}\xi\xi^\top)\bullet Y^* \mid \Omega\right\} - \mathrm{Prob}\left\{\min_{1\leq q\leq n_1}\eta^\top B_q\eta < \gamma_2 \mid \Omega\right\} \\
&\geq \frac{3}{100} - n_1\max\left\{\sqrt{\gamma_2}, \frac{2(r_{Y^*}-1)\gamma_2}{\pi-2}\right\},
\end{aligned} \tag{5.2.9}$$

where the first inequality comes from the fact that

$$\mathrm{Prob}(U\cap V) \geq 1-\mathrm{Prob}(U^c)-\mathrm{Prob}(V^c)$$

for any two random events U and V, where U^c stands for the contrary event of U, etc. Noticing the relation that

$$\begin{aligned}
&\left\{\min_{1\leq p\leq m_1}\xi^\top A_p\xi \geq \gamma_1, \min_{1\leq q\leq n_1}\eta^\top B_q\eta \geq \gamma_2, \eta^\top(\mathcal{F}\xi\xi^\top)\eta \leq \mu_1\mu_2(\mathcal{F}X^*)\bullet Y^*\right\} \\
&\qquad\supseteq \left\{\eta^\top(\mathcal{F}\xi\xi^\top)\eta \leq \mu_2(\mathcal{F}\xi\xi^\top)\bullet Y^*, \min_{1\leq q\leq n_1}\eta^\top B_q\eta \geq \gamma_2,\right\}\bigcap\Omega,
\end{aligned}$$

it follows from (5.2.7) and (5.2.9) that

$$\begin{aligned}
&\mathrm{Prob}\left\{\min_{1\leq p\leq m_1}\xi^\top A_p\xi \geq \gamma_1, \min_{1\leq q\leq n_1}\eta^\top B_q\eta \geq \gamma_2, \eta^\top(\mathcal{F}\xi\xi^\top)\eta \leq \mu_1\mu_2(\mathcal{B}X^*)\bullet Y^*\right\} \\
&\geq \left(\frac{3}{100} - m_1\max\left\{\sqrt{\gamma_1}, \frac{2(r_{X^*}-1)\gamma_1}{\pi-2}\right\}\right)\left(\frac{3}{100} - n_1\max\left\{\sqrt{\gamma_2}, \frac{2(r_{Y^*}-1)\gamma_2}{\pi-2}\right\}\right).
\end{aligned}$$

Let $\gamma_1 = \frac{1}{10^4 m_1^2}$, $\gamma_2 = \frac{1}{10^4 n_1^2}$, $\mu_1 = 1$ and $\mu_2 = 1$. By (5.2.6), we have

$$\sqrt{\gamma_1} \geq \frac{2(r_{X^*} - 1)\gamma_1}{\pi - 2} \quad \text{and} \quad \sqrt{\gamma_2} \geq \frac{2(r_{Y^*} - 1)\gamma_2}{\pi - 2}.$$

Thus, it holds that

$$\begin{aligned} &\text{Prob}\left\{ \min_{1\leq p\leq m_1} \xi^\top A_p \xi \geq \gamma_1, \min_{1\leq q\leq n_1} \eta^\top B_q \eta \geq \gamma_2, \eta^\top (\mathcal{F}\xi\xi^\top)\eta \leq \mu_1\mu_2 (\mathcal{F}X^*) \bullet Y^* \right\} \\ &\geq \frac{1}{2500}, \end{aligned}$$

which implies that there exists a vector pair $(x, y) \in \Re^m \times \Re^n$ such that

$$\min_{1\leq p\leq m_1} x^\top A_p x \geq \gamma_1, \quad \min_{1\leq q\leq n_1} y^\top B_q y \geq \gamma_2 \tag{5.2.10}$$

and

$$y^\top (\mathcal{F}xx^\top) y \leq \mu_1\mu_2 (\mathcal{F}X^*) \bullet Y^*. \tag{5.2.11}$$

Let $\bar{x} = \dfrac{x}{\sqrt{\gamma_1}}$ and $\bar{y} = \dfrac{y}{\sqrt{\gamma_2}}$. Then, by (5.2.10), we know that $(\bar{x}, \bar{y})$ is a feasible solution pair of (1.2.7), i.e.,

$$\bar{x}^\top A_p \bar{x} \geq 1 \ (p = 1, \cdots, m_1) \quad \text{and} \quad \bar{y}^\top B_q \bar{y} \geq 1 \ (q = 1, \cdots, n_1).$$

Furthermore, by (5.2.5) and (5.2.11), we have

$$f(\bar{x}, \bar{y}) \leq \frac{\mu_1\mu_2}{\gamma_1\gamma_2} (\mathcal{F}X^*) \bullet Y^* \leq \frac{\mu_1\mu_2}{\gamma_1\gamma_2} (\mathcal{F}\bar{X}) \bullet \bar{Y}. \tag{5.2.12}$$

Since $(\bar{X}, \bar{Y})$ is an r-bound approximation solution of (5.1.1), one has

$$(\mathcal{F}\bar{X}) \bullet \bar{Y} \leq \frac{1}{r} g_{\min}^{\text{sdp}} \leq \frac{1}{r} f_{\min},$$

where the second inequality due to the fact that (5.1.1) is a relaxation of (1.2.7). This implies, together with (5.2.12), that

$$F(\bar{x}, \bar{y}) \leq \frac{\mu_1\mu_2}{\gamma_1\gamma_2} (\mathcal{F}\bar{X}) \bullet \bar{Y} \leq \frac{10^8 m_1^2 n_1^2}{r} F_{\min}.$$

Thus the desired result follows. □

In the case where $m_1, n_1 \leq 2$, we have the following result, which is a generalization of Theorem 2.4 in [33].

Proposition 5.2.1 *Suppose that $m_1, n_1 \leq 2$. Then, the bi-quadratic optimization problem (1.2.7) and its bi-linear SDP relaxation (5.1.1) are equivalent.*

Proof. Without loss of generality, we assume that $m_1 = n_1 = 2$. Suppose that $(\bar{X}, \bar{Y})$ is an optimal solution pair of (5.1.1). Similar to the proof of the theorem above, we can find a matrix pair (X^*, Y^*) such that

$$(\mathcal{F}X^*) \bullet Y^* \leq (\mathcal{F}X^*) \bullet \bar{Y} \leq (\mathcal{F}\bar{X}) \bullet \bar{Y} \tag{5.2.13}$$

and

$$\frac{r_{X^*}(r_{X^*}+1)}{2} \leq 2, \ \frac{r_{Y^*}(r_{Y^*}+1)}{2} \leq 2. \tag{5.2.14}$$

By (5.2.13) and (5.2.14), we know that (X^*, Y^*) is an optimal solution matrix pair of (5.1.1), which satisfies $r_{X^*} = r_{Y^*} = 1$. Hence, there exist $x^* \in \Re^m$ and $y^* \in \Re^n$ such that $X^* = x^*(x^*)^\top$ and $Y^* = y^*(y^*)^\top$. Then, we have

$$(x^*)^\top A_p x^* \geq 1 \ (p = 1, 2), \quad (y^*)^\top B_q y^* \geq 1 \ (q = 1, 2) \tag{5.2.15}$$

and

$$F(x^*, y^*) = g(X^*, Y^*). \tag{5.2.16}$$

By (5.2.15), we know that (x^*, y^*) is feasible for (1.2.7). Furthermore, by (5.2.16), it follows that

$$F(x^*, y^*) = F_{\min}.$$

We obtain the desired result and complete the proof. □

In the rest of this section, we discuss the approximation bound for the maximization problem (1.2.8).

Theorem 5.2.2 *Suppose that $(\bar{X}, \bar{Y})$ is an r-bound approximation solution of (5.1.2). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (1.2.8) such that*

$$\frac{r}{4\,(1 + 2\ln(100m_1^2))\ln(100n_1)} F_{\max} \leq F(\bar{x}, \bar{y}) \leq F_{\max}.$$

Proof. Without loss of generality, we assume that the ranks of matrices $\bar{X}$ and $\bar{Y}$ satisfy $r_{\bar{X}} \leq \sqrt{2(m_1+1)}$, $r_{\bar{Y}} \leq \sqrt{2n_1}$, respectively. Let $\bar{X} = ZZ^\top$ with $Z \in \Re^{m \times r_{\bar{X}}}$. Since $Z^\top(\bar{Y}\mathcal{F})Z$ is symmetric, there exists an orthogonal matrix Q such that $Q^\top Z^\top(\bar{Y}\mathcal{F})ZQ$ is diagonal. Let $\xi_k, k = 1, 2, \cdots, r_{\bar{X}}$ be i.i.d random variables taking values -1 and 1 with equal probabilities, and let

$$x(\xi) := \frac{1}{\sqrt{\max\limits_{0 \leq p \leq m_1} (\xi^\top \bar{A}_p \xi + 1)}} ZQ\xi,$$

where $\bar{A}_p = Q^\top Z^\top A_p Z Q$ $(p = 0, 1, \cdots, m_1)$ and $\xi = (\xi_1, \cdots, \xi_{r_{\bar{X}}})^\top$.

It is easy to see that the random vector $x(\xi)$ is always well-defined from the positive semidefinition of A_i for $i = 1, 2, \cdots, m_1$, and $x(\xi)^\top A_p x(\xi) \leq 1$ for all $p = 0, 1, \cdots, m_1$. From the definition of $x(\xi)$, it holds that

$$\begin{aligned} x(\xi)^\top (\bar{Y}\mathcal{F}) x(\xi) &= \frac{1}{\max\limits_{0 \leq p \leq m_1} (\xi^\top \bar{A}_p \xi + 1)} \xi^\top Q^\top Z^\top (\bar{Y}\mathcal{F}) Z Q \xi \\ &= \frac{1}{\max\limits_{0 \leq p \leq m_1} (\xi^\top \bar{A}_p \xi + 1)} (\bar{Y}\mathcal{F}) \bullet \bar{X}. \end{aligned}$$

It is ready to verify that $\mathrm{tr}(\bar{A}_p) = A_p \bullet \bar{X} \leq 1$ $(p = 0, 1, \cdots, m_1)$ and $\bar{A}_p \succeq 0$ for $p = 1, \cdots, m_1$. Therefore, from the process of the proof of Theorem 4.2, Lemma 4.1 in [23] and (12) in [42], it follows that for any $\alpha > 2$,

$$\mathrm{Prob}(\Theta) \geq \frac{3}{100} - 2m_1^2 e^{-\frac{\alpha-1}{2}}, \tag{5.2.17}$$

where

$$\Theta = \left\{ x(\xi)^\top (\bar{Y}\mathcal{F}) x(\xi) \geq \frac{1}{\alpha} (\bar{Y}\mathcal{F}) \bullet \bar{X} \right\}.$$

Let $\eta \sim \mathcal{N}(0, Y^*)$ be an normal random variable with the covariance matrix Y^*. From the fact that $x(\xi)$ and η are independent, by a similar way to that used in the proof of Theorem 5.2.1, we can prove that the conditional probability

$$\mathrm{Prob}\left\{ \eta^\top (\mathcal{F} x(\xi) x(\xi)^\top) \eta < \nu (\mathcal{F} x(\xi) x(\xi)^\top) \bullet \bar{Y} \mid \Theta \right\} < 1 - \frac{3}{100} \tag{5.2.18}$$

for any $0 \leq \nu \leq 1$.

On the other hand, since $E[\eta^\top B_q \eta] = B_q \bullet \bar{Y} \leq 1$ for $q = 1, \cdots, n_1$, it is ready to see that $\{\eta^\top B_q \eta > \beta\} \subseteq \{\eta^\top B_q \eta > \beta E[\eta^\top B_q \eta]\}$, where $\beta > 0$. Consequently, by Lemma 5.2.3, we have that for $q = 1, \cdots, n_1$,

$$\mathrm{Prob}\left\{ \eta^\top B_q \eta > \beta \right\} \leq \mathrm{Prob}\left\{ \eta^\top B_q \eta > \beta E[\eta^\top B_q \eta] \right\} \leq e^{\frac{1}{2}(1 - \beta + \ln \beta)}.$$

Therefore, from the independence of $x(\xi)$ and η, we have

$$\begin{aligned} \mathrm{Prob}\left\{ \max_{1 \leq q \leq n_1} \eta^\top B_q \eta > \beta \mid \Theta \right\} &= \mathrm{Prob}\left(\bigcup_{q=1}^{n_1} \left\{ y^\top B_q y > \beta \right\} \right) \\ &\leq n_1 e^{\frac{1}{2}(1 - \beta + \ln \beta)}. \end{aligned} \tag{5.2.19}$$

By (5.2.18) and (5.2.19), it follows that

$$\begin{aligned}&\text{Prob}\left\{\max_{1\le q\le n_1}\eta^\top B_q\eta\le\beta,\eta^\top(\mathcal{F}x(\xi)x(\xi)^\top)\eta\ge\nu(\mathcal{F}x(\xi)x(\xi)^\top)\bullet\bar{Y}\mid\Theta\right\}\\&\ge\frac{3}{100}-n_1e^{\frac{1}{2}(1-\beta+\ln\beta)}.\end{aligned}\tag{5.2.20}$$

Noticing that

$$\begin{aligned}&\left\{\eta^\top(\mathcal{F}x(\xi)x(\xi)^\top)\eta\ge\frac{1}{\alpha}\nu(\mathcal{F}\bar{X})\bullet\bar{Y},\max_{1\le q\le n_1}\eta^\top B_q\eta\le\beta\right\}\\&\quad\supseteq\left\{\eta^\top(\mathcal{F}x(\xi)x(\xi)^\top)\eta\ge\nu(\mathcal{F}x(\xi)x(\xi)^\top)\bullet\bar{Y},\max_{1\le q\le n_1}\eta^\top B_q\eta\le\beta\right\}\bigcap\Theta,\end{aligned}$$

it follows from (5.2.17) and (5.2.20) that

$$\begin{aligned}&\text{Prob}\left\{\eta^\top(\mathcal{F}x(\xi)x(\xi)^\top)\eta\ge\frac{1}{\alpha}\nu(\mathcal{F}\bar{X})\bullet\bar{Y},\max_{1\le q\le n_1}\eta^\top B_q\eta\le\beta\right\}\\&\ge\left(\frac{3}{100}-2m_1^2e^{-\frac{\alpha-1}{2}}\right)\left(\frac{3}{100}-n_1e^{\frac{1}{2}(1-\beta+\ln\beta)}\right).\end{aligned}$$

Let $\alpha=1+2\ln(100m_1^2)$ and $\beta=4\ln(100n_1)$, we have

$$\text{Prob}\left\{\eta^\top(\mathcal{F}x(\xi)x(\xi)^\top)\eta\ge\frac{1}{\alpha}\nu(\mathcal{F}\bar{X})\bullet\bar{Y},\ \max_{1\le q\le n_1}\eta^\top B_q\eta\le\beta\right\}\ge\frac{1}{10^4}>0,$$

which implies that there exist vectors $\bar{x}=x(\xi)\in\Re^m$ and $y\in\Re^n$ such that

$$\bar{x}^\top A_p\bar{x}\le 1\ (p=0,1,\cdots,m_1),\quad y^\top B_q y\le\beta\ (q=1,\cdots,n_1)$$

and

$$y^\top(\mathcal{F}\bar{x}\bar{x}^\top)y\ge\frac{1}{\alpha}\nu(\mathcal{F}\bar{X})\bullet\bar{Y}.$$

Let $\bar{y}=\dfrac{y}{\sqrt{\beta}}$ and $\nu=1$. Then $(\bar{x},\bar{y})$ is a feasible solution of (1.2.8) satisfying

$$\frac{1}{\alpha\beta}(\mathcal{F}\bar{X})\bullet\bar{Y}\le\bar{y}^\top(\mathcal{F}\bar{x}\bar{x}^\top)\bar{y}\le f_{\max}.$$

Furthermore, by the definition of r-bound approximation solution, we obtain the desired result and complete the proof. □

Similar to Proposition 5.2.1, we have

Proposition 5.2.2 *Suppose that the numbers of constraints on x and y are not larger than 2, respectively. Then, the bi-quadratic optimization problem (1.2.8) and its bi-linear SDP relaxation (5.1.2) are equivalent.*

Remark Notice that the computational effort required for solving the bi-linear SDP relaxations of (1.2.7) and (1.2.8) can be significantly large. Therefore, it is very interesting to analyze the size of the resulted SDP relaxations, which will be our future research topic.

5.3 Approximation Solution of Bi-quadratic Problems

Our main goal in this chapter is to design polynomial time approximation algorithms for (1.2.7) and (1.2.8). Theorems 5.2.1 and 5.2.2 show that this task depends strongly on our ability to approximately solve the relaxed problems (5.1.1) and (5.1.2), which by themselves are also NP-hard. However, it is possible to derive approximation solution of the relaxed problems. In this section, we consider some forms of optimization problems whose approximation solution of their SDP relaxation problem can be solved in polynomial time. We first give an approximation result for the general model (1.2.8) under some mild assumptions. Then we investigate the bi-quadratic optimization problems with two constraints.

5.3.1 The Bi-quadratic Maximization Model

In this subsection, we consider the maximization problem (1.2.8). To this end, we make the following assumptions.

(A1) $|\mathrm{tr}(A_0)| < m$, $\mathrm{tr}(A_p) < m$ for every $p = 1, \cdots, m_1$, and $\mathrm{tr}(B_q) < n$ for every $q = 1, \cdots, n_1$.

(A2) There exist nonnegative numbers α_p $(p = 0, 1, \cdots, m_1)$ with $\sum\limits_{p=0}^{m_1} \alpha_p = 1$ and β_q $(q = 1, \cdots, n_1)$ with $\sum\limits_{q=1}^{n_1} \beta_q = 1$, such that

$$\sum_{p=0}^{m_1} \alpha_p A_p - I_m \succeq 0 \quad \text{and} \quad \sum_{q=1}^{n_1} \beta_q B_q - I_n \succeq 0.$$

(A3) $A_0 + I_m \succeq 0$.

Now we are ready to present some properties of feasible solutions of (5.1.2) under Assumptions (A1)-(A3).

Lemma 5.3.1 *Suppose that (A1)-(A3) holds. Then for any feasible solution pair (X, Y) of (5.1.2), there hold*

(1) $(\frac{1}{m}I_m, \frac{1}{n}I_n)$ is a feasible solution pair of (5.1.2).

(2) $\mathrm{tr}(X) \leq 1$, $\mathrm{tr}(Y) \leq 1$ and $(A_0 \bullet X)^2 \leq 1$.

Proof. It is easy to see that (1) holds from Assumption (A1).

Since (X, Y) is feasible, $X \succeq 0$ and $Y \succeq 0$. Consider Assumption (A3), $A_0 \bullet X \geq -\mathrm{tr}(X)$. Combining with Assumption (A2), we have

$$\mathrm{tr}(X) \leq (\sum_{p=0}^{m_1} \alpha_p A_p) \bullet X = \sum_{p=0}^{m_1} \alpha_p (A_p \bullet X) \leq \sum_{p=0}^{m_1} \alpha_p = 1.$$

So, we can assert that $A_0 \bullet X \geq -1$ and $\mathrm{tr}(X) \leq 1$.

By a similar way, we can show that $\mathrm{tr}(Y) \leq 1$, which complete the proof. □

To obtain an approximation solution of bi-linear relaxation problem, we further need the following lemma, which generalizes the result used in [33].

Lemma 5.3.2 *For any $X \in \mathcal{S}^m$, the following statements hold.*

(1) If $\|X\|_F \leq \frac{1}{m}$, then $\bar{X} := X + \frac{1}{m}I_m \succeq 0$.

(2) Suppose $m \geq 2$. If $\mathrm{tr}(X) \leq 0$ and $X \succeq -\frac{1}{m}I_m$, then $\|X\|_F \leq \sqrt{1 - \frac{1}{m}}$.

Proof. (1) Since $\|X\|_F \leq \frac{1}{m}$, it follows that $|x_{ii}| \leq \frac{1}{m}$ for every $i = 1, \cdots, m$. This implies that

$$\mathrm{tr}(\bar{X}) = \mathrm{tr}(X) + 1 = \sum_{i=1}^{m} x_{ii} + 1 \geq 0. \tag{5.3.21}$$

To show that $\bar{X} \succeq 0$, by Lemma 2.1 in [6], we only need to show that

$$\sqrt{m-1}\|\bar{X}\|_F \leq \mathrm{tr}(\bar{X}). \tag{5.3.22}$$

It is easy to see that

$$\|\bar{X}\|_F^2 = \|X\|_F^2 + \frac{2}{m}\mathrm{tr}(X) + \frac{1}{m}, \tag{5.3.23}$$

which implies, together with (5.3.21), that

$$\begin{array}{rcl}(m-1)\|\bar{X}\|_F^2 - (\mathrm{tr}(\bar{X}))^2 & = & (m-1)\left[\|X\|_F^2 + \frac{2}{m}\mathrm{tr}(X) + \frac{1}{m} - \frac{1}{m-1}\left(\mathrm{tr}(X)+1\right)^2\right] \\ & \leq & (m-1)\left[\frac{1}{m^2} + \frac{2}{m}\mathrm{tr}(X) + \frac{1}{m} - \frac{1}{m-1}\left(\mathrm{tr}(X)+1\right)^2\right] \\ & = & -\left(\mathrm{tr}(X) + \frac{1}{m}\right)^2 \leq 0.\end{array}$$

Therefore, (5.3.22) holds. This shows that $\bar{X} \succeq 0$.

(2) Since $\bar{X} = X + \frac{1}{m}I_m \succeq 0$, it follows that

$$-1 \leq \mathrm{tr}(X) \leq 0, \tag{5.3.24}$$

from the given condition that $\mathrm{tr}(X) \leq 0$. Moreover, it holds that

$$\begin{array}{rcl}\|\bar{X}\|_F^2 & \leq & (\mathrm{tr}(\bar{X}))^2 \\ & = & (\mathrm{tr}(X))^2 + 2\mathrm{tr}(X) + 1,\end{array}$$

where the inequality is due to the positive semidefiniteness of $\bar{X}$. This implies, together with (5.3.23), that

$$\|X\|_F^2 \leq (\mathrm{tr}(X))^2 + 2\left(1 - \frac{1}{m}\right)\mathrm{tr}(X) + 1 - \frac{1}{m}. \tag{5.3.25}$$

Consider the optimization problem as follows

$$\begin{array}{rl}p_{\max} := \max & p(t) = t^2 + 2bt + c \\ \text{s.t.} & l \leq t \leq u.\end{array}$$

It is easy to verify that $p_{\max} = \max\{p(l), p(u)\}$. Consequently, by this, (5.3.24) and (5.3.25), we know that $\|X\|_F^2 \leq 1 - \frac{1}{m}$ and complete the proof. □

Considering linear transformations $X := X - \frac{1}{m}I_m$, $Y := Y - \frac{1}{n}I_n$, we know that based on Lemma 5.3.1 and Lemma 5.3.2, a restriction and a relaxation for (5.1.2) can be written in a unified form as

$$\begin{array}{rl}p_\lambda := \max & \Phi(X,Y) = (\mathcal{F}X) \bullet Y + \frac{1}{m}(\mathcal{F}I_m) \bullet Y + \frac{1}{n}(\mathcal{F}X) \bullet I_n + \frac{1}{mn}(\mathcal{F}I_m) \bullet I_n \\ \text{s.t.} & \left(A_p \bullet X + \frac{1}{m}\mathrm{tr}(A_p)\right)^2 \leq 1,\ p = 0, 1, \cdots, m_1, \\ & \left(B_q \bullet Y + \frac{1}{n}\mathrm{tr}(B_q)\right)^2 \leq 1,\ q = 1, \cdots, n_1, \\ & \|X\|_F \leq \lambda, \\ & \|Y\|_F \leq \lambda,\end{array} \tag{5.3.26}$$

where $\lambda = \frac{1}{\max\{m,n\}}$ and $\lambda = \sqrt{1 - \frac{1}{\max\{m,n\}}}$ correspond to the restriction and the relaxation, respectively. It is easy to see that matrix pair $(0,0) \in \mathcal{S}^m \times \mathcal{S}^n$ is a feasible solution of (5.3.26) for any $\lambda \geq 0$. Furthermore $p_0 = \frac{1}{mn}(\mathcal{F}I_m) \bullet I_n$.

By stacking up the entries of a symmetric matrix (ignoring the symmetric part) into a vector, denoted by $vec_S(\cdot)$, there exists a suitable quadratic function $q_0(u,v)$ such that (5.3.26) can be rewritten into the following form

$$\begin{array}{rll} p_\lambda := \max & q_0(u,v) & \\ \text{s.t.} & \left(vec_S(A_p)^\top u + \frac{1}{m}\mathrm{tr}(A_p)\right)^2 \leq 1, & p = 0,1,\cdots,m_1, \\ & \left(vec_S(B_q)^\top v + \frac{1}{n}\mathrm{tr}(B_q)\right)^2 \leq 1, & q = 1,\cdots,n_1, \\ & \|u\| \leq \lambda, \ \|v\| \leq \lambda, & \end{array} \tag{5.3.27}$$

where $u = vec_S(X), v = vec_S(Y)$. It is well-known that for a quadratic function $q(x) = c + 2b^\top x + x^\top Ax$, the homogenized version of $q(x)$ can be represented by the matrix denoted by

$$M(q(\cdot)) = \begin{pmatrix} c & b^\top \\ b & A \end{pmatrix}.$$

Hence, a standard SDP relaxation for the homogenized version of (5.3.27) is

$$\begin{array}{rll} z(\lambda^2) := \max & \bar{Q}_0 \bullet Z & \\ \text{s.t.} & \bar{C}_p \bullet Z \leq 1, & p = 0,1,\cdots,m_1+n_1, \\ & \bar{C} \bullet Z \leq \lambda^2, & \\ & \bar{D} \bullet Z \leq \lambda^2, & \\ & Z = \begin{pmatrix} 1 & u^\top & v^\top \\ u & W & U^\top \\ v & U & V \end{pmatrix} \succeq 0, & \end{array} \tag{5.3.28}$$

where $\bar{Q}_0$, $\bar{C}_p$, $(p = 0,1,\cdots,m_1+n_1)$, $\bar{C}$ and $\bar{D}$ are some suitable matrices, which correspond to the matrix representations of the homogenized version of the quadratic functions with respect to (u,v) in problem (5.3.27), respectively. Note that (5.3.28) can be solved in polynomial time.

Based upon the analysis above, we arrive at the following conclusion.

Theorem 5.3.1 *Suppose that Assumptions (A1)-(A3) hold and* $(\mathcal{F}I_m) \bullet I_n \geq 0$. *Then a* $\frac{(1-\gamma)^2}{(\sqrt{m_1+n_1+3}+\gamma)^2\rho(\rho-1)}$*-bound approximation solution of (5.1.2) can be found in polynomial*

time, where $\rho = \max\{m, n\}$ *and*

$$\gamma = \max\left\{\frac{1}{m}|\mathrm{tr}(A_0)|,\ \frac{1}{m}\mathrm{tr}(A_p), p = 1, \cdots, m_1,\ \frac{1}{n}\mathrm{tr}(B_q), q = 1, 2, \cdots, n_1\right\}.$$

Proof. We consider the problem (5.3.27) with $\lambda = \frac{1}{\rho}$. By Theorem 1 in [67], there exists a feasible solution (u, v) of problem (5.3.27) satisfying

$$q_0(u, v) \geq \frac{(1-\gamma)^2}{(\sqrt{m_1 + n_1 + 3} + \gamma)^2} z\left(\frac{1}{\rho^2}\right).$$

On the other hand, it is easy to see that $z(\lambda)$ is concave on $\lambda \geq 0$, and hence

$$\begin{aligned} z\left(\tfrac{1}{\rho^2}\right) &\geq \left(1 - \tfrac{1}{\rho(\rho-1)}\right) z(0) + \tfrac{1}{\rho(\rho-1)} z\left(1 - \tfrac{1}{\rho}\right) \\ &\geq \tfrac{1}{\rho(\rho-1)} z\left(1 - \tfrac{1}{\rho}\right) \\ &\geq \tfrac{1}{\rho(\rho-1)} g_{\max}^{\mathrm{sdp}}, \end{aligned}$$

where the second inequality is due to $z(0) = p_0 = \frac{1}{mn}(\mathcal{B}I_m) \bullet I_n \geq 0$, and the last inequality comes from the fact that $z\left(1 - \frac{1}{\rho}\right) \geq p_{\sqrt{1-\frac{1}{\rho}}} \geq g_{\max}^{\mathrm{sdp}}$. Therefore,

$$q_0(u, v) \geq \frac{(1-\gamma)^2}{(\sqrt{m_1 + n_1 + 3} + \gamma)^2 \rho(\rho-1)} g_{\max}^{\mathrm{sdp}}. \tag{5.3.29}$$

By the obtained (u, v) and the stack relation between the vector and the matrix, we can find a feasible matrix pair $(\bar{X}, \bar{Y})$ for (5.3.26) with $\lambda = \frac{1}{\rho}$ such that $\Phi(\bar{X}, \bar{Y}) = q_0(u, v)$. Denote $X^* = \bar{X} + \frac{1}{m} I_m$ and $Y^* = \bar{Y} + \frac{1}{n} I_n$. By Lemma 5.3.2 (1), it holds that (X^*, Y^*) is a feasible solution of (5.1.2), satisfying

$$(\mathcal{F}X^*) \bullet Y^* \geq \frac{(1-\gamma)^2}{(\sqrt{m_1 + n_1 + 3} + \gamma)^2 \rho(\rho-1)} g_{\max}^{\mathrm{sdp}}.$$

Therefore, we can assert that (X^*, Y^*) is a $\frac{(1-\gamma)^2}{(\sqrt{m_1+n_1+3}+\gamma)^2\rho(\rho-1)}$-bound approximation solution of (5.1.2). Combining with the fact that $0 \leq \gamma < 1$, the desired result follows. □

5.3.2 The Bi-quadratic Optimization Problems with Two Constraints

In this subsection, we first consider the following problem

$$\begin{aligned} F_{\max} := \max\ & \mathcal{F}xxyy \\ \text{s.t.}\ & x^\top A x \leq 1, \\ & y^\top B y \leq 1, \end{aligned} \tag{5.3.30}$$

where $A \in \mathcal{S}^m$ and $B \in \mathcal{S}^n$ are positive definite. We assume that $(\mathcal{F}I_m) \bullet I_n \geq 0$. Without loss of generality, we further assume that $A = I_m$ and $B = I_n$.

Notice that the optimal solution must satisfy the constraints with equality. Therefore, the bi-linear SDP relaxation of (5.3.30) can be written equivalently as follows

$$\begin{array}{rll} g_{\max}^{\rm sdp} := \max & (\mathcal{F}X) \bullet Y & \\ \text{s.t.} & \mathrm{tr}(X) = 1, & \\ & \mathrm{tr}(Y) = 1, & \\ & X \succeq 0,\ Y \succeq 0. & \end{array} \tag{5.3.31}$$

By a similar procedure used in Subsection 5.3.1, a restriction and a relaxation of (5.3.31) can be written in a unified form as

$$\begin{array}{rll} p_\lambda := \max & (\mathcal{F}X) \bullet Y + \frac{1}{m}(\mathcal{F}I_m) \bullet Y + \frac{1}{n}(\mathcal{F}X) \bullet I_n + \frac{1}{mn}(\mathcal{F}I_m) \bullet I_n & \\ \text{s.t.} & \mathrm{tr}(X) = 0, & \\ & \mathrm{tr}(Y) = 0, & \\ & \|X\|_F \leq \lambda, & \\ & \|Y\|_F \leq \lambda, & \end{array} \tag{5.3.32}$$

where $\lambda = \frac{1}{\max\{m,n\}}$ and $\lambda = \sqrt{1 - \frac{1}{\max\{m,n\}}}$ correspond to the restriction and the relaxation, respectively. Hence, it follows that $p_{\sqrt{1-\frac{1}{\max\{m,n\}}}} \geq g_{\max}^{\rm sdp} \geq p_{\frac{1}{\max\{m,n\}}} \geq p_0 = \frac{1}{mn}(\mathcal{F}I_m) \bullet I_n \geq 0$. Furthermore, for $vec_S(X)$ and $vec_S(Y)$, we can eliminate two variables, say X_{11} and Y_{11}, by their linear relation with the other variables. For convenience, let

$$u = vec_S(X) \backslash X_{11} \ \text{ and } \ v = vec_S(Y) \backslash Y_{11}.$$

Then, there exist $Q_0 \in \Re^{L_m \times L_n}$, $Q_1 \in \mathcal{S}^{L_m}$, $Q_2 \in \mathcal{S}^{L_n}$, $b_0 \in \Re^{L_m}$, $c_0 \in \Re^{L_n}$ and $d_0 = \frac{1}{mn}(\mathcal{F}I_m) \bullet I_n \in \Re$ such that the above problem is equivalent to

$$\begin{array}{rll} p_\lambda := \max & q(u,v) = u^\top Q_0 v + 2b_0^\top u + 2c_0^\top v + d_0 & \\ \text{s.t.} & q_1(u,v) = u^\top Q_1 u \leq \lambda^2, & \\ & q_2(u,v) = v^\top Q_2 v \leq \lambda^2, & \end{array} \tag{5.3.33}$$

where $L_m = m(m+1)/2 - 1$, $L_n = n(n+1)/2 - 1$ and Q_1, Q_2 are positive definite. Furthermore, it is easy to see that the SDP relaxation of the homogenized version of

(5.3.33) is

$$\begin{array}{rl} z(\lambda^2) := \max & \bar{Q}_0 \bullet Z \\ \text{s.t.} & \bar{Q}_1 \bullet Z \le \lambda^2 \\ & \bar{Q}_2 \bullet Z \le \lambda^2, \\ & Z = \begin{pmatrix} 1 & u^\top & v^\top \\ u & W & U^\top \\ v & U & V \end{pmatrix} \succeq 0, \end{array} \tag{5.3.34}$$

where $\bar{Q}_0, \bar{Q}_1, \bar{Q}_2$ are three matrices which correspond to the homogenized version of the quadratic functions $q(u,v)$, $q_1(u,v)$ and $q_2(u,v)$, respectively.

Consider problem (5.3.34) with $\lambda_0 = \frac{1}{\sqrt{2}\rho}$ and $\rho = \max\{m,n\}$. Since this SDP has three constraints, so that an optimal solution Z^* can be computed in polynomial time such that its rank equals 2 (e.g., see [72]). Let us denote by I_{11} the $(L_m + L_n + 1) \times (L_m + L_n + 1)$ symmetric matrix with 1 at its $(1,1)$th position and 0 elsewhere. It is clear that $I_{11} \bullet Z^* = 1$. Hence, by Corollary 4 in [65], one can always find two vectors $z^i = (t_i, (u^i)^\top, (v^i)^\top)^\top (i = 1,2) \in \Re^{1+L_m+L_n}$ such that $Z^* = z^1(z^1)^\top + z^2(z^2)^\top$ and

$$I_{11} \bullet z^i(z^i)^\top = I_{11} \bullet Z^*/2 = 1/2, \quad \text{for } i = 1,2,$$

which implies that $t_1^2 = t_2^2 = 1/2$. From the structure of the constraints of (5.3.33), it is ready to know that both $\bar{Q}_1$ and $\bar{Q}_2$ are positive semidefinite. Consequently, since Z^* is feasible for (5.3.34), it holds that

$$(z^i)^\top \bar{Q}_1 z^i \le \lambda_0^2 \ \text{ and } \ (z^i)^\top \bar{Q}_2 z^i \le \lambda_0^2, \quad \text{for } i = 1,2,$$

which implies that $(\bar{u}^i, \bar{v}^i) = (u^i/t_i, v^i/t_i)$, $i = 1,2$, are feasible solutions of (5.3.33) with $\lambda = \frac{1}{\rho}$. Furthermore, we have

$$q(\bar{u}^1, \bar{v}^1) + q(\bar{u}^2, \bar{v}^2) = \left(\bar{Q}_0 \bullet z^1(z^1)^\top + \bar{Q}_0 \bullet z^2(z^2)^\top\right)/t_1^2 = 2\bar{Q}_0 \bullet Z^* = 2z(\lambda_0^2),$$

which implies that either $(\bar{u}^1, \bar{v}^1)$ or $(\bar{u}^2, \bar{v}^2)$, denoted by $(\bar{u}, \bar{v})$, satisfies

$$q(\bar{u}, \bar{v}) \ge z(\lambda_0^2).$$

On the other hand, it is easy to see that $z(\cdot)$ is concave, and hence

$$z(\lambda_0^2) \ge \left(1 - \frac{1}{2\rho(\rho-1)}\right) z(0) + \frac{1}{2\rho(\rho-1)} z(1 - 1/\rho) \ge \frac{1}{2\rho(\rho-1)} z(1 - 1/\rho),$$

where the last inequality due to the assumption that $z(0) \geq d_0 \geq 0$. Therefore,

$$q(\bar{u}, \bar{v}) \geq z(\lambda_0^2) \geq \frac{1}{2\rho(\rho-1)} z(1-1/\rho) \geq \frac{1}{2\rho(\rho-1)} F_{\max}, \tag{5.3.35}$$

where the last inequality comes from the fact that $z(1-1/\rho) \geq p_{\sqrt{1-1/\rho}} \geq g_{\max}^{\text{sdp}} \geq F_{\max}$. Similar to the process of the proof of Theorem 5.3.1, from the obtained $(\bar{u}, \bar{v})$, we can find a feasible matrix pair $(\bar{X}, \bar{Y})$ of (5.3.31) such that $(\mathcal{F}\bar{X}) \bullet \bar{Y} = q(\bar{u}, \bar{v})$. Consequently, by using a similar procedure to that used in Theorem 2.4 in [33], we can get a vector pair $(\bar{x}, \bar{y})$ such that $\|\bar{x}\| = \|\bar{y}\| = 1$ and $\bar{y}^\top(\mathcal{F}\bar{x}\bar{x}^\top)\bar{y} \geq q(\bar{u}, \bar{v})$. This shows that $(\bar{x}, \bar{y})$ is a feasible solution of (5.3.30), and hence $F_{\max} \geq \bar{y}^\top(\mathcal{F}\bar{x}\bar{x}^\top)\bar{y} \geq q(\bar{u}, \bar{v})$. Together with (5.3.35), we can assert that $(\bar{x}, \bar{y})$ is a $\frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}$-bound approximation solution of (5.3.30). Therefore, the following assertion is established.

Theorem 5.3.2 *If $(\mathcal{F}I_m) \bullet I_n \geq 0$, then a $\frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}$-bound approximation solution of (5.3.30) can be found in polynomial time.*

In fact, from above procedure, we can see that assumption $(\mathcal{F}I_m) \bullet I_n \geq 0$ is used to guarantee that $z(0) \geq 0$. Therefore, if we replace $\mathcal{F}$ by $\mathcal{F} - cI_m \otimes I_n$ with constant $c \leq \frac{1}{mn}(\mathcal{F}I_m) \bullet I_n$, then $z(0) \geq 0$ is guaranteed. By Theorem 5.3.2, there exists a feasible solution pair $(\bar{x}, \bar{y})$ such that

$$\mathcal{F}\bar{x}\bar{x}\bar{y}\bar{y} - c \geq \frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}(F_{max} - c).$$

Let $c = \bar{g}_{\min}$, where $\bar{g}_{\min}$ is the minimum value of the objective in (5.3.31), then $c \leq \frac{1}{mn}(\mathcal{F}I_m) \bullet I_n$. This leads to the following result.

Theorem 5.3.3 *There exists a $\left(1 - \frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}\right)$-relative approximation solution for (5.3.30) in polynomial time.*

We conclude this subsection by considering the following minimization problem

$$\begin{array}{ll} \min & \mathcal{F}xxyy \\ \text{s.t.} & x^\top x \geq 1, \\ & y^\top y \geq 1. \end{array} \tag{5.3.36}$$

It is easy to see that the optimal solution must satisfy the constraints with equality if $f_{\min}$ is attainable. Thus, the bi-linear SDP relaxation can be written as (5.3.31) with tensor $-\mathcal{B}$, which leads to the following result.

Theorem 5.3.4 *Suppose that the optimal value is attainable. If $(\mathcal{F}I_m) \bullet I_n \leq 0$, then a $\frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}$-bound approximation solution of (5.3.36) can be found in polynomial time. Otherwise, there exists a $\left(1 - \frac{1}{2\max\{m,n\}(\max\{m,n\}-1)}\right)$-relative approximation solution for (5.3.36) in polynomial time.*

5.4 Extensions and Discussions

Motivated by the aforementioned work on complex SDP in [38], our analysis can be extended to the so-called complex bi-quadratic optimization problems. In this section, we further consider the minimization model

$$\begin{array}{ll} \min & F(x,y) := \mathcal{F}xxyy \\ \text{s.t.} & x^H A_p x \geq 1,\ p = 1, \cdots, m_1, \\ & y^H B_q y \geq 1,\ q = 1, \cdots, n_1, \\ & x \in \mathbf{C}^m,\ y \in \mathbf{C}^n \end{array} \tag{5.4.37}$$

and the maximization model

$$\begin{array}{ll} \max & F(x,y) = \mathcal{F}xxyy \\ \text{s.t.} & x^H A_p x \leq 1,\ p = 0, 1, \cdots, m_1, \\ & y^H B_q y \leq 1,\ q = 1, \cdots, n_1, \\ & x \in \mathbf{C}^m,\ y \in \mathbf{C}^n, \end{array} \tag{5.4.38}$$

where $A_p \in \mathcal{H}^m$ $(p = 1, \cdots, m_1)$ and $B_q \in \mathcal{H}^n$ $(q = 1, \cdots, n_1)$ are positive semidefinite, whereas $A_0 \in \mathcal{H}^m$ is indefinite.

A similar procedure to that in Section 5.2 can be applied to yield approximation bounds for the complex bi-quadratic optimization problems above. To this end, we need the following probability estimation results, which come from [23] and [38], respectively.

Lemma 5.4.1 *Let A, Z be two Hermitian matrices satisfying $Z \succeq 0$ and $\mathrm{tr}(AZ) \geq 0$. Let $\xi \sim N_{\boldsymbol{C}}(0, Z)$ be a complex normal random vector. Then,*

(a) For any $0 \leq \gamma \leq 1$, it holds that

$$\mathrm{Prob}\left\{\xi^H A\xi < \gamma E(\xi^H A\xi)\right\} < 1 - \frac{1}{20}.$$

(b) For any $\beta \geq 1$, it holds that

$$\mathrm{Prob}\left\{\xi^H A\xi > \beta E(\xi^H A\xi)\right\} < 1 - \frac{1}{20}.$$

Lemma 5.4.2 *Let A, Z be two Hermitian positive semidefinite matrices. Suppose that ξ is a random vector generated from the complex-valued normal distribution $N_{\boldsymbol{C}}(0, Z)$. Then for any $\gamma > 0$, the following probability estimation hold.*

(a) $\mathrm{Prob}\left\{\xi^H A\xi < \gamma E(\xi^H A\xi)\right\} \leq \max\left\{\frac{4}{3}\gamma, 16(r-1)^2\gamma^2\right\}$,

(b) $\mathrm{Prob}\left\{\xi^H A\xi > \gamma E(\xi^H A\xi)\right\} \leq re^{-\gamma}$,

where $r := \min\{\mathrm{rank}(A), \mathrm{rank}(Z)\}$.

The following main result in this section can be proved in the similar ways to that used in the proofs of Theorems 5.2.1 and 5.2.2.

Theorem 5.4.1 *Let $(\bar{X}, \bar{Y})$ be an r-bound approximation solution of the bi-linear SDP relaxation of (5.4.37). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (5.4.37) and the probability that*

$$\frac{r}{1600 m_1 n_1} F(\bar{x}, \bar{y}) \leq F_{\min} \leq F(\bar{x}, \bar{y})$$

is at least $\frac{1}{3600}$.

Suppose that $(\bar{X}, \bar{Y})$ be an r-bound approximation solution of the bi-linear SDP relaxation for (5.4.38). Then we have a feasible solution $(\bar{x}, \bar{y})$ of (5.4.38) and the probability that

$$\frac{r}{(1 + 2\ln 100 m_1^2)\ln\left(40\sqrt{2} n_1^{\frac{3}{2}}\right)} F_{\max} \leq F(\bar{x}, \bar{y}) \leq F_{\max}$$

is at least $\frac{1}{4000}$.

It is well-known that if the number of constraints in the considered complex SDP problem is at most 3, then its rank-one optimal solution can be found, see Theorem 2.1, Proposition 5.1 in [25]. As a consequence, we get the following proposition which can be proved by similar ways to that used in the proofs of Proposition 5.2.1 and Proposition 5.2.2.

Proposition 5.4.1 *Suppose that the numbers of constraints on x and y are less than* 4*, respectively. Then, the bi-quadratic optimization problems (5.4.37), (5.4.38) and their relaxations are equivalent, respectively.*

5.5 Some Remarks

For bi-quadratic optimization problems with quadratic constraints, we show that some approximation solutions can be obtained in randomize polynomial time via the corresponding bi-linear SDP relaxation problems. Then we present some approximation solutions for some special cases. However, there are some issues may be addressed in future study.

Because approximation solutions are obtained in randomize polynomial time, it is hoped that determinable approximation solutions may be obtained by matrix decomposition. In addition, some possible approximation solutions with improved quality bound may be designed for some special cases of considered problems.

Chapter 6

Conclusions and Suggestions for Future Studies

In this thesis, we studied some polynomial optimization problems arising in practical applications.

In Chapter 2, we establish an application of cubic one-spherical optimization problem in magnetic resonance imaging of medical engineering.

In Chapter 3, we first present complexity analysis of the cubic two-spherical/three-spherical optimization problems. We discuss some special cases which are solvable in polynomial time and have PTAS. We then establish a quality bound for general case of the cubic three-spherical optimization problem. Furthermore, some approximation solution methods for one-spherical optimization problem are presented. Finally, a practical method for three-spherical optimization is proposed.

In Chapter 4, for bi-quadratic optimization problem over unit spheres, we reformulate the problem as the largest M-eigenvalue computation problem and then generalize the power method to obtain the reformulated problem. To make the proposed method more effective, an initialization technique is introduced.

In Chapter 5, for bi-quadratic optimization problems with quadratic constraints, we first present the relationship between the problems and their corresponding SDP relaxations. We then propose some approximation solutions for some special cases.

In previous chapters, we have already raised some questions. Here, the following is a list of some interesting and challenging problems for future research.

1. Since multivariate polynomials have simple expressions with aid of tensors, it is of great significance to establish properties and definitions of related tensors.

2. Develop some new applications of polynomial optimization problems in academic researches and practical applications.

3. For the cubic one-spherical optimization problem, to propose some approximation solution methods with quality bounds. In addition, it is interesting to establish the conditions under which the problem is solvable in polynomial time.

4. For the cubic two-spherical/three-spherical optimization problems, propose method to improve the quality bound obtained in this thesis.

5. To consider the following general polynomial optimization problem:

$$\begin{array}{ll} \min & \mathcal{A}(x^{(1)})^{m_1}(x^{(2)})^{m_2}\cdots(x^{(m_l)})^{m_l} \\ \text{s.t.} & \mathcal{B}^{(p)}(x^{(1)})^{p_1}(x^{(2)})^{p_2}\cdots(x^{(m_l)})^{p_l} \geq 0, \quad \text{for} \quad p = 1, 2, \cdots, q. \end{array}$$

where $\mathcal{A}$ is a $(m_1, m_2, \cdots, m_l)$-order $(n_1 \times n_2 \times \cdots \times n_l)$-dimensional tensor, $\mathcal{B}^{(p)}$ is a $(p_1, \cdots, p_l)$-order tensor and $x^{(i)} \in \Re^{n_i}$.

We will continue work on related topics.

Bibliography

[1] P.J. Basser and D.K. Jones, "Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review," *NMR in Biomedicine,* 15 (2002) 456-467.

[2] P.J. Basser, J. Mattiello and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophysical Journal,* 66 (1994) 259-267.

[3] P.J. Basser, J. Mattiello and D. LeBihan, "Estimation of the effective selfdiffusion tensor from the NMR spin echo," *Journal of Magnetic Resonance, B,* 103 (1994) 247-254.

[4] P.J. Basser and S. Pajevic, "Spectral decomposition of a 4th-order covariance tensor: Applications to diffusion tensor MRI," *Signal Processing,* 87 (2007) 220-236.

[5] A. Ben-Tal, A. Nemirovski and C. Roos, "Robust solutions of uncertain quadratic and conic quadratic problems," *SIAM Journal on Optimization,* 13 (2002) 535-560.

[6] A.B. Berkelaar, J.F. Sturm and S.Z. Zhang, "Polynomial primal-dual cone affine scaling for semidefinite programming," *Applied Numerical Mathematics,* 29 (1999) 317-333.

[7] D.P. Bertekas, "Nonlinear Programming," *Athena Scientific,* 1999.

[8] P. Biswas, T.C. Liang, T.C. Wang and Y.Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Transatction on Sensor Networks,* 2 (2006) 188-220.

[9] I.M. Bomze and E. de Klerk, "Solving standard quadratic optimization problems via linear, semidefinite and copositive programming," *Journal of Global Optimization,* 24 (2002) 163-185.

[10] P.H. Calamai and J.J. Moré, "Projected gradient methods for linearly constrained problems," *Mathematical Programming,* 39 (1987) 93-116.

[11] J.F. Cardoso, "Blind signal separation: statistical Principles," *Proceedings of the IEEE,* 90 (1998) 2009-2026.

[12] J.F. Cardoso, "High-order contrasts for independent component analysis," *Neural Computation,* 11 (1999) 157-192.

[13] S. Chiriță, A. Danescu and M. Ciarletta, "On the srtong ellipticity of the anisotropic linearly elastic materials," *Journal of Elasticity,* 87 (2007) 1-27.

[14] P. Comon, "Independent component analysis, a new concept?" *Signal Processing,* 36 (1994) 287-314.

[15] P. Comon, G.H. Golub, L.H. Lim, and B. Mourrain, "Symmetric tensor and symmetric tensor rank," *SIAM Journal on Matrix Analysis and Applications,* 30 (2008) 1254-1279.

[16] D. Cox, J. Little and D. O'Shea, "Using Algebraic Geometry," *Springer-Verlag, New York,* 1998.

[17] G. Dahl, J.M. Leinaas and J. Myrheim, E. Ovrum,"A tensor product matrix approximation problem in quantum physics," *Linear Algebra and its Applications,* 420 (2007) 711-725.

[18] A.C. Doherty, P.A. Parillo and F.M. Spedalieri "Distinguishing separable and entangled states," *Physical Review letter,* 88 (2002) 187904.

[19] G.H. Golub and C.F. Van Loan, "Matrix Computations," *3rd ed., Johns Hopkins University Press, Baltimore, MD,* 1996.

[20] L. Gurvits "Classical deterministic complexity of Edmonds' problem and quantum entanglement," *Proceeding of the Thirty-Fifth ACM symposum on theory of computing, ACM: New York.* (2003) 10-19.

[21] D.R. Han, H.H. Dai and L.Q. Qi, "Conditions for strong ellipicity of anisotropy elastic materials," *Journal of Elasticity,* 97 (2009) 1-13.

[22] D.R. Han, L.Q. Qi and Ed.X. Wu, "Extreme diffusion values for non-Gaussian diffusions," *Optimization Methods and Software,* 23 (2008) 703-716.

[23] S.M. He, Z.Q. Luo, J.W. Nie and S.Z. Zhang,"Semidefinite relaxation bounds for indefinite homogeneous quadratic optimization," *SIAM Journal on Optimization,* 19 (2008) 503-523.

[24] R. Horst, P.M. Pardalos and N.V. Thoai, "Introduction to Global Optimization," *Kluwer Academic Publishers, Dordrecht/Boston/London,* 2000.

[25] Y.M. Huang and S.Z. Zhang,"Complex matrix decomposition and quadratic programming, " *Mathematics of Operations Research,* 32 (2007) 758-768.

[26] E.S. Hui, M.M. Cheung, L.Q. Qi and Ed X. Wu,"Towards better MR charactrization of neural tissues using directional diffusion kurtosis analysis," *NeuroImage,* 42 (2008) 122-134.

[27] J.H. Jensen, J.A. Helpern, A. Ramani, H. Lu and K. Kaczynski, " Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging," *Magnetic Resonance in Medicine,* 53 (2005) 1432-1440.

[28] D. Jibetean and M. Laurent, "Semidefinite approximations for global unconstrained polynomial optimization," *SIAM Journal on Optimization,* 16 (2005) 490-514.

[29] J.K. Knowles and E. Sternberg, "On the ellipticity of the equations of non-linear elastostatics for a special material," *Journal of Elasticity,* 5 (1975) 341-361.

[30] J.K. Knowles and E. Sternberg, "On the failure of ellipticity of the equalitions for finite elastostatic plane strain," *Archive for Rational Mechanics Analysis,* 63 (1977) 321-326.

[31] E. Kofidis and P.A. Regalia,"On the best rank-1 approximation of higher-order supersymmetric tensors," *SIAM Journal on Matrix Analysis and Applications,* 23 (2002) 863-884.

[32] L. De Lathauwer, B. De Moor and J. Vandewalle,"On the best rank-1 and rank-$(R_1, R_2, \cdots, R_N)$ approximation of higher-order tensor," *SIAM Journal on Matrix Analysis and Applications,* 21 (2000) 1324-1342.

[33] C. Ling, J.W. Nie, L.Q. Qi and Y.Y. Ye, "Bi-quadratic optimization over unit spheres and semidefinite programming relaxations," *SIAM Journal on Optimization,* 20 (2009) 1286-1310.

[34] C.L. Liu, R. Bammer, B. Acar and M.E. Moseley, "Characterizing non-gaussian diffusion by using generalized diffusion tensors," *Magnetic Resonance in Medicine,* 51 (2004) 924-937.

[35] C.L. Liu, R. Bammer and M.E. Moseley, "Limitations of apparent diffusion coefficient-based models in characterizing non-gaussian diffusion, " *Magnetic Resonance in Medicine.,* 54 (2005) 419-428.

[36] H. Lu, J.H. Jensen, A. Ramani and J.A. Helpern, "Three-dimensional characterization of non-Gaussian water diffusion in humans using diffusion kurtosis imaging," *NMR in Biomedicine,* 19 (2006) 236C247.

[37] W.S. Lu and S.C. Pei, "On optimal low-rank approximation of multidimensional discrete signals," *IEEE Transacations on Circuits and Systems II,* 45 (1998) 417-422.

[38] Z.Q. Luo, N. Sidiropoulos, P. Tseng and S.Z. Zhang. "Approximation bounds for quadratic optimization with homogeneous quadratic constraints." *SIAM Journal on Optimization,* 18 (2007) 1-28.

[39] Z.L. Luo and S.Z. Zhang, "A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints," *SIAM Journal on Optimization,* 20 (2010) 1716-1736.

[40] B. Maricic, Z.Q. Luo and T.N. Davidson,"Blind constant modulus equalization via convex optimization," *IEEE Transactions on Signal Processing,* 51 (2003) 805-818.

[41] M. Mignotte, "Mathematics for computer algebra," *Springer-Verlag, New York, NY,* 1991.

[42] A. Nemirovski, C. Roos and T. Terlaky, "On maximization of quadratic form over intersection of ellipsoids with common center" *Mathematical programming,* 86 (1999) 463-473.

[43] Yu. Nesterov, "Random walk in a simplex and quadratic optimization over convex polytopes,"CORE Discussion Paper 2003/71, CORE, Catholic University of Louvain, Louvain-la-Neuve, Belgium, 2003.

[44] Yu. Nesterov, "Global quadratic optimization via conic relaxation," Working paper, CORE, Catholic University of Louvaine, 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium, 1998.

[45] Q. Ni, L.Q. Qi and F. Wang, "An egienvalue method for testing positive definiteness of a multivariate form,"*IEEE transaction on automatic control,* 53 (2008) 1096-1107.

[46] Q. Ni, L.Q. Qi, F. Wang and Y.J. Wang, "The degree of the E-characteristic polynomial of an even order tensor,"*Journal of Mathematical Analysis and Applications,* 329 (2007) 1218-1229.

[47] J. Nocedal and S.J. Wright, "Numerical Optimization," *Springer Press* 1999.

[48] P.M. Pardalos and H. Wolkowicz, "Topics in semidefinite and interior-point methods,"*Fields Institute Communications,* Vol. 18, American Mathematical Society, 1998.

[49] G. Pataki, "On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues", *Mathematics of Operations Research,* 23 (1998) 339-358.

[50] A. Peres, " Separabilities criterion for density matrices,"*Physical Review Letters,* 77 (1966) 1413-1415.

[51] L.Q. Qi, "Eigenvalues of a real supersymmetric tensor."*Journal of Symbolic Computation,* 40 (2005) 1302-1324.

[52] L.Q. Qi, "Rank and eigenvalues of a supersymmetric tensor, the multivariate homogeneous polynomial and the algebraic hypersurface it defines,"*Journal of Symbolic Computation,* 41 (2006) 1309-1327.

[53] L.Q. Qi, "Eigenvalues and invariants of tensors,"*Journal of Symbolic Computation,* 325 (2007) 1363-1377.

[54] L.Q. Qi, H.H. Dai and D.R. Han,"Conditions for strong ellipticity and M-eigenvalue," *Frontiers of Mathematics in China,* 4 (2009) 349-364.

[55] L.Q. Qi, D.R. Han and Ed X. Wu,"Principal invariants and inherent parameters of diffusion kurtosis tensors," *Journal of Mathematical Analysis and Applications,* 349 (2009) 165-180.

[56] L.Q. Qi, W.Y. Sun and Y.J. Wang, "Numerical multilinear algebra and its applications," *Frontiers of Mathematics in China,* 2 (2007) 501-526.

[57] L.Q. Qi, F. Wang and Y.J. Wang, "Z-Eigenvalue methods for a global optimization polynomial optimization problem," *Mathematical Programming,* 118 (2009) 301-316.

[58] L.Q. Qi, Y.J. Wang and Ed X.Wu, "D-eigenvalues of diffusion kurtosis tensors," *Journal of Computational and Applied Mathematics,* 221 (2008) 150-157.

[59] P.A. Regalia and E. Kofidis "The higher-order power method revisited: Convergence proofs and effective initialization," *Proceedings of the Acoustics, Speech, and Signal Processing,* 5 (2000) 2709-2712.

[60] B. Reznick, "Some concrete aspects of Hilbert's 17^{th} problem," *Contemporary Mathematics,* 253 (2000) 251-272.

[61] P. Rosakis, "Ellipticity and deformations with discontinuous deformation gradients in finite dlastostatics," *Archive for Rational Mechanics and Analsis,* 109 (1990) 1-37.

[62] T. Terlaky eds, "Interior Point methods of Mathematical Programming," *Kluwer Academic Publishers,* (1996) 369-398.

[63] H.C. Simpson and S.J. Spector, "On copositive matrices and strong ellipticity for isotropic elastic materials," *Archive for Rational Mechanics and Analsis,* 84 (1983) 55-68.

[64] A.M.-C. So, Y.Y. Ye and J.W. Zhang, "A unified theorem on SDP rank reduction," *Mathematics of Operations Research,* 33 (2008) 910-920.

[65] J.F. Sturm and S.Z. Zhang, "On cones of nonnegative quadratic functions," *Mathematics of Operations Research,* 28 (2003) 246-267.

[66] W. (Lord Kelvin) Thomson, "Elements of a mathematical theory of elasticity," *Philosophical Transactions of the Royal Society of London,* 146 (1856) 481-498.

[67] P. Tseng,"Further results on approximating nonconvex quadratic optimization by semidefinite progremming relaxation," *SIAM Journal on Optimization,* 14 (2003) 268-283.

[68] F. Verstraete, J. Dehaene and B.De Moor, "On the geometry of entangled states," *Journal of Modern Optics,* 49 (2002) 1277-1287.

[69] Y. Wang and M. Aron, "A reformulation of the strong ellipticity conditions for unconstrained hyperelastic media," *Journal of Elasticity,* 44 (1996) 89-96.

[70] Y.J. Wang and L.Q. Qi, "On the successive supersymmetric rank-1 decomposition of higher-order supersymmetric tensors," *Numerical Linear Algebra with Applications,* 14 (2007) 503-519.

[71] Y.J. Wang, L.Q. Qi and X.Z. Zhang "A practical method for computing the largest M-eigenvalue of a fourth-order partially symmetric tensor," *Numerical Linear Algebra with Applications,* 16 (2009) 589-5601.

[72] Y.Y. Ye and S.Z. Zhang, "New results on quadratic minimization," *SIAM Journal on Optimization,* 14 (2003) 245-267.

[73] X.Z. Zhang, C. Ling, L.Q. Qi, Ed.X. Wu, "The measure of diffusion skewness and kurtosis in magnetic resonance imaging," *Pacific Journal of Optimization,* 6 (2010) 391-404.

[74] T. Zhang and G.H. Golub, "Rank-1 approximation of higher-order tensors," *SIAM Journal on Matrix Analysis and Applications,* 23 (2001) 534-550.