School of Electrical Engineering, Computing and Mathematical Sciences

Numerical Algorithms for Polynomial Optimisation Problems with Applications

Mohammed Aeyed M Alqahtani

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

January 2019

Declaration

To the best of my knowledge and belief, this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

.....
Mohammed Aeyed M Alqahtani
30 January 2019

Abstract

A tensor is a multi-dimensional array of real (or complex) numbers. It is a natural extension of a matrix with many applications in science and engineering fields. Over the last few years, there have been significant developments on the spectral theory of tensors. Many important characterizations of nonnegative matrices have been extended to some special classes of tensors such as nonnegative tensors. In particular, the Perron-Frobenius theory and the minmax theorem have been extended from nonnegative matrices to the class of nonnegative tensors. Furthermore, the eigenvalue problem for matrices has been generalized to high order tensors. Eigenvalue problems of nonnegative tensors find applications or links with higher order Markov chains, spectral hypergraph theory, and the quantum entanglement. In Chapter 3, we present a fast algorithm for computing the spectral radii of symmetric nonnegative tensors. In particular, by this proposed algorithm we are able to obtain the spectral radii of weakly reducible symmetric nonnegative tensors without requiring the partition of the tensors. As we know, it is very costly to determine the partition for large-size weakly reducible tensors. Numerical results are reported to show that the proposed algorithm in Chapter 3 is efficient and also it is able to compute the spectral radii of large-size tensors. As an application, we present an algorithm for testing the positive definiteness of Z-tensors. By this algorithm, it is guaranteed to determine the positive definiteness for any Z-tensor. In Chapter 4, we study homogenous polynomial optimization problems under the unit sphere constraints. These optimization problems have wide applications e.g., in signal and image processing, high order statistics, and computer vision. Since these problems are, in general, NP-hard, we are interested in studying approximation algorithms. In particular, we propose some polynomial-time approximation algorithms with new approximation bounds. In addition, based on these approximation algorithms, some efficient algorithms are presented and numerical results are reported to show the efficiency of our proposed algorithms. In Chapter 5, we study the maximum clique problem (MCP) which is a well-known example of combinatorial optimization. The MCP has been proven to be NP-complete, so it is difficult to determine the global solutions

of this problem. We show that the MCP can be equivalently formulated into homogeneous polynomial optimization problems under the unit spheres which are studied in Chapter 4. Then, we develop a computational algorithm that is called the alternating direction method of multipliers (ADMM) to solve for the MCP. In particular, we apply the ADMM method to the most popular DIMICS benchmark graphs of different dimensions, and the numerical results show that proposed method performs well for some of these graphs.

List of publications during PhD candidature

The following material is based upon work presented in this thesis.

- Zhou G, Wang G, Qi L, Alqahtani M. A fast algorithm for the spectral radii of weakly reducible nonnegative tensors. *Numer Linear Algebra Appl.* 2018;25:e2134. https://doi.org/10.1002/nla.2134

  (Note: Chapter 3 is based on this paper)

- Zhang X, Zhou G, Caccetta L, Alqahtani M. Approximation algorithms for nonnegative polynomial optimization problems over unit spheres. *Frontiers of Mathematics in China*, 12(6), 1409-1426, 2017.

  (Note: Chapter 4 is based on this paper)

- Alqahtani M, Zhou G. Maximum clique problem. Appeared in 40^{th} *Australian Combinatorial Mathematics and Combinatorial Computing*, Newcastle, Australia 12-16 Dec 2016, P.13, https://40accmcc.newcastle.edu.au.

  (Note: Chapter 5 is based on this technical report)

Acknowledgements

The research reported in this thesis was carried out from Aug 2014 to Aug 2018. During this period, I was a PhD student at the Department of Mathematics and Statistics, Curtin University.

First of all, I would like to express my deeply grateful appreciations and thanks to my supervisors, Associate Professor Guanglu Zhou and Professor Yong Hong Wu for their incredible insight and knowledge, and invaluable guidances during my studying journey at Curtin University. Without their support and encouragements, I would never be able to complete this work. I truly appreciate their shared knowledge, excellence and bright ideas.

Furthermore, I would like to extend my appreciations to my sponsor, the Ministry of Higher Education of Saudi Arabia and the Saudi Cultural Mission in Australia, Canberra, for their financial support to pursue a PhD degree. With their support and my advisor's encouragements, I had the opportunity to participate in academic exchanges and academic conferences, such as in Australia (Newcastle university, NSW) and internationally in China (Changsha university, Hunan State), which helped me improve my research greatly.

Very special thanks are also due to my parents, for their deep love, limitless supports and inspirations. In addition, I would like to extend my appreciations to my wife, Sharifah Sultan, for being so patient and supportive, and for helping me with everything. Many special thanks are also due to all of my family members (brothers and sisters) who encouraged and supported me throughout my studies. Their infinite supports and inspirations helped me accomplish great things.

Contents

CHAPTER 1

Introduction

This chapter consists of three main parts. In the first part of this chapter, we introduce some basic definitions, concepts, notations and properties about tensor theory and polynomial optimisation problems. In addition, the second part reviews some important results and developments regarding tensor eigenvalue problems. The last section of this chapter briefly outlines the contents of this thesis.

1.1 Preliminaries

Matrix theory is a vast field of fundamental importance to both theoretical and practical aspects in a various area of mathematics, science and engineering. However, some scientific and engineering applications are of high dimensional sets (i.e. multi-indexed data sets) hence the matrix analysis is extended from a second order case to higher orders. A higher order generalization of a matrix is called a tensor which used to describe the multi-arrays with indices more than two. The origin of tensor concepts can be traced back to the 18th century, in particular, to great mathematicians such as Carl Gauss, Bernhard Riemann, William Hamilton, and Elwin Christoffel. Many relations relating to vectors, matrices and tensors can be described geometrically by tensor theory. For instance, in 1846, William Hamilton used tensors in both algebraic systems and differential geometry. Further in 1898, Woldemar Voigt a German physicist who extensively studied the nature of crystals used tensor theory in order to describe the properties of crystals. Also, tensors treated as physical quantities and the current meaning of tensors in Mathematical physics were commonly refereed to him. Gregorio Ricci developed tensor notation as an extension of vectors in absolute differential calculus while Einstein used tensor analysis for general relativity.

Recently, tensors have become increasingly common tool to treat multidimensional arrays as they in general are an adequate generalization of both matrices and vectors. Tensors have a wide range of interesting applications in various fields, including natural science, engineering, physics, and mathematics. In particular, tensors arise in solid mechanics, fluid dynamics, electromagnetism, diffusion tensor imaging, tensorial elastography analysis, computer vision, and diffusion filters. There has been an increased in theoretical work on tensor theory and its spectral properties and a detailed survey that explains relative basic definitions and properties of tensors can be found in [103]. This section presents some definitions and notations relating to nonnegative tensors and recalls some well-known results.

Throughout this thesis, we use $\Re$, $\Re^n$ and $\mathbb{C}^n$, to denote the set of all real numbers, the n-dimensional real space, and the n-dimensional complex space, respectively. In addition, let $m, n \geq 2$, we use $\mathcal{T}^{m,n}$, $\mathcal{T}_+^{m,n}$, $\mathcal{S}^{m,n}$, and $\mathcal{S}_+^{m,n}$ to denote the set of all real tensors of order m and dimension n, the set of all real nonnegative tensors of order m and dimension n, the set of all real symmetric tensors of order m and dimension n, and the set of all real symmetric nonnegative tensors of order m and dimension n, respectively. Also, it is very important to point out that the italic capitals $(A, B, ...)$ are used to denote matrices, and calligraphic capitals $(\mathcal{A}, \mathcal{B}, ...)$ are used to denote higher-order tensors. Let

$$P_n = \{x_i \geq 0 : x \in \Re^n, 1 \leq i \leq n\}$$

and

$$int(P_n) = \{x_i > 0 : x \in \Re^n, 1 \leq i \leq n\}.$$

A tensor (i.e. hyper-matrix) $\mathcal{A} = (a_{i_1 \ldots i_m})$ is simply defined as a multi-dimensional array of entries. It is a generalization of vectors and matrices. For instance, scalars are simply zero-order tensors, vectors are tensors of first-order, matrices are tensors of second-order, and tensors of order three or higher ($m \geq 3$) are called higher-order tensors.

In this thesis, we use $\mathcal{A}$ to denote a real $m-$order $n-$dimensional square real tensor that is composed of n^m entries in the field $\Re$, and has the form such that,

$$\mathcal{A} := (a_{i_1 \ldots i_m}), \ where \ a_{i_1 \ldots i_m} \in \Re, \quad \forall i_j \in \{1, ..., n\} \ and \ j \in \{1, ..., m\}. \quad (1.1)$$

A tensor $\mathcal{A} = (a_{i_1 \ldots i_m})$ is said to be *nonnegative*, if its entries are nonnegative (*i.e.* $a_{i_1 \ldots i_m} \geq 0$), *positive*, if its entries are positive (*i.e.* $a_{i_1 \ldots i_m} > 0$), and *symmetric*, if its entries $a_{i_1 \ldots i_m}$ are invariant under any permutation of their indices

$\{i_1, ..., i_m\}$ [144], (i.e. $a_{j_1 \ldots j_m} = a_{i_1 \ldots i_m}$, among all the permutations $j_1 \ldots j_m$ of $i_1 \ldots i_m$, $1 \leq i_1 \ldots i_m \leq n$). Clearly, $\mathcal{A}$ is a symmetric matrix if $m = 2$ and a higher order tensor if $m \geq 3$.

A symmetric tensor $\mathcal{A}$ represents the homogeneous polynomial $F_{\mathcal{A}}(x)$ of degree m with real coefficient: $x = (x_1, \ldots, x_n)^T \in \Re^n$ as

$$F_{\mathcal{A}}(x) := \mathcal{A}x^m = \sum_{i_1, i_2, \ldots, i_m = 1}^{n} a_{i_1\, i_2 \ldots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}, \tag{1.2}$$

where x^m can be regarded as a rank one tensor of order m and dimension n with its entries as $x_{i_1}, \ldots, x_{i_m}$ [89]. Let $\mathcal{A} \in \mathcal{T}^{m,n}$, for an n-dimensional column vector $x = [x_1, x_2, ..., x_n]^T$, real or complex, we let $\mathcal{A}x^{m-1} \in \mathbb{C}^n$, be an n-dimensional column vector as:

$$\mathcal{A}x^{m-1} := \left(\sum_{i_2, \ldots, i_m = 1}^{n} a_{i\, i_2 \ldots i_m} x_{i_2} \cdots x_{i_m} \right)_{1 \leq i \leq n}. \tag{1.3}$$

Definition 1.1. *Let tensor $\mathcal{A}$ defined in (1.1) be symmetric, and let m be even, then:*

(i) *$\mathcal{A}$ is said to be positive definite iff its homogeneous polynomial $F_{\mathcal{A}}(x) := \mathcal{A}x^m > 0$ for all $x \in \Re^n$ and $x \neq 0$ [144],*

(ii) *$\mathcal{A}$ is said to be positive semi-definite iff $\mathcal{A}x^m \geq 0$ for all $x \in \Re^n$ [144],*

(iii) *$\mathcal{A}x^m$ is said to be a copositive tensor iff $x \in P_n$, then $\mathcal{A}x^m \geq 0$ [147],*

(iv) *$\mathcal{A}$ is said to be a strictly copositive tensor iff $x \in P_n$, then $\mathcal{A}x^m > 0$ [147].*

It is very important to point out that the symmetric tensor and homogeneous polynomial are bijectively associated [47, 89]. When the order m is even, the positive definiteness of (1.2) plays a prominent role in the stability analysis of nonlinear autonomous system by the direct method of Lyapunov in automatic control [4,24,71,91,170]. In [24], by using Sturm theorem, a method was suggested to check the positive definiteness of (1.2) if $n \leq 3$, but it is not practically useful if $n > 3$ and $m \geq 4$. Recently, Qi [144] was motivated to study this issue and defined eigenvalue concepts, in particular, the H-eigenvalues and Z-eigenvalues of symmetric tensor $\mathcal{A}$ and used them to check for the positive definiteness of tensor $\mathcal{A}$ [see Theorem (2.16)]. Furthermore, authors in [48, 65] established the close relationship of eigenvalues and eigenvectors of higher order tensor with the theory of resultants. Independently, Lim [101] defined the same notion of eigenvalues,

eigenvectors, singular values and singular vectors for real tensors. In particular, l^2-eigenvalues and l^k-eigenvalues. The l^2-eigenvalues and l^k-eigenvalues in [101] are Z-eigenvalues and H-eigenvalues in [144], respectively.

For a symmetric tensor $\mathcal{A}$, the gradient and the Hessian matrix of $F_{\mathcal{A}}(x)$ are as follows:

$$\begin{aligned}\nabla F_{\mathcal{A}}(x) &:= m\mathcal{A}x^{m-1} \\ &= m\left(\sum_{i_2,\ldots,i_m=1}^{n} a_{i\,i_2\ldots i_m}x_{i_2}\cdots x_{i_m}\right)_{1\le i\le n}, \\ \nabla^2 F_{\mathcal{A}}(x) &:= m(m-1)\mathcal{A}x^{m-2} \\ &= m(m-1)\left(\sum_{i_3,\ldots,i_m=1}^{n} a_{ij\,i_3\ldots i_m}x_{i_3}\cdots x_{i_m}\right)_{1\le i\le n,\ 1\le j\le n}.\end{aligned}$$

By this observation, we present a definition of weakly symmetric tensor.

Definition 1.2. *[38] A tensor $\mathcal{A}\in\mathcal{T}^{m,n}$ is said to be weakly symmetric if its associated homogeneous polynomial $F_{\mathcal{A}}(x)$ satisfies*

$$\nabla F_{\mathcal{A}}(x) = m\mathcal{A}x^{m-1}, \quad \forall x\in\Re^n$$

and the right-hand side of the above equation is not identical to zero.

1.2 Literature Review

Over the last few years, there has been a flurry of work on the spectral tensor theory in which many researchers have been attracted to develop different numerical studies on tensors of higher order. This is due to its significant roles to address issues in various fields. In particular, the spectral theory of nonnegative tensors has found many related applications or links in automatical control [144], spectral hypergraph theory [32,81,138], higher order Markov chains [37,82], magnetic resonance imaging [151, 152], algebraic geometry [33, 96], Finsler geometry [9], quantum entanglement [83,146], image authenticity verification [188], multilinear page rank [101], polynomial optimizations [122] and others. Nonnnegative tensors form a singularly important class of tensors that has gained a lot of attention recently due to some shared intrinsic characteristics with those of nonnegative matrices. In specific, the Perron-Frobenius theorem and the Minmax theorem are both two illustrations of those explored properties that play a significant role in eigenvalue's computation for nonnegative tensor. Recently, Qi and Lim [101,144]

expanded the eigenvalue problem from matrices to tensors, and further proved that many properties of eigenvalues can be generalized to tensors. However, the tensor eigenvalue problems are almost invariably computationally difficult. Unlike the matrix eigenvalue problem, the eigenvalue problems for tensors of higher orders are nonlinear. It has been shown [76] that computing eigenvalues (i.e. the simplest multilinear generalization) of general higher order tensor (i.e. $m \geq 3$) is an NP-hard problem. Although, for some special classes of tensors, it was also noted that eigenvalue problem is computable, however, all these methods quickly become impractical when the tensor order becomes large. A detailed reference on the computational complexity of tensor eigenvalue is provided in [76]. This poses a challenging task for the spectral theory of tensors and thus, many researchers have been encouraged to develop the theory and algorithms for eigenvalues of nonnegative tensors. In recent years, there have been many research works that concentrating on eigenvalues of nonnegative tensors. This is in part due to its significant role in many practical applications, including higher order Markov chains [82], spectral hypergraph theory [100] and the quantum entanglement [146], and also in part due to its strong foothold in theoretical concepts. We refer to [35] for a detailed research study of the spectral theory and computational algorithms of eigenvalues of nonnegative tensors with some applications. Qi [144] defined the concepts of eigenvalues for tensors of higher order, and proved that there exists real and complex eigenvalues and eigenvectors. Also, in [144], it was proven that eigenvalues exist for an even real symmetric tensor $\mathcal{A}$ with practical importance as indicators to determine the positive definiteness of an even degree multivariate form and to find the best rank-one approximation of a supersymmetric tensor. Lim [101] independently introduced eigenvalues and eigenvectors, but restricted them to be real. Many classical concepts and results for eigenvalues of square matrices have been extended to tensors, as they have a wide range of interesting applications including mechanics, physics, and the classification of hyper-surfaces and hyper-graphs.

Nonnegative tensors have been classified into seven categories including strictly nonnegative tensor [78], weakly irreducible nonnegative tensor [63], weakly primitive tensor [63], irreducible nonnegative tensors [39], primitive tensor [39], essentially positive tensor [136], and weakly positive tensors [190]. The relations among these classes have been studied in [78], and it was found that the weakly irreducible tensors are strictly nonnegative tensors but not conversely. Furthermore, in [137], it was discovered that essentially positive tensors are primitive tensors, primitive tensors are irreducible tensors and weakly primitive tensors are

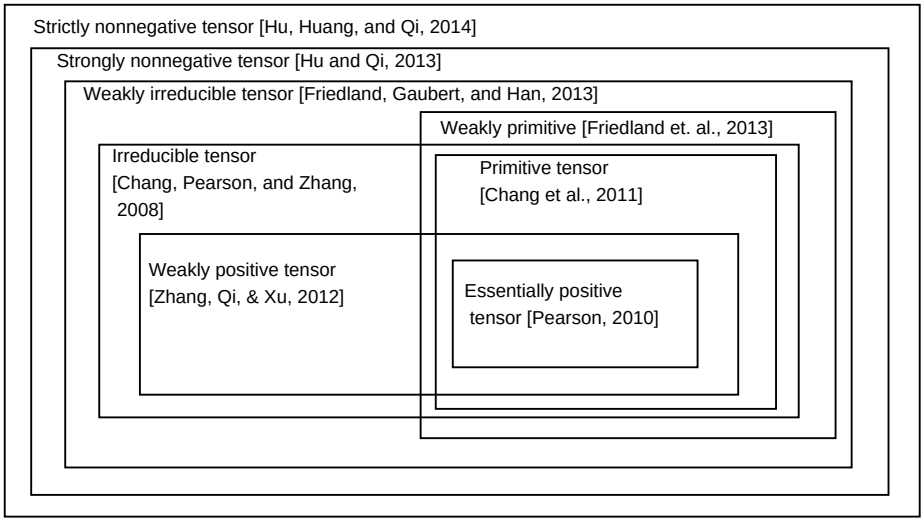

Figure 1.1: Relations among different types of nonnegative tensors

weakly irreducible but not necessarily vice versa. It was also shown that essentially positive tensors are weakly positive tensors, but the converse is false, and weakly positive tensors are irreducible nonnegative tensors, but the converse is not valid [190]. The relationships among these classes are shown in Figure 1.1 and a detailed description of the nonnegative tensors is given in [78].

The field of nonnegative tensors has blossomed, especially since the introduction of eigenvalues and eigenvectors of higher order tensors in [101, 144]. Since then, the spectral theory of tensors has been developing rapidly and many results have been established. In particular, the Perron-Frobenius theorem has been generalized from nonnegative matrices to higher order nonnegative tensors in [39]. In addition, Yang et al. [185,186] had extended the weak Perron-Frobenius theorem to general tensors. Also, other research studies on Perron-Frobenius theorem exist in [79, 193]. In the theory of nonnegative tensors, examples where the Perron-Frobenius theorem is useful arise in measuring higher order connectivity in linked objects [101] and hypergraphs [100]. Furthermore, in [39], the well known minmax theorem of Collatz [46] was expanded from irreducible nonnegative matrices to irreducible nonnegative tensors. Recent work presented in [98] on the Perron-Frobenius theorem indicated that the eigenvalues with modulus (i.e. the maximum eigenvalue) have the same geometric multiplicity. On the basis of graph theory, new bounds were established for the clique number of graphs and their calculations were based on solving the maximum eigenvalue of a $\{0, 1\}$ nonnegative tensor [**?**]. Lim [101] introduced singular values of non-square tensors, and its related properties have been studied in [36, 101]. Recently, in [36], the Perron-Frobenius theorem was extended to nonnegative rectangular tensors

and an iterative method was designed to compute the maximum singular value of a nonnegative rectangular tensor. Zhou et al. [196] has improved a power type method proposed in [36] such that it is convergent for any irreducible nonnegative rectangular tensor. Also, more results were established for the singular values of nonnegative rectangular tensors in [185]. In particular, they expanded the weak Perron-Frobenius result (i.e Theorem 2.3, [186]) to nonnegative rectangular tensor and found out that a generalization of Perron -Frobenius result of positive tensors does not hold for positive rectangular tensor, but a weaker conclusion was proved such that it may be hold only for the positive situation. Some other related theoretical results of nonnegative tensors can be found in [35,36,100,122,126,148,187].

In recent years, tensor eigenvalue problems have received considerable attention in numerical multi-linear algebra. Ng, Qi, and Zhou [120] developed an iterative power type algorithm in order to calculate the highest eigenvalue (i.e. the spectral radius) and the corresponding eigenvector of an irreducible nonnegative tensor. This computational algorithm is called the NQZ algorithm, which is an extension of Collatz's result in [46,179] that has been used to determine the largest magnitude eigenvalue and its corresponding eigenvector of an irreducible nonnegative matrix. The NQZ method is an effective method but its convergence is not guaranteed for weakly reducible nonnegative tensors. In [40], the notion of primitive tensors is introduced and the linear convergence of the NQZ algorithm is established for primitive nonnegative tensors. Friedland et al. [63] proved that the power algorithm in [120] converges for weakly primitive nonnegative tensors. In [136], a class called an essentially positive tensors is defined which is a special class of primitive tensors, and Pearson conjectured that the NQZ method could be convergent for essentially positive tensors of even order. A remarkable result was made by Zhang and Qi [189], as they studied the convergence of the NQZ algorithm for the class of essentially positive tensors. Liu, Zhou and Ibrahim [107] modified the NQZ algorithm to present the LZI method, which is always convergent for solving the largest eigenvalue of nonnegative irreducible tensors. In [191], the linear convergence rate of the LZI method was established for the class of weakly positive tensors. Also in [136] for an even order essentially positive tensor, it has been proven that the unique positive eigenvalue is real-geometrically simple (i.e. the corresponding real eigenvector is unique up to a scaling constant). Friedland, Gaubert and Han [63] identified the concepts of weakly irreducible nonnegative tensors and derived the Perron-Frobenius theorem and its fundamental properties for such tensors. Moreover, they discovered that there is a very close link between the Perron-Frobenius theorem for

nonnegative tensors and the Perron-Frobenius theorem for homogenous monotone maps. Friedland et al. [63] proposed a power algorithm for polynomial eigenvalues problems, and established the R-linear convergence of this algorithm under the weak primitivity condition. Recently, Hu, Huang and Qi [78] modified the NQZ method and proved that this updated version has a global R-linear convergence for weakly irreducible nonnegative tensors. In addition, there are many existing efficient numerical algorithms that are well developed for computing eigenvalues of nonnegative tensors, (see e.g. [38–40, 63, 78, 107, 120, 147, 185, 190, 191, 199] for additional details).

Furthermore, some researchers have devised a number of power iteration methods to calculate the spectral radius and its corresponding positive eigenvector for the class of weakly irreducible nonnegative tensors. Recently, the NQZ method was modified in [107] such that the updated version is convergent for any weakly irreducible nonnegative tensor. Also, the global linear convergence of NQZ method and the updated version in [107] had been established in [79, 189, 190] under weakly irreducible conditions. However, if a nonnegative tensor is weakly reducible, then these iterative power methods are not guaranteed to converge as shown in [120] for the NQZ method. Consequently, it is suggested in [78, 199] that if a nonnnegative tensor is weakly reducible, then there always exists a partition $\{I_1, I_2, \ldots, I_m\}$ of the index set $\{1, 2, \ldots, n\}$ such that every tensor in $\{\mathcal{A}_{Ij}|j \in (1, 2, \ldots, m)\}$ is weakly irreducible; and from these induced tensors, we are able to determine the largest eigenvalue of tensor $\mathcal{A}$. Therefore, these iterative methods can be applied to weakly irreducible tensors to find its spectral radius. Hu, Huang and Qi [78] modified the power algorithm in [63] such that it has a global R-linear convergence under the condition of weakly irreducibility. Furthermore, by (Algorithm 4.1, [78]), it has been shown that the spectral radius of a general nonnegative tensor can be found by calculating the spectral radii of weakly irreducible sub-tensors. However, this algorithm does not work if we consider reducible tensor instead of weakly reducible tensor.

It is important to point out that there is a vital link between symmetric tensor approximation and polynomial optimization problems. For instance, in [199], it has been shown that there is a closed connection between solving the maximum eigenvalue of a symmetric nonnegative tensor and the global solution of a convex polynomial optimization problem. In particular, it is found that solving the maximum eigenvalue of a symmetric nonnegative tensor is equivalent to solving the global solution of a convex optimization problems [199]. This feature provides some generalized forms that are so useful in solving some hardly tractable

combinatorial optimization problems such as the maximum clique problem and the labelling problem. More recently in [199], Zhou et al. proved some important characteristics for symmetric nonnegative tensors. In [186], the geometric programming method has been used to determine the maximum eigenvalue for nonnegative tensors, which has an advantage of a polynomial time complexity. Numerical studies in [186,197] show that the power type method outperforms the geometric programming method. Furthermore, a recent treatise on structure tensors was made in [42] to study the so called essentially nonnegative tensor, and it is shown that the maximum H-eigenvalue can be computed by using a semi-definite programming to solve for a polynomial optimization problem. In addition, numerical experiments show that their proposed method is efficient and it can be 10 times faster than the NQZ method. In [80], another iterative method was proposed for calculating the maximum Z-eigenvalue of an even-order symmetric tensor. This method is regarded as an improvement of methods in [150,152], and their numerical results indicate its effectiveness. For more details about other numerical methods for computing the real eigenvalues of a symmetric tensors, tensor decompositions and applications, we refer readers to [47,50,78,90,123,124].

Polynomial optimization simply means that there is a polynomial function that needs to be optimized subject to a set of equality and inequality constraints. Recently, polynomial optimization problems have received growing interest both theoretically as well as numerically due to its extensive applications in a wide variety of domains such as biomedical engineering, signal processing, material science, speech recognition, radar waveform design and many more. Polynomial optimization problems are notoriously challenging as they are typically nonconvex and highly nonlinear. Nestrov [119] showed that maximizing a cubic polynomial over a sphere is NP-hard. For a comprehensive survey of the computational complexity for the polynomial optimization problems over some simple constraint sets, we refer the reader to De Klerk [51]. In order to cope with this difficulty, researchers have centred their focus on designing such approximation methods to solve for polynomial optimization problems. In [109], the first polynomial-time approximation algorithm was designed to optimize a multivariate quartic polynomial over a region defined by quadratic inequalities. Later in [104], a polynomial time approximation algorithm was proposed to optimize a biquadratic function over two spheres. These findings have triggered a number of scholarly research in recent years. Specifically, He et al. [74] extended the results in [109], and studied the homogeneous polynomial optimization problems of any fixed degree over quadratic constraints.

In terms of computational methods in the literature of polynomial optimization, there are various approaches for solving polynomial optimization problems that are mostly based on non-linear programming and global optimizations. In particular, there are some available solver programs like KNITRO, BARON, MINOS, SNOPT, and Matlab optimization toolbox. Furthermore, in [94], the sum of square method (SOS) is one of the most important mathematical tool that has been designed for solving the general polynomial optimization problem by considering it as a semidefinite program. Also, in [143], another interesting programming solver called SOSTOOLS is designed based on decomposition of multivariate polynomials to solve for SOS. Based on SOS, Henrion et al. [75] developed several specialized Matlab toolboxes, in particular, the GloptiPoly and its improved version, GloptiPoly 3. These toolboxes are useful for solving global optimal solution of general polynomial optimization problems. Many other researchers have made significant advances into polynomial optimization problems with spherical constraints. In [66, 168], the method of Lagrange multipliers is one of such typical solution methods that has been used in order to reach a set of multivariate polynomial equations, namely the Kuhn-Karush-Tucker (KKT) system that provides the necessary conditions for optimality. Some other researchers aim to enumerate all solutions of a KKT system and hence design special methods such as the subdivision methods [116] and generalized normal forms methods [117]. However, these methods can be less effective if the encountered polynomial degree is high. Recently, another entirely different approach was proposed in [150] to study a tensor eigenvalue-based method for a global polynomial optimization problem. Particularly, this approach is called the Z-eigenvalue method and it aims to solve the homogeneous polynomial functions with a degree of at most three. In addition, in [111], diffusion-based methods are used to solve non-convex polynomial optimization models that arise in portfolio selection problems. Due to concavity of polynomial optimization problems, many researchers have focused on designing efficient algorithms to find good KKT solutions for polynomial optimization problems. Based on tensor optimization, Chen et al. [41] developed an optimization method that is called the maximum block improvement (MBI) method which guarantees that any cluster point converges to a stationary solution (hence a KKT point). Also in practice, the MBI method has the feature that it solves any optimization problem with separate block constraints. Also, it has been used as a local improvement scheme for polynomial optimization by starting from any good initial solutions in order to achieve better performance. Recently, Zhening et al. [99] proved the global and local linear convergence of

the MBI method under certain conditions. Moreover, another different yet efficient method for finding the KKT point is to use alternating direction method of multiplier (ADMM). Specifically, in [86], the classical ADMM was used for solving polynomial optimization problems and proved to be convergent under some conditions.

1.3 Outline of the thesis

In this research, we conduct a systematic study for tensor optimization problems which arise widely in many application areas. Issues discussed in this thesis are considered to be NP-hard problems, hence, they are difficult to solve both theoretically and numerically for many reasons. First, most of these problems are large-scale in size, and hence they require much more efficient algorithms with less computing time to solve them. In addition, the problems are nonlinear and nonconvex and cover a wide range of applications in different domains. Motivated by these challenges, we are interested to develop efficient and effective solution techniques for tensor eigenvalue problem and tensor optimization problems.

This thesis is divided into six chapters and organized as follows: Chapters 1 and 2 summarize the key points and also provide very brief recapitulation of nonnegative matrices, nonnegative tensors and their spectral properties.

Chapter 3 is devoted to study the tensor eigenvalue problems, in particular, we focus on finding the spectral radius of a symmetric nonnegative tensor. In this chapter, we develop an efficient and effective computational Algorithm (i.e. Algorithm 1 in [200]) for computing the spectral radius of a symmetric weakly reducible nonnegative tensor. One nice spectral property of this proposed algorithm is that we can obtain the spectral radius without requiring the partition of the tensors. This Algorithm 1 in [200] is an improved version of the iterative power method (i.e. NQZ method) proposed in [120] for calculating the maximum eigenvalue of a nonnegative tensor. The NQZ method has been regarded as one of the most important developments in the realm of numerical multi-linear algebra. Numerical results showed that Algorithm 1 in [200] is very efficient and also can compute the spectral radii of large size tensors. In addition, we design an algorithm that is guaranteed and useful for identifying the positive definiteness of a Z-tensor.

Chapter 4 is dedicated to study the tensor optimization problem over a unit sphere. We consider polynomial optimization problems with nonnegative coefficients. In particular, we study homogeneous polynomial optimization problems

over a single spherical constraint (P1) or over product of spherical constraints (P2). The aim of this chapter is to present some new approximation bounds for problems (P1) and (P2), which improve the current bound derived for general polynomial optimization problem in the literature. Particular attention is also paid to establish some polynomial-time approximation methods for these new approximation bounds. Numerical results reported in Section 4.3 of chapter 4 show that the proposed approximation methods are practical and able to produce very high quality solutions.

Chapter 5 is devoted to a special problem in graph theory called the maximum clique problem. One of the equivalent reformulations of the maximum clique problem is an optimization problem ($P1$) discussed in Chapter 4. However, a new continuous reformulation based on the Motzkin-Strauss QP is presented for the maximum clique problem. This new formulation is equivalent to the multi-linear functions with spherical constraints. We used Alternating Direction Method (ADM) to solve for the new relaxed problem as in general, ADM is a local search procedure and found to be efficient for many optimization problems. Furthermore, we showed extensive computational results on 66 DIMACS benchmark graphs with different dimensions. The overall results are reported in Tables 5.1 and 5.2, by comparing ADM performance with the proposed method in [22] which is based on Penalty formulation.

Finally, Chapter 6 summarizes all key findings and results in this thesis and provides further research direction and discussions.

CHAPTER 2

Nonegative matrices and nonnegative tensors

In this chapter, we briefly discuss both nonnegative matrices and nonnegative tensors, as well as their applications. The primary aim is to recall well known results and important spectral properties about nonnegative matrices and nonnegative tensors.

2.1 Nonnegative matrices

Over the last few decades, Perron [141] and Frobenius [64] introduced the concepts of nonnegative matrix theory and laid the foundations for several fundamental facts about their spectral properties. Since then, nonnegative matrix theory has been a very active area of research due to its increasingly important roles in diverse areas including, numerical analysis, linear complementarity problems, Markov chains [161], the input-output model in economics [95], queuing theory, geometric convexity [156], transition probabilities in finite Markov chains [161], machine learning, strategic market games [3], signal processing, and social network analysis for competitions [157].

Recall that a square matrix $A = (a_{ij})$ with real elements is called nonnegative (resp. positive) if its elements are called nonnegative (resp. positive), (i.e., $A \geq 0$ if $a_{ij} \geq 0$, and $A > 0$ if $a_{ij} > 0$ for all $i, j = 1, ..., n$). This concept and notation extends to vectors such that a real n-tuple $x = (x_1, ..., x_n)$ is called nonnegative (resp. positive) if $x_i \geq 0$ (resp. $x_i > 0), i = 1, 2, ..., n$. In the following, we will discuss some important combinatorial properties such as irreducibility and primitivity of nonnegative matrices. In 1912, Frobenius [64] originally introduced the term irreducible; it is also called unreduced or indecomposable in the literature.

Definition 2.1. *[169] A nonnegative matrix $A \in \Re^{n\times n}$ is called reducible if there*

exits a permutation matrix P such that

$$PAP^T = \begin{bmatrix} U & B \\ 0 & L \end{bmatrix},$$

where the diagonal blocks U and L are non-empty square. Moreover, A is reducible if an ordering $(i_1, \ldots, i_r, j_{r+1}, \ldots, j_n)$ of $(1, \ldots, n)$ exists such that $A[i_1, \ldots, i_r | j_{r+1}, \ldots, j_n]$ is equal to zero. Otherwise, A is called irreducible.

Let $G(A)$ be a directed graph that is associated to $A = (a_{ij})$ such that $G(A) = (V, E)$, with vertices $1, 2, \ldots, n$, and an edge from i to j iff $a_{ij} \neq 0$ [169]. A directed graph $G(A)$ is said to be strongly connected if for any pair of distinct vertices, i and j, there is some path between i and j ([169], p.20), for more details about graph theory concepts, we refer to [23, 108]. The notions of irreducibility and strong-connectivity are also related. The strong connectedness of $G(A)$ is equivalent to requiring that A is an irreducible matrix, which is stated in the following result.

Theorem 2.1. *[169] Let A be nonnegative matrix of order $n \times n$. Then A is irreducible iff it has a strongly connected graph $G(A)$.*

Further important characterizations of irreducibility of the nonnegative matrices are presented in the next theorem as follows:

Theorem 2.2. *[59] If a matrix A is a nonnegative, then the following statements are equivalent:*

(1) A is called irreducible.

(2) $(I + A)^{n-1} > 0$.

(3) For any pair i, j $(1 \leq i, j \leq n)$ there exists a natural number $k = k(i, j) \leq n$ such that $(A^k)_{ij} = a_{ij}^{(k)} > 0$.

It follows from Theorem 2.2 that A^T is irreducible whenever A is irreducible. According to Theorem 2.1, one can observe that $a_{ij}^{(k)} > 0$ iff there is an edge of length k form vertex i to vertex j in $G(A)$. Also, according to the equivalence of (1) and (3) in Theorem 2.2, it follows that A is irreducible iff $G(A)$ is strongly connected. The following corollary is an easy consequence of Theorem 2.2.

Corollary 2.1. *[114] If A is an irreducible nonnegative matrix of order n, then $A^{n-1} > 0$.*

Definition 2.2. *Let $A \in \Re^{n\times n}$. We use a nonzero vector $x \in \Re^n$ to denote the eigenvector of A and $\lambda \in \mathbb{C}$ to denotes its corresponding eigenvalue, satisfying*

$$Ax = \lambda x. \tag{2.1}$$

Computing eigenvalues and eigenvectors of a nonnegative matrix is useful in various modern applications, including age-specific population growth, animal population harvesting models, genetics context, and also in determining website page rankings as per Google's PageRank. Given a matrix A, let $\sigma(A)$ denote the spectrum of A, the set of all eigenvalues of A, and let $\rho(A)$ denote its spectral radius, which is the maximum distance of an eigenvalue from the origin. Spectral radius plays a major role in investigations of the rapidity of convergence of different iterative methods. Its calculation is useful in the identification of M-matrices [178] and in determining the unique positive solutions of a linear system in Leontief input-output models [176] in Economics. Other applications include computing the 2-norm of a matrix (that is, $\|A\|_2 = \sqrt{\rho(A^TA)}$) and also verifying a matrix is convergent (*i.e.* $\lim_{n\to\infty} A^n = 0$, which occurs iff $\rho(A) < 1$).

Definition 2.3. *[169] Let $A = (a_{ij})$ be an arbitrary $n \times n$ complex matrix, and let eigenvalues $\lambda_i, 1 \le i \le n$. Then*

$$\rho(A) \equiv \max_{1\le i\le n}\{|\lambda_i|, \lambda_i \textit{ is an eigenvalue of } A\}. \tag{2.2}$$

This can be reformulated based on the formula of Gelfand as follows,

$$\rho(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}},$$

where $\|.\|$ denotes the operator norm. Therefore, $\rho(A)$ is entirely found by A itself and hence is an intrinsic property. The Perron-Frobenius Theorem provides a simple characterization in the case of the eigenvalues and eigenvectors of irreducible nonnegative matrices. It is very important for two reasons; first, it is a most useful concept in many real world applications such as economics, population biology, and other areas of science and engineering, including theory of dynamical systems, Markov chains, economics, statistics and optimizations. Second, it plays a crucial role in the convergence analysis of some iterative methods for the eigenvalue problems. Positive matrices are a special type of irreducible nonnegative matrices. In 1907, Perron [141] discovered remarkable spectral properties for positive matrices. In particular, if $A > 0$, then the its spectral radius is positive with a corresponding positive eigenvector. Later, in 1912, Frobenius [64]

extended and generalized Perron's result to the class of irreducible nonnegative matrices. The classical Perron-Frobenius theorem for nonnegative matrices is stated in the following theorem.

Theorem 2.3. *(Perron-Frobenius Theorem, see Varga [169])*
Let A be an $n \times n$ irreducible and nonnegative matrix, then the following assertions hold:

(1) A has a positive real eigenvalue equal to its spectral radius $\rho(A)$,

(2) there exists an eigenvector (i.e., the Perron vector) with strictly positive entries (i.e., $x > 0$) corresponding to the largest positive eigenvalue of modulus $\rho(A)$ such that $Ax = \rho(A)x$,

(3) the spectral radius $\rho(A)$ is an algebraically and (hence geometrically) simple eigenvalue,

(4) If λ is an eigenvalue, then $\rho(A) \geq |\lambda|$. Moreover, if $Q \geq 0$ is a primitive matrix, then

$$\rho(Q) > |\lambda|, \forall \lambda \in \sigma(Q) \setminus \{\rho(Q)\},$$

where $\sigma(Q)$ denote the spectrum of Q

(5) If any entry of A increases (resp. decreases), then the value of $\rho(A)$ increases (resp. decreases).

(6) Let $m(> 1)$ be the index of a nonnegative irreducible cyclic matrix A, and let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the eigenvalues of A of modulus $\rho(A)$ which have the form

$$\rho(A)e^{2\pi i.j/m},\ 0 \leq j \leq m-1.$$

Furthermore, the entire spectrum of A of index m is invariant under a rotation of the plane about the origin through the angle of $2\pi/m$, but not through a positive angle smaller than $2\pi/m$. The m eigenvalues of modulus $\rho(A)$ are the distinct mth roots of the equation $\lambda^m - \rho^m(A) = 0$. If $m \geq 2$, there exists a permutation matrix P such that PAP^T has the cyclic form as follows:

$$PAP^T = \begin{bmatrix} 0 & A_{12} & 0 & \cdots & 0 & 0 \\ 0 & 0 & A_{23} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & A_{m-1,m} \\ A_{m1} & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

where there are zero square blocks along the main diagonal.

It is very important to point that if A in Theorem 2.3 is just nonnegative and not irreducible, then $\rho(A)$ is an eigenvalue of A and there exists a nonnegative vector ($x \geq 0,\ x \neq 0$) such that $Ax = \rho(A)x$.

Submatrices of nonnnegative matrices are also nonnegative, and hence there exists a relation between the maximal eigenvalues of a nonnegative matrix and its principle submatrices as follows:

Theorem 2.4. *[64] Let A be an $n \times n$ irrreducible nonnegative matrix, and let B be any principal submatrix of A, then $\rho(A) \geq \rho(B)$. That is, $\max_{1 \leq i \leq n} a_{ii} \leq \rho(A)$.*

A criterion for principal submatrix to be reducible is given in the next theorem:

Theorem 2.5. *[114] A nonnegative matrix A with spectral radius $\rho(A)$ is reducible iff $\rho(A)$ is an eigenvalue of a principle submatrix of A.*

Primitive matrices have been studied in [14], in particular, its shared spectral properties with positive matrices. Also, the following result introduced a criterion to check that if a matrix is primitive without calculating its eigenvalues.

Definition 2.4. *[14] Let $A = (a_{ij}) \geq 0$ be an irreducible matrix, and let m denote the number of eigenvalues of A of modulus $\rho(A)$. Then A is a primitive matrix if $m = 1$, and a cyclic matrix of index m if $m > 1$.*

According to Definition 2.4, a primitive matrix is necessarily both nonnegative and irreducible. Some other elementary properties about primitive matrices are stated in the following theorem.

Theorem 2.6. *[14] Suppose that a nonnegative matrix A is irreducible, then the following properties are equivalent:*

(1) A is called a primitive matrix. Furthermore, A^T is a primitive.

(2) $A^k > 0$ for some $k \geq 1$.

(3) A^k is irreducible for all $k \geq 1$.

(4) $\lim_{k \to \infty} \left[\rho(A)^{-1} A\right]^k$ exists.

The following theorems follow immediately from Theorem2.1.

Theorem 2.7. *[169] Suppose that $A = (a_{ij}) \geq 0$ is irreducible, and let $G(A)$ denote its associated directed graph. Then A is primitive if the largest common divisor of the lengths of all its closed paths in $G(A)$ equal one.*

The converse of Theorem 2.7 also holds.

As in [169], the spectral norm of A is defined as,

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|},$$

where $\|.\|$ is the vector norm on the vector space $\Re^n$. The relationship between the spectral radius and the spectral norm is as follows:

Proposition 2.1. *[69, 77] Let A be an $n \times n$ matrix, then its spectral radius can be described by*

$$\rho(A) = \inf_{\|.\| \in N} \|A\|,$$

where N is the set of spectral norms. In particular, for every $\epsilon > 0$, there is $\|.\|_\epsilon \in N$ such that $\|A\| \leq \rho(A) + \epsilon$.

It is generally difficult to determine precisely the spectral radius of a given matrix. Nevertheless, upper bounds can be easily found from the following theorem by Gerschgorin (1931).

Theorem 2.8. *(Gerschgorin disks) [169] Let $A = (a_{ij})$ be an arbitrary $n \times n$ complex matrix, and let*

$$\lambda_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, 1 \leq i \leq n. \tag{2.3}$$

Then, every eigenvalue λ of A lies in the union of the disks

$$|s - a_{ii}| \leq \lambda_i, 1 \leq i \leq n.$$

Based on irreducibility concept, Theorem 2.8 can be refined as follows:

Theorem 2.9. *[169] Let $A = (a_{ij})$ be an irreducible $n \times n$ complex matrix, and assume that λ, an eigenvalue of A, is a boundary point of the union of the disks $|s - a_{ii}| \leq \lambda_i$. Then, all the n circles $|s - a_{ii}| = \lambda_i$ pass through the point λ.*

In the following, we shall discuss some existing bounds for the spectral radius of nonnegative matrices. Primary results due to Frobenius [64] on the bound for the spectral radius of a nonnegative matrix is known as Frobenius' bound as follows:

Theorem 2.10. *[64] Let $A = (a_{ij})$ be an $n \times n$ nonnegative matrix with $n \geq 2$,*

and x an n-tuple positive vector. Then

$$\min_{1\le i\le n}\frac{(Ax)_i}{x_i}\le\rho(A)\le\max_{1\le i\le n}\frac{(Ax)_i}{x_i}, \tag{2.4}$$

where x_i is the i-th component of x. Moreover, if A is irreducible, then equality holds in 2.4 if and only if x is an eigenvector which corresponds to $\rho(A)$.

Another bound for the spectral radius of A is given in [26] and known as Brauer-Gentry's bound as follows:

$$\min_{i\neq j}M_A(i,j)\le\rho(A)\le\max_{i\neq j}M_A(i,j), \tag{2.5}$$

where

$$M_A(i,j)=\frac{1}{2}\left\{a_{ii}+a_{jj}+\left((a_{ii}+a_{jj})^2+4\sum_{k\neq i}|a_{ik}|\sum_{k\neq j}|a_{jk}|\right)^{\frac{1}{2}}\right\}.$$

Other results on the bound of the spectral radius of nonnegative matrices can be seen in [25, 26, 114, 128]. The following theorem is a well-known minimax characterization of the spectral radius for irreducible nonnegative matrices.

Theorem 2.11. *[46] Let $A\ge 0$ be an irreducible matrix, then*

$$\min_{x\in int(P_n)}\max_{\{i|x_i>0\}}\frac{(Ax)_i}{x_i}=\rho(A)=\max_{x\in int(P_n)}\min_{\{i|x_i>0\}}\frac{(Ax)_i}{x_i}.$$

The importance of the spectral radius $\rho(A)$ in various applications have encouraged many researchers to focus on its calculations. In some applications, the purpose is to calculate all the eigenvalues of matrices. For this kind of application, QR method [61, 62, 69, 92] is found to be useful for calculating all the eigenvalues. However, if the purpose is to only determine the largest eigenvalue of a large sparse nonnegative matrix, then QR method is required to take prohibitively expensive steps, in particular, the subject matrix will first need to be reduced to the 'almost triangular' or to Hessenberg form which can be done in $O(n^3)$ time, and this would very likely damage the sparsity of the subject matrix. In this regard, many studies indicated that a more promising technique is to use a method that computes only the largest eigenvalue (i.e. the dominant), or the first few largest eigenvalues of the matrix while preserving the sparsity of the matrix at the same time. For instance, the invariant sub-space method appears to be more promising, in particular, the Arnoldi [5], and other methods suggested

in [155, 167]. Furthermore, Saad [154] presented a variation of Arnoldi's method in order to calculate large eigenproblems. Another promising way is to further improve Collatz's result [46] for bounding the spectral radius of an irreducible nonnegative matrix. This approach received more attentions and contributions to the problem of bounding the spectral radius; see e.g. [25, 64, 128]. Furthermore, Wood and O'Neill [179] presented a variation of the method of Collatz [46] and also compared it with the Arnoldi method [5] for calculating the maximum eigenvalue of a non-negative matrix.

The power method is one of the most popular iterative methods that has been used for approximating the dominant eigenvalue (i.e. the largest eigenvalue) of a nonnegative matrix and its associated eigenvector. By relation (2.1), we have $A^{(k)}x = \lambda^{(k)}x$ for all $k \in \mathbb{N}$. Assume that $\{x_i\}$ is the set of eigenvectors of A and $\{\lambda_i\}$ the corresponding eigenvalue set such that $|\lambda_1| > |\lambda_2| \geq ... \geq |\lambda_n|$. Also, let $z^{(0)}$ be an arbitrary vector with $\|z^{(0)}\| = 1$, so, for some $c_1, ..., c_n \in \Re$, $z^{(0)}$ can be written as a linear combination of the eigenvectors of A as follows:

$$z^{(0)} = c_1x_1 + c_2x_2 + ... + c_nx_n, \quad \text{where } c_1 \neq 0.$$

Then

$$\begin{aligned}
z^{(1)} &= Az^{(0)} = A(c_1x_1 + c_2x_2 + ... + c_nx_n) \\
&= c_1Ax_1 + c_2Ax_2 + ... + c_nAx_n \\
&= c_1\lambda_1x_1 + c_2\lambda_2x_2 + ... + c_n\lambda_nx_n \quad (\text{since } Ax = \lambda x) \\
z^{(2)} &= A^2z^{(0)} = c_1\lambda_1^2x_1 + c_2\lambda_2^2x_2 + ... + c_n\lambda_n^2x_n \\
&\vdots \\
z^{(k)} &= A^kz^{(0)} = c_1\lambda_1^kx_1 + c_2\lambda_2^kx_2 + ... + c_n\lambda_n^kx_n \\
&= \lambda_1^k\left(c_1x_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k + ... + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k x_n\right), \\
&= \lambda_1^k\sum_{i=1}^{n} c_i\left(\frac{\lambda_i}{\lambda_1}\right)^k x_i.
\end{aligned}$$

Since $|\lambda_1| > |\lambda_2| \geq ... \geq |\lambda_n|$, it follows that

$$\lim_{k\to\infty}\left(\frac{\lambda_i}{\lambda_1}\right)^k = 0.$$

When k increases, $A^k z^{(0)}$ approaches $c_1 \lambda_1^k x_1$, and therefore for large values of k,

$$x_1 \approx \frac{A^k z^{(0)}}{\|A^k z^{(0)}\|}.$$

Moreover, if we apply the Rayleigh Quotient to the resulting $z^{(k)}$, we have the eigenvalue as follows:

$$\begin{aligned} S &= \lim_{k\to\infty} \frac{z^{(k)^T} A z^{(k)}}{z^{(k)^T} z^{(k)}} \\ &= \lim_{k\to\infty} \frac{\left(\lambda_1^k c_1 z_1^T\right)\left(\lambda_1^{k+1} c_1 z_1\right)}{\left(\lambda_1^k c_1 z_1^T\right)\left(\lambda_1^k c_1 z_1\right)} \\ &= \lambda_1, \end{aligned}$$

where S is the Rayleigh Quotient. Briefly, the power iterative procedure is stated as follows:

Step 0: Choose an initial vector $z^{(0)}$ with $\|z^{(0)}\| = 1$.

Step 1: For $k = 1, 2, ...$, compute iteratively

$$\begin{aligned} y &= A z^{(k-1)}, \\ z^{(k)} &= \frac{y}{\|y\|}. \end{aligned}$$

Step 2: Set $k = k - 1$ until convergence.

Wood and O'Neill [177] used two approaches to find the dominant eigenvalue as follows:

(1) Calculate $\lambda^{(k)} = u^T A z^k / u^T z^{(k)}$, such that $u^T z^{(0)} \neq 0$.

(2) The differences of these non-zero components $z^{(k)}$ and $z^{(k+1)}$ such that its ratio tends to λ_1.

The power method has a convergence rate of $|\lambda_2/\lambda_1|$, with λ_1 being the dominant eigenvalue and λ_2 being the sub-dominant eigenvalue. This iteration method is simple and efficient at approximating dominant eigenvalues. However, some practical challenges exist by consider the following cases; if the matrix A has two eigenvalues of maximum magnitude, or if $|\lambda_1| = |\lambda_2|$ and if the initial vector $z^{(0)}$ does not have a component in the direction of the dominant eigenvector, or if $c_1 = 0$. As a result, the power method will not converge to the dominant eigenvalue. To overcome these challenges, rounding errors are usually used to

ensure that the method converges to λ_1 [167]. In [177], the power method has been tested on several irreducible matrices of sizes 2×2, 3×3 and 4×4. In these tested cases, three conclusions are found, in particular, the power method does not converge at all, it converges very slowly, or a premature convergence to a wrong value. In [177], the Colltaz method has been extended to provide bounds for the limit of the convergent sequence for the spectral radius of an irreducible nonnegative matrix as follows:

Theorem 2.12. *[177] Assume that $A \geq 0$ is irreducible matrix and $z^{(0)}$ is an arbitrary positive n-dimensional column vector. Define*

$$z^{(k)} = Az^{(k-1)} = \cdots = A^{(k)}z^{(0)}, k \geq 1, \tag{2.6}$$

suppose

$$\underline{\lambda}_k = \min_{1 \leq i \leq n} \left\{ \frac{z_i^{(k+1)}}{z_i^{(k)}} \right\} \text{ and } \overline{\lambda}_k = \max_{1 \leq i \leq n} \left\{ \frac{z_i^{(k+1)}}{z_i^{(k)}} \right\}, \tag{2.7}$$

where the superscript i denotes the ith component of a vector. Then, it follows that

$$\underline{\lambda}_0 \leq \underline{\lambda}_1 \leq \underline{\lambda}_2 \leq \cdots \leq \rho(A) \leq \cdots \leq \overline{\lambda}_2 \leq \overline{\lambda}_1 \leq \overline{\lambda}_0. \tag{2.8}$$

This method is adopted from a method of Collatz [46] and is closely related to the iterative power method and the inverse power method. By applying the Collatz method for matrix A, the convergence rate is $|\lambda_1/\lambda_0|$ with λ_0 be the dominant eigenvalue and λ_1 be the subdominant eigenvalue. If we suppose that the matrix A is primitive, then a result from [177] is considered as follows:

Theorem 2.13. *[177] Both generated sequences $\{\underline{\lambda}_k\}_{k=0}^{\infty}$ and $\{\overline{\lambda}_k\}_{k=0}^{\infty}$, produced by the power method in Theorem 2.12 converge to the spectral radius of A iff the nonnegative irreducible matrix A is primitive.*

Thus, the method of Collatz is always a convergent method for the case that A is primitive. However, for an irreducible matrix, the method of Collatz is not guaranteed to converge because irreducible matrices can be either primitive or cyclic. If it is observed that A is cyclic, of course, then there exists a permutation matrix P such that PAP^T has diagonal blocks which are square zero matrices [169]. The diagonal elements of an irreducible matrix are entirely zero, and a simple positive diagonal shift can ensure that a cyclic matrix is converted to a primitive matrix as follows:

Theorem 2.14. *[177] If $A \geq 0$ is irreducible, then matrix $A + \varepsilon I$ is primitive, where $\varepsilon > 0$, and I is the identity matrix.*

Theorem 2.14 provides an irreducible matrix A with a positive spectral shift expressed as $D = A + \varepsilon I$, which is primitive. Furthermore, by using the method of Collatz on matrix D then the convergence of Theorem 2.12 is guaranteed for an irreducible matrix A. If λ_0 is the dominant eigenvalue of D, then $\lambda_0 - \varepsilon$ is the dominant eigenvalue of A and both matrices D and A have the same associated eigenvector.

In [177], the Collatz's method in Theorem 2.12 has been applied to matrix $(\varepsilon I - A)^{-1}$, and in order to ensure the convergence of this method, the matrix $(\varepsilon I - A)^{-1}$ is very required to be primitive first as follows:

Theorem 2.15. *[177] If A is an $n \times n$ nonnegative irreducible matrix with $\rho(A) < \varepsilon$, then the matrix $(\varepsilon I - A)^{-1} \geq 0$ is irreducible. Moreover, it is primitive.*

Interestingly, Theorem 2.15 has similarities to the inverse power method. It guarantees $(\varepsilon I - A)^{-1}$ to be both nonnegative and primitive with convergence ratio of $|(\varepsilon - \lambda_0)/(\varepsilon - \lambda_1)|$. It has several merits, in particular, if the matrix A is an irreducible then convergence is guaranteed. Also, it provides a reliable estimate for the error at each step. Note that this method has a merit over the method of Collatz in Theorem 2.12 as it can be used for any nonnegative irreducible matrix, not only primitive matrices. This is in fact because $\rho(A) \leq \|A\|_\infty$, by letting $\varepsilon > \|A\|_\infty$ will provide $\rho(A) < \varepsilon$.

If A is reducible matrix, then the method of Collatz is not guaranteed to be convergent. In order ensure irreducibility, a perturbation of a matrix A was suggested in [177] as follows:

$$E = \begin{pmatrix} 0 & \varepsilon & 0 \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & & \ddots & \varepsilon \\ \varepsilon & 0 & \cdots & 0 \end{pmatrix},$$

where ε is a small positive quantity, $\|E\|_2 = \varepsilon$. This will ensure that $A + E$ is irreducible matrix, and the method of Collatz can be applied. A related concern is about the impacts of the perturbation matrix on the eigenvalues. In [179], it is suggested that the spectral radius of a nonnegative matrix A is not significantly affected if the values in E are appropriately small.

2.2 Nonnegative tensors

A nonnegative tensor is a tensor with nonnegative entries, which has applications in many areas, including spectral hypergraph theory [**?**, 32], higher order Markov chains [37], numerical analysis [102], signal processing [149], game theory [3], and so on. Notably, many spectral properties and concepts pertaining to nonnegative matrices have been successfully extended to nonnegative tensors. In particular, the classical Perron-Frobenius theory has been extended to tensors; for example, see [39, 40, 63, 101, 107, 120, 148, 186]. In literature, two types of extensions exist based on irreducible nonnegative tensors [39, 186] and weakly irreducible nonnegative tensors [63, 78, 153]. Recently, in 2005, the concepts of square matrix eigenvalue have been generalized to higher order tensors by Qi [144] and Lim [101]. Tensor eigenvalues problems have received a great deal of interest in the field of numerical multi-linear algebra and various definitions of tensor eigenvalues exist. They are applicable to a wide range of potential practical applications in areas such as blind source separation [89], hypergraphs [102, 148] automatical control [144], magnetic resonance imaging [151, 160], image authenticity verification [188], higher order Markov chains [37, 82, 97], molecular conformation [53], algebraic geometry [33, 96], quantum entanglement [83, 146], Finsler geometry [9] and for more recent developments see [100, 145, 148]. This section is devoted to summarize the Perron-Frobenius theorem for tensors with nonnegative entries, as well as study some existing computing algorithms for the maximum eigenvalue and their convergence results.

Qi [144] introduced two kinds of eigenvalues (i.e., H-eigenvalues and E-eigenvalues) for symmetric tensors and described some related results and properties similar to matrix eigenvalues. Lim [101] independently developed alternate notion for tensor eigenpair, and tensor singular values and vectors by using a constrained variational approach. For $\mathcal{A} \in \mathcal{T}^{m,n}$, let $\mathcal{A}x^{m-1}$ be an n-dimensional vector in $\Re^n$ such that

$$(\mathcal{A}x^{m-1})_i = \sum_{i_2,\ldots,i_m=1}^{n} a_{i_1\, i_2\ldots i_m} x_{ii_2} x_{i_2} \cdots x_{i_m},\ i = 1, 2, ..., n. \tag{2.9}$$

Qi [144] gave the following definition which extends the idea of eigenvalues and eigenvectors of matrices to higher order tensors.

Definition 2.5. *[144] Let $\mathcal{A} \in \mathcal{T}^{m,n}$, and let $\mathbb{C}$ is a complex field, and $x^{[\alpha]} = (x_i^{\alpha})_{1\le i\le n}$. Then, a number $\lambda \in \mathbb{C}$ and a nonzero vector $x \in \mathbb{C}^n \setminus \{0\}$ are called*

eigenvalue and eigenvector of $\mathcal{A}$ respectively, associated with λ if they satisfy

$$\mathcal{A}x^{m-1} = \lambda x^{[m-1]}, \tag{2.10}$$

Also, note that the pair (λ, x) is called an H-eigenpair if they are both real.

Qi [144] introduced Definition 2.5 by considering an even m and a symmetric $\mathcal{A}$ and also gave an extension to the general case in [39]. Qi also discussed the concept of positive semi-definite symmetric tensors and showed that an even order symmetric tensor is positive semi-definite if its H-eigenvalues and (Z-eigenvalues) are positive (nonnegative), respectively. Lim [101] independently defined tensor eigenvalues and restricted x and λ to be real. Additionally, Lim studied that when m is odd, and consider $\left[x_1^{m-1}sign(x_1), ..., x_n^{m-1}sign(x_n)\right]^T$ on the right-hand side instead. Chang et al. [39] gave a slightly different definition when m is odd as follows.

Definition 2.6. *Let $\mathcal{A} \in \mathcal{T}^{m,n}$, a pair $(\lambda, x) \in \mathbb{C} \times (\mathbb{C}^n \setminus \{0\})$ is called an E-eigenvalue-E-eigenvector of $\mathcal{A}$ if*

$$\begin{cases} \mathcal{A}x^{m-1} = \lambda x \\ x^T x = 1, \end{cases} \tag{2.11}$$

The pair (λ, x) is called a Z-eigenpair if they are both real.

Here, a Z-eigenpair is called sometimes a Z_2-eigenpair and If the real eigenpair (λ, x) satisfies

$$Ax^{m-1} = \lambda x^{[m-1]}, \ \|x\|_1 = 1,$$

then it is called a Z_1-eigenpair [37]. Furthermore, Chang and Zhang [37] proved that (λ, x) is a Z_1-eigenpair if and only if $(\frac{\lambda}{\|x\|_2^{m-2}}, \frac{x}{\|x\|_2})$ is a Z_2-eigenpair, where by $\|x\|_1$ and $\|x\|_2$ we denote the l_1 and l_2-norms respectively for a vector x.

Definition 2.7. *[186] Let $\mathcal{A} \in \mathcal{T}^{m,n}$. Let $\sigma(\mathcal{A})$ denote the spectrum set of all eigenvalues of $\mathcal{A}$, and let $\rho(\mathcal{A})$ denote the spectral radius, which is the maximum modulus of the eigenvalues of $\mathcal{A}$, i.e.,*

$$\rho(\mathcal{A}) = \max\{|\lambda| : \lambda \in \sigma(\mathcal{A})\}. \tag{2.12}$$

Unlike matrices, eigenvalue problems for general higher order tensors are nonlinear and considered as NP-hard problems. Eigenvalue problems have gained great insights and special importance in recent years in the realm of numerical

multilinear algebra. Several methods have been established to compute the eigenvalues of nonnegative tensors. In particular, many researchers have centred their focus on studying the Perron-Frobenius theorem and its related algorithms for nonnegative tensors in order to find the maximum eigenvalue (i.e., the spectral radius). The following theorem summarized some basic features with respect to the eigenvalues of a symmetric tensor $\mathcal{A}$. Proof of the following theorem can be found in [144].

Theorem 2.16. *[144]*

(1) If $\lambda \in \mathbb{C}$ is a root of the characteristic polynomial $P_{\mathcal{A}}(\lambda) = det(\mathcal{A} - \lambda\mathcal{I})$, then it is an eigenvalue of $\mathcal{A}$. An $\mathcal{I} = (I_{i_1 \ldots i_m})$ is the identity tensor such that

$$I_{i_1 i_2 \ldots i_m} = \begin{cases} 1 & \text{if } i_1 = i_2 = \ldots = i_m, \\ 0 & \text{otherwise.} \end{cases}$$

(2) The tensor $\mathcal{A}$ has $n(m-1)^{n-1}$ eigenvalues and their product is equal to the determinant of $\mathcal{A}$.

(3) The sum of all the eigenvalues of $\mathcal{A}$ is $(m-1)^{n-1}tr(\mathcal{A})$, where $tr(\mathcal{A})$ denote the sum of the diagonal elements of $\mathcal{A}$.

(4) All the eigenvalues λ of $\mathcal{A}$ lie in the union of the disks:

$$|\lambda - a_{i,i,\ldots,i}| \leq \sum \{|a_{i,i_2,\ldots,i_m}| : 1 \leq i_2, \ldots, i_m \leq n, \{i_2, \ldots, i_m\} \neq \{i, \ldots, i\}\},$$

for $i = 1, \ldots, n$.

(5) Furthermore, if $\mathcal{A}$ is an even order symmetric tensor, then

(i) H-eigenvalues always exist and $\mathcal{A}$ is said to be positive definite (positive semidefinite) if and only if all of its H-eigenvalues are positive (nonnegative).

(ii) Z-eigenvalues always exist and $\mathcal{A}$ is said to be positive definite (positive semidefinite) if and only if all of its Z-eigenvalues are positive (nonnegative).

There are two types of irreducibility for nonnegative tensors, namely, irreducible nonnegative tensor [39, 101] and weakly irreducible nonnegative tensor [63]. First, let us recall the notion of irreducibility for higher order tensors as follows.

Definition 2.8. *[39] Suppose that $\mathcal{A} \in \mathcal{T}^{m,n}$, then $\mathcal{A}$ is said to be reducible if there exists a non-empty proper index subset $I \subset \{1, ..., n\}$ such that*

$$a_{i_1 \ldots i_m} = 0, \ \forall i_1 \in I, \ \forall i_2, ..., i_m \notin I.$$

A tensor $\mathcal{A}$ is called irreducible if it is not reducible.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed graph that is consisting of a vertex set $\mathcal{V}$ and an edge set $\mathcal{E}$. Then, $\mathcal{G}$ is said to be strongly connected if for any ordered pair of vertices i and j, there exists a directed path from i to j. Let $\mathcal{A} = (a_{i_1,...,i_m}) \in \mathcal{T}_+^{[m,n]}$, then its associated directed graph is $\mathcal{G}(\mathcal{A}) = (\mathcal{V}, \mathcal{E})$, with vertices $\mathcal{V} = \{1, 2, ..., n\}$ and a directed edge $(i, j) \in \mathcal{E}$ if and only if there exists indices $\{i_2, ..., i_m\}$ such that $j \in \{i_2, ..., i_m\}$ and $a_{ii_2,...,i_m} \neq 0$, i.e., $\sum_{j \in \{i_2,...,i_m\}} |a_{ii_2,...,i_m}| > 0$. Friedland et al. [63] used the feature of strongly connected graph to introduce weakly irreducible tensor as follows.

Definition 2.9. *[63] A tensor $\mathcal{A} \in \mathcal{T}^{m,n}$ is weakly irreducible if its associated graph $\mathcal{G}(\mathcal{A})$ is strongly connected. Otherwise, $\mathcal{A}$ is called weakly reducible. Furthermore, $\mathcal{A}$ is weakly primitive if $\mathcal{G}(\mathcal{A})$ is strongly connected and its greatest common divisor of the length of its circuits equals one.*

In [63], it has been shown that an irreducible nonnegative tensor $\mathcal{A}$ is weakly irreducible. Moreover, if a tensor is of order $m = 2$, then $\mathcal{A}$ is irreducible iff $\mathcal{A}$ is weakly irreducible.

Lemma 2.1. *If $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is irreducible. Then*

(1) [39] $\sum_{i_2,...,i_m}^{n} a_{ii_2,...,i_m} > 0 \quad \forall \quad 1 \leq i \leq n,$

(2) [120] for any positive vector $x > 0, x \in \Re^n$, $\mathcal{A}x^{m-1}$ is a positive vector; i.e., $\mathcal{A}x^{m-1} > 0$.

Lemma 2.2. *[120] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$, with two column vectors x and y are both nonnegative. Then, the following hold:*

(1) If $x \geq y$, then $\mathcal{A}x^{m-1} \geq \mathcal{A}y^{m-1}$;

(2) $\mathcal{A}(tx)^{m-1} = t^{m-1}\mathcal{A}x^{m-1}$, where $t > 0$.

Furthermore, Chang et al. [39] defined the majorization concept of nonnegative tensors as follows.

Definition 2.10. *[39] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$. Its associated nonnegative matrix $M(\mathcal{A})$ is called the majorization, which is defined as,*

$$[M(\mathcal{A})]_{ij} = a_{ij\ldots j}, \forall\ i, j \in \{1, ..., n\}.$$

Note that if $[M(\mathcal{A})]_{ij} > 0$ for all $i \neq j$, then tensor $\mathcal{A}$ is called weakly positive.

Proposition 2.2. *[162] Suppose that $\mathcal{A} \in \mathcal{T}_+^{m,n}$, then it is a primitive tensor if there exists some positive integer k such that $\mathcal{A}^k$ is essentially positive. In addition, the smallest k is the primitive degree of $\mathcal{A}$. Also, if $\mathcal{A}^k$ is essentially positive, then $\mathcal{A}^{k+1}$ is also essentially positive.*

Remark 2.1. *[162] A tensor $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is a primitive if and only if there exists some positive integer k such that $M(\mathcal{A}^k) > 0$.*

Corollary 2.2. *[63] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is irreducible, then it is weakly irreducible.*

Clearly, irreducible tensors contains weakly irreducible tensor, but the converse may not hold.

Theorem 2.17. *[40] If $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is irreducible with $a_{ii\ldots i} > 0,\ \forall\ 1 \leq i \leq n$, then $\mathcal{A}$ is a primitive tensor.*

In [40], it has been shown that a primitive nonnegative tensor $\mathcal{A}$ is an irreducible nonnegative tensor, but not conversely.

Corollary 2.3. *[40] If $\mathcal{A} \in \mathcal{T}_+^{m,n}$, is essentially positive, then it is a primitive tensor.*

Proposition 2.3. *[79] Suppose that a nonnegative tensor $\mathcal{A}$ is irreducible. Then, $\mathcal{A} + \mathcal{I}$ is weakly primitive.*

Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$. Then let $\{\mathcal{A}^{(k)}x\}$, be a sequence that is defined for any vector $x \in int(P_n)$, as follows:

$$\begin{aligned}
y^{(1)} &= \left(\mathcal{A}^{(1)}x\right)^{\left[\frac{1}{m-1}\right]},\ \mathcal{A}^{(1)}x = \mathcal{A}(x)^{m-1},\\
y^{(2)} &= \left(\mathcal{A}^{(2)}x\right)^{\left[\frac{1}{m-1}\right]},\ \mathcal{A}^{(2)}x = \mathcal{A}\left(y^{(1)}\right)^{m-1},\\
&\vdots\\
y^{(k)} &= \left(\mathcal{A}^{(k-1)}x\right)^{\left[\frac{1}{m-1}\right]},\ \mathcal{A}^{(k+1)}x = \mathcal{A}\left(y^{(k)}\right)^{m-1},\quad k \geq 1.
\end{aligned}$$

Definition 2.11. *For a tensor* $\mathcal{A} = (a_{i_1\ldots i_m})$*, we call* $\mathcal{B} = (b_{i_1\ldots i_m})$ *the diagonal of* $\mathcal{A}$*, denoted by* $\mathcal{B} = \text{diag}(\mathcal{A})$*, if*

$$b_{i_1\ldots i_m} = \begin{cases} a_{i\ldots i} & \text{if } i_1 = \ldots i_m = i, i = 1, \ldots, n, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 2.18. *[107] Suppose that* $\mathcal{A} \in \mathcal{T}_+^{m,n}$ *is irreducible. If* $\text{diag}(\mathcal{A}) > 0$*, then* $\mathcal{A}$ *is primitive.*

Theorem 2.19. *[40] Let* $\mathcal{A} \in \mathcal{T}_+^{m,n}$ *is primitive. If* λ *is an eigenvalue of* $\mathcal{A}$ *such that* $|\lambda| = \rho(\mathcal{A})$*, then* $\lambda = \rho(\mathcal{A})$*, that is, its cyclic index,* $k = 1$*.*

2.2.1 The Perron-Frobenius Theorem

The Perron-Frobenius theory for nonnegative tensors is well developed as many important spectral properties have been extended from nonnegative matrices to nonnegative tensors. It plays a key role to derive many spectral characterizations of nonnegative tensor that are essential for eigenvalue computations and its convergence conditions as well as establishing the relations among nonnegative tensor classes. In [39], the Perron-Frobenius theorem has been generalized from nonnegative matrices to irreducible nonnegative tensors. Later in [186], the weak Perron- Frobenius theorem is extended to general nonnegative tensors, and the spectral radius of a nonnegative tensor is shown as an eigenvalue. In [63], the Perron-Frobenius theorem of nonnegative tensors is shown to have a very close connection with the Perron-Frobenius theorem for homogeneous monotone maps. Furthermore, in [63], the Perron-Frobenius theorem has been established for weakly irreducible nonnegative tensors. Chang, Pearson, and Zhang [39] studied and established some theoretical properties based on the Perron-Frobenius theorem for eigenvalues of nonnegative tensors. In the following, we recall the weak form of the Perron-Frobenius theorem for tensors.

Theorem 2.20. *[39] If* $\mathcal{A} \in \mathcal{T}_+^{m,n}$*, then there exists* $\lambda_0 \geq 0$ *and a nonnegative vector* $x_0 \neq 0$ *such that*

$$\mathcal{A}x^{m-1} = \lambda_0 x_0^{[m-1]}. \tag{2.13}$$

The pair (λ_0, x_0) in Theorem 2.20 is a real number and real vector. Further in [39], the strong form of the Perron-Frobenius theorem is introduced as following:

Theorem 2.21. *[39] If* $\mathcal{A} \in \mathcal{T}_+^{m,n}$ *is an irreducible tensor, then there exists* $\lambda_0 > 0$ *and a nonnegative vector* $x_0 > 0, x_0 \in \Re^n$ *such that*

$$\mathcal{A}x_0^{m-1} = \lambda_0 x_0^{[m-1]}. \tag{2.14}$$

Furthermore, it is important to note that if λ is an eigenvalue with nonnegative eigenvector, then $\lambda = \lambda_0$. Also, If λ is an eigenvalue of tensor $\mathcal{A}$, then $|\lambda| \leq \lambda_0$. However for tensors in general, we note that such λ_0 is not necessarily a simple eigenvalue. Now, let us introduce the geometric multiplicity of an eigenvalue of $\mathcal{A}$ as follows:

Definition 2.12. *[40] Suppose that λ is an eigenvalue of $\mathcal{A}$. Then, λ has geometric multiplicity q if the largest number of linearly independent eigenvectors corresponding to λ equals q, and it is called simple if its geometrically multiplicity is equal to one.*

The minimax characterization in Theorem 2.11, for the spectral radius of irreducible nonnegative matrices has been extended to irreducible nonnegative tensors as follows:

Theorem 2.22. *[39] Suppose that $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is irreducible tensor. Then, the unique positive eigenvalue λ_0 such that there is a positive vector x satisfies:*

$$\min_{x \in int(P_n)} \max_{x_i > 0} \frac{(\mathcal{A}x^{m-1})_i}{x_i^{m-1}} = \lambda_0 = \max_{x \in int(P_n)} \min_{x_i > 0} \frac{(\mathcal{A}x^{m-1})_i}{x_i^{m-1}}, \tag{2.15}$$

Yang and Yang [186] proved some spectral results for nonnegative tensors as follows.

Lemma 2.3. *[186] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$, Then*

$$\min_{1 \leq i \leq n} \sum_{i_2,\ldots,i_m=1} a_{ii_2\ldots i_m} \leq \rho(\mathcal{A}) \leq \max_{1 \leq i \leq n} \sum_{i_2,\ldots,i_m=1} a_{ii_2\ldots i_m}. \tag{2.16}$$

Furthermore, it has been asserted that the spectral radius of a nonnegative tensor is an eigenvalue as follows.

Lemma 2.4. *Suppose $\mathcal{A} \in \mathcal{T}_+^{m,n}$, then $\rho(\mathcal{A})$ is an eigenvalue of $\mathcal{A}$ with a nonzero nonnegative eigenvector. Moreover, for any $x \in int(P_n)$, we have*

$$\min_{1 \leq i \leq n} \frac{(\mathcal{A}x^{m-1})_i}{x_i^{m-1}} \leq \rho(\mathcal{A}) \leq \max_{1 \leq i \leq n} \frac{(\mathcal{A}x^{m-1})_i}{x_i^{m-1}}. \tag{2.17}$$

Theorem 2.23. *[186] Suppose that $\mathcal{A} \in \mathcal{T}_+^{m,n}$ where $m, n \geq 2$. Then*

$$\rho(\mathcal{A}) = \max_{x \in P_n, x \neq 0} \min_{x_i > 0} \frac{(\mathcal{A}x^{m-1})}{x_i^{m-1}}. \tag{2.18}$$

The following theorem summarizes some important and basic consequences of the Perron-Frobenius theorem.

Theorem 2.24. *Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$. Then, the following facts hold:*

- *The spectral radius $\rho(\mathcal{A})$ is an eigenvalue with a nonnegative eigenvector, [186];*
- *If $\mathcal{A}$ is strictly nonnegative tensor, then the spectral radius of $\mathcal{A}$ is positive, i.e. $\rho(\mathcal{A}) > 0$, [78];*
- *If $\mathcal{A}$ is weakly irreducible, then $\rho(\mathcal{A})$ has a unique positive eigenvector, [63];*
- *If $\mathcal{A}$ is irreducible and if λ is an eigenvalue with a nonnegative eigenvector, then $\lambda = \rho(\mathcal{A})$, [39];*
- *If $\mathcal{A}$ is irreducible, and $\mathcal{A}$ has k distinct eigenvalues of modulus $\rho(\mathcal{A})$, then the eigenvalues are $\rho(\mathcal{A})e^{2\pi i.j/k}$ $\forall j = 0, 1, \ldots, k-1$, and $i^2 = -1$, [186];*
- *If furthermore $\mathcal{A}$ is primitive, then $k = 1$, [40];*
- *If $\mathcal{A}$ is further essentially positive and m is even, $\rho(\mathcal{A})$ is real geometrically simple, [136];*
- *For a nonempty subset $I \subset \{1, ..., n\}$, we let $\mathcal{A}_I$ be the induced tensor such that $\{\mathcal{A}_{i_1,\ldots,i_m} | i_1, ..., i_m \in I\}$, then $\rho(\mathcal{A}_I) \leq \rho(\mathcal{A})$ [78].*

CHAPTER 3

Algorithms for computing the spectral radius of nonnegative tensors

In this chapter, we first present the state-of-art of the spectral theory of nonnegative tensors. This includes the NQZ algorithm [120], the LZI algorithm [107], and the ZQW algorithm [199]. Then, we propose a fast algorithm for computing the spectral radii of symmetric nonnegative tensors. In particular, by this proposed algorithm we are able to obtain the spectral radii of weakly reducible symmetric nonnegative tensors without requiring the partition of the tensors. As we know, it is very costly to determine the partition for large-size weakly reducible tensors. Numerical results are reported to show that the proposed algorithm is efficient and also it is able to compute the spectral radii of large-size tensors. As an application, we present an algorithm for testing the positive definiteness of Z-tensors. By this algorithm, it is guaranteed to determine the positive definiteness for any Z-tensor.

3.1 Introduction

In this chapter, we consider an m-order n-dimensional tensor $\mathcal{A}$ consisting of n^m entries in $\Re$ defined in (1.1). For $\mathcal{A} \in \mathcal{T}^{m,n}$, Qi [144] introduced the nonnegative tensor eigenvalue problem defined in (2.5) when m is even and $\mathcal{C}$ is symmetric. Independently, Lim [101] defined the l_k-eigenvalue which is equivalent to Definition (2.5).

As we know, eigenvectors of matrices are central to linear algebra. Eigenvectors of tensors are a natural generalization of eigenvectors of matrices, and have applications in higher order Markov chains, spectral hypergraph theory, and the quantum entanglement; See the survey paper [35] for details. Recently,

a number of algorithms for computing tensor eigenpairs have been proposed; See [35, 44, 78, 90, 122, 192, 199]. In particular, Ng, Qi and Zhou [120] proposed a power type method for computing the largest eigenvalue of non-negative tensors. The convergence results of this algorithm have been established in [40, 63, 107] for weakly irreducible tensor. For weakly reducible tensors, the proposed algorithm in [120] may not be convergent. It is shown in [78, 199] that weakly reducible tensors can be decomposed into some weakly irreducible tensors. Based on this result, some algorithms for the largest eigenvalue of weakly reducible tensors have been presented in [78,199]. For the algorithms in [78,199], it is needed to compute the partition for reducible tensors. As we know, it is very costly to compute the partition of a weakly reducible tensor $\mathcal{A} \in \mathcal{T}_+^{m,n}$ when n and m are large.

In this chapter, we propose an algorithm for computing the spectral radii of symmetric nonnegative tensors. This algorithm is a modified version of the algorithm in [120]. In particular, this algorithm has the following nice properties: (1) It does not require the partition for weakly reducible tensors; (2) It is convergent for any symmetric nonnegative tensor; and (3) It is able to compute the spectral radius for large-size tensors. As an application, an algorithm for testing the positive definiteness of Z-tensors is presented.

3.2 Computing methods for the maximum H-eigenvalue of a nonnegative tensor

3.2.1 Existing Algorithms for weakly irreducible tensor

In fact, there are three existing ways for calculating the maximum eigenvalue of weakly irreducible tensors as follows.

1. Power-type algorithms
   In [39], the classical Perron-Frobenius theorem for nonnegative tensors was developed. Furthermore, by using Theorems 2.21 and 2.22, Ng, Qi, and Zhou [120] developed an efficient iterative power method called the NQZ method in order to calculate the maximum eigenvalue of an irreducible nonnegative tensor, which is an extension of Collatz's method for finding the spectral radius of an irreducible nonnegative matrix. The NQZ algorithm proceeds as follows:

Algorithm 3.1. ***NQZ Method [120]***
Given a tensor $\mathcal{A} \in \mathcal{T}_+^{m,n}$, not necessarily be symmetric.

Step 1. *Initialization: choose $x^{(0)} \in int(P_n)$. Let $y^{(0)} = \mathcal{A}(x^{(0)})^{m-1}$. Set $k := 0$.*

Step 2. *Compute*

$$y^{(k+1)} := \mathcal{A}\left(x^{(k+1)}\right)^{m-1},$$
$$\bar{\lambda}_{k+1} := \max_{1\le i\le n} \frac{\left(y^{(k+1)}\right)_i}{\left(x_i^{(k+1)}\right)^{m-1}},$$
$$\underline{\lambda}_{k+1} := \min_{1\le i\le n} \frac{\left(y^{(k+1)}\right)_i}{\left(x_i^{(k+1)}\right)^{m-1}}.$$

Step 3. *If $\bar{\lambda}_{k+1} = \underline{\lambda}_{k+1}$, then the iteration stops. Otherwise, calculate*

$$x^{(k+1)} := \frac{\left(y^{(k)}\right)^{\left[\frac{1}{m-1}\right]}}{\left\|\left(y^{(k)}\right)^{\left[\frac{1}{m-1}\right]}\right\|},$$

replace k by $k+1$, and go to Step 2.

Next, the convergence of Algorithm 3.1 was established in [120] for irreducible tensor as follows.

Theorem 3.1. *[120] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$ be irreducible and suppose that λ_0 is the unique positive eigenvalue with a corresponding nonnegative eigenvector. Then, Algorithm 3.1 produced the value of λ_0 in a finite number of steps or generate two sequences $\{\underline{\lambda}_k\}$ and $\{\bar{\lambda}_k\}$. Furthermore, let $\underline{\lambda} = \lim_{k\to+\infty} \underline{\lambda}$ and $\bar{\lambda} = \lim_{k\to+\infty} \bar{\lambda}$. Then, $\underline{\lambda}$ and $\bar{\lambda}$ are the lower and upper bounds of λ_0 respectively. If $\underline{\lambda} = \bar{\lambda}$, then $\lambda_0 = \underline{\lambda} = \bar{\lambda}$.*

From Theorem 3.1, we can assure that the limit exists as the sequence $\{\underline{\lambda}_k\}$ is monotonically non-decreasing and with an upper bound. Hence, the sequence $\{x^{(k)}\}$ converges to a vector x. Note that Algorithm 3.1 can only generate two convergent sequences $\{\bar{\lambda}_k\}$ and $\{\underline{\lambda}_k\}$, both of which are not guaranteed to converge to λ_0 if the underlying tensor $\mathcal{A}$ is an irreducible. In [120], an example of irreducible tensor was given to demonstrate this case. However, in [40], the convergence of Algorithm 3.1 was established if the underlying tensor $\mathcal{A}$ is primitive and an example was given to show this for primitive tensor but not for essentially positive tensor.

For $\mathcal{A} \in \mathcal{T}^{m,n}$, its associated nonlinear map $T_{\mathcal{A}} : P_n \to P_n$ was defined in [40] as

$$T_{\mathcal{A}}(x) := \left(\mathcal{A}x^{m-1}\right)^{\left[\frac{1}{(m-1)}\right]} \text{ for any } x \in P_n,$$

then, the following statements hold.

Theorem 3.2. *[40, 107]*

(1) For all $k \in \mathbb{N}, \bar{\lambda}_{k+1} \leq \bar{\lambda}_k$ and $\underline{\lambda}_{k+1} \geq \underline{\lambda}_k$.

(2) If $\mathcal{A}$ is irreducible, then $\underline{\lambda}_{k+1} \nearrow \underline{\lambda}, \bar{\lambda}_k \searrow \bar{\lambda}$, and $\underline{\lambda} \leq \rho(\mathcal{A}) \leq \bar{\lambda}$.

(3) From $\|x^{(k)}\| = 1$, there exists a subsequence $x^{(k_j)} \to x^*$ with $\|x^*\| = 1$.

(4) $(\underline{\lambda}_k)^{\left[\frac{1}{(m-1)}\right]} x^{(k)} \leq y^{(k)} = T_{\mathcal{A}} x^{(k)} \leq \left(\bar{\lambda}_k\right)^{\left[\frac{1}{(m-1)}\right]} x^{(k)}$; hence, $\underline{\lambda}^{\frac{1}{(m-1)}} x^* \leq T_{\mathcal{A}} x^* \leq \bar{\lambda}^{\frac{1}{(m-1)}} x^*$.

(5) For all $k \in \mathbb{N}$, there exists $1 \leq i_0 \leq n$ such that $\left(T_{\mathcal{A}}^{k+1} x^*\right)_{i_0} = \underline{\lambda}^{\frac{1}{(m-1)}} \left(T_{\mathcal{A}}^{k} x^*\right)_{i_0}$.

The proofs of Theorem 3.2 are given in [40, 107].

The following results find the convergence of the NQZ algorithm if the underlying tensor $\mathcal{A}$ is primitive or weakly primitive tensor.

Lemma 3.1. *[40] If $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is primitive, and $\{\bar{\lambda}_k\}$ is monotonically decreasing and $\{\underline{\lambda}_k\}$ is monotonically increasing. Then, both of the sequences converge to $\rho(\mathcal{A})$.*

Theorem 3.3. *[40] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$ be a primitive and the sequences $\{\underline{\lambda}_k, \bar{\lambda}_k, x^{(k)}\}$ are generated by Algorithm 3.1. Then $\{\underline{\lambda}_k\}$ and $\{\bar{\lambda}_k\}$ converge to $\rho(\mathcal{A})$, and $\{x^{(k)}\}$ converges to x^*, satisfying that $\|x^*\| = 1$, where x^* is the unique positive eigenvector corresponding to $\rho(\mathcal{A})$.*

Furthermore, Friedland, Gaubert, and Han [63] established the convergence of Algorithm 3.1 for weakly primitive tensor in the following result.

Theorem 3.4. *[63] Suppose that $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is a weakly primitive. Then the NQZ algorithm generates the sequences $\{\underline{\lambda}_k, \bar{\lambda}_k, x^{(k)}\}$ such that $\{\underline{\lambda}_k\}$ and $\{\bar{\lambda}_k\}$ converge to $\rho(\mathcal{A})$, and $\{x^{(k)}\}$ converges to x^* satisfying $\|x^*\| = 1$, where x^* is the unique positive eigenvector corresponding to $\rho(\mathcal{A})$.*

Corollary 3.1. *[40] If $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is essentially positive, then Algorithm 3.1 converges.*

Corollary 3.2. *[40] Suppose $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is irreducible. If $\mathcal{A} = \mathcal{B} + \mathcal{I}$, then Algorithm 3.1 converges.*

Zhang and Qi [189] established the linear convergence rate of Algorithm 3.1 for essentially positive tensors as follows.

Theorem 3.5. *[189] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is essentially positive, then*

$$\bar{\lambda}_k - \underline{\lambda}_k \leq \alpha \left(\bar{\lambda}_{k-1} - \underline{\lambda}_{k-1} \right), \quad k = 1, 2, \cdots,$$

where

$$\alpha = 1 - \frac{\beta}{\bar{R}} \in (0, 1),$$
$$\beta = \min_{i,j \in \{1,2,\ldots,n\}} a_{ij\ldots j},$$
$$\bar{R} = \max_{1 \leq i \leq n} R_i$$

and

$$R_i = \sum_{i_2,\ldots,i_m=1}^{n} a_{ii_2\ldots i_m}.$$

According to [107, 186], Algorithm 3.1 is not always convergent for some irreducible nonnegative tensors and hence tensor $\mathcal{A}$ can be modified by adding $\rho\mathcal{I}$ to tensor $\mathcal{A}$. By Theorems 2.17, 2.21, and Corollary 3 [144], we have the following Theorem.

Theorem 3.6. *[107] Let $\mathcal{A} \in \mathcal{T}_+^{m,n}$ be an irreducible tensor and suppose $\mathcal{B} = \mathcal{A} + \rho\mathcal{I}$ for any $\rho > 0$. Then,*

(i) $\mathcal{B}$ is primitive tensor.

(ii) If λ is the maximum eigenvalue of $\mathcal{B}$, then $\lambda - \rho$ is the maximum eigenvalue of $\mathcal{A}$.

Algorithm 3.1 is efficient but not always guaranteed to converge for some irreducible nonnegative tensors. A counterexample was given in [120] to show that Algorithm 3.1 may not converge. Later, Liu et al. [107] modified Algorithm 3.1 such that it is always a convergent method for solving the maximum eigenvalue of irreducible tensor. We give the details of the refined algorithm as follows.

Algorithm 3.2. ***LZI Method [107]***

Step 1. *Choose* $x^{(1)} \in int(P_n)$, *and suppose that* $\mathcal{B} = \mathcal{A} + \rho\mathcal{I},\ \rho > 0$.

Step 2. *For* $k = 1, 2, ...,$ *calculate*

$$y^{(k)} = \mathcal{B}\left(x^{(k)}\right)^{m-1},$$
$$\bar{\lambda}_k = \max_{x_i^{(k)}>0} \frac{\left(y^{(k)}\right)_i}{\left(x_i^{(k)}\right)^{m-1}},$$
$$\underline{\lambda}_k = \min_{x_i^{(k)}>0} \frac{\left(y^{(k)}\right)_i}{\left(x_i^{(k)}\right)^{m-1}}.$$

Step 3. *If* $\bar{\lambda}_k = \underline{\lambda}_k$, *then let* $\lambda = \bar{\lambda}_k$ *and stop. Otherwise, calculate*

$$x^{(k+1)} = \frac{\left(y^{(k)}\right)^{\left[\frac{1}{m-1}\right]}}{\left\|\left(y^{(k)}\right)^{\left[\frac{1}{m-1}\right]}\right\|},$$

replace k *by* $k+1$, *and go to step 2.*

Furthermore, it was shown in [107] that the two sequences $\{\underline{\lambda}_k\}$ and $\{\bar{\lambda}_k\}$ converge to λ in the next theorem as follows.

Theorem 3.7. *[107] Assume that* $\mathcal{A} \in \mathcal{T}_+^{m,n}$ *is irreducible, and let* $\mathcal{B} = \mathcal{A} + \rho\mathcal{I}$, *where* $\rho > 0$. *Then, Algorithm 3.2 produces a value of the spectral radius* λ, *or generates the sequence* $\{\underline{\lambda}_k, \bar{\lambda}_k\}$, *which converges to* λ, *where* λ *is the maximum eigenvalue of* $\mathcal{B}$. *Moreover,* $\lambda - \rho$ *is the dominant eigenvalue of* $\mathcal{A}$.

In addition, Zhang et al. [190] established the linear convergence rate of Algorithm 3.2 under the weakly positive condition as follows.

Theorem 3.8. *[190] Suppose that* $\mathcal{A} \in \mathcal{T}_+^{m,n}$, *and choose* $x^{(0)}$ *as the vector of ones in Algorithm 3.2, then it generates two sequences* $\{\underline{\lambda}_k\}$ *and* $\{\bar{\lambda}_k\}$. *If*

$\mathcal{A}$ is weakly positive, then

$$\bar{\lambda}_{k+1} - \underline{\lambda}_{k+1} \leq \alpha(\bar{\lambda}_k - \underline{\lambda}_k), \quad k = 1, 2, \cdots,$$

where

$$\alpha = 1 - \frac{\beta}{\bar{\mu}} \in (0,1), \quad \bar{\mu} = \rho + \max_{1\leq i\leq n} \mu_i,$$
$$\beta = \min\left\{ \min_{i,j\in\{1,2,\ldots,n\}, i\neq j} A_{ij\ldots j},\ \rho + \min_{i\leq i\leq n} A_{ii\ldots i} \right\},$$
$$\mu_i = \sum_{i_2,\ldots,i_m=1}^{n} A_{ii_2\ldots i_m}.$$

Furthermore, Zhou et al. [199] designed a variant of Algorithm 3.1 such that this algorithm is convergent for all weakly irreducible nonnnegative tensors. For any nonnegative column vector $x \in \Re^n$, we defined $\phi : P_n \to P_1$ by

$$\phi(x) = \sum_{i=1}^{n} x_i. \tag{3.1}$$

Algorithm 3.3. ***ZQW Method [199]***
Step 1. *Choose $x^{(1)} \in int(P_n)$. Let $\mathcal{B} = \mathcal{A} + \alpha\mathcal{I},\ \alpha > 0$ and set $k = 1$.*
Step 2. *For $k = 2, 3, ...,$ compute*

$$y^{(k)} = \mathcal{B}\left(x^{(k)}\right)^{m-1},$$
$$\bar{\lambda}_k = \max_{x_i^{(k)}>0} \frac{\left(y^{(k)}\right)_i}{\left(x_i^{(k)}\right)^{m-1}},$$
$$\underline{\lambda}_k = \min_{x_i^{(k)}>0} \frac{\left(y^{(k)}\right)_i}{\left(x_i^{(k)}\right)^{m-1}}.$$

Step 3. *If $\bar{\lambda}_k = \underline{\lambda}_k$, then let $\lambda = \bar{\lambda}_k$ and stop. Otherwise, compute*

$$x^{(k+1)} = \frac{\left(y^{(k)}\right)^{\left[\frac{1}{m-1}\right]}}{\phi\left(\left(y^{(k)}\right)^{\left[\frac{1}{m-1}\right]}\right)},$$

replace k by $k+1$, and go to step 2.

Theorem 3.9. *[198] Suppose that $\mathcal{A} \in \mathcal{T}_+^{m,n}$ is weakly irreducible. Assume tensor $\mathcal{B} = \mathcal{A} + \mathcal{I}$, then by Algorithm 3.3, the maximum eigenvalue λ of $\mathcal{B}$ and the corresponding eigenvector u are produced, or three conver-*

gent sequences $\{\underline{\lambda}_k\}, \{\bar{\lambda}_k\}$ *and* $\{x^{(k)}\}$ *are generated such that* $\lim_{k\to\infty} \underline{\lambda}_k = \lim_{k\to\infty} \bar{\lambda}_k = \lambda$, $\lim_{k\to\infty} x^{(k)} = u$. *Moreover, it follows that* $\lambda - 1$ *is the maximum eigenvalue of* $\mathcal{A}$ *associated with the eigenvector* u.

Define

$$F(x) = \mathcal{B}x^{m-1}, \tag{3.2}$$

$$H(x) = \frac{F(x)^{\left[\frac{1}{m-1}\right]}}{\phi\left(F(x)^{\left[\frac{1}{m-1}\right]}\right)}. \tag{3.3}$$

Clearly, it can be observed that the sequence $\{x^{(k)}\}$ in Theorem 3.9 can be generated by

$$x^{(k+1)} = H\left(x^{(k)}\right),\ k = 1, 2, ..., \tag{3.4}$$

and $\phi\left(x^{(k)}\right) = 1 \quad \forall\ k = 1, 2, \cdots$.

The following theorem was established in [198] to show that the ZQW algorithm is Q-linear convergent under the weakly irreducibility conditions.

Theorem 3.10. *[198] Suppose that* $\mathcal{A}, \mathcal{B}$, *and* $\{x^{(k)}\}$ *defined as in Theorem 3.9. As the sequence* $\{x^{(k)}\}$ *converges to* u, *then its convergence rate is Q-linear, that is, there exists a vector norm* $\|.\|$ *such that*

$$\limsup_{k\to\infty} \frac{\|x^{(k+1)} - u\|}{\|x^{(k)} - u\|} < 1. \tag{3.5}$$

In the light of Theorems 3.5 and 3.8, the linear convergence rate is established for the sequences $\{\underline{\lambda}_k\}$ and $\{\bar{\lambda}_k\}$, while Theorem 3.10 gives the Q-linear convergence rate for the sequence $\{x^{(k)}\}$ for weakly irreducible tensors. Furthermore, in [78], it has been shown that Algorithm 3.2 is globally R-linearly convergent under weakly irreducibility conditions.

2. Newton-type algorithms

In 1971, Noda [125] developed a method called the Noda iteration which has the feature of preserving positivity. This method can compute the maximal eigenvalue of an irreducible nonnegative matrix, and its convergence is proved to be quadratic [56]. For irreducible nonnegative second order tensors (i.e., matrices), there are some existing methods [56, 85, 125] such that they maintain faster convergence and positivity preserving for computing the Perron pairs of tensor $\mathcal{A}$. The Newton type iteration algorithm is one of

most important computational techniques that has been used to calculate the eigenpairs of weakly irreducible tensors. In particular, in [121], a local quadratic convergent method was presented to compute the eigenpair of irreducible tensor. This method does not preserve positivity, and hence the line search with additional conditions has been applied to achieve global convergence, but still not preserve positivity. Recently, Liu et al. [105, 106] combined Newton's method [121] and Noda's iteration [125] to present a method called Newton-Noda iteration (i.e., NNI) which is a positivity preserving method for nonnegative third order tensors. The NNI iteration is an inverse method with variable shifts that naturally maintains the positivity of approximate eigenvectors at all iterations. The advantage of the NNI method is that it is always quadratically convergent for any starting vector, and able to find eigenpair of a tensor. Moreover, it always shows global convergence for any positive initial vector and can determine the desired eigenpair [133]. This method shows positivity preserving and requires the selection of a positive parameter θ_k in the kth iteration. In addition, a halving procedure is designed in [106] in order to determine θ_k, starting with $\theta_k = 1$ for each k. Then, such a sequence $(\bar{\lambda}_k, x^{(k)})$ produced by NNI method converges monotonically to $(\rho(\mathcal{A}), x^*)$ as long as the sequence $\{\theta_k\}$ is bounded below by a fixed positive constant. Thus, it always produces a monotonically non-increasing sequence of approximate eigenvalues whose convergence to the maximum eigenvalue is guaranteed. In [106], a comparison was made between the NNI method and the NQZ method by using flop counts and found that the NNI method has less iterations and faster convergence.

3. Homotopy methods

   Among the various computational iterative methods for solving tensor eigenpair problem of an irreducible nonnegative tensor, one can use the classical approach called the homotopy continuation type algorithm in order to calculate the maximum eigenvalue and its corresponding eigenvector. In particular, in [44], a homotopy continuation method was used to calculate all real eigenpairs of a general symmetric tensors. It is also able to find all equivalent classes of isolated eigenpairs and some generalized eigenpairs contained in the positive dimensional components. Kuo, Lin and Liu [93] developed a homotopy continuation method for computing nonnegative Z/H-eigenpairs of $\mathcal{A}$. It has been guaranteed that the homotopy method can calculate a nonnegative eigenpair and also can ensure the global convergence for non-

negative Z-eigenpairs. Furthermore, they were able to determine that the number of positive Z-eigepairs of an irreducible nonnegative tensor is odd by using degree analysis. Over the last decade, there have been significant advances on homotopy techniques for polynomial systems of eigenvalues of tensors, see e.g, [43, 44, 93]. Recently, in [73], homotopy solution techniques are found to be useful to study tensor decomposition and perfect identification problems.

3.2.2 Existing Algorithms for weakly reducible tensor

Below are two existing ways for calculating the maximum eigenvalue of weakly reducible nonnegative tensors.

1. Approximation algorithms
   In [192], a counterexample was given to show that the Algorithm 3.2 may not be convergent for some reducible nonnegative tensor. Thus, an approximation method (Algorithm 4.1, [192]) was designed to find the maximum eigenvalue for reducible nonnegative tensors. Some of the existing approximation algorithms in [107,192] have been used to check the positive definiteness for the following multivariate form:

$$f(x) = \mathcal{A}x^m = \sum_{i_1,i_2,\ldots,i_m=1}^{n} a_{i_1 i_2 \ldots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}$$

   where $\mathcal{A} = (a_{i_1 i_2 \ldots i_m})$ is a symmetric Z-tensor. In [145], it is proved that $f(x)$ is called positive definite if its real eigenvalues are positive. In addition, for square tensors, it is well-known that Z-eigenvalues are parallel to M-eigenvalues as shown in [38, 101, 144, 150]. A tensor $\mathcal{A}$ is called a strong M-tensor if its smallest real eigenvalue of $\mathcal{A}$ is positive [192], so, a criterion is established to check the positive definiteness of $f(x)$.

   Theorem 3.11. *[192] Suppose that $\mathcal{A} = (a_{i_1 i_2 \ldots i_m})$ is a symmetric Z-tensor with even m. If tensor $\mathcal{A}$ is strong M-tensor, then $f(x) = \mathcal{A}x^m$ is called positive definite.*

   Theorem 3.12. *[144] Let tensor $\mathcal{B} = u(\mathcal{A} + v\mathcal{I})$, where $\mathcal{A}$ is a symmetric tensor, and u and v are real numbers. If λ is an eigenvalue of $\mathcal{A}$, then $\mu = u(\lambda + v)$ is an eigenvalue of $\mathcal{B}$.*

Lemma 3.2. *[192] Let $\mathcal{A} \in \mathcal{T}^{m,n}$. Then, define*

$$L_{\mathcal{A}} = \min_{1\leq i\leq n} \{a_{ii\ldots i} - C_i\}, \quad U_{\mathcal{A}} = \max_{1\leq i\leq n} \{a_{ii\ldots i} + C_i\}, \tag{3.6}$$

where

$$C_i = \sum_{(i,i_2,\ldots,i_m)\neq(i,i,\ldots,i)} |a_{ii_2\ldots i_m}|, \ i = 1, 2, ..., n,$$

where $L_{\mathcal{A}}$ and $U_{\mathcal{A}}$ are used to denote the lower and upper bounds of real eigenvalues of $\mathcal{A}$, respectively. For a Z-tensor $\mathcal{A}$, a tensor $\mathcal{C}$ is defined as,

$$\mathcal{C} = U_{\mathcal{A}}\mathcal{I} - \mathcal{A}. \tag{3.7}$$

Recently, Zhang et al. [192] proposed an iterative method (i.e. Algorithm 4.1) in order to calculate the maximum eigenvalue of the tensor $\mathcal{C}$ defined in (3.7). We should note that for any nonnegative tensor $\mathcal{C}$, $U_{\mathcal{A}} - \rho(\mathcal{C})$ is the smallest real eigenvalue of $\mathcal{A}$. Moreover, if the multivariate form $f(x)$ is positive definite, then it follows that $U_{\mathcal{A}} - \rho(\mathcal{C})$ is positive too. It is important to observe that Algorithm 3.2 is not always guaranteed for some reducible nonnegative tensors. A counterexample was given in [192] to demonstrate this issue. Hence, a perturbation term is added to tensor $\mathcal{C}$ in [192] and Algorithm 3.2 has been modified such that it can compute the maximum eigenvalue of the tensor

$$\mathcal{B} = \mathcal{C} + \gamma\mathcal{I} + \mathcal{E}, \tag{3.8}$$

where the parameter γ is positive and $\mathcal{E}$ is a positive tensor with positive entries, that is, ε is a very small positive number.

Lemma 3.3. *[186] Suppose that $\mathcal{A} \in \mathcal{T}_{+}^{m,n}$, and $\varepsilon > 0$ is a sufficiently small number. If $\mathcal{A}_\varepsilon = \mathcal{A} + \mathcal{E}$ where $\mathcal{E}$ is a positive tensor with every entry being ε, then*

$$\lim_{\varepsilon\to 0} \rho(\mathcal{A}_\varepsilon) = \rho(\mathcal{A}). \tag{3.9}$$

It is important to note that for any tensor $\mathcal{C} \geq 0$, we have tensor $\mathcal{B}$ defined as above is an irreducible nonnegative tensor. According to Lemma 3.3 and Theorem 3.7, the convergence of the proposed algorithm in [192] was established for any nonnegative tensor as follows.

Theorem 3.13. *[192] Suppose that tensor $\mathcal{C} \geq 0$, and tensor $\mathcal{B}$ be defined as (3.8). Then Algorithm 4.1 in [192] terminates in a finite number of steps*

and produces a value of $\rho(\mathcal{B})$ *or yields two sequences* $\{\underline{\lambda}_k\}$ *and* $\{\bar{\lambda}_k\}$ *which converge to* $\rho(\mathcal{B})$. *Moreover,* $\rho(\mathcal{C}) = \lim_{\varepsilon\to 0}\rho(\mathcal{B}) - \gamma$.

The following theorem can be used to determine that the error bound between the maximum eigenvalues of $\mathcal{C}+\mathcal{E}$ and $\mathcal{C}$ for symmetric nonnegative tensors.

Theorem 3.14. *[192] Let* $\mathcal{C} \in \mathcal{S}_+^{m,n}$ *and let* $\mathcal{C}_\varepsilon = \mathcal{C} + \mathcal{E}$. *Then,*

$$0 \le \rho(\mathcal{C}_\varepsilon) - \rho(\mathcal{C}) \le \varepsilon n^{m-1}. \tag{3.10}$$

Furthermore, in [192], another iterative algorithm is proposed for testing the positive definiteness of $f(x)$ with a Z-tensor. This algorithm works as follow.

Algorithm 3.4. *[192]*

Step 1. Input a Z-tensor $\mathcal{A} \in T^{m,n}$ *with even* m. *Calculate* $U_{\mathcal{A}} = \max_{1\le i\le n}\{a_{ii\ldots i} + C_i\}$ *and let* $\mathcal{C} = U_{\mathcal{A}}\mathcal{I} - \mathcal{A}$ *be as in (3.7).*

Step 2. By using [Algorithm 4.1, [192]], find the spectral radius $\rho(\mathcal{C})$ *of* $\mathcal{C}$.

Step 3. Let $\mu = U_{\mathcal{A}} - \rho(\mathcal{C})$. *If* $\mu > 0$, *then* $f(x) = \mathcal{A}x^m$ *is said to be positive definite, or otherwise.*

However, Algorithm 3.4 may fail to test the positive definiteness for some Z-tensors. More details about this issue are given in the next Chapter.

2. Decomposition algorithms

Some of the existing iterative algorithms discussed in the previous subsection do not guarantee convergence if $\mathcal{A}$ is weakly reducible. For weakly reducible tensors, we recall the following results.

Theorem 3.15. *[78] Suppose that* $\mathcal{A} \in T_+^{m,n}$. *If* $\mathcal{A}$ *is weakly reducible, then there is a partition* $\{I_1, ..., I_k\}$ *of* $\{1, ..., n\}$ *such that every induced tensor in* $\{\mathcal{A}_{I_j} | j \in \{1, ..., k\}\}$ *is weakly irreducible.*

Theorem 3.16. *[78] Suppose that* $\mathcal{A} \in T_+^{m,n}$ *is weakly reducible, and there is a partition* $\{I_1, ..., I_k\}$ *of* $\{1, ..., n\}$ *that is determined by Theorem 3.15. Then,* $\rho(\mathcal{A}) = \rho(\mathcal{A}_{I_i})$ *for some* $i \in \{1, ..., k\}$.

In [199], several important spectral properties have been proven for symmetric nonnegative tensor $\mathcal{A}$. These properties are summarized as follows: (1) the spectral radius of $\mathcal{A}$ is zero if and only if $\mathcal{A} = 0$, but this is not true if $\mathcal{A}$ is not symmetric; (2) $\mathcal{A}$ is a weakly irreducible if and only if there exists a unique positive eigenvalue with a positive eigenvector; (3) $\mathcal{A}$ is irreducible if and only if there exists a unique nonnegative eigenvalue with a nonnegative eigenvector; (4) weakly reducible $\mathcal{A}$ is decomposed into some weakly irreducible tensors; and finally the minimax theorem is satisfied without requiring the weak irreducibility condition.

Let us recall the minimax theorem without the weak irreducibility condition for symmetric nonnegative tensors as follows.

Theorem 3.17. *[199] Suppose that $\mathcal{A} \in \mathcal{S}_+^{m,n}$. Then,*

$$\min_{x \in int(P)} \max_{1 \le i \le n} \frac{(\mathcal{A}x^{m-1})_i}{x_i^{m-1}} = \rho(\mathcal{A}) = \max_{x \in P \setminus \{0\}} \min_{x_i \neq 0, 1 \le i \le n} \frac{(\mathcal{A}x^{m-1})_i}{x_i^{m-1}}. \tag{3.11}$$

Let $\mathcal{A} \in \mathcal{S}_+^{m,n}$, Zhou et al. [199] gave a procedure to calculate the partition $\{I_1, I_2, ..., I_k\}$ of $\{1, 2, ..., n\}$ such that each induced tensor $\mathcal{A}_{I_i}, i = 1, 2, ..., k$ is either weakly irreducible tensor or a zero tensor. The following Algorithm is based upon the result that the graph of a weakly irreducible tensor is strongly connected.

Algorithm 3.5. *[199]*

1. *Let $J_1 = \{1, 2, ..., n\}$ and set $j = 1$. If $J_j = \phi$, then stop. Otherwise, choose an element i_1 from J_j, and let $I_j := \{i_1\}$.*

2. *Check all elements $a_{i_1, i_2, ..., i_m}$ such that $i_d \in J_j, d = 2, 3, ..., m$. If all these elements are zero, then, let $J_{j+1} := J_j \setminus I_j$, set $j := j + 1$, and go to Step 1. Otherwise, for all non-zero elements $a_{i_1, i_2, ..., i_m}$ and all $d = 2, 3, ..., m, I_j := I_j \cup \{i_d\}$ if $i_d \notin I_j$. Let $K := \{i_1\}$.*

3. *If $\phi = I_j \setminus K$, then, let $J_{j+1} := J_j \setminus I_j$, set $j := j + 1$, and go to Step 1. Otherwise, go to Step 4.*

4. *Choose an $i_k \in I_j \setminus K$ and check all elements $a_{i_k, i_2, ..., i_m}$ such that $i_d \in J_j, d = 2, 3, ..., m$. For all nonzero element $a_{i_k, i_2, ..., i_m}$ and all $d = 2, 3, ..., m, I_j := I_j \cup \{i_d\}$ if $i_d \notin I_j$. Let $K := K \cup \{i_k\}$ and go to Step 3.*

Furthermore, a power-type algorithm for finding the largest eigenvalue of symmetric nonnegative tensors was proposed as follows.

Algorithm 3.6. *[199]*

Step 0. *Compute the partition* $\{I_1, I_2, ..., I_k\}$ *of the index set* $\{1, 2, ..., n\}$ *using Algorithm 3.5 such that each induced tensor* $\mathcal{A}_{I_i}, i = 1, 2, ..., k$ *is either a zero tensor or weakly irreducible tensor.*

Step 1. *For* $i = 1, ..., k$, *compute*
If $\mathcal{A}_{I_i}$ *is a zero tensor, then let* $\lambda^{(i)} = 0$. *Otherwise, compute the largest eigenvalue* $\lambda^{(i)}$ *of* $\mathcal{A}_{I_i}$ *and a corresponding eigenvector* $u^{(i)}$ *by using Algorithm 3.1*
End

Step 2. *Output the spectral radius* $\rho(\mathcal{A}) = \max_{i=1}^{k} \lambda^{(i)}$. *Assume that* $\lambda^{(i)} = \rho(\mathcal{A})$, *and* u *is defined as,*

$$u_j = (u^{(i)})_j \text{ if } j \in I_i. \text{ Otherwise, } u_j = 0.$$

Then, $\rho(\mathcal{A})$ *is the largest eigenvalue of* $\mathcal{A}$ *and* u *is its corresponding eigenvector.*

Clearly as noted in [199], based on Theorem 3.9, if the underlying tensor $\mathcal{A}$ is weakly irreducible, then by using Algorithm 3.3, we are able to compute the maximum eigenvalue and the corresponding eigenvector of $\mathcal{A}$. Similarly, in Step 1 of Algorithm 3.6, we can use Algorithm 3.3 to find the maximum eigenvalue $\lambda^{(i)}$ of induced weakly irreducible tensor $\mathcal{A}_{I_i}$ and the corresponding eigenvector $u^{(i)}$. In the light of Theorem 3.15, it is indicated that weakly reducible symmetric tensors are partitioned into some induced weakly irreducible tensors. Hence, by using Algorithm 3.6, we are able to obtain the eigenvalue, which is the largest in modulus, together with its associated eigenvector for any symmetric nonnegative tensor. Furthermore, in [199], it has been shown that Algorithm 3.6 performs better than Algorithm 3.2 for weakly reducible tensors.

3.3 An algorithm for the spectral radius of a symmetric nonnegative tensor

"Publication has been removed due to copyright restrictions". For more details, we refer the reader to [200].

3.4 Testing the positive definiteness of Z-tensors

"Publication has been removed due to copyright restrictions". For more details, we refer the reader to [200].

3.5 Conclusion

In this chapter, we proposed algorithm 1 in [200] for computing the spectral radii of a symmetric nonnegative tensor. This new algorithm is a modified version of the algorithm proposed by Ng. Qi and Zhou in [120]. In particular, algorithm 1 in [200] has the following important properties. First, it does not require the partition for weakly reducible tensors. Second, it is convergent for any symmetric nonnegative tensors (cf. Theorem 1 in [200]). Third, it is able to compute the spectral radius for large-size tensors. Numerical results reported in Table 1 and Figure 1 in [200] have shown that our new algorithm 1 in [200] works very efficiently. It is able to compute the spectral radii of large-size tensors and used much less cpu time than Algorithm 4.1 [199].

CHAPTER 4

Approximation methods for nonnegative polynomial optimization problems

In this chapter, we consider approximation algorithms for non-negative polynomial optimization problems over unit spheres. These optimization problems have wide applications e.g., in signal and image processing, high order statistics, and computer vision. Since these problems are NP-hard, we are interested in studying on approximation algorithms. In particular, we propose some polynomial-time approximation algorithms with new approximation bounds. In addition, based on these approximation algorithms, some efficient algorithms are presented and numerical results are reported to show the efficiency of our proposed algorithms.

4.1 Introduction

"Publication has been removed due to copyright restrictions". For more details, we refer the reader to [195].

4.2 Approximation solutions for (P1) and (P2)

"Publication has been removed due to copyright restrictions". For more details, we refer the reader to [195].

4.3 Efficient algorithms for (P1) and (P2)

"Publication has been removed due to copyright restrictions". For more details, we refer the reader to [195].

4.4 Conclusion

In this chapter, we studied nonnegative polynomial optimization problems over unit spheres, in particular (P1) and (P2). Both are NP-hard problems, hence, they are difficult to solve theoretically as well as numerically. Motivated by this, we presented some new approximation algorithms for (P1) and (P2) with some improved bounds. The new approximation bounds obtained in this chapter are summarized in Table 1 in [195]. We proposed some new efficient algorithms for (P1) and (P2), and our numerical results indicated the efficiency of the proposed methods for tested problems.

CHAPTER 5

Maximum Clique Problem

In this chapter, the maximum clique problem is formulated into a polynomial optimization problem which has been studied in Chapter 4. Based on this formulation, we develop a computational algorithm to solve the maximum clique problem. We then apply the proposed method to compute the largest clique for the most popular DIMICS benchmark graphs.

5.1 Introduction

The combinatorial optimization problem is a very important research area that has many interesting different optimization models arise in operation research, graph theory, management, engineering and computer science, amongst many others. In general, many areas of these disciplines need to be designed using some forms of combinatorial optimization. Even though, many classical problems in combinatorial optimization are quite challenging and computationally difficult even to approximate due to theoretical and practical implications of such issues, in particular, the maximum clique problem, the maximum graph matching and vertex cover problems. The maximum clique problem (MCP for short) is a well-known optimization problem that finds numerous applications in different domains, including coding theory [29,57], computer vision [12,135], fault diagnosis [15], pattern recognition [127], printed circuit board testing [54, 175], location problems [28], and social network analysis [11, 134]. Furthermore, it plays a central role in many combinatorial optimizations such as clique partitioning [55], graph clustering [60,159], graph vertex colouring [34,57,180], maximum diversity [52], optimal winner determination [7,183], set packing [8,184], and sum colouring [181]. These problems are associated with MCP and can either be directly formulated as MCP or could be as a sub-problem of the clique problem. Also, it is important to point out that several intractable combinatorial optimiza-

tion problems such as the independent set problem, the vertex cover problem and others can be easily reducible to the MCP.

In general, the MCP is a very challenging problem because it has been proven to be NP-complete [113], unless $P = NP$. Hence, due to its computational complexity, exact algorithms (such as explicit and implicit enumerative methods) run in a time that grows exponentially with the dimensionality of the problem, and return infeasible solutions. Furthermore, it is also found that the MCP remains intractable to solve even in detecting satisfactory approximate solutions within a factor of $n/2^{(logn)^{1-\epsilon}}$ for any $\epsilon > 0$ [58]. This has motivated many researchers to focus on characterizing the approximation properties of the clique problem and study its complexity, for instance, [6, 16, 17, 58, 129, 142]. In the light of these theoretical complexity studies, much recent research efforts have been directed to develop highly effective heuristics to find the MCP with no guarantee of performance, but still viable in practical applications. Despite the inherent difficulty of the MCP, the last decade has witnessed remarkable developments in problem formulations, algorithm's complexity, bounds, and approximation methods. We refer interested readers to [20,87,182] for comprehensive surveys that include the exact algorithms and local search heuristics as well as some applications of the MCP and its links with other combinatorial problems.

Historically, Motzkin and Straus [115] discovered a remarkable result by establishing the connection between a global maximum of quadratic optimization problem and the largest clique size in a certain graph. Recently, this result has been successfully generalized in numerous different directions [19,67,140]. Lately, there has been a growing interest centered around the Motzkin-Straus continuous formulation and its related generalizations to solve the MCP. In [130], the MCP has been successfully formulated as an indefinite quadratic program (QP) with linear constraints, and a global optimization method was designed to determine the largest clique sizes by implementing an iterative clique retrieval process, as well as obtained sharper lower bound for clique size. However, it is found that computing a largest clique from the optimal objective value of QP is difficult because of local solutions of QP lack strictness. In [67, 140], different strategies have been presented to overcome the spurious solutions exist in Motzkin-Straus formulation. Gibbson et al. [67] developed a continuous-based heuristic (CBH) based on a parametrization of the Motzkin-Straus program in order to avoid spurious solution. In particular, the sign constraints $x \geq 0$ of the Motzkin-Straus program is replaced with $x^T x = 1/s$, where a parameter $s \in [1, n]$. Their approach is to optimize a quadratic function over a spherical constraint, which turns the

problem to be solvable in a polynomial time. Therefore, relaxing from positivity, rounding, and a careful choice of the parameter can help to achieve good cliques. Pelillo [139] proposed a new approximation method for the MCP based on relaxation method, which is closely related to replicator equations. In [18], replicator equations, a class of dynamic principles of evolutionary game theory, have been successfully used to treat Motzkin-Straus quadratic program, and further applied them to determine the maximal cliques of a graph. The drawback of this method is that its inherent inability to escape local solutions, and hence they incorporate some block pivoting methods to overcome this issue. Another approximation method was developed in [112], which is based on complementary pivoting method and known as (PBH) for the clique problem.

Some other results available in the literature concerning the approximation methods of the MCP on arbitrary or special graphs [45, 49]. If the focus is on graphs with special structure, then the MCP is polynomially solvable in many cases. For instance, Balas and Yu [10] studied many classes of graphs and indicated that most of these graph problems can be solved in polynomial time with many maximal cliques, in particular, perfect graphs. In addition, there are some other common combinatorial optimization heuristics that have been successfully used to tackle the MCP in order to obtain good solutions within a reasonable time. Many of these optimization heuristics are related to greedy construction heuristics, and stochastic local search algorithms. In particular, some existing stochastic local search heuristics, including simulated annealing [34], neural networks [84], genetic algorithms [118], and tabu search [164]. Recently, other developments have been directed towards devising efficient methods such as penalty-evaporation approach embedded into a decomposition methods in [166] and the augmentation method based on the edge projection in [110]. Other approximation algorithms include the Variable Neighbourhood Search heuristic [70], the k-opt local search method [88], the Augmentation algorithm [110], the hybrid evolutionary method [163], genetic algorithm [163], multi-neighbourhood tabu search algorithm [184], the breakout local search method [13], and many more.

The structure of this chapter: In Section 5.1, we review some definitions and notations that will be used in this chapter. We present an equivalent formulation based on Motzkin-Straus QP formulation in Section 5.2. Based on this formulation, in Section 5.3, we propose a computational algorithm for the MCP. Finally, we show the effectiveness of our proposed method by evaluating it on popular 66 DIMACS benchmark graphs in Section 5.4.

5.1.1 Notations and Terminologies

Suppose that a graph $G = (V, E)$ is undirected, which composed of a finite set of vertices or nodes $V = \{1, \ldots, n\}$, and a finite set of edges or arcs $E \subseteq V \times V$. We denote by $A_G \in \{0, 1\}^{n \times n}$, the adjacency matrix of G, which is a symmetric matrix defined as

$$a_{ij} = \begin{cases} 1, & \text{if } (i, j) \in E, \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

Let $d(v)$ to denote the degree of the vertex $v \in V$, and $|V|$ to indicate the size of V. The neighbourhood of a vertex v is denoted by $N(i) = \{j \in V : \{i, j\} \in E\}$ (i.e. the set of all vertices adjacent to v). Let $\overline{G} = (V, \overline{E})$ be the complementary graph of G, where $\overline{E} = \{(i, j) | i, j \in V, i \neq j \text{ and } (i, j) \notin E\}$. A graph $G = (V, E)$ is said to be complete if every vertex is connected to every other vertex of G by an edge such that $\forall\ i, j \in V, with\ i \neq j$,we have $(i, j) \in E$. A clique C of G is a subset of vertices such that every two distinct vertices in the clique are adjacent. The size of the largest clique in G (i.e. maximum clique) is denoted by $\omega(G)$, which is the number of vertices in the largest clique. The MCP seeks to find a clique C (i.e. complete subgraph) in a graph G of the largest cardinality among all the cliques. A clique is said to be maximal, if it is not possible to add an additional vertex to C and C remains a clique. Also we use this formula $\frac{2|E|}{(|V|*(|V|-1))}$ to find the density of each graph, where $|V|$ and $|E|$ are the number of nodes and arcs of G, respectively.

5.2 Modeling formulation

The MCP has many equivalent formulations available in literature, and most of these formulations are based on the integer programming formulation and non-convex continuous optimization. In this section, we mainly focus on the continuous-based formulation of the MCP. In particular, the Motzkin-Strauss formulation of the MCP is as follows:

$$\begin{aligned} &\text{maximize} && f(x) = x^T A_G x \\ &\text{subject to} && x \in \Delta_n. \end{aligned} \tag{5.2}$$

The superscript x^T stands for transportation; A_G denote the adjacency matrix of G as in (5.1); and Δ_n is the standard simplex of $\Re^n$ defined as

$$\Delta_n = \{x \in \Re^n : x_i \geq \ 0 \ \forall \ i \in V, e^T x = \sum_{i=1}^{n} x_i = 1\},$$

where $e = \sum_{i \in V} e_i = [1, \ldots, 1]^T$, (i.e. e is the n-vector of the appropriate length of all ones, consisting of unit entries) and e_i used here to indicate the i-th standard basis vector in $\Re^n$. In mid sixties, Motzkin and Straus [115] found that a connection exists between the MCP and an indefinite quadratic programming (5.2).

Theorem 5.1. *[115] The global optimal value of (5.2) is determined by this following relation* $\frac{1}{2}(1 - \frac{1}{\omega(G)})$*, where* $\omega(G)$ *is the clique number of* G*.*

Specifically, if x^* is a global solution of (5.3), then it is proved that the clique number $\omega(G)$ is related to the optimal value $f(x^*)$ as follows

$$\omega(G) = \frac{1}{1 - f(x^*)} \geq \frac{1}{1 - f(x)}, \ \forall \ x \in \Delta_n. \tag{5.3}$$

Its proof can be found in [1]. In addition, let S be a subset of vertices and x^S be its characteristic vector, which is defined as

$$x_i^S = \begin{cases} \frac{1}{|S|}, & \text{if } \ i \in S, \\ 0, & \text{otherwise.} \end{cases}$$

As proved in [115], a subset $S \subseteq V$ is the largest clique of G if and only if its x^S is a global solution of f on Δ_n. Gibbons et al. [68], and Pelillo et al. [140], extended the results of [115] in order to provide a characterization of maximal cliques in terms of local solutions of f on Δ_n, however, not all local solutions were in the form of a characteristic vector. Additionally, the first and second-order optimality conditions of the Motzkin-Straus quadratic program were described in [68]. Furthermore, the Motzkin-Straus result has been expanded to hypergraphs [165], and recently been generalized to r-graphs by considering a continuous characterization of maximal cliques other than Lagrangians of hypergraphs [31, 32].

The Motzkin-Straus theorem has become increasingly important due to its intriguing computational significance in many combinatorial problems. It provides a new technique to tackle the MCP, by allowing a shift from the discrete to continuous setting. It is noted in [131] that the continuous formulations of discrete problems have greatly caught some attention, because they not only allow

us to discover the full arsenal of continuous optimization methods, thereby leading to the novel development of new and effective strategies, but would also lead to explore unexpected theoretical aspects and properties. Despite its drawback, the Motzkin-Straus and its related theorems have played a fundamental role in the development of many great optimization algorithms and heuristics especially the largest clique finding techniques [18, 19, 21, 67, 130, 132, 165]. It has also been successfully used to determine several and better bounds for the largest clique size of graphs [2, 30, 171–174]. In addition, the Motzkin-Straus theorem has been extended to some other applications and problems of combinatorial optimization such as the vertex weighted graphs [67], and the edge weighted graphs [135].

The formulation of Motzkin and Straus has a major drawback associated with it due to the existence of spurious solutions, namely, local or global solutions of f which are not in the form of characteristic vectors, i.e., the local solutions that are not in a one-to-one correspondence with maximal cliques of the original combinatorial problems [140]. This has been detected empirically by Pardalos and Philips [132] and later in 1996, Pelillo and Jagota [140] confirmed this finding. Fortunately, in an attempt to circumvent the Motzkin-Straus formulation drawback, Bomze [19] solved the spurious solution problem and introduced the following regularized variant formulation of (5.2) for the MCP.

$$\begin{aligned} \max \quad & g(x) = x^T \left(A_G + \frac{1}{2} I \right) x \\ \text{s.t.} \quad & x \in \Delta_n, \end{aligned} \tag{5.4}$$

where I is the $n \times n$ identity matrix. This is just a standard quadratic optimization problem (StQP) with matrix $A = -2(A_G + \frac{1}{2}I)$. More importantly, Bomze [19] proved the following fundamental result for the spurious free counterpart of the Motzkin-Straus theorem.

Theorem 5.2. *[19] For a graph G, let S and x^S to denote a subset of vertices, and its characteristic vector respectively, then the following assertions hold:*

(i) S is said to be the largest clique of G iff x^S is a global solution of problem (5.4). Thus, we have $\omega(G) = \frac{1}{2(1-g(x^S))}$.

(ii) S is said to be a maximal clique of G iff x^S is a local solution of problem (5.4).

(iii) All local and global solutions x of problem (5.4) over the standard simplex Δ_n are strict and of the form $x = x^S$ for some $S \subseteq V$.

Theorem 5.2 guarantees that all local and global solutions of g on Δ_n are strict, and are characteristic vectors of maximal/maximum cliques of G. More precisely, it establishes a one-to-one correspondence between local (global) solutions of problem (5.4) and maximal (maximum) cliques of G.

Lemma 5.1. *[119] With appropriate change of variables, Motzkin-Straus quadratic problem (5.2) can be reformulated in terms of quartic or cubic maximization problem over the Euclidean ball as follows:*

(a) Quartic maximization problem:

$$\max_{u\in\Re^n}\left\{\sum_{k=1}^{m}(u^TA_ku)^2:\ \|u\|=1\right\}. \tag{5.5}$$

where $A_k=\frac{e_{i_k}^Te_{i_k}+e_{j_k}^Te_{j_k}}{\sqrt{2}}$, $(i_k,j_k)\in E, 1\le k\le m$, *and* $e_i, 1\le i\le n$, *are the basis vectors of* $\Re^n$.

(b) Cubic maximization problem:

$$\max_{u\in\Re^n,w\in\Re^m}\left\{\sum_{k=1}^{m}w_k(u^TA_ku):\|u\|=1,\|w\|=1\right\}. \tag{5.6}$$

Lemma 5.2. *[197] Assume that* (u^*,w^*) *is a global solution of problem (5.6). Then, it follows that* (u^*,u^*,w^*) *is a global solution of the following optimization problem:*

$$\max_{u,v\in\Re^n,w\in\Re^m}\left\{\sum_{k=1}^{m}w_k(u^TA_kv):\|u\|=1,\|v\|=1,\|w\|=1\right\}. \tag{5.7}$$

From the Motzkin-Straus formulation (5.2), the second constraint $x_i\ge 0$ can be eliminated by considering the square of the variable x (*i.e.* $x_i=y_i^2$, $y^2=(y_1^2,y_2^2,...,y_n^2)^T\in\Re^n$) in the optimization problem. Hence, an equivalent formulation of (5.2) is as follows:

$$\begin{aligned}\text{maximize}\quad & (y^2)^TA(y^2)\\ \text{s.t.}\quad & \sum_{i=1}^{n}y_i^2=1,\end{aligned} \tag{5.8}$$

Furthermore, we use the following notation:

$$uv = \begin{pmatrix} u_1v_1 \\ u_2v_2 \\ \vdots \\ u_nv_n \end{pmatrix}, \quad wz = \begin{pmatrix} w_1z_1 \\ w_2z_2 \\ \vdots \\ w_nz_n \end{pmatrix}. \tag{5.9}$$

Then, (5.8) is equivalent to the following multi-linear optimization problem:

$$\begin{aligned} \max \quad & (uv)^T A(wz) \\ \text{s.t.} \quad & \sum_{i=1}^{n} u_i^2 = 1,\ u \geq 0, \\ & \sum_{i=1}^{n} v_i^2 = 1,\ v \geq 0, \\ & \sum_{i=1}^{n} w_i^2 = 1,\ w \geq 0, \\ & \sum_{i=1}^{n} z_i^2 = 1,\ z \geq 0, \\ & \text{where } u, v, z, w \in \Re^n. \end{aligned} \tag{5.10}$$

Moreover, since A is a nonnegative symmetric matrix, there exists a nonnegative symmetric fourth order tensor $\mathcal{A} = (a_{ijkl}), 1 \leq i, j, k, l \leq n$, such that (5.10) can be rewritten into the following optimization problem:

$$\begin{aligned} \max \quad & F(u, v, w, z) = \mathcal{A}uvwz = \sum_{1 \leq i,j,k,l \leq n} a_{ijkl}u_iv_jw_kz_l, \\ \text{s.t.} \quad & \|u\| = \|v\| = \|w\| = \|z\| = 1,\ u, v, w, z \geq 0, \\ & \text{where } u, v, w, z \in \Re^n. \end{aligned} \tag{5.11}$$

Simply, we compute the gradients of F with respect to u, v, w, and z as follows:

$$\begin{aligned} \nabla_u F(u, v, w, z) &= \mathcal{A}vwz, \\ \nabla_v F(u, v, w, z) &= \mathcal{A}uwz, \\ \nabla_w F(u, v, w, z) &= \mathcal{A}uvz, \\ \nabla_z F(u, v, w, z) &= \mathcal{A}uvw. \end{aligned}$$

Then, we have

$$\begin{aligned}
&(\nabla_u F(u,v,w,z))^T u = F(u,v,w,z),\\
&(\nabla_v F(u,v,w,z))^T v = F(u,v,w,z),\\
&(\nabla_w F(u,v,w,z))^T w = F(u,v,w,z),\\
&(\nabla_z F(u,v,w,z))^T z = F(u,v,w,z).
\end{aligned}$$

Theorem 5.3. *Problems (5.8), (5.10) and (5.11) have the same optimal values.*

Proof. By Theorem 2.1 [194] or Theorem 4.1 [41], this theorem holds. □

5.3 An algorithm for problem (5.11)

In this section, we present an alternating direction method (ADM) to solve problem (5.11) which is stated as follows.

Algorithm 5.1.

- ***Initialization****: Given $u^0 > 0, v^0 > 0, w^0 > 0$ and $z^0 > 0$, set $k = 0$;*
- ***Iteration****: For $k = 1, 2, \ldots,$ do,*

$$\begin{aligned}
&u^{(k)} := \arg\max F\left(u, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right)\\
&s.t.\ \|u\| = 1, u \geq 0,\\
&v^{(k)} := \arg\max F\left(u^{(k)}, v, w^{(k-1)}, z^{(k-1)}\right)\\
&s.t.\ \|v\| = 1, v \geq 0,\\
&w^{(k)} := \arg\max F\left(u^{(k)}, v^{(k)}, w, z^{(k-1)}\right)\\
&s.t.\ \|w\| = 1, w \geq 0,\\
&z^{(k)} := \arg\max F\left(u^{(k)}, v^{(k)}, w^{(k)}, z\right)\\
&s.t.\ \|z\| = 1, z \geq 0.
\end{aligned}$$

Remark 5.1.

(i) In Algorithm (5.1), $u^{(k)}, v^{(k)}, w^{(k)}$ and $z^{(k)}$ can be calculated easily as follows:

$$
\begin{aligned}
u^{(k)} &= \frac{\nabla_u F\left(u, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right)}{\|\nabla_u F\left(u, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right)\|} = \frac{\mathcal{A}v^{(k-1)}w^{(k-1)}z^{(k-1)}}{\|\mathcal{A}v^{(k-1)}w^{(k-1)}z^{(k-1)}\|},\\
v^{(k)} &= \frac{\nabla_v F\left(u^{(k)}, v, w^{(k-1)}, z^{(k-1)}\right)}{\|\nabla_v F\left(u^{(k)}, v, w^{(k-1)}, z^{(k-1)}\right)\|} = \frac{\mathcal{A}u^{(k)}w^{(k-1)}z^{(k-1)}}{\|\mathcal{A}u^{(k)}w^{(k-1)}z^{(k-1)}\|},\\
w^{(k)} &= \frac{\nabla_w F\left(u^{(k)}, v^{(k)}, w, z^{(k-1)}\right)}{\|\nabla_w F\left(u^{(k)}, v^{(k)}, w, z^{(k-1)}\right)\|} = \frac{\mathcal{A}u^{(k)}v^{(k)}z^{(k-1)}}{\|\mathcal{A}u^{(k)}v^{(k)}z^{(k-1)}\|},\\
z^{(k)} &= \frac{\nabla_z F\left(u^{(k)}, v^{(k)}, w^{(k)}, z\right)}{\|\nabla_z F\left(u^{(k)}, v^{(k)}, w^{(k)}, z\right)\|} = \frac{\mathcal{A}u^{(k)}v^{(k)}w^{(k)}}{\|\mathcal{A}u^{(k)}v^{(k)}w^{(k)}\|}.
\end{aligned}
$$

(ii) Note that $u^{(k)}, v^{(k)}, w^{(k)}$, and $z^{(k)}$ satisfy the following systems.

$$
\begin{aligned}
&\nabla_u F\left(u, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right) = \lambda_u^{(k)} u^{(k)}\\
&s.t.\ \|u^{(k)}\| = 1,\\
&\nabla_v F\left(u^{(k)}, v, w^{(k-1)}, z^{(k-1)}\right) = \lambda_v^{(k)} v^{(k)}\\
&s.t.\ \|v^{(k)}\| = 1,\\
&\nabla_w F\left(u^{(k)}, v^{(k)}, w, z^{(k-1)}\right) = \lambda_w^{(k)} w^{(k)}\\
&s.t.\ \|w^{(k)}\| = 1,\\
&\nabla_z F\left(u^{(k)}, v^{(k)}, w^{(k)}, z\right) = \lambda_z^{(k)} z^{(k)}\\
&s.t.\ \|z^{(k)}\| = 1.
\end{aligned}
$$

(iii) We terminate Algorithm (5.1), when

$$
\max\left\{\|u^k - u^{k-1}\|, \|v^k - v^{k-1}\|, \|w^k - w^{k-1}\|, \|z^k - z^{k-1}\|\right\} \le \epsilon,
$$

where ϵ is very small number (i.e., 10^{-6}).

Pre-multiply the above with $u^{(k)}, v^{(k)}, w^{(k)}$ and $z^{(k)}$ respectively, thus we have

$$
\begin{aligned}
F\left(u^{(k)}, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right) &= \lambda_u^{(k)}\\
F\left(u^{(k)}, v^{(k)}, w^{(k-1)}, z^{(k-1)}\right) &= \lambda_v^{(k)}\\
F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k-1)}\right) &= \lambda_w^{(k)}\\
F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\right) &= \lambda_z^{(k)}.
\end{aligned}
\tag{5.12}
$$

Furthermore,

$$\begin{aligned} F\left(u^{(k-1)}, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right) &\leq F\left(u^{(k)}, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right) \\ &\leq F\left(u^{(k)}, v^{(k)}, w^{(k-1)}, z^{(k-1)}\right) \\ &\leq F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k-1)}\right) \\ &\leq F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\right). \end{aligned} \tag{5.13}$$

Next, we state the convergence of Algorithm 5.1 as follows:

Theorem 5.4. *Let* $\{u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\}$ *be a sequence generated by Algorithm 5.1. Suppose* $\lim_{k\to\infty}\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\right) = (u^*, v^*, w^*, z^*)$. *Then,* (u^*, v^*, w^*, z^*) *is a KKT point of problem (5.11).*

Proof. Since the sequence $\left\{F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\right)\right\}$ is non-decreasing together with the fact that $F(u, v, w, z)$ is bounded over unit spheres, we know that the sequence $\left\{F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\right)\right\}$ converges and

$$\lim_{k\to\infty}\left(F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\right) - F\left(u^{(k-1)}, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right)\right) = 0. \tag{5.14}$$

From (5.12), (5.13), (5.14), $\{\lambda_u^{(k)}\}, \{\lambda_v^{(k)}\}, \{\lambda_w^{(k)}\}$ and $\{\lambda_z^{(k)}\}$ are convergent. Suppose that (u^*, v^*, w^*, z^*) is a limit point of $(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)})$. Then, we have

$$\begin{aligned} \lim_{k\to\infty} u^{(k)} = u^*, \lim_{k\to\infty} v^{(k)} = v^*, \lim_{k\to\infty} w^{(k)} = w^*, \lim_{k\to\infty} z^{(k)} = z^*, \\ \lim_{k\to\infty} \lambda_u^{(k)} = \lim_{k\to\infty} \lambda_v^{(k)} = \lim_{k\to\infty} \lambda_w^{(k)} = \lim_{k\to\infty} \lambda_z^{(k)} = \lambda^*. \end{aligned}$$

From the iterative step of Algorithm 5.1, we have

$$\begin{cases} \nabla_u F\left(u, v^{(k-1)}, w^{(k-1)}, z^{(k-1)}\right) = \lambda_u^{(k)} u^{(k)}, \\ \nabla_v F\left(u^{(k)}, v, w^{(k-1)}, z^{(k-1)}\right) = \lambda_v^{(k)} v^{(k)}, \\ \nabla_w F\left(u^{(k)}, v^{(k)}, w, z^{(k-1)}\right) = \lambda_w^{(k)} w^{(k)}, \\ \nabla_z F\left(u^{(k)}, v^{(k)}, w^{(k)}, z\right) = \lambda_z^{(k)} z^{(k)}, \\ \|u^{(k)}\| = \|v^{(k)}\| = \|w^{(k)}\| = \|z^{(k)}\| = 1. \end{cases} \tag{5.15}$$

Then, it holds that

$$\begin{cases} \nabla_u F\left(u^*, v^*, w^*, z^*\right) = \lambda^* u^*, \\ \nabla_v F\left(u^*, v^*, w^*, z^*\right) = \lambda^* v^*, \\ \nabla_w F\left(u^*, v^*, w^*, z^*\right) = \lambda^* w^*, \\ \nabla_z F\left(u^*, v^*, w^*, z^*\right) = \lambda^* z^*, \\ \|u^*\| = \|v^*\| = \|w^*\| = \|z^*\| = 1. \end{cases} \tag{5.16}$$

Thus, u^*, v^*, w^* *and* z^* are KKT points of the problem (5.11), which completes the proof. □

Remark 5.2. *In Theorem (5.4), we assume that the sequence* $\left\{F\left(u^{(k)}, v^{(k)}, w^{(k)}, z^{(k)}\right)\right\}$ *is convergent. In our future research, we will study under which conditions this assumption will be satisfied.*

5.4 Numerical Experiment

To properly evaluate the alternating iterative algorithm for optimization problem (5.11), it is important to carry out practical experiments to test the efficiency of the proposed method. Hence, this section is devoted to show the performance of Algorithm 5.1 for DIMACS benchmark graphs. The proposed method is tested extensively on various instances (i.e. 66 graphs) obtained from DIMACS challenge benchmarks due to Johnson and Trick [87].

DIMACS benchmark: There are many different graphs and instances presented at the second DIMACS Implementation challenge. These instances are the most frequently used to evaluate the performance of MCP algorithms in order to achieve the optimum/best values for $\omega(G)$. We will conduct our experiment on the following popular families of DIMACS benchmark. These instances have varying degree of complexity in terms of the graph size, and here is the description as follows:

- Brock- instances are generated by Brockington and Culberson [27] which have large cliques hidden among nodes where the expected clique size is much smaller. In brock- graph family, there are about 12 graphs of different sizes, the smallest (brock200-1) consists of 200 vertices and 14834 edges and the largest (brock800-4) has 800 vertices and 207643 edges.

- p-Hat graphs, these are random instances generated by Patrick Soriano and Michel Gendreau [164] which have wider vertex degree spread and larger cliques. The p-hat generator of the random instances is a generalization of the classical uniform random graphs generator. In pHat- graph family, there are about 15 graph instances of different sizes, the smallest (i.e. p-Hat-300-1) consists of 300 vertices and 10933 edges, and the largest (i.e.p-Hat1500-3) has 1500 vertices and 847244 edges.

- MANN instances: Clique formulation of the Steiner Triple problem, and generated by Carlo Mannino, translated from the set covering formulation.

The MANN instances range from 45 vertices and 918 edges up to 3321 vertices and 5506380 edges.

- Hamming graph problems defined as $H(n,d)$, of size n and hamming distance d. Each graph $H(n,d)$ consists of 2^n vertices, $2^{n-1}\sum_{i=d}^{n}\begin{pmatrix} n \\ i \end{pmatrix}$ edges and the degree of each vertex is $\sum_{i=d}^{n}\begin{pmatrix} n \\ w \end{pmatrix}$. Its generating procedure can be found in [72]. In hamming- instances, there are about 6 graphs of different dimensions, the smallest (i.e. hamming6-2) consists of 64 vertices and 1824, and the largest (i.e. hamming10-4) has 1024 vertices and 434176 edges.

- Keller graph instances are created by Peter Shor which based on Keller's conjecture on tilings using hypercubes. There are about three keller graph instances of different dimensions, the smallest (i.e. keller4) consists of 171 vertices and 9435 edges and the largest (i.e. keller6) has 3361 vertices and 4619898 edges.

- Johnson graphs: Johnson graphs $J(n,w,d)$ are similar to Hamming graphs $H(n,d)$, with additional parameter w used to indicate the weighted vertex. $J(n,w,d)$ graphs consist of n vertices, $\frac{1}{2}\begin{pmatrix} n \\ w \end{pmatrix}\sum_{k=\lceil\frac{d}{2}\rceil}^{w}\begin{pmatrix} w \\ k \end{pmatrix}\begin{pmatrix} n-w \\ k \end{pmatrix}$ edges and the degree of each vertex is defined by $\sum_{k=\lceil\frac{d}{2}\rceil}^{w}\begin{pmatrix} w \\ k \end{pmatrix}\begin{pmatrix} n-w \\ k \end{pmatrix}$, and we say that two vertices are adjacent if their Hamming distance is at least d. More details about its generated procedure can be found in [72].

- Sanchis graphs: (san for short) Sanchis [158] developed the generation procedures to produce the graphs prefixed with *san*.

- Sanr: are random graphs with dimensions similar to those in Sanchis graphs, and developed by Laura Sanchis [158].

In the following, we will compare ADM algorithm performance with the competing method based on unconstrained non-convex formulation that was presented by (Bomze, Grippo, & Palagi, 2012) [22]. They used the non-monotone Barzilai-Borwein gradient method to solve for the penalty formulation of the StQPs and for the MCP. All solution techniques are implemented in MATLAB (R2015b) and all the numerical computations are conducted using personal laptop with processor *Intel (R) Core (TM) i7-4510U CPU@ 2.00 GHz*, with installed

RAM of 8.00 GB. However, We use the most popular-tested problems of DIMACS clique benchmark with different types of problems ranging from easy to difficult in order to compare the efficiency of the proposed ADM algorithm and the penalty-based unconstrained formulation presented in [22].

The results of our proposed method are shown in Tables 5.1 and 5.2, each table contains graph's name, $|V(G)|$ "the number of vertices in each graph", $|E(G)|$ "the number of edges", "density" , $\omega(G)$ "the best known clique", "ADM", and finally penalty-based relaxation results extracted from [22]. We perform 150 random runs for most of these problems, and each problem has been solved starting with randomly generated points $u^0, v^0, w^0,$ *and* z^0 satisfying $\|u^0\| = \|v^0\| = \|w^0\| = \|z^0\| = 1$ respectively. In these tables, we compare the ADM algorithm with the results presented in [22], and the best performance for the MCP are presented in a bold face. We found that both methods are efficient and return best cliques in $c - fat$, $Hamming$ and $Johnson$ graph categories in most instances of different sizes. Furthermore, the ADM algorithm returns better maximum cliques for $hamming10-2$, and $hamming10-4$ graphs, whereas proposed method in [22] finds better maximum cliques for some of *brock* and *keller* graphs. However, both methods do not have a performance guarantee for solutions provided for other classes such as *brock*, *keller*, *San*, and *Sanr* in most dimensions.

5.5 Concluding Remarks

In this chapter, we presented a continuous reformulation for the maximum clique problem (MCP) based on the Motzkin-Strauss QP. This formulation is equivalent to the multi-linear function with spherical constraints. We used an Alternating Direction Method (ADM) to solve the new relaxed problem as in general ADM has been shown to be efficient for many optimization problems. Finally, we showed extensive computational results on 66 DIMACS benchmark graphs with different dimensions. The overall results are reported in Tables 5.1 and 5.2, to compare the performance of ADM (i.e., Algorithm 5.1) with a method presented in [22] based on Penalty formulation. In fact, the superiority of any computational method for the MCP is measured by the best value obtained. Hence, in terms of clique size found, it seems that none of these proposed solution methods can be considered better than the other one.

Table 5.1: Numerical results and comparison of ADM and Penalty formulation [22] on the 66 DIMACS instances

.

Class	Graph	$\|V(G)\|$	$\|E(G)\|$	Density	$\omega(G)$	ADM	[22]
Brock	brock200-1	200	14834	74.5	21	20	20
	brock200-2	200	9876	49.6	12	10	10
	brock200-3	200	12048	60.5	15	13	13
	brock200-4	200	13089	65.8	17	15	15
	brock400-1	400	59723	74.8	27	23	**24**
	brock400-2	400	59786	74.9	29	**23**	22
	brock400-3	400	59681	74.8	31	21	21
	brock400-4	400	59765	74.9	33	22	**23**
	brock800-1	800	207505	64.9	23	18	**19**
	brock800-2	800	208166	65.1	24	18	18
	brock800-3	800	207333	64.9	25	18	18
	brock800-4	800	207643	65.0	26	18	18
c-fat	c-fat200-1	200	1534	7.7	12	12	12
	c-fat200-2	200	3235	16.3	24	24	24
	c-fat200-5	200	8473	42.6	58	58	58
	c-fat500-1	500	4459	3.6	14	14	14
	c-fat500-2	500	9139	7.3	26	26	26
	c-fat500-5	500	23191	18.6	64	64	64
	c-fat500-10	500	46627	37.4	126	126	126
hamming	hamming6-2	64	1824	90.5	32	32	32
	hamming6-4	64	704	34.9	4	4	4
	hamming8-2	256	31616	96.9	128	128	128
	hamming8-4	256	20864	63.9	16	16	16
	hamming10-2	1024	518656	99.0	512	**499**	453
	hamming10-4	1024	434176	82.9	40	**35**	34
johnson	johnson8-2-4	28	210	55.6	4	4	4
	johnson8-4-4	70	1855	76.8	14	14	14
	johnson16-2-4	120	5460	76.5	8	8	8
	johnson32-2-4	496	107880	87.9	16	16	16
keller	keller4	171	9435	69.9	11	9	**10**
	keller5	776	225990	75.2	27	15	**19**
	keller6	3361	4619898	81.8	59	31	**36**

Table 5.2: Numerical results and comparison of ADM and Penalty formulation [22] on the 66 DIMACS instances

.

Class	Graph	$\|V(G)\|$	$\|E(G)\|$	Density	$\omega(G)$	ADM	[22]
MANN	MANN-a9	45	918	92.7	16	16	16
	MANN-a27	378	70551	99.0	126	117	**119**
	MANN-a45	1035	533115	99.6	345	330	330
	MANN-a81	3321	5506380	99.8	1100	1064	**1080**
p-hat	p-hat300-1	300	10933	24.4	8	8	8
	p-hat300-2	300	21928	48.9	25	25	25
	p-hat300-3	300	33390	74.5	36	**36**	33
	p-hat500-1	500	31569	25.3	9	9	9
	p-hat500-2	500	62946	50.5	36	**36**	34
	p-hat500-3	500	93800	75.2	50	48	48
	p-hat700-1	700	60999	24.9	11	9	9
	p-hat700-2	700	121728	49.8	44	43	43
	p-hat700-3	700	183010	74.8	62	60	60
	p-hat1000-1	1000	122253	24.5	10	8	**10**
	p-hat1000-2	1000	244799	49.0	46	**46**	45
	p-hat1000-3	1000	371746	74.4	68	**64**	63
	p-hat1500-1	1500	284923	25.3	12	10	10
	p-hat1500-2	1500	568960	50.6	65	**65**	62
	p-hat1500-3	1500	847244	75.4	94	**92**	91
San	san200-0.7-1	200	13930	70.0	30	15	**17**
	san200-0.7-2	200	13930	70.0	18	12	12
	san200-0.9-1	200	17910	90.0	70	**49**	46
	san200-0.9-2	200	17910	90.0	60	39	**47**
	san200-0.9-3	200	17910	90.0	44	**35**	34
	san400-0.5-1	400	39900	50.0	13	7	7
	san400-0.7-1	400	55860	70.0	40	20	20
	san400-0.7-2	400	55860	70.0	30	15	**16**
	san400-0.7-3	400	55860	70.0	22	12	12
	san400-0.9-1	400	71820	90.0	100	50	**69**
	san1000	1000	250500	50.2	15	8	8
Sanr	sanr200-0.7	200	13868	69.7	18	17	17
	sanr200-0.9	200	17863	89.8	42	**40**	37
	sanr400-0.5	400	39984	50.1	13	12	12
	sanr400-0.7	400	55869	70.0	22	19	**20**

CHAPTER 6

Conclusions and Further Research Work

In this chapter, we summarize our theoretical approach and methodology that have been used and tested in this thesis. We also discuss future research directions for the tensor eigenvalue problems and tensor optimization problems.

6.1 Summary of the results

In this thesis, we have conducted a systematic study on the tensor eigenvalue problems and tensor optimization problems which arise widely in many application areas.

In Chapter 3, we developed an efficient and effective power type algorithm (i.e., Algorithm 1 in [200]), which is able to calculate the spectral radius of any nonnegative super-symmetric tensor. This new computational algorithm is an improvement over the NQZ method (i.e., Algorithm 3.1), which consists of several important properties. Firstly, it can find the spectral radius of weakly reducible nonnegative symmetric tensor without partitioning the tensor. Next, it converges under mild conditions for any symmetric nonnegative tensor, (cf. In particular, algorithm 1 in [200] produces a sequence of numbers that always converges monotonically to the spectral radius of the tensor. Finally, it is useful for finding the largest eigenvalue of large-size tensors and for testing the positive definiteness of a Z-tensor particularly when the size of the tensor is large. Numerical results reported in Table 1 and Figure 1 in [200] show that our new Algorithm 1 in [200] outperforms Algorithm 3.6.

In Chapter 4, we considered the following two polynomial optimization problems with nonnegative coefficients:

(1) Homogeneous polynomial problem over a single spherical constraint

$$(P1) \qquad \max_{x\in\Re^n} \sum_{i_1,i_2,\ldots,i_d=1}^{n} a_{i_1\, i_2\ldots i_d} x_{i_1} x_{i_2} \cdots x_{i_d}$$
$$\text{s.t.} \qquad \|x\|_2 = 1,$$

where $a_{i_1\, i_2\ldots i_d} \geq 0$.

(2) Bi-quadratic homogeneous polynomial optimization problem over a product of spherical constraints

$$(P2) \qquad \max_{x\in\Re^n, y\in\Re^m} \sum_{i,j=1}^{n}\sum_{k,l=1}^{m} b_{ijkl} x_i x_j y_k y_l$$
$$\text{s.t.} \qquad \|x\|_2 = 1, \quad \|y\|_2 = 1,$$

where $b_{ijkl} \geq 0$.

In this chapter, we presented some new approximation bounds for problems (P1) and (P2). In addition, we proposed polynomial-time approximation algorithms with new approximation bounds for those optimization problems. Numerical results reported in Section 4.3 show that the proposed approximation algorithms are efficient in terms of speed and efficiency and may produce global solutions for some testing problems.

In Chapter 5, we systematically studied the maximum clique problem which plays an important role in graph theory. One of the equivalent reformulations of the maximum clique problem is the optimization problem (P1) discussed in Chapter 4. We proposed an alternating direction method (ADM) to solve this reformulated optimization problem and reported our extensive computational results on 66 DIMACS benchmark graphs with different dimensions. Our numerical results show that the proposed method is encouraging.

6.2 Directions for further research

In this section, we will give some research directions that are worthy to investigate in future.

1. In Chapter 3, we proposed an algorithm for computing the spectral radii of nonnegative tensors and we proved that it is convergent for any symmetric nonnegative tensors. However, when the tensors are not symmetric, we do

not know if the proposed algorithm is convergent. This is worthy to be investigated in future.

2. In Chapter 4, we proposed some approximation algorithms for nonnegative polynomial optimization problems. Numerical results reported in this chapter show that the proposed algorithms are efficient and may produce global solutions for some testing problems. For these algorithms, an important issue to be investigated in future is how to choose a "good" starting point. To this end, we may try to find a way to generate a set of points on the unit sphere. We expect this set of points will be evenly distributed on the unit sphere. Among these points, we will choose a "good" starting point. In our future research, we will study how to generate this set of points.

3. In Chapter 5, we studied the maximum clique problem (MCP) and proposed an alternating direction method (ADM) to solve the MCP. However, like the approximation algorithms in Chapter 4, the ADM algorithm proposed in Chapter 5 may not lead to a global solution. In our future research, we will investigate some global algorithms for the MCP by combining these local algorithms with some global strategies used in [41]. Furthermore, we will consider solving the MCP using combinatorial optimization techniques and compare with ADM and the penalty variation as well.

Bibliography

[1] R. Aharoni, P. Erdős, and N. Linial. Optima of dual integer linear programs. *Combinatorica*, 8(1):13–20, 1988.

[2] A. Amin and S. Hakimi. Upper bounds on the order of a clique of a graph. *SIAM Journal on Applied Mathematics*, 22(4):569–573, 1972.

[3] R. Amir, S. Sahi, M. Shubik, and S. Yao. A strategic market game with complete markets. *Journal of economic theory*, 51(1):126–143, 1990.

[4] B. Anderson, N. Bose, and E. Jury. Output feedback stabilization and related problems-solution via decision methods. *IEEE Transactions on Automatic control*, 20(1):53–66, 1975.

[5] W. E. Arnoldi. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of applied mathematics*, 9(1):17–29, 1951.

[6] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. On the intractability of approximation problems. *Preliminary Draft, University of Rochester, NY*, 1992.

[7] L. M. Ausubel, P. Cramton, P. Milgrom, et al. The clock-proxy auction: A practical combinatorial auction design. *Combinatorial auctions*, 115:138, 2006.

[8] S. M. Baas, M. Bonvanie, and A. J. Verschoor. A relaxation method for the set packing problem using polyhedral characteristics. 1988.

[9] V. Balan and N. Perminov. Applications of resultants in the spectral m-root framework. *Applied Sciences*, 12:20–29, 2010.

[10] E. Balas and C. S. Yu. Finding a maximum clique in an arbitrary graph. *SIAM Journal on Computing*, 15(4):1054–1068, 1986.

[11] B. Balasundaram, S. Butenko, and I. V. Hicks. Clique relaxations in social network analysis: The maximum k-plex problem. *Operations Research*, 59(1):133–142, 2011.

[12] D. H. Ballard and C. M. Brown. Computer vision, article, 4 pages prentice-hall. *Englewood Cliffs, New Jersey, believed to be published more than one year prior to the filing date of the present application*, 1982.

[13] U. Benlic and J.-K. Hao. Breakout local search for maximum clique problems. *Computers & Operations Research*, 40(1):192–206, 2013.

[14] A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

[15] P. Berman and A. Pelc. Distributed probabilistic fault diagnosis for multiprocessor systems. In *Fault-Tolerant Computing, 1990. FTCS-20. Digest of Papers., 20th International Symposium*, pages 340–346. IEEE, 1990.

[16] P. Berman and G. Schnitger. On the complexity of approximating the independent set problem. *STACS 89*, pages 256–268, 1989.

[17] P. Berman and G. Schnitger. On the complexity of approximating the independent set problem. *Information and Computation*, 96(1):77–94, 1992.

[18] I. Bomze, M. Pelillo, and R. Giacomini. Evolutionary approach to the maximum clique problem: Empirical evidence on a larger scale. In *Developments in global optimization*, pages 95–108. Springer, 1997.

[19] I. M. Bomze. Evolution towards the maximum clique. *Journal of Global Optimization*, 10(2):143–164, 1997.

[20] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo. The maximum clique problem. In *Handbook of combinatorial optimization*, pages 1–74. Springer, 1999.

[21] I. M. Bomze, M. Budinich, M. Pelillo, and C. Rossi. Annealed replication: a new heuristic for the maximum clique problem. *Discrete Applied Mathematics*, 121(1):27–49, 2002.

[22] I. M. Bomze, L. Grippo, and L. Palagi. Unconstrained formulation of standard quadratic optimization problems. *Top*, 20(1):35–51, 2012.

[23] J. A. Bondy, U. S. R. Murty, et al. *Graph theory with applications*, volume 290. Citeseer, 1976.

[24] N. Bose and P. Kamat. Algorithm for stability test of multidimensional filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 22(5):307–314, 1974.

[25] A. Brauer et al. The theorems of ledermann and ostrowski on positive matrices. *Duke Mathematical Journal*, 24(2):265–274, 1957.

[26] A. Brauer and I. C. Gentry. Bounds for the greatest characteristic root of an irreducible nonnegative matrix. *Linear Algebra and Its Applications*, 8(2):105–107, 1974.

[27] M. Brockington and J. C. Culberson. Camouflaging independent sets in quasi-random graphs. *Cliques, coloring, and satisfiability: second DIMACS implementation challenge*, 26:75–88, 1996.

[28] L. Brotcorne, G. Laporte, and F. Semet. Fast heuristics for large scale covering-location problems. *Computers & Operations Research*, 29(6):651–665, 2002.

[29] A. E. Brouwer, L. B. Shearer, N. Sloane, et al. A new table of constant weight codes. In *IEEE Trans Inform Theory*. Citeseer, 1990.

[30] M. Budinich. Exact bounds on the order of the maximum clique of a graph. *Discrete Applied Mathematics*, 127(3):535–543, 2003.

[31] S. R. Bulo and M. Pelillo. A continuous characterization of maximal cliques in k-uniform hypergraphs. In *International Conference on Learning and Intelligent Optimization*, pages 220–233. Springer, 2007.

[32] S. R. Bulò and M. Pelillo. A generalization of the motzkin–straus theorem to hypergraphs. *Optimization Letters*, 3(2):287–295, 2009.

[33] D. Cartwright and B. Sturmfels. The number of eigenvalues of a tensor. *Linear algebra and its applications*, 438(2):942–952, 2013.

[34] M. Chams, A. Hertz, and D. De Werra. Some experiments with simulated annealing for coloring graphs. *European Journal of Operational Research*, 32(2):260–266, 1987.

[35] K. Chang, L. Qi, and T. Zhang. A survey on the spectral theory of nonnegative tensors. *Numerical Linear Algebra with Applications*, 20(6):891–912, 2013.

[36] K. Chang, L. Qi, and G. Zhou. Singular values of a real rectangular tensor. *Journal of mathematical analysis and applications*, 370(1):284–294, 2010.

[37] K. Chang and T. Zhang. On the uniqueness and non-uniqueness of the positive z-eigenvector for transition probability tensors. *Journal of Mathematical Analysis and Applications*, 408(2):525–540, 2013.

[38] K.-C. Chang, K. Pearson, and T. Zhang. On eigenvalue problems of real symmetric tensors. *Journal of Mathematical Analysis and Applications*, 350(1):416–422, 2009.

[39] K.-C. Chang, K. Pearson, T. Zhang, et al. Perron-frobenius theorem for nonnegative tensors. *Commun. Math. Sci*, 6(2):507–520, 2008.

[40] K.-C. Chang, K. J. Pearson, and T. Zhang. Primitivity, the convergence of the nqz method, and the largest eigenvalue for nonnegative tensors. *SIAM Journal on Matrix Analysis and Applications*, 32(3):806–819, 2011.

[41] B. Chen, S. He, Z. Li, and S. Zhang. Maximum block improvement and polynomial optimization. *SIAM Journal on Optimization*, 22(1):87–107, 2012.

[42] H. Chen, Y. Chen, G. Li, and L. Qi. A semidefinite program approach for computing the maximum eigenvalue of a class of structured tensors and its applications in hypergraphs and copositivity test. *Numerical Linear Algebra with Applications*, 25(1):e2125, 2018.

[43] L. Chen, L. Han, H. Yin, and L. Zhou. A homotopy method for computing the largest eigenvalue of an irreducible nonnegative tensor. *arXiv preprint arXiv:1701.07534*, 2017.

[44] L. Chen, L. Han, and L. Zhou. Computing tensor eigenvalues via homotopy methods. *SIAM Journal on Matrix Analysis and Applications*, 37(1):290–319, 2016.

[45] N. Chiba, T. Nishizeki, and N. Saito. An approximation algorithm for the maximum independent set problem on planar graphs. *SIAM Journal on Computing*, 11(4):663–675, 1982.

[46] L. Collatz. Einschließungssatz für die charakteristischen zahlen von matrizen. *Mathematische Zeitschrift*, 48(1):221–226, 1942.

[47] P. Comon. Tensor decompositions. *Mathematics in Signal Processing V*, pages 1–24, 2002.

[48] D. A. Cox, J. Little, and D. O'shea. *Using algebraic geometry*, volume 185. Springer Science & Business Media, 2006.

[49] P. Crescenzi, C. Fiorini, and R. Silvestri. A note on the approximation of the max clique problem. *Information Processing Letters*, 40(1):1–5, 1991.

[50] C.-F. Cui, Y.-H. Dai, and J. Nie. All real eigenvalues of symmetric tensors. *SIAM Journal on Matrix Analysis and Applications*, 35(4):1582–1601, 2014.

[51] E. De Klerk. The complexity of optimizing over a simplex, hypercube or sphere: a short survey. *Central European Journal of Operations Research*, 16(2):111–125, 2008.

[52] F. Della Croce, A. Grosso, and M. Locatelli. A heuristic approach for the max–min diversity problem based on max-clique. *Computers & Operations Research*, 36(8):2429–2433, 2009.

[53] R. Diamond. A note on the rotational superposition problem. *Acta Crystallographica Section A: Foundations of Crystallography*, 44(2):211–216, 1988.

[54] G. A. Dirac. On rigid circuit graphs. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 25, pages 71–76. Springer, 1961.

[55] U. Dorndorf, F. Jaehn, and E. Pesch. Modelling robust flight-gate scheduling as a clique partitioning problem. *Transportation Science*, 42(3):292–301, 2008.

[56] L. Elsner. Inverse iteration for calculating the spectral radius of a non-negative irreducible matrix. *Linear algebra and its applications*, 15(3):235–242, 1976.

[57] T. Etzion and P. R. Ostergard. Greedy and heuristic algorithms for codes and colorings. *IEEE Transactions on Information Theory*, 44(1):382–388, 1998.

[58] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating clique is almost np-complete. In *Foundations of Computer Science, 1991. Proceedings., 32nd Annual Symposium on*, pages 2–12. IEEE, 1991.

[59] M. Fiedler. Eigenvalues of nonnegative symmetric matrices. *Linear Algebra and its Applications*, 9:119–142, 1974.

[60] A. Figueroa, J. Borneman, and T. Jiang. Clustering binary fingerprint vectors with missing values for dna array data analysis. *Journal of Computational biology*, 11(5):887–901, 2004.

[61] J. G. Francis. The qr transformation a unitary analogue to the lr transformationpart 1. *The Computer Journal*, 4(3):265–271, 1961.

[62] J. G. Francis. The qr transformationpart 2. *The Computer Journal*, 4(4):332–345, 1962.

[63] S. Friedland, S. Gaubert, and L. Han. Perron–frobenius theorem for nonnegative multilinear forms and extensions. *Linear Algebra and its Applications*, 438(2):738–749, 2013.

[64] F. G. Frobenius. Über matrizen aus nicht negativen elementen. 1912.

[65] I. M. Gelfand, M. Kapranov, and A. Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Springer Science & Business Media, 2008.

[66] A. Ghosh, E. P. Tsigaridas, M. Descoteaux, P. Comon, B. Mourrain, and R. Deriche. A polynomial based approach to extract the maxima of an antipodally symmetric spherical function and its application to extract fiber directions from the orientation distribution function in diffusion mri. In *11th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pages pages–237, 2008.

[67] L. E. Gibbons, D. W. Hearn, and P. M. Pardalos. *A continuous based heuristic for the maximum clique problem*. University of Florida. Department of Industrial & Systems Engineering, 1994.

[68] L. E. Gibbons, D. W. Hearn, P. M. Pardalos, and M. V. Ramana. Continuous characterizations of the maximum clique problem. *Mathematics of Operations Research*, 22(3):754–768, 1997.

[69] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. Baltimore: Johns Hopkins University Press, 2012.

[70] P. Hansen, N. Mladenović, and D. Urošević. Variable neighborhood search for the maximum clique. *Discrete Applied Mathematics*, 145(1):117–125, 2004.

[71] M. A. Hasan and A. A. Hasan. A procedure for the positive definiteness of forms of even order. *IEEE Transactions on Automatic Control*, 41(4):615–617, 1996.

[72] J. Hasselberg, P. M. Pardalos, and G. Vairaktarakis. Test case generators and computational results for the maximum clique problem. *Journal of Global Optimization*, 3(4):463–482, 1993.

[73] J. D. Hauenstein, L. Oeding, G. Ottaviani, and A. J. Sommese. Homotopy techniques for tensor decomposition and perfect identifiability. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2016.

[74] S. He, Z. Li, and S. Zhang. Approximation algorithms for homogeneous polynomial optimization with quadratic constraints. *Mathematical Programming*, 125(2):353–383, 2010.

[75] D. Henrion, J.-B. Lasserre, and J. Löfberg. Gloptipoly 3: moments, optimization and semidefinite programming. *Optimization Methods & Software*, 24(4-5):761–779, 2009.

[76] C. J. Hillar and L.-H. Lim. Most tensor problems are np-hard. *Journal of the ACM (JACM)*, 60(6):45, 2013.

[77] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 1990.

[78] S. Hu, Z. Huang, and L. Qi. Strictly nonnegative tensors and nonnegative tensor partition. *Science China Mathematics*, 57(1):181–195, 2014.

[79] S. Hu, Z.-H. Huang, and L. Qi. Finding the spectral radius of a nonnegative tensor. *arXiv preprint arXiv:1111.2138*, 2011.

[80] S. Hu, Z.-H. Huang, and L. Qi. Finding the extreme z-eigenvalues of tensors via a sequential semidefinite programming method. *Numerical Linear Algebra with Applications*, 20(6):972–984, 2013.

[81] S. Hu and L. Qi. Algebraic connectivity of an even uniform hypergraph. *Journal of Combinatorial Optimization*, pages 1–16, 2012.

[82] S. Hu and L. Qi. Convergence of a second order markov chain. *Applied Mathematics and Computation*, 241:183–192, 2014.

[83] S. Hu, L. Qi, and G. Zhang. The geometric measure of entanglement of pure states with nonnegative amplitudes and the spectral theory of nonnegative tensors. *arXiv preprint arXiv:1203.3675*, 2012.

[84] A. Jagota, L. Sanchis, and R. Ganesan. Approximately solving maximum clique using neural networks and related heuristics. *Cliques, Coloring, and Satisfiability*, 26:169–204, 1996.

[85] Z. Jia, W.-W. Lin, and C.-S. Liu. A positivity preserving inexact noda iteration for computing the smallest eigenpair of a large irreducible m-matrix. *Numerische Mathematik*, 130(4):645–679, 2015.

[86] B. Jiang, S. Ma, and S. Zhang. Alternating direction method of multipliers for real and complex polynomial optimization models. *Optimization*, 63(6):883–898, 2014.

[87] D. S. Johnson and M. A. Trick. *Cliques, coloring, and satisfiability: second DIMACS implementation challenge, October 11-13, 1993*, volume 26. American Mathematical Soc., 1996.

[88] K. Katayama, A. Hamamoto, and H. Narihisa. An effective local search for the maximum clique problem. *Information Processing Letters*, 95(5):503–511, 2005.

[89] 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(3):863–884, 2002.

[90] T. G. Kolda and J. R. Mayo. Shifted power method for computing tensor eigenpairs. *SIAM Journal on Matrix Analysis and Applications*, 32(4):1095–1124, 2011.

[91] W. Ku. Explicit criterion for the positive definiteness of a general quartic form. *IEEE Transactions on Automatic Control*, 10(3):372–373, 1965.

[92] V. N. Kublanovskaya. On some algorithms for the solution of the complete eigenvalue problem. *USSR Computational Mathematics and Mathematical Physics*, 1(3):637–657, 1962.

[93] Y.-C. Kuo, W.-W. Lin, and C.-S. Liu. Continuation methods for computing z-/h-eigenpairs of nonnegative tensors. *arXiv preprint arXiv:1702.05841*, 2017.

[94] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[95] W. W. Leontief. *Input-output economics*, volume 2. JSTOR, 1951.

[96] A.-M. Li, L. Qi, and B. Zhang. E-characteristic polynomials of tensors. *arXiv preprint arXiv:1208.1607*, 2012.

[97] W. Li and M. K. Ng. On the limiting probability distribution of a transition probability tensor. *Linear and Multilinear Algebra*, 62(3):362–385, 2014.

[98] Y. Li, Q. Yang, and Y. Yang. A new definition of geometric multiplicity of eigenvalues of tensors and some results based on it. *Frontiers of Mathematics in China*, 10(5):1123, 2015.

[99] Z. Li, A. Uschmajew, and S. Zhang. On convergence of the maximum block improvement method. *SIAM Journal on Optimization*, 25(1):210–233, 2015.

[100] L. Lim. Multilinear pagerank: measuring higher order connectivity in linked objects. *The Internet: Today and Tomorrow*, 2005.

[101] L.-H. Lim. Singular values and eigenvalues of tensors: a variational approach. In *Computational Advances in Multi-Sensor Adaptive Processing, 2005 1st IEEE International Workshop on*, pages 129–132. IEEE, 2005.

[102] L.-H. Lim. *Foundations of numerical multilinear algebra: decomposition and approximation of tensors*. Stanford University, 2007.

[103] L.-H. Lim. Tensors and hypermatrices. *Handbook of Linear Algebra, 2nd Ed., CRC Press, Boca Raton, FL*, pages 231–260, 2013.

[104] C. Ling, J. Nie, L. Qi, and Y. Ye. Biquadratic optimization over unit spheres and semidefinite programming relaxations. *SIAM Journal on Optimization*, 20(3):1286–1310, 2009.

[105] C.-S. Liu, C.-H. Guo, and W.-W. Lin. A positivity preserving inverse iteration for finding the perron pair of an irreducible nonnegative third order tensor. *SIAM Journal on Matrix Analysis and Applications*, 37(3):911–932, 2016.

[106] C.-S. Liu, C.-H. Guo, and W.-W. Lin. Newton–noda iteration for finding the perron pair of a weakly irreducible nonnegative tensor. *Numerische Mathematik*, 137(1):63–90, 2017.

[107] Y. Liu, G. Zhou, and N. F. Ibrahim. An always convergent algorithm for the largest eigenvalue of an irreducible nonnegative tensor. *Journal of Computational and Applied Mathematics*, 235(1):286–292, 2010.

[108] L. Lovász. Graph theory and integer programming. *Annals of Discrete Mathematics*, 4:141–158, 1979.

[109] Z.-Q. Luo and S. Zhang. A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints. *SIAM Journal on Optimization*, 20(4):1716–1736, 2010.

[110] C. Mannino and E. Stefanutti. An augmentation algorithm for the maximum weighted stable set problem. *Computational Optimization and Applications*, 14(3):367–381, 1999.

[111] D. Maringer and P. Parpas. Global optimization of higher order moments in portfolio selection. *Journal of Global optimization*, 43(2-3):219–230, 2009.

[112] A. Massaro, M. Pelillo, and I. M. Bomze. A complementary pivoting approach to the maximum weight clique problem. *SIAM Journal on Optimization*, 12(4):928–948, 2002.

[113] R. G. Michael and S. J. David. Computers and intractability: a guide to the theory of np-completeness. *WH Free. Co., San Fr*, pages 90–91, 1979.

[114] H. Minc. Nonnegative matrices, 1988.

[115] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of turán. *Canad. J. Math*, 17(4):533–540, 1965.

[116] B. Mourrain and J. P. Pavone. Subdivision methods for solving polynomial equations. *Journal of Symbolic Computation*, 44(3):292–306, 2009.

[117] B. Mourrain and P. Trebuchet. Generalized normal forms and polynomial system solving. In *Proceedings of the 2005 international symposium on Symbolic and algebraic computation*, pages 253–260. ACM, 2005.

[118] A. S. Murthy, G. Parthasarathy, and V. Sastry. Clique finding-a genetic approach. In *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on*, pages 18–21. IEEE, 1994.

[119] Y. Nesterov et al. *Random walk in a simplex and quadratic optimization over convex polytopes*. CORE, 2003.

[120] M. Ng, L. Qi, and G. Zhou. Finding the largest eigenvalue of a nonnegative tensor. *SIAM Journal on Matrix Analysis and Applications*, 31(3):1090–1099, 2009.

[121] Q. Ni and L. Qi. A quadratically convergent algorithm for finding the largest eigenvalue of a nonnegative homogeneous polynomial map. *Journal of Global Optimization*, 61(4):627–641, 2015.

[122] Q. Ni, L. Qi, and F. Wang. An eigenvalue method for testing positive definiteness of a multivariate form. *IEEE Transactions on Automatic Control*, 53(5):1096–1107, 2008.

[123] J. Nie. Generating polynomials and symmetric tensor decompositions. *Foundations of Computational Mathematics*, 17(2):423–465, 2017.

[124] J. Nie and L. Wang. Semidefinite relaxations for best rank-1 tensor approximations. *SIAM Journal on Matrix Analysis and Applications*, 35(3):1155–1179, 2014.

[125] T. Noda. Note on the computation of the maximal eigenvalue of a nonnegative irreducible matrix. *Numerische Mathematik*, 17(5):382–386, 1971.

[126] R. D. Nussbaum. *Hilbert's projective metric and iterated nonlinear maps*, volume 391. American Mathematical Soc., 1988.

[127] H. Ogawa. Labeled point pattern matching by delaunay triangulation and maximal cliques. *Pattern Recognition*, 19(1):35–40, 1986.

[128] A. Ostrowski. Bounds for the greatest latent root of a positive matrix. *Journal of the London Mathematical Society*, 1(2):253–256, 1952.

[129] C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *Journal of computer and system sciences*, 43(3):425–440, 1991.

[130] P. Pardalos and Y. Li. Global quadratic optimization and the maximum clique problem. *Ann. Discr. Math., submitted*, 1991.

[131] P. M. Pardalos. Continuous approaches to discrete optimization problems. In *Nonlinear optimization and applications*, pages 313–325. Springer, 1996.

[132] P. M. Pardalos and A. Phillips. A global optimization approach for solving the maximum clique problem. *International Journal of Computer Mathematics*, 33(3-4):209–216, 1990.

[133] B. N. Parlett. *The symmetric eigenvalue problem*, volume 20. siam, 1998.

[134] J. Pattillo, N. Youssef, and S. Butenko. Clique relaxation models in social network analysis. *Handbook of Optimization in Complex Networks*, pages 143–162, 2012.

[135] M. Pavan and M. Pelillo. Generalizing the motzkin-straus theorem to edge-weighted graphs, with applications to image segmentation. In *EMMCVPR*, pages 485–500. Springer, 2003.

[136] K. Pearson. Essentially positive tensors. *Int. J. Algebra*, 4:421–427, 2010.

[137] K. Pearson. Primitive tensors and convergence of an iterative process for the eigenvalues of a primitive tensor. *arXiv preprint arXiv:1004.2423*, 2010.

[138] K. J. Pearson and T. Zhang. On spectral hypergraph theory of the adjacency tensor. *Graphs and Combinatorics*, 30(5):1233–1248, 2014.

[139] M. Pelillo. Relaxation labeling networks for the maximum clique problem. *J. Artif. Neural Networks*, 2(4):313–328, 1995.

[140] M. Pelillo and A. Jagota. Feasible and infeasible maxima in a quadratic program for maximum clique. *J. Artif. Neural Networks*, 2(4):411–420, 1995.

[141] O. Perron. Zur theorie der matrices. *Mathematische Annalen*, 64(2):248–263, 1907.

[142] A. Popanconesi and D. Ranjan. Quantifiers and approximation. In *Structure in Complexity Theory Conference, 1990, Proceedings., Fifth Annual*, page 244. IEEE, 1990.

[143] S. Prajna, A. Papachristodoulou, and P. A. Parrilo. Introducing sostools: A general purpose sum of squares programming solver. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 1, pages 741–746. IEEE, 2002.

[144] L. Qi. Eigenvalues of a real supersymmetric tensor. *Journal of Symbolic Computation*, 40(6):1302–1324, 2005.

[145] L. Qi. Eigenvalues and invariants of tensors. *Journal of Mathematical Analysis and Applications*, 325(2):1363–1377, 2007.

[146] L. Qi. The minimum hartree value for the quantum entanglement problem. *arXiv preprint arXiv:1202.2983*, 2012.

[147] L. Qi. Symmetric nonnegative tensors and copositive tensors. *Linear Algebra and its Applications*, 439(1):228–238, 2013.

[148] L. Qi, W. Sun, and Y. Wang. Numerical multilinear algebra and its applications. *Frontiers of Mathematics in China*, 2(4):501–526, 2007.

[149] L. Qi and K. L. Teo. Multivariate polynomial minimization and its application in signal processing. *Journal of Global Optimization*, 26(4):419–433, 2003.

[150] L. Qi, F. Wang, and Y. Wang. Z-eigenvalue methods for a global polynomial optimization problem. *Mathematical Programming*, 118(2):301–316, 2009.

[151] L. Qi, Y. Wang, and E. X. Wu. D-eigenvalues of diffusion kurtosis tensors. *Journal of Computational and Applied Mathematics*, 221(1):150–157, 2008.

[152] L. Qi, G. Yu, and E. X. Wu. Higher order positive semidefinite diffusion tensor imaging. *SIAM Journal on Imaging Sciences*, 3(3):416–433, 2010.

[153] M. Rajesh Kannan, N. Shaked-Monderer, and A. Berman. On weakly irreducible nonnegative tensors and interval hull of some classes of tensors. *Linear and Multilinear Algebra*, 64(4):667–679, 2016.

[154] Y. Saad. Variations on arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear algebra and its applications*, 34:269–295, 1980.

[155] Y. Saad. *Numerical methods for large eigenvalue problems*. Manchester University Press, New York, 1992.

[156] S. Sahi. Logarithmic convexity of perron-frobenius eigenvectors of positive matrices. *Proceedings of the American Mathematical Society*, 118(4):1035–1036, 1993.

[157] S. Sahi. A note on the resolvent of a nonnegative matrix and its applications. *Linear Algebra and Its Applications*, 432(10):2524–2528, 2010.

[158] L. A. Sanchis. Test case construction for the vertex cover. In *Computational Support for Discrete Mathematics: DIMACS Workshop, March 12-14, 1992*, volume 15, page 315. American Mathematical Soc., 1994.

[159] S. E. Schaeffer. Graph clustering. *Computer science review*, 1(1):27–64, 2007.

[160] T. Schultz and H.-P. Seidel. Estimating crossing fibers: A tensor decomposition approach. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), 2008.

[161] E. Seneta. *Non-negative matrices and Markov chains*. Springer Science & Business Media, 2006.

[162] J.-Y. Shao. A general product of tensors with applications. *Linear Algebra and its applications*, 439(8):2350–2366, 2013.

[163] A. Singh and A. K. Gupta. A hybrid heuristic for the maximum clique problem. *Journal of Heuristics*, 12(1):5–22, 2006.

[164] P. Soriano and M. Gendreau. Tabu search algorithms for the maximum clique problem. *Cliques, Coloring, and Satisfiability*, 26:221–244, 1996.

[165] V. T. Sós and E. G. Straus. Extremals of functions on graphs with applications to graphs and hypergraphs. *Journal of Combinatorial Theory, Series B*, 32(3):246–257, 1982.

[166] P. St-Louis, B. Gendron, and J. Ferland. A penalty-evaporation heuristic in a decomposition method for the maximum clique problem. *Optimization Days*, 2003.

[167] G. W. Stewart. *Matrix algorithms volume 2: eigensystems*, volume 2. Siam, 2001.

[168] P. M. Van den Hof, C. Scherer, and P. S. Heuberger. *Model-Based Control:: Bridging Rigorous Theory and Advanced Technology*. Springer Science & Business Media, 2009.

[169] R. S. Varga and J. Gillis. Matrix iterative analysis. *Physics Today*, 16:52, 1963.

[170] F. Wang and L. Qi. Comments on" explicit criterion for the positive definiteness of a general quartic form". *IEEE Transactions on Automatic Control*, 50(3):416–418, 2005.

[171] V. Wei. A lower bound on the stability number of a simple graph. Technical report, Bell Laboratories Technical Memorandum 81-11217-9, Murray Hill, NJ, 1981.

[172] H. S. Wilf. The eigenvalues of a graph and its chromatic number. *J. London Math. Soc*, 42(1967):330, 1967.

[173] H. S. Wilf. Spectral bounds for the clique and independence numbers of graphs. *Journal of Combinatorial Theory, Series B*, 40(1):113–117, 1986.

[174] H. S. Wilf and H. S. Wilf. *Algorithms and complexity*, volume 986. Prentice-Hall Englewood Cliffs, NJ, 1986.

[175] S. Wimer, R. Y. Pinter, and J. A. Feldman. Optimal chaining of cmos transistors in a functional cell. *IEEE transactions on computer-aided design of integrated circuits and systems*, 6(5):795–801, 1987.

[176] R. Wood and M. O'Neill. Using the spectral radius to determine whether a leontief system has a unique positive solution. *Asia-Pacific Journal of Operational Research*, 19(2):233, 2002.

[177] R. Wood and M. O'Neill. An always convergent method for finding the spectral radius of an irreducible non-negative matrix. *ANZIAM Journal*, 45:474–485, 2004.

[178] R. Wood and M. O'Neill. A faster algorithm for identification of an m-matrix. *ANZIAM Journal*, 46:732–743, 2005.

[179] R. J. Wood and M. O'Neill. Finding the spectral radius of a large sparse non-negative matrix. *ANZIAM Journal*, 48:330–345, 2007.

[180] Q. Wu and J.-K. Hao. Coloring large graphs based on independent set extraction. *Computers & Operations Research*, 39(2):283–290, 2012.

[181] Q. Wu and J.-K. Hao. An effective heuristic algorithm for sum coloring of graphs. *Computers & Operations Research*, 39(7):1593–1600, 2012.

[182] Q. Wu and J.-K. Hao. A review on algorithms for maximum clique problems. *European Journal of Operational Research*, 242(3):693–709, 2015.

[183] Q. Wu and J.-K. Hao. Solving the winner determination problem via a weighted maximum clique heuristic. *Expert Systems with Applications*, 42(1):355–365, 2015.

[184] Q. Wu, J.-K. Hao, and F. Glover. Multi-neighborhood tabu search for the maximum weight clique problem. *Annals of Operations Research*, 196(1):611–634, 2012.

[185] Q. Yang and Y. Yang. Further results for perron–frobenius theorem for nonnegative tensors ii. *SIAM Journal on Matrix Analysis and Applications*, 32(4):1236–1250, 2011.

[186] Y. Yang and Q. Yang. Further results for perron–frobenius theorem for nonnegative tensors. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2517–2530, 2010.

[187] Y. Yang and Q. Yang. A note on the geometric simplicity of the spectral radius of nonnegative irreducible tensors. *arXiv preprint arXiv:1101.2479*, 2011.

[188] F. Zhang, B. Zhou, and L. Peng. Gradient skewness tensors and local illumination detection for images. *Journal of Computational and Applied Mathematics*, 237(1):663–671, 2013.

[189] L. Zhang and L. Qi. Linear convergence of an algorithm for computing the largest eigenvalue of a nonnegative tensor. *Numerical Linear Algebra with Applications*, 19(5):830–841, 2012.

[190] L. Zhang, L. Qi, and Y. Xu. Finding the largest eigenvalue of an irreducible nonnegative tensor nad linear convergence for weakly positive tensors. *Journal of Computational Mathematics*, 30(1):24–33, 2012.

[191] L. Zhang, L. Qi, and Y. Xu. Weakly positive tensors and linear convergence of the lzi algorithm. *J. Comput. Math*, 30:24–33, 2012.

[192] L. Zhang, L. Qi, and G. Zhou. m-tensors and some applications. *SIAM Journal on Matrix Analysis and Applications*, 35(2):437–452, 2014.

[193] T. Zhang. Existence of real eigenvalues of real tensors. *Nonlinear Analysis: Theory, Methods & Applications*, 74(8):2862–2868, 2011.

[194] X. Zhang, C. Ling, and L. Qi. The best rank-1 approximation of a symmetric tensor and related spherical optimization problems. *SIAM Journal on Matrix Analysis and Applications*, 33(3):806–821, 2012.

[195] X. Zhang, G. Zhou, L. Caccetta, and M. Alqahtani. Approximation algorithms for nonnegative polynomial optimization problems over unit spheres. *Frontiers of Mathematics in China*, 12(6):1409–1426, 2017.

[196] G. Zhou, L. Caccetta, and L. Qi. Convergence of an algorithm for the largest singular value of a nonnegative rectangular tensor. *Linear Algebra and its Applications*, 438(2):959–968, 2013.

[197] G. Zhou, L. Caccetta, K. L. Teo, and S.-Y. Wu. Nonnegative polynomial optimization over unit spheres and convex programming relaxations. *SIAM Journal on Optimization*, 22(3):987–1008, 2012.

[198] G. Zhou, L. Qi, and S.-Y. Wu. Efficient algorithms for computing the largest eigenvalue of a nonnegative tensor. *Frontiers of mathematics in China*, 8(1):155–168, 2013.

[199] G. Zhou, L. Qi, and S.-Y. Wu. On the largest eigenvalue of a symmetric nonnegative tensor. *Numerical Linear Algebra with Applications*, 20(6):913–928, 2013.

[200] G. Zhou, G. Wang, L. Qi, and M. Alqahtani. A fast algorithm for the spectral radii of weakly reducible nonnegative tensors. *Numerical Linear Algebra with Applications*, 25(2):e2134, 2018.