NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2013; **00**:1–28
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nla

The Z-eigenvalues of a Symmetric Tensor and its Application to Spectral Hypergraph Theory

Guoyin Li [1, *], Liqun Qi[2] and Gaohang Yu [3]

[1] *Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia.*
[2] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.*
[3] *School of Mathematics and Computer Sciences, Gannan Normal University, Ganzhou, 341000, China.*

SUMMARY

In this paper, using variational analysis and optimization techniques, we examine some fundamental analytic properties of Z-eigenvalues of a real symmetric tensor with even order. We first establish that the maximum Z-eigenvalue function is a continuous and convex function on the symmetric tensor space, and so, provide formulas of the convex conjugate function and ϵ-subdifferential of the maximum Z-eigenvalue function. Consequently, for an mth order n-dimensional tensor $\mathcal{A}$, we show that the normalized eigenspace associated with maximum Z-eigenvalue function is ρth-order Hölder stable at $\mathcal{A}$ with $\rho = \frac{1}{m(3m-3)^{n-1}-1}$. As a by-product, we also establish that the maximum Z-eigenvalue function is always at least ρth-order semismooth at $\mathcal{A}$. As an application, we introduce the characteristic tensor of a hypergraph and show that the maximum Z-eigenvalue function of the associated characteristic tensor provides a natural link for the combinatorial structure and the analytic structure of the underlying hypergraph. Finally, we establish a variational formula for the second largest Z-eigenvalue for the characteristic tensor of a hypergraph and use it to provide lower bounds for the bipartition width of a hypergraph. Some numerical examples are also provided to show how one can compute the largest/2nd-largest Z-eigenvalue of a medium size tensor, using polynomial optimization techniques and our variational formula. Copyright © 2013 John Wiley & Sons, Ltd.

Received ...

KEY WORDS: Symmetric Tensor; Maximum Z-eigenvalue; Semismoothness; Spectral Graph Theory; Characteristic Tensor; Polynomial Optimization

1. INTRODUCTION

An mth-order n-dimensional tensor $\mathcal{A}$ consists of n^m entries in real number:

$$\mathcal{A} = (\mathcal{A}_{i_1 i_2 \cdots i_m}), \;\; \mathcal{A}_{i_1 i_2 \cdots i_m} \in \mathbb{R}, \;\; 1 \le i_1, i_2, \cdots, i_m \le n. \tag{1.1}$$

We say a tensor $\mathcal{A}$ is symmetric if the value of $\mathcal{A}_{i_1 i_2 \cdots i_m}$ is invariant under any permutation of its index $\{i_1, i_2, \cdots, i_m\}$. Clearly, when $m = 2$, a symmetric tensor is nothing but a symmetric matrix. A symmetric tensor uniquely defines an mth degree homogeneous polynomial function f with real coefficient: for all $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$,

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

*Correspondence to: Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au (G. Li), liqun.qi@polyu.edu.hk (L. Qi), maghyu@163.com (G. Yu)

Over the last few years, there has been a realization that there is a reasonably complete and consistent theory of eigenvalues and singular values for tensors of higher order, proposed by Lim and Qi independently, that generalizes the theory of matrix eigenvalues and singular values in various manners and extent. Recently, numerical study on tensors also has attracted a lot of researchers due to its wide applications in polynomial optimization [35], hypergraph theory[5, 13], Higher order Markov chain [27], signal processing [40] and image science [42]. In particular, various efficiently numerical schemes have been proposed to find the low rank approximations of a tensor and the eigenvalues/eigenvectors of a tensor with specific structure (cf. [16, 17, 7, 8, 25, 33]).

Among the various definitions of an eigenvalue of a symmetric tensor, there are two particular interesting definitions called Z-eigenvalues and H-eigenvalues (see the definition later on). Recall that a tensor is said to be positive semidefinite if the corresponding homogeneous polynomial function the tensor uniquely determined always takes nonnegative values. As shown in [38], a tensor is positive semidefinite if and only if its Z-eigenvalues (resp. H-eigenvalues) are all nonnegative. So, the Z-eigenvalues and H-eigenvalues play an important role in determining whether a symmetric tensor is positive semidefinite or not. On the other hand, Z-eigenvalues and H-eigenvalues can be fundamentally different as investigated in [38]. For example, finding an H-eigenvalue of a symmetric tensor is equivalent to solving a homogeneous polynomial equation while calculating a Z-eigenvalue is equivalent to solving nonhomogeneous polynomial equations. Moreover, a diagonal symmetric tensor A has exactly n many H-eigenvalues and may have more than n Z-eigenvalues (for more details see [38]).

Very recently, in our preceding paper [23], we investigated the analytic properties of the maximum H-eigenvalue function of a symmetric tensor. In particular, we showed that, for an mth-order n-dimensional symmetric tensor $\mathcal{A}$, the H-maximum eigenvalue function is $\frac{1}{(2m-1)^n}$th-order semismooth at $\mathcal{A}$ when the geometric multiplicity of $\mathcal{A}$ is one. As an application, we proposed a generalized Newton method to solve the space tensor conic linear programming problem which arises in medical imaging area. Local convergence rate of this method was established by using the semismooth property of the maximum H-eigenvalue function. In this paper, we continue our study and examine the analytic properties of Z-eigenvalues. We first show that the maximum Z-eigenvalue function is continuous, convex and differentiable almost everywhere, extending the fundamental analytic properties of the maximum eigenvalue of a symmetric matrix. Then, we establish that the normalized eigenspace associated with maximum Z-eigenvalue function is ρth-order Holder stable at $\mathcal{A}$ with $\rho = \frac{1}{m(3m-3)^{n-1}-1}$. As a by-product, without the geometric multiplicity assumption, we also establish that the maximum Z-eigenvalue function is always at least ρth-order semismooth at $\mathcal{A}$. As an application, we introduce the characteristic tensor of a hypergraph and show that the maximum Z-eigenvalue function of the associated characteristic tensor provides a natural link for the combinatorial structure and the analytic structure of the underlying hypergraph. A variational formula for the second largest Z-eigenvalue for the characteristic tensor of a hypergraph is also provided.

The organization of this paper is as follows. We first fix the notations and collect some basic definitions in Section 2. In Section 3, by using the variational analysis techniques, we show that the maximum Z-eigenvalue function is continuous and convex, and hence differentiable almost everywhere. In particular, we obtain the formula for calculating the convex conjugate of and ϵ-convex subdifferential for the maximum Z-eigenvalue function. In Section 4, for an mth order n-dimensional tensor $\mathcal{A}$, we show that the normalized eigenspace associated with maximum Z-eigenvalue function is ρth-order Hölder stable at $\mathcal{A}$ with $\rho = \frac{1}{m(3m-3)^{n-1}-1}$. We also show that the maximum Z-eigenvalue function is always at least ρth-order semismooth at $\mathcal{A}$. Sufficient condition ensuring the strong semismoothness of the maximum Z-eigenvalue function is also provided. In Section 5, we introduce the characteristic tensor of a hypergraph and show that the maximum Z-eigenvalue function of the associated characteristic tensor provides a natural link for the combinatorial structure and the analytic structure of the underlying hypergraph. Moreover, we establish a variational formula for the second largest Z-eigenvalue for the characteristic tensor of a hypergraph. In Section 6, numerical examples are provided to show how one can compute the largest

DOI: 10.1002/nla

and the second largest Z-eigenvalue of a real symmetric tensor using polynomial optimization technique. Finally, we conclude our paper and present some future research topics in Section 7.

2. PRELIMINARIES

In this section, we fix the notations and collect some basic definitions and facts which we will use later on. Let X, Y be finite dimensional inner product spaces. We use $\mathbb{B}_X$ (resp. $\mathbb{B}_Y$) to denote the unit open ball in X (resp. Y). Denote the space of all linear map from X to Y by $L(X, Y)$. The norm of X is defined by $\|x\| = \sqrt{\langle x, x\rangle_X}$ for all $x \in X$ where $\langle \cdot, \cdot \rangle_X$ is the inner product in X. Consider a locally Lipschitz function $G : X \to Y$. By the Rademacher's Theorem, G is differentiable almost everywhere on X. Let D_G be the set consists of all the points where G is differentiable. Then, for any $x \in D_G$, the derivative of G, $\nabla G(x)$ exists. Denote

$$J_B G(x) = \{V \in L(X, Y) : V = \lim_{x_k \to x} \nabla G(x_k), x_k \in D_G\}.$$

Then, its Clarke's generalized Jacobian is defined by $J_C G(x) = \mathrm{conv} J_B G(x)$. In particular, if $Y = \mathbb{R}$ and $G = g$ where $g : X \to \mathbb{R}$ is locally Lipschitz, by identifying X^* as X, then the Clarke's generalized Jacobian reduces to the Clarke's subdifferential defined by

$$\partial_C g(x) = \{\xi \in X : \langle \xi, v\rangle_X \leq g^\circ(x; v) \text{ for all } v \in X\}.$$

where $\langle \cdot, \cdot \rangle_X$ is the inner product in X and $g^\circ(x; v)$ is the Clarke directional derivative of g at the point x in the direction v given by

$$g^\circ(x; v) = \limsup_{y \to x,\ t \downarrow 0} \frac{g(y + tv) - g(y)}{t}.$$

The Clarke subdifferential $\partial_C g(x)$ is a nonempty, convex and compact subset of X for each $x \in X$. Recall that g is convex on X if

$$g(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda g(x_1) + (1 - \lambda)g(x_2),\ \forall\, \lambda \in [0, 1] \text{ and } x_1, x_2 \in X.$$

For each $\epsilon \geq 0$, define the ϵ-convex subdifferential of g at x by

$$\partial_\epsilon g(x) := \{\xi \in X : \langle \xi, z - x\rangle_X \leq g(z) - g(x) + \epsilon \text{ for all } z \in X\}.$$

If $\epsilon = 0$, we simply call it convex subdifferential of g and denote it by $\partial g(x)$. If g is convex on X, then $\partial_C g(x) = \partial g(x)$ for all $x \in X$. The Fenchel conjugate function of a convex function g on X, is denoted by g^*, and is defined by

$$g^*(\xi) = \sup_{x \in X} \{\langle \xi, x\rangle_X - g(x)\} \text{ for all } \xi \in X.$$

An important property relating the Fenchel conjugate and the ϵ-subdifferntial is the so-called generalized Fenchel inequality (cf. [47, Theorem 2.4.2 (ii)])

$$g(x) + g^*(\xi) \leq \langle \xi, x\rangle_X + \epsilon \ \Leftrightarrow\ \xi \in \partial_\epsilon g(x).$$

We are now ready to state the definitions of semismooth functions and ρth-order semismooth functions.

Definition 2.1
Let $G : X \to Y$ be a locally Lipschitz and directionally differentiable function. Then, the function G is said to be semismooth at x if

$$G(x + \Delta x) - G(x) - V\Delta x = o(\|\Delta x\|),\ \forall\, V \in J_C G(x + \Delta x).$$

DOI: 10.1002/nla

Moreover, G is said to be ρth-order semismooth function at x for some $\rho \in (0,1]$ if

$$G(x+\Delta x) - G(x) - V\Delta x = O(\|\Delta x\|^{1+\rho}),\ \forall V \in J_C G(x+\Delta x).$$

In particular, if $\rho = 1$, we say G is strongly semismooth at x. We also say $G : X \to Y$ is a semismooth (resp. ρth-order semismooth, strongly semismooth) function if G is semismooth (resp. ρth-order semismooth, strongly semismooth) at x for all $x \in X$.

The concept of a semismooth function was originally given by Mifflin [30] when $Y = \mathbb{R}$. Later on, Qi and Sun [39] extended the definition to vector value functions and showed that semismooth functions play an important role in establishing the local convergence rate of the generalized Newton method for solving nonsmooth equations. From the definitions of the semismooth functions, it is clear that scalar multiplication and sums of semismooth (resp. ρth-order semismooth) functions are still semismooth (resp. ρth-order semismooth) functions. An important example of the strongly semismooth function is the eigenvalue function of a symmetric matrix [45]. The next result [45, Theorem 3.7] provides a convenient tool for proving ρth-order semismoothness.

Lemma 2.1
Suppose that $G : \mathbb{R}^n \to \mathbb{R}^m$ is locally Lipschitzian and directionally differentiable in a neighborhood of x. Let $p \in (0,1]$. Then, G is ρth-order semismooth if and only if for any $x + \Delta x \in D_G$,

$$G(x+\Delta x) - G(x) - \nabla G(x+\Delta x)\Delta x = O(\|\Delta x\|^{1+\rho}).$$

Next, we recall some basic definitions and facts of tensor and its eigenvalues. Let $n \in \mathbb{N}$ and let m be an even number. Consider

$$S = \{\mathcal{A} : \mathcal{A} \text{ is an } m\text{th-order } n\text{-dimensional symmetric tensor}\}.$$

Clearly, S is a vector space under the addition and multiplication defined as below: for any $t \in \mathbb{R}$, $\mathcal{A} = (\mathcal{A}_{i_1,\dots,i_m})_{1\le i_1,\dots,i_m\le n}$ and $\mathcal{B} = (\mathcal{B}_{i_1,\dots,i_m})_{1\le i_1,\dots,i_m\le n}$

$$\mathcal{A}+\mathcal{B} = (\mathcal{A}_{i_1,\dots,i_m} + \mathcal{B}_{i_1,\dots,i_m})_{1\le i_1,\dots,i_m\le n} \text{ and } t\mathcal{A} = (t\mathcal{A}_{i_1,\dots,i_m})_{1\le i_1,\dots,i_m\le n}.$$

For each $\mathcal{A}, \mathcal{B} \in S$, we define the inner product by

$$\langle \mathcal{A}, \mathcal{B}\rangle_S = \sum_{i_1,\dots,i_m=1}^{n} \mathcal{A}_{i_1,\dots,i_m}\mathcal{B}_{i_1,\dots,i_m}.$$

The corresponding norm is defined by $\|\mathcal{A}\|_S = (\langle \mathcal{A}, \mathcal{A}\rangle_S)^{1/2} = \big(\sum_{i_1,\dots,i_m=1}^{n} \mathcal{A}^2_{i_1,\dots,i_m}\big)^{1/2}$. The unit ball in S is denoted by $\mathbb{B}_S$. For a vector $x \in \mathbb{R}^n$, we use x_i to denotes its ith component. We use $x^{[m-1]}$ to denote a vector in $\mathbb{R}^n$ such that $x_i^{[m-1]} = (x_i)^{m-1}$. Moreover, for a vector $x \in \mathbb{R}^n$, we use x^m to denote the mth-order n-dimensional symmetric rank one tensor induced by x, i.e.,

$$(x^m)_{i_1\dots i_m} = x_{i_1}\dots x_{i_m},\ \forall\, i_1,\dots,i_m \in \{1,\dots,n\}.$$

Let $\mathcal{A} \in S$. By the tensor product (cf [40]), $\mathcal{A}x^m$ is a real number defined as

$$\mathcal{A}x^m = \sum_{i_1,\dots,i_m=1}^{n} \mathcal{A}_{i_1,\dots,i_m} x_{i_1}\dots x_{i_m} = \langle \mathcal{A}, x^m\rangle_S$$

and $\mathcal{A}x^{m-1}$ is a vector in $\mathbb{R}^n$ whose ith component is

$$\sum_{i_2\cdots i_m=1}^{n} A_{i\,i_2\cdots i_m} x_{i_2}\cdots x_{i_m}. \tag{2.2}$$

DOI: 10.1002/nla

Definition 2.2
Let $\mathcal{A}$ be an mth-order n-dimensional real symmetric tensor. We say $\lambda \in \mathbb{R}$ is a Z-eigenvalue of $\mathcal{A}$ and $x \neq 0, x \in \mathbb{R}^n$ is a Z-eigenvector corresponding to λ if (x, λ) satisfies

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

Moreover, $\lambda \in \mathbb{R}$ is an H-eigenvalue of $\mathcal{A}$ and $x \neq 0, x \in \mathbb{R}^n$ is an H-eigenvector corresponding to λ if (x, λ) satisfies

$$\mathcal{A}x^{m-1} = \lambda x^{[m-1]}.$$

This definition of Z-eigenvalues and H-eigenvalues was introduced by Qi in [38]. Independently, Lim [25] also gave the definitions via a variational approach and established an interesting Perron-Frobenius theorem for tensors with nonnegative entries. From [38] and [3], both Z-eigenvalues and H-eigenvalues for an even order symmetric tensor always exist. Moreover, from the definitions, we can see that finding an H-eigenvalue of a symmetric tensor is equivalent to solving a homogeneous polynomial equation while calculating a Z-eigenvalue is equivalent to solving nonhomogeneous polynomial equations. In general, the behaviors of Z-eigenvalues and H-eigenvalues can be quite different. For example, a diagonal symmetric tensor A has exactly n many H-eigenvalues and may have more than n many Z-eigenvalues (for more details see [38]). Recently, by reducing a symmetric tensor to a pseudo-canonical form, Qi, Wang and Wang [43] proposed a direct method for finding all the Z-eigenvalues in the case of order three and dimension three. More recently, Kolda and Mayo [18] provided a shifted power method for computing a Z-eigenvalue and its associated eigenvector for a symmetric tensor. For numerical methods of finding H-eigenvalues for tensors with nonnegative entries see [33]. Recently, a new method for finding the maximum Z-eigenvalue of a weakly nonnegative symmetric tensor using sum-of-squares programming problem is also proposed in [14].

3. THE MAXIMUM Z-EIGENVALUE FUNCTION

In this section, we examine the continuity and differentiability of the maximum Z-eigenvalue function. To do this, we first formally define the maximum Z-eigenvalue function. Since any real symmetric tensor with even order always has a Z-eigenvalue (cf [38, 3]), it then makes sense to define the maximum Z-eigenvalue function $\lambda_1^Z : S \to \mathbb{R}$ as follows:

$$\lambda_1^Z(\mathcal{A}) = \{\lambda \in \mathbb{R} : \lambda \text{ is the largest Z-eigenvalue of } \mathcal{A}\}.$$

We first recall the following simple lemma which will be useful for our later analysis. Its proof can be found in [25, 38]. However, for the completeness of the paper, we present the proof here.

Lemma 3.1
Let $\mathcal{A}$ be an mth-order n-dimensional real symmetric tensor where m is even. Then, we have

$$\lambda_1^Z(\mathcal{A}) = \max_{\|x\|=1} \mathcal{A}x^m.$$

Proof
Consider the following optimization problem (P)

$$\begin{array}{lll} (P) & \max_{x\in\mathbb{R}^n} & \mathcal{A}x^m \\ & \text{s.t.} & \|x\|^m = 1. \end{array}$$

Let $f(x) := \mathcal{A}x^m$ and $g(x) := (x^T x)^{\frac{m}{2}} = \|x\|^m$. Since f is continuous and the feasible set $\{x : g(x) = 1\}$ is compact, a global maximizer of (P) exists. Denote a maximizer of (P) by x_0. Clearly, $x_0 \neq 0$. Note that g is a homogeneous polynomial with degree m. The Euler identity implies that $\nabla g(x)^T x = mg(x)$. Thus, for any x with $g(x) = 1$, $\nabla g(x) \neq 0$. So, the standard KKT theory

DOI: 10.1002/nla

implies that there exists $\lambda_0 \in \mathbb{R}$ such that

$$m\mathcal{A}x_0^{m-1} - m\lambda_0\|x_0\|^{m-2}x_0 = \nabla f(x_0) - \lambda_0 \nabla g(x_0) = 0.$$

This implies that λ_0 is a real eigenvalue of $\mathcal{A}$ and so, $\lambda_0 \leq \lambda_1^Z(\mathcal{A})$. Note that $v(P) = \mathcal{A}x_0^m = x_0^T(\mathcal{A}x_0^{m-1}) = x_0^T(\lambda_0 x_0) = \lambda_0$, where $v(P)$ is the optimal value of (P). It follows that $v(P) \leq \lambda_1^Z(\mathcal{A})$, that is, $\max_{\|x\|_m=1} \mathcal{A}x^m \leq \lambda_1^Z(\mathcal{A})$. Finally, noting that, for any eigenvector u corresponds to $\lambda_1^Z(\mathcal{A})$ with $\|u\| = 1$, we have

$$\mathcal{A}u^m = u^T(\mathcal{A}u^{m-1}) = \lambda_1^Z(\mathcal{A})u^T u = \lambda_1^Z(\mathcal{A})\|u\|^2 = \lambda_1^Z(\mathcal{A}).$$

Thus, $\lambda_1^Z(\mathcal{A}) = \max_{\|x\|=1} \mathcal{A}x^m$, and so, the conclusion follows . □

Remark 3.1
Define the normalized eigenspace associated with $\lambda_1^Z(\mathcal{A})$ by $E_1^Z(\mathcal{A}) := \{u : \mathcal{A}u^{m-1} = \lambda_1^Z(\mathcal{A})u, \ \|u\| = 1\}$. From the proof of the Lemma 3.1, we see that

$$E_1^Z(\mathcal{A}) = \{u : \mathcal{A}u^m = \lambda_1^Z(\mathcal{A}), \ \|u\| = 1\}.$$

Next, we show that the maximum Z-eigenvalue function is continuous and convex.

Theorem 3.1
The function λ_1^Z is a continuous and convex function on S. Moreover, its conjugate $(\lambda_1^Z)^*$ can be calculated as

$$(\lambda_1^Z)^*(\mathcal{B}) = \begin{cases} 0, & \text{if} \quad \mathcal{B} \in \mathrm{conv}\{u^m : \|u\| = 1\}, \\ +\infty, & \text{else}, \end{cases}$$

where $\mathrm{conv}A$ denotes the convex hull of A, and is defined by

$$\mathrm{conv}A = \{\sum_{i=1}^{s} \mu_i a_i : \mu_i \geq 0, \sum_{i=1}^{s} \mu_i = 1, a_i \in A, s \in \mathbb{N}\}.$$

Proof
Let $T := \{u^m : \|u\| = 1\} \subseteq S$. Since $\lambda_1^Z(\mathcal{A}) = \max_{\|x\|=1} \mathcal{A}x^m$, we have $\lambda_1^Z(\mathcal{A}) = \max_{\mathcal{B}\in T}\langle \mathcal{B}, \mathcal{A}\rangle_S$. Note that $\mathcal{B} \mapsto \langle \mathcal{B}, \mathcal{A}\rangle_S$ is affine and the supremum of a series of affine functions is convex. It follows that λ_1^Z is a finite-valued convex function on S and so, is continuous and convex. From the definition, it can be verified that $\max_{\mathcal{B}\in T}\langle \mathcal{B}, \mathcal{A}\rangle_S = \max_{\mathcal{B}\in \mathrm{conv}T}\langle \mathcal{B}, \mathcal{A}\rangle_S$, and so,

$$\lambda_1^Z(\mathcal{A}) = \max_{\mathcal{B}\in \mathrm{conv}T} \langle \mathcal{B}, \mathcal{A}\rangle_S. \tag{3.3}$$

It follows that, for each $\mathcal{B} \in S$,

$$\begin{aligned}(\lambda_1^Z)^*(\mathcal{B}) = \sup_{\mathcal{A}\in S}\{\langle \mathcal{B}, \mathcal{A}\rangle_S - \lambda_1^Z(\mathcal{A})\} &= \sup_{\mathcal{A}\in S}\{\langle \mathcal{B}, \mathcal{A}\rangle_S - \lambda_1^Z(\mathcal{A})\} \\ &= \sup_{\mathcal{A}\in S}\{\langle \mathcal{B}, \mathcal{A}\rangle_S - \sup_{\mathcal{C}\in \mathrm{conv}T}\langle \mathcal{C}, \mathcal{A}\rangle\} \\ &= \sup_{\mathcal{A}\in S}\min_{\mathcal{C}\in \mathrm{conv}T}\{\langle \mathcal{B}, \mathcal{A}\rangle_S - \langle \mathcal{C}, \mathcal{A}\rangle_S\} \\ &= \min_{\mathcal{C}\in \mathrm{conv}T}\sup_{\mathcal{A}\in S}\{\langle \mathcal{B}, \mathcal{A}\rangle_S - \langle \mathcal{C}, \mathcal{A}\rangle_S\},\end{aligned}$$

where the last equality follows from the standard convex-concave minimax theorem (cf. [47, Theorem 2.10.2]). Note that for each $\mathcal{C} \in \mathrm{conv}T$,

$$\sup_{\mathcal{A}\in S}\{\langle \mathcal{B}, \mathcal{A}\rangle_S - \langle \mathcal{C}, \mathcal{A}\rangle_S\} = \begin{cases} 0, & \text{if} \quad \mathcal{B} = \mathcal{C}, \\ +\infty, & \text{else}. \end{cases}$$

Thus, the conclusion follows. □

DOI: 10.1002/nla

As λ_1^Z is continuous and convex, its convex ϵ-subdifferential ($\epsilon \geq 0$) always exists where the convex ϵ-subdifferential (cf [12]) $\partial_\epsilon \lambda_1^Z$ is defined by

$$\partial_\epsilon \lambda_1^Z(\mathcal{A}) = \{\mathcal{B} \in S : \langle \mathcal{B}, \mathcal{A}' - A\rangle_S \leq \lambda_1^Z(\mathcal{A}') - \lambda_1^Z(\mathcal{A}) + \epsilon \text{ for all } \mathcal{A}' \in S\}.$$

We are now ready to state the formula for the ϵ-subdifferential of maximum Z-eigenvalue function.

Theorem 3.2
Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, for all $\epsilon \geq 0$, we have

$$\partial_\epsilon \lambda_1^Z(\mathcal{A}) = \{B \in S : B = \sum_{i=1}^{s} \mu_i u_i^m, \mu_i \geq 0, \sum_{i=1}^{s} \mu_i = 1, \|u_i\| = 1, s \in \mathbb{N} \text{ and } \lambda_1^Z(\mathcal{A}) \leq \sum_{i=1}^{s} \mu_i \left(\mathcal{A} u_i^m\right) + \epsilon\}.$$

Proof
From the generalized Fenchel inequality, we have

$$\mathcal{B} \in \partial_\epsilon \lambda_1^Z(\mathcal{A}) \Leftrightarrow \lambda_1^Z(\mathcal{A}) + (\lambda_1^Z)^*(\mathcal{B}) \leq \langle \mathcal{A}, \mathcal{B}\rangle_S + \epsilon.$$

Note that

$$(\lambda_1^Z)^*(\mathcal{B}) = \begin{cases} 0, & \text{if} \quad \mathcal{B} \in \operatorname{conv}\{u^m : \|u\| = 1\}, \\ +\infty, & \text{else.} \end{cases}$$

So, $\mathcal{B} \in \partial_\epsilon \lambda_1^Z(\mathcal{A})$ if and only if $\mathcal{B} \in \operatorname{conv}\{u^m : \|u\| = 1\}$ and

$$\lambda_1^Z(\mathcal{A}) \leq \langle \mathcal{A}, \mathcal{B}\rangle_S + \epsilon.$$

This is further equivalent to the fact that there exist $s \in \mathbb{N}$, $\mu_i \geq 0$, $i = 1, \ldots, s$ with $\sum_{i=1}^s \mu_i = 1$ and $\|u_i\| = 1$ such that

$$\mathcal{B} = \sum_{i=1}^{s} \mu_i u_i^m \text{ and } \lambda_1^Z(\mathcal{A}) \leq \langle \mathcal{A}, \sum_{i=1}^{s} \mu_i u_i^m\rangle_S + \epsilon = \sum_{i=1}^{s} \mu_i \left(\mathcal{A} u_i^m\right) + \epsilon.$$

Thus, the conclusion follows. □

When $\epsilon = 0$, the convex subdifferential formula can be simplified as follows.

Corollary 3.1
Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, we have

$$\partial \lambda_1^Z(\mathcal{A}) = \operatorname{conv}\{u^m : u \in E_1^Z(\mathcal{A})\}.$$

Proof
Let $\epsilon = 0$. Then the preceding theorem shows that

$$\partial \lambda_1^Z(\mathcal{A}) = \{B \in S : B = \sum_{i=1}^{s} \mu_i u_i^m, \mu_i \geq 0, \sum_{i=1}^{s} \mu_i = 1, \|u_i\| = 1, s \in \mathbb{N} \text{ and } \lambda_1^Z(\mathcal{A}) \leq \sum_{i=1}^{s} \mu_i \left(\mathcal{A} u_i^m\right)\}$$

Note that $\lambda_1^Z(\mathcal{A}) = \max_{\|x\|=1} \mathcal{A} x^m$. So, $\lambda_1^Z(\mathcal{A}) \leq \sum_{i=1}^s \mu_i \left(\mathcal{A} u_i^m\right)$ with $\mu_i \geq 0$, $\sum_{i=1}^s \mu_i = 1$ and $\|u_i\| = 1$ is equivalent to $\lambda_1^Z(\mathcal{A}) = \mathcal{A} u_i^m$, for all $i = 1, \ldots, s$. It follows that

$$\partial \lambda_1^Z(\mathcal{A}) = \{B \in S : B = \sum_{i=1}^{s} \mu_i u_i^m, \mu_i \geq 0, \sum_{i=1}^{s} \mu_i = 1, \|u_i\| = 1, s \in \mathbb{N} \text{ and } \lambda_1^Z(\mathcal{A}) = \mathcal{A} u_i^m\}$$

Therefore, the conclusion follows Remark 3.1. □

Prepared using nlaauth.cls
DOI: 10.1002/nla

Remark 3.2
If $m = 2$, our subdifferential formula for λ_1^Z reduces to

$$\partial\lambda_1^Z(\mathcal{A}) = \mathrm{conv}\{uu^T : (\lambda_1^Z(\mathcal{A}), u) \text{ is an eigenpair of } \mathcal{A} \text{ and } \|u\| = 1\},$$

which is the classical subdifferential formula of the maximum eigenvalue function in the matrix case (cf. [20]).

Definition 3.1
Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Recall that the normalized eigenspace associated with $\lambda_1^Z(\mathcal{A})$ is given by $E_1^Z(\mathcal{A}) := \{u : \mathcal{A}u^{m-1} = \lambda_1^Z(\mathcal{A})u,\ \|u\| = 1\}$. Then, the eigenspace associated with $\lambda_1^Z(\mathcal{A})$ is $\mathrm{span}E_1^Z(\mathcal{A})$ where $\mathrm{span}E_1^Z(\mathcal{A})$ is the subspace generated by $E_1^Z(\mathcal{A})$, i.e., $\mathrm{span}E_1^Z(\mathcal{A}) := \bigcup_{t\in\mathbb{R}}\{t\,\mathrm{conv}E_1^Z(\mathcal{A})\}$. We now define the geometric multiplicity of $\lambda_1^Z(\mathcal{A})$ as the dimension of the subspace $\mathrm{span}E_1^Z(\mathcal{A})$.

Corollary 3.2
Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, the maximum Z-eigenvalue function λ_1^Z is locally Lipschitz, and is (Fréchet) differentiable almost everywhere. Moreover, λ_1^Z is differentiable at $\mathcal{A} \in S$ if and only if the geometric multiplicity of $\lambda_1^Z(\mathcal{A})$ is one.

Proof
From the preceding theorem, the maximum Z-eigenvalue function λ_1^Z is continuous and convex. So, λ_1^Z is locally Lipschitz. Then, the Radamecher theorem implies that it is (Fréchet) differentiable almost everywhere. To see the last assertion, as m is even, we see that $u^m = (-u)^m$. So, the geometric multiplicity of $\lambda_1^Z(\mathcal{A})$ is one is equivalent to the fact that the set

$$\partial\lambda_1^Z(\mathcal{A}) = \{u^m : (\lambda_1^Z(\mathcal{A}), u) \text{ is an eigenpair of } \mathcal{A} \text{ and } \|u\| = 1\}$$

is a singleton. Note that a continuous convex function on a finite dimensional space is Fréchet differentiable if and only if its subdifferential is a singleton. Thus, the conclusion follows. □

Perturbation Bound

Consider $\mathcal{A}(v) = \mathcal{A} + \sum_{j=1}^r v_j\mathcal{B}_j$ where $v = (v_1, \ldots, v_r) \in \mathbb{R}^r$ and $\mathcal{A}, \mathcal{B}_j \in S$, $j = 1, \ldots, r$. Define the map $h : \mathbb{R}^r \to \mathbb{R}$ by

$$h(v) = \lambda_1^Z(\mathcal{A}(v)).$$

Then, we see that h is a continuous and convex function on $\mathbb{R}^r$. Below, we present the following sensitivity result of the maximum Z-eigenvalue function. In the matrix case, this result collapses to the classical sensitivity result derived in [46].

Proposition 3.1
Let $\mathcal{A}, \mathcal{B}_j \in S$, $j = 1, \ldots, r$. Consider the map $h : \mathbb{R}^r \to \mathbb{R}$ defined by

$$h(v) = \lambda_1^Z(\mathcal{A} + \sum_{j=1}^r v_j\mathcal{B}_j),$$

where $v = (v_1, \ldots, v_r) \in \mathbb{R}^r$. Then, we have

$$\sup_{u\in E_1^Z(\mathcal{A})} \sum_{j=1}^r v_j\langle\mathcal{B}_j, u^m\rangle_S \le h(v) - h(0) \le \lambda_1^Z(\sum_{j=1}^r v_j\mathcal{B}_j).$$

Proof
First of all, from (3.3), we have

$$\begin{aligned} h(v) - h(0) &= \max_{\mathcal{C}\in\mathrm{conv}T}\langle\mathcal{C}, \mathcal{A} + \sum_{j=1}^r v_j\mathcal{B}_j\rangle_S - \max_{\mathcal{C}\in\mathrm{conv}T}\langle\mathcal{C}, \mathcal{A}\rangle_S \\ &\le \max_{\mathcal{C}\in\mathrm{conv}T}\langle\mathcal{C}, \sum_{j=1}^r v_j\mathcal{B}_j\rangle_S = \lambda_1^Z(\sum_{j=1}^r v_j\mathcal{B}_j). \end{aligned}$$

On the other hand, from the convexity of h, we have

$$h(v) - h(0) \geq \sup_{a \in \partial h(0)} a^T v.$$

Note from the chain rule of the convex subdifferential (cf. [47]) that

$$a \in \partial h(0) = \{ \begin{pmatrix} \langle \mathcal{B}_1, \mathcal{D} \rangle_S \\ \vdots \\ \langle \mathcal{B}_r, \mathcal{D} \rangle_S \end{pmatrix} : \mathcal{D} \in \partial \lambda_1^Z(\mathcal{A}) \}.$$

So,

$$\begin{aligned}
\sup_{a \in \partial h(0)} a^T v = \sup_{\mathcal{D} \in \partial \lambda_1^Z(\mathcal{A})} \sum_{j=1}^{r} \langle \mathcal{B}_j, \mathcal{D} \rangle_S v_j &= \sup_{\substack{\mu_i \geq 0, \sum_{i=1}^{s} \mu_i = 1, s \in \mathbb{N}, \\ \|u_i\| = 1, \lambda_1^Z(\mathcal{A}) = \mathcal{A} u_i^m}} \sum_{j=1}^{r} \langle \mathcal{B}_j, \sum_{i=1}^{s} \mu_i u_i^m \rangle_S v_j \\
&= \sup_{\substack{\mu_i \geq 0, \sum_{i=1}^{s} \mu_i = 1, s \in \mathbb{N}, \\ \|u_i\| = 1, \lambda_1^Z(\mathcal{A}) = \mathcal{A} u_i^m}} \sum_{i=1}^{s} \mu_i \langle \sum_{j=1}^{r} v_j \mathcal{B}_j, u_i^m \rangle_S \\
&= \sup_{\substack{\mu_i \geq 0, \sum_{i=1}^{s} \mu_i = 1, s \in \mathbb{N}, \\ \|u_i\| = 1, \lambda_1^Z(\mathcal{A}) = \mathcal{A} u_i^m}} \langle \sum_{j=1}^{r} v_j \mathcal{B}_j, \sum_{i=1}^{s} \mu_i u_i^m \rangle_S \\
&= \sup_{\|u\| = 1, \lambda_1^Z(\mathcal{A}) = \mathcal{A} u^m} \langle \sum_{j=1}^{r} v_j \mathcal{B}_j, u^m \rangle_S \\
&= \sup_{u \in E_1^Z(\mathcal{A})} \sum_{j=1}^{r} v_j \langle \mathcal{B}_j, u^m \rangle_S,
\end{aligned}$$

where the last equality follows from Remark 3.1. □

4. STABILITY ANALYSIS OF THE NORMALIZED EIGENSPACE

In this section, we study the stability of the normalized eigenspace $E_1^Z(\mathcal{A})$, that is, how the normalized eigenspace $E_1^Z(\mathcal{A})$ changes when the corresponding symmetric tensor $\mathcal{A}$ perturbs. To achieve the Hölder stability, we need the following two results. The first result gives an effective estimate for the growth rate (Łojasiewicz exponent) of a polynomial with real coefficients (for sharper exponent under specific conditions, see [10, Theorem 2.3] and [22, Lemma 4.3]). Then second result is a local error bound result which estimates how far a point to the lower level set $S_f = \{x : f(x) \leq 0\}$ is, in terms of its function value (for related error bound result see [24, 26]).

Lemma 4.1
(cf. [6, Theorem 4.2] Let f be a polynomial with real coefficients on $\mathbb{R}^n$ with degree $m \geq 2$. Suppose that $f(0) = 0$ and $\nabla f(0) = 0$. Then, there exist $\epsilon, r, c > 0$ such that

$$\|\nabla f(x)\| \geq c|f(x)|^\tau \text{ for all } \|x\| \leq \epsilon \text{ with } f(x) \leq r,$$

where $\tau \leq 1 - \big(m(3m-3)^{n-1}\big)^{-1}$.

Lemma 4.2
(cf. [34, Corollary 2.1]) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function and let $\overline{x} \in \mathbb{R}^n$ such that $\overline{x} \in \mathrm{bdry} S_f$ where $S_f = \{x : f(x) \leq 0\}$ and $\mathrm{bdry} C$ denotes the boundary of the set C. Suppose that there exist $\epsilon, c > 0$ such that $\|\nabla f(x)\| f(x)^{\theta - 1} \geq c$ for all x with $\|x - \overline{x}\| \leq \epsilon$ and $x \notin S_f$. Then,

$$d(x, S_f) \leq \frac{1}{c}(\max\{f(x), 0\})^\theta \text{ for all } x \text{ with } \|x - \overline{x}\| \leq \frac{\epsilon}{2}.$$

Recall that, for an $(n \times n)$ matrix M, $M \prec 0$ (resp. $M \preceq 0$) means that M is negative definite (resp. negative semi-definite). Moreover, we use $\mathbf{I_n}$ to denote the $(n \times n)$ identity matrix. We now provide the stability result of the normalized eigenspace. We achieve this by noting that the normalized eigenspace $E_1^Z(\mathcal{A})$ is just the optimal solution of the constraint polynomial optimization problem

$$(P_{\mathcal{A}})\ \max\{\mathcal{A}x^m : \|x\| = 1\},$$

and the stability of the optimal solution of the parameterized optimization problem $(P_{\mathcal{A}})$ can be approached by examining the growth property of a related real polynomial which can be regarded as a generalized Lagrangian function of the constraint optimization problem $(P_{\mathcal{A}})$.

Theorem 4.1
Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even and $m \geq 4$.

(i) Then, the normalized eigenspace is Hölder stable at $\mathcal{A}$ with the order $\rho = \frac{1}{m(3m-3)^{n-1}-1}$, that is, there exist $\epsilon > 0$ and $\alpha > 0$ such that for any tensor $\mathcal{B}$ with $\|\mathcal{B} - \mathcal{A}\|_S \leq \epsilon$,

$$E_1^Z(\mathcal{B}) \subseteq E_1^Z(\mathcal{A}) + \alpha\big(\|\mathcal{B} - \mathcal{A}\|_S\big)^{\frac{1}{d}}\, \mathbb{B}_{\mathbb{R}^n}, \tag{4.4}$$

where $d \in \mathbb{N}$ with $d \leq m(3m-3)^{n-1} - 1$ and $\mathbb{B}_{\mathbb{R}^n}$ is the unit ball in $\mathbb{R}^n$.

(ii) If we further assume that the following second-order condition holds: $\forall\, u \in E_1^Z(\mathcal{A})$

$$(m-1)\mathcal{A}u^{m-2} - \lambda_1^Z(\mathcal{A})\big((m-2)uu^T + \mathbf{I_n}\big) \prec 0 \text{ on } C_u = \{h \in \mathbb{R}^n : h^T u = 0\}, \tag{4.5}$$

then the integer d in (4.4) can be set as 1, i.e., there exist $\epsilon > 0$ and $\alpha > 0$ such that for any tensor $\mathcal{B}$ with $\|\mathcal{B} - \mathcal{A}\|_S \leq \epsilon$,

$$E_1^Z(\mathcal{B}) \subseteq E_1^Z(\mathcal{A}) + \alpha\|\mathcal{B} - \mathcal{A}\|_S\, \mathbb{B}_{\mathbb{R}^n}.$$

Proof

[Proof of (i)] Fix any $u \in E_1^Z(\mathcal{A})$. Let $\gamma > 0$ and let

$$f(x) := \lambda_1^Z(\mathcal{A})\|x+u\|^m - \mathcal{A}(x+u)^m + (\|x+u\|^2 - 1)^2.$$

It can be seen that f is a real polynomial on $\mathbb{R}^n$ with degree m. Moreover, it can be verified that f is nonnegative, $f(0) = 0$ and $\nabla f(0) = 0$ (as 0 is the minimizer of f). So, Lemma 4.1 implies that there exist $\gamma_u, r_u, c_u > 0$ such that

$$\|\nabla f(x)\| \geq c_u |f(x)|^\tau = f(x)^\tau \text{ for all } \|x\| \leq \gamma_u \text{ with } f(x) \leq r_u,$$

where $\tau \leq 1 - \big(m(3m-3)^{n-1}\big)^{-1}$. Choose $\delta_u < \gamma_u/2$ such that $f(x) \leq r_u$ for all $\|x\| \leq \delta_u$ (this can be done as $f(0) = 0$ and f is continuous). Then, we see that

$$\|\nabla f(x)\|\, |f(x)|^{(1-\tau)-1} = \|\nabla f(x)\|\, |f(x)|^{-\tau} \geq c_u \text{ for all } \|x\| \leq 2\delta_u.$$

Letting $h(x) := f(x-u) = \lambda_1^Z(\mathcal{A})\|x\|^m - \mathcal{A}(x)^m + (\|x\|^2 - 1)^2$, we obtain that

$$\|\nabla h(x)\|\, |h(x)|^{(1-\tau)-1} \geq c_u \text{ for all } \|x - u\| \leq 2\delta_u.$$

As $E_1^Z(\mathcal{A}) \subseteq \{x : \|x\| = 1\}$, $E_1^Z(\mathcal{A})$ has no interior point and hence $u \in \mathrm{bdry} E_1^Z(\mathcal{A})$. Note that h is nonnegative and $\{x : h(x) \leq 0\} = \{x : h(x) = 0\} = E_1^Z(\mathcal{A})$. Applying Lemma 4.2 with $f = h$ and $\overline{x} = u$, for each $u \in E_1^Z(\mathcal{A})$ we can find $c_u > 0$ such that

$$d(x, E_1^Z(\mathcal{A})) \leq \frac{1}{c_u} h(x)^\theta \text{ for all } x \text{ with } \|x - u\| \leq \delta_u, \tag{4.6}$$

where $\theta = 1-\tau \geq \big(m(3m-3)^{n-1}\big)^{-1}$. Now, using standard compactness argument, we show that there exist $c>0$ and $\delta>0$ such that

$$d(x, E_1^Z(\mathcal{A})) \leq \frac{1}{c}h(x)^{\theta} \text{ for all } x \text{ with } d(x, E_1^Z(\mathcal{A})) \leq \delta. \tag{4.7}$$

As $E_1^Z(\mathcal{A})$ is compact and

$$E_1^Z(\mathcal{A}) \subseteq \bigcup_{u \in E_1^Z(\mathcal{A})} \{x : \|x-u\| \leq \frac{\delta_u}{2}\},$$

there exist $l \in \mathbb{N}$ and $\{u_1, \ldots, u_l\} \subseteq E_1^Z(\mathcal{A})$ such that

$$E_1^Z(\mathcal{A}) \subseteq \bigcup_{i=1}^{l} \{x : \|x-u_i\| \leq \frac{\delta_{u_i}}{2}\}. \tag{4.8}$$

Take $\delta = \frac{1}{2}\min\{\delta_{u_1}, \ldots, \delta_{u_l}\} > 0$ and $c = \min\{c_{u_1}, \ldots, c_{u_l}\} > 0$. Then, for each x with $d(x, E_1^Z(\mathcal{A})) \leq \delta$, we can find $a \in E_1^Z(\mathcal{A})$ such that $\|x-a\| \leq \delta$. By (4.8), there exists $i_0 \in \{1, \ldots, l\}$ such that $\|a-u_{i_0}\| \leq \frac{\delta_{u_{i_0}}}{2}$. So,

$$\|x-u_{i_0}\| \leq \|x-a\| + \|a-u_{i_0}\| \leq \delta + \frac{\delta_{u_{i_0}}}{2} \leq \delta_{u_{i_0}}.$$

Then, (4.6) implies that

$$d(x, E_1^Z(\mathcal{A})) \leq \frac{1}{c_{u_{i_0}}}h(x)^{\theta} \leq \frac{1}{c}h(x)^{\theta}.$$

Thus, (4.7) holds. Letting $\beta := c^{\frac{1}{\theta}}$, this shows that for any x with $d(x, E_1^Z(\mathcal{A})) \leq \delta$,

$$\lambda_1^Z(\mathcal{A})\|x\|^m - \mathcal{A}(x)^m + (\|x\|^2-1)^2 = h(x) \geq c^{\frac{1}{\theta}} d(x, E_1^Z(\mathcal{A}))^{\frac{1}{\theta}} = \beta d(x, E_1^Z(\mathcal{A}))^{\frac{1}{\theta}}. \tag{4.9}$$

For the above δ, there exists $\epsilon > 0$ such that for any $\mathcal{B} \in S$ with $\|\mathcal{B}-\mathcal{A}\|_S \leq \epsilon$

$$E_1^Z(\mathcal{B}) \subseteq E_1^Z(\mathcal{A}) + \delta\, \mathbb{B}_{\mathbb{R}^n}. \tag{4.10}$$

(Otherwise, there exists a sequence of symmetric tensors B_n with $\|\mathcal{B}_n - \mathcal{A}\|_S \to 0$ and $y_n \in S_{\mathcal{B}_n}$ such that $d(y_n, E_1^Z(\mathcal{A})) \geq \delta$. As $\|y_n\| = 1$, by passing to subsequence, we may assume that $y_n \to y$ for some y with $\|y\| = 1$. Then, we have $d(y, E_1^Z(\mathcal{A})) \geq \delta > 0$. Now, as λ_1^Z is continuous, we have $\mathcal{B}_n(y_n)^m = \lambda_1^Z(\mathcal{B}_n) \to \lambda_1^Z(\mathcal{A})$. So, passing to limit, we see that $\mathcal{A}y^m = \lambda_1^Z(\mathcal{A})$. Note that $\|y\| = 1$. So, $y \in E_1^Z(\mathcal{A})$ which makes contradiction.) Now, fix any arbitrary $\mathcal{B}$ with $\|\mathcal{B}-\mathcal{A}\|_S \leq \epsilon$. From (4.10), we have $E_1^Z(\mathcal{B}) \subseteq E_1^Z(\mathcal{A}) + \delta\mathbb{B}_{\mathbb{R}^n}$. Let $r > 0$ be a constant such that

$$\|x^m - y^m\|_S \leq r\|x-y\| \text{ for all } x, y \in K := \{x : \|x\| = 1\}. \tag{4.11}$$

Let $d = \frac{1}{\theta} - 1 \leq m(3m-3)^{n-1} - 1$ and $\alpha = (\beta^{-1}r)^d > 0$. For any arbitrary $v \in E_1^Z(\mathcal{B})$, take $u \in E_1^Z(\mathcal{A})$ be such that $\|v-u\| = d(v, E_1^Z(\mathcal{A}))$. To finish the proof, it suffices to show that

$$\|v-u\| \leq \alpha\|\mathcal{B}-\mathcal{A}\|_S^{\frac{1}{d}}. \tag{4.12}$$

To see this, we first note that $\|v-u\| \leq \delta$ as $E_1^Z(\mathcal{B}) \subseteq E_1^Z(\mathcal{A}) + \delta\, \mathbb{B}_{\mathbb{R}^n}$. So, (4.9) and $\|v\| = 1$ (as $v \in E_1^Z(\mathcal{B})$) imply that

$$\lambda_1^Z(\mathcal{A}) - \mathcal{A}v^m = \lambda_1^Z(\mathcal{A})\|v\|^m - \mathcal{A}v^m + (\|v\|^2-1)^2 \geq \beta\|v-u\|^{\frac{1}{\theta}}.$$

As u is an optimal solution of $P_{\mathcal{A}}$, we have $\|u\| = 1$ and $\mathcal{A}u^m = \lambda_1^Z(\mathcal{A})$ and so,

$$\|v-u\|^{\frac{1}{\theta}} \leq \beta^{-1}\big(\lambda_1^Z(\mathcal{A}) - \mathcal{A}v^m\big) = \beta^{-1}\big(\mathcal{A}u^m - \mathcal{A}v^m\big). \tag{4.13}$$

Then, as v is optimal for $(P_{\mathcal{B}})$ and so, $\|v\|=1$ and $\mathcal{B}u^m \le \mathcal{B}v^m$. It follows from (4.11) that

$$\begin{aligned} \mathcal{A}u^m - \mathcal{A}v^m &= \big(\mathcal{B}u^m - \mathcal{B}v^m\big) + \big((\mathcal{A}-\mathcal{B})u^m - (\mathcal{A}-\mathcal{B})v^m\big) \\ &\le (\mathcal{A}-\mathcal{B})u^m - (\mathcal{A}-\mathcal{B})v^m \\ &\le \|u^m - v^m\|_S \, \|\mathcal{A}-\mathcal{B}\|_S \\ &\le r\|u-v\| \, \|\mathcal{B}-\mathcal{A}\|_S \, . \end{aligned}$$

This together with (4.13) implies that

$$\|v-u\|^{\frac{1}{\theta}} \le \beta^{-1}\big(\mathcal{A}u^m - \mathcal{A}v^m\big) \le \beta^{-1} r \|u-v\| \, \|\mathcal{B}-\mathcal{A}\|_S \, .$$

So, we have $\|v-u\|^d = \|v-u\|^{\frac{1}{\theta}-1} \le \beta^{-1} r \|\mathcal{B}-\mathcal{A}\|_S$. Note that $\alpha = (\beta^{-1}r)^{\frac{1}{d}}$. Then,

$$d(v, E_1^Z(\mathcal{A})) = \|v-u\| \le \alpha \|\mathcal{B}-\mathcal{A}\|_S^{\frac{1}{d}} \, .$$

Therefore, (4.12) holds, and so, statement (i) follows.

[Proof of (ii)] Suppose that the second order condition (4.5) holds. Fix an arbitrary $u \in E_1^Z(\mathcal{A})$ and consider the minimization problem

$$\begin{array}{lll} (P_0) & \min_{x\in\mathbb{R}^n} & f(x) := -\mathcal{A}x^m \\ & \text{s.t.} & \|x\|^m = 1. \end{array}$$

Clearly, u satisfy the KKT condition of (P_0) with Lagrange multiplier $\lambda_1^Z(\mathcal{A})$. Note that the usual second order sufficient condition for this problem reduces to (4.5). So, the following second order growth condition holds at u (for example see [9, Corollary 1]): there exist $\beta>0$ and $\delta>0$ such that

$$-\mathcal{A}x^m + \lambda_1^Z(\mathcal{A}) = f(x) - f(u) \ge \beta \, \|x-u\|^2 \text{ for all } x \in \mathbb{R}^n \text{ with } \|x-u\| \le \delta.$$

Now, using the same method of the proof as in part (1), we see that the conclusion holds. □

Semismoothness of the Maximum Z-Eigenvalue Function

In this subsection, as an application of the preceding stability result of the normalized eigenspace, we examine the semismoothness of the maximum Z-eigenvalue function. Now, consider a function $f : S \to \mathbb{R}$ where S is the symmetric tensor space on $\mathbb{R}^n$. Note that the symmetric tensor space S can be identified as a finite dimensional space with an appropriate dimension. The definition of semismoothness of f can be translated as follows:

Definition 4.1
Let $f : S \to \mathbb{R}$ be a locally Lipschitz and directionally differentiable function. Then, the function $f : S \to \mathbb{R}$ is said to be semismooth at $\mathcal{A} \in S$ if ,

$$f(\mathcal{A}+\Delta\mathcal{A}) - f(\mathcal{A}) - \langle V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S = o(\|\Delta\mathcal{A}\|_S), \ \forall V(\Delta\mathcal{A}) \in \partial_C f(\mathcal{A}+\Delta\mathcal{A}).$$

Moreover, $f : S \to \mathbb{R}$ is said to be ρth-order semismooth function for some $\rho \in (0,1]$ at $\mathcal{A} \in S$ if

$$f(\mathcal{A}+\Delta\mathcal{A}) - f(\mathcal{A}) - \langle V(\Delta\mathcal{A}), \Delta\mathcal{A}\rangle_S = O(\|\Delta\mathcal{A}\|_S^{1+\rho}), \ \forall V(\Delta\mathcal{A}) \in \partial_C f(\mathcal{A}+\Delta\mathcal{A}).$$

In particular, if $\rho=1$, we say f is a strongly semismooth function. We also say $f : S \to \mathbb{R}$ is a semismooth (resp. ρth-order semismooth) function on S if f is semismooth (resp. ρth-order semismooth) at $\mathcal{A}$ for all $\mathcal{A} \in S$.

To achieve this, we first observe from Lemma 3.1 that the maximum eigenvalue of tensors $\lambda_1^Z(\mathcal{A})$ is indeed the optimal value of a polynomial optimization problem $(P_{\mathcal{A}})$ $\max\{\mathcal{A}x^m : \|x\|=1\}$. Thus, the semismooth properties of the maximum Z-eigenvalue function can be approached by examining the stability of the parameterized optimization problem $(P_{\mathcal{A}})$ which we have already studied earlier. We now study the ρth-order semismooth properties of the maximum Z-eigenvalue function and establish some explicit estimation of ρ.

DOI: 10.1002/nla

Theorem 4.2
Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor (m is even). Then, we have

(i) The maximum Z-eigenvalue function λ_1^Z is at least ρth-order semismooth at $\mathcal{A}$ with $\rho = \frac{1}{m(3m-3)^{n-1}-1}$.
(ii) Moreover, if we further assume that the following second-order condition holds: for all $u \in E_1^Z(\mathcal{A})$

$$(m-1)\mathcal{A}u^{m-2} - \lambda_1^Z(\mathcal{A})\big((m-2)uu^T + \mathbf{I_n}\big) \prec 0 \text{ on } C_u = \{h \in \mathbb{R}^n : h^T u = 0\}, \quad (4.14)$$

then the maximum Z-eigenvalue function λ_1^Z is strongly semismooth at $\mathcal{A}$.

Proof

[Proof of (i)] Suppose that $m = 2$. Then, $\mathcal{A}$ is an $(n \times n)$ symmetric matrix. From [45], we see that the maximum eigenvalue function is strongly semismooth. So, without loss of generality, we may assume that $m \geq 4$. Let $\mathcal{A}$ be an arbitrary mth-order n-dimensional symmetric tensor and let $\rho = \frac{1}{m(3m-3)^{n-1}-1}$. Let $\Delta\mathcal{A}$ be an mth-order n-dimensional symmetric tensor such that $\|\Delta\mathcal{A}\|_S > 0$ and λ_1 is differentiable at $\mathcal{A} + \Delta\mathcal{A}$. So, Corollary 3.2 implies that $\nabla\lambda_1(\mathcal{A} + \Delta\mathcal{A}) = (w_{\Delta\mathcal{A}})^m$ where $w_{\Delta\mathcal{A}} \in E_1^Z(\mathcal{A} + \Delta\mathcal{A})$. Note that λ_1 is continuous and convex (and so, is directionally differentiable), and the Clarke subdifferential and the convex subdifferential of λ_1 coincide. To see the conclusion, we only need to show that

$$\lambda_1^Z(\mathcal{A} + \Delta\mathcal{A}) - \lambda_1^Z(\mathcal{A}) - \langle (w_{\Delta\mathcal{A}})^m, \Delta\mathcal{A}\rangle_S = O(\|\Delta\mathcal{A}\|_S^{1+\rho}). \quad (4.15)$$

To see (4.15), let $r > 0$ be a constant such that

$$\|x^m - y^m\|_S \leq r\|x - y\| \text{ for all } x, y \in K := \{x : \|x\| = 1\}. \quad (4.16)$$

Let $v \in E_1^Z(\mathcal{A})$ be such that $\|w_{\Delta\mathcal{A}} - v\| = d(w_{\Delta\mathcal{A}}, E_1^Z(\mathcal{A}))$. Then, the preceding proposition implies that there exists $c > 0$ such that $\|w_{\Delta\mathcal{A}} - v\| \leq c\|\Delta\mathcal{A}\|_S^\rho$. Clearly, $v^m \in \partial\lambda_1^Z(\mathcal{A})$. It follows from (4.16) that

$$\begin{aligned}
\lambda_1^Z(\mathcal{A} + \Delta\mathcal{A}) - \lambda_1^Z(\mathcal{A}) - \langle (w_{\Delta\mathcal{A}})^m, \Delta\mathcal{A}\rangle_S &\geq \langle v^m, \Delta\mathcal{A}\rangle_S - \langle (w_{\Delta\mathcal{A}})^m, \Delta\mathcal{A}\rangle_S \\
&\geq -\|v^m - (w_{\Delta\mathcal{A}})^m\|_S \, \|\Delta\mathcal{A}\|_S \\
&\geq -r\|v - w_{\Delta\mathcal{A}}\| \, \|\Delta\mathcal{A}\|_S \\
&\geq -rc\|\Delta\mathcal{A}\|_S^{1+\rho}.
\end{aligned}$$

On the other hand, as $\nabla\lambda_1(\mathcal{A} + \Delta\mathcal{A}) = (w_{\Delta\mathcal{A}})^m$ and λ_1^Z is convex, we see that

$$\langle (w_{\Delta\mathcal{A}})^m, -\Delta\mathcal{A}\rangle_S = \langle (w_{\Delta\mathcal{A}})^m, \mathcal{A} - (\mathcal{A} + \Delta\mathcal{A})\rangle_S \leq \lambda_1^Z(\mathcal{A}) - \lambda_1^Z(\mathcal{A} + \Delta\mathcal{A}).$$

This implies that

$$\lambda_1^Z(\mathcal{A} + \Delta\mathcal{A}) - \lambda_1^Z(\mathcal{A}) - \langle (w_{\Delta\mathcal{A}})^m, \Delta\mathcal{A}\rangle_S \leq \langle (w_{\Delta\mathcal{A}})^m, \Delta\mathcal{A}\rangle_S - \langle (w_{\Delta\mathcal{A}})^m, \Delta\mathcal{A}\rangle_S = 0.$$

Therefore, we have

$$\lambda_1^Z(\mathcal{A} + \Delta\mathcal{A}) - \lambda_1^Z(\mathcal{A}) - \langle (w_{\Delta\mathcal{A}})^m, \Delta\mathcal{A}\rangle_S = O(\|\Delta\mathcal{A}\|_S^{1+\rho}),$$

and so, the conclusion follows.

[Proof of (ii)] Using the same method of proof as in part (i) and using Proposition 4.1(ii) (instead of Proposition 4.1(i)), we see that the conclusion follows. □

Remark 4.1
In Theorem 4.2, for an mth order n-dimensional tensor $\mathcal{A}$, we showed that the maximum Z-eigenvalue function is always at least ρth-order semismooth at $\mathcal{A}$ with $\rho = \frac{1}{m(3m-3)^{n-1}-1}$. In

DOI: 10.1002/nla

our preceding paper [23], ρth-order semismoothness of the maximum H-eigenvalue function at $\mathcal{A}$ with $\rho = \frac{1}{(2m-1)^n}$ was shown under an additional assumption that the geometric multiplicity at $\mathcal{A}$ is one. At this moment, it is not clear whether our method of proof here can be used to relax the geometric multiplicity assumption in our previous ρth-order semismoothness result for maximum-H eigenvalue function with some appropriate exponent ρ (as we made use of the fact that an eigenvector associated with the maximum Z-eigenvalue is of norm one in proposition 4.1). Moreover, one can observe that the degree of the semismooth property is different for the maximum H-eigenvalue function and the maximum Z-eigenvalue function. At this moment, it is not clear for us that which type of maximum eigenvalue function has a better analytic properties.

Nonsmooth Newton Method for Space Tensor Problems

Let $S(4,3)$ be the space consisting of all the 4th-order 3-dimensional symmetric tensors. It is shown [42] that $S(4,3)$ is of dimension 15, and so, there exists a one-to-one mapping $L : \mathbb{R}^{15} \to S(4,3)$ (indeed the mapping L can be explicitly constructed see [42, 41] for details). Let n be a natural number, $\mathcal{A}_i \in S(4,3)$, $i = 0, 1, \ldots, n$ and $b_i \in \mathbb{R}$, $i = 1, \ldots, n$. Consider the space tensor conic linear programming (STCLP) problem which was proposed and studied in [41]:

$$(STCLP) \quad \min_{\mathcal{X} \in S(4,3)} \ \langle \mathcal{A}_0, \mathcal{X} \rangle_S \quad \text{s.t.} \ \langle \mathcal{A}_i, \mathcal{X} \rangle_S \le b_i,\ i = 1, \ldots, n, \quad \mathcal{X} \in -\mathcal{C}(4,3),$$

where $\mathcal{C}(4,3)$ is the cone of all negative semidefinite 4th-order 3-dimensional symmetric tensor, i.e., $\mathcal{C}(4,3) := \{\mathcal{A} \in S(4,3) : \mathcal{A}x^4 \le 0,\ \forall\, x \in \mathbb{R}^3\}$. The problem (STCLP) arises from the medical imaging area where a high order tensor is used to describe the non-Gaussian diffusion feature [42]. Recently, [23] proposed a nonsmooth Newton method based on the maximum H-eigenvalue for solving the space tensor problem, and established the corresponding nonsmooth Newton method converges superlinearly to a solution with order $1+\rho$ for some unknown constant $\rho > 0$. Here, we propose a new nonsmooth Newton method based on the maximum Z-eigenvalue. An advantage of the new method is that: we are now able to obtain the superlinear convergence with an explicit estimate of the order $1+\rho$, using the semismooth property of the maximum-Z eigenvalue established in Theorem 4.2.

Note that $\mathcal{C}(4,3) = \{\mathcal{A} \in S(4,3) : \lambda_1^Z(\mathcal{A}) \le 0\}$ and λ_1^Z is convex. We see that problem (STCLP) is a convex programming problem. By assuming the Slater constraint qualification, solving (STCLP) is equivalent to solving its KKT system. As shown in [44] (see also [41]), $\mathcal{C}(4,3)$ is a self-dual cone and $\big(\mathcal{C}(4,3)\big)^* = \mathcal{C}(4,3) = \mathcal{U}(4,3)$ where $(\mathcal{C}(4,3))^*$ is the usual dual cone of $\mathcal{C}(4,3)$, $\mathcal{U}(4,3)$ is the rank one tensor space in $S(4,3)$ defined by $\mathcal{U}(4,3) = \{\mathcal{A} \in S(4,3) : \mathcal{A} = \sum_{j=1}^r (a_j)^4,\ a_j \in \mathbb{R}^3, r \in \mathbb{N}\}$, where a^4 is the 4th-order 3-dimensional symmetric rank one tensor defined by $(a^4)_{i_1 i_2 i_3 i_4} = a_{i_1} a_{i_2} a_{i_3} a_{i_4}$ for each $i_1, i_2, i_3, i_4 \in \{1,2,3\}$. So, its KKT system is to find $y_1, \ldots, y_n \in \mathbb{R}$ and $\mathcal{X} \in S(4,3)$ such that

$$(KKT) \left\{ \begin{array}{ll} \mathcal{X} \in -\mathcal{C}(4,3) & \text{(Primal Feasible)} \\ -\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i \in \mathcal{C}(4,3) & \text{(Dual Feasible)} \\ 0 \le y_i \perp \big(\langle \mathcal{A}_i, \mathcal{X} \rangle_S - b_i\big) \le 0 & \text{(Complementary Slackness I)} \\ \langle \mathcal{A}_0 + \sum_{i=1}^n y_i \mathcal{A}_i, \mathcal{X} \rangle_S = 0 & \text{(Complementary Slackness II)} \end{array} \right. . \tag{4.17}$$

The following proposition establishes a useful observation: solving the KKT problem is equivalent to solving the nonsmooth equation $F(x) = 0$ where $F : \mathbb{R}^{n+15} \to \mathbb{R}^{n+2}$ is defined by

$$F(x) = \begin{pmatrix} \max\{\lambda_1^Z(-\mathcal{A}_0 - \sum_{i=1}^n y_i \mathcal{A}_i), \lambda_1^Z(-Lz)\} \\ \langle \mathcal{A}_0 + \sum_{i=1}^n y_i \mathcal{A}_i, Lz \rangle_S \\ \max\{-y_1, \langle \mathcal{A}_1, Lz \rangle_S - b_1\} \\ \vdots \\ \max\{-y_n, \langle \mathcal{A}_n, Lz \rangle_S - b_n\} \end{pmatrix}, \ x = (y,z) \in \mathbb{R}^n \times \mathbb{R}^{15} \tag{4.18}$$

where $L : \mathbb{R}^{15} \to S(4,3)$ is the one-to-one linear mapping between $\mathbb{R}^{15}$ and $S(4,3)$. Its proof is similar to [23, Proposition 5.1], so we omit the proof here.

Proposition 4.1
Let $y \in \mathbb{R}^n$, $\mathcal{X} \in S(4,3)$ and $x := (y, L^{-1}(\mathcal{X})) \in \mathbb{R}^n \times \mathbb{R}^{15}$. Then, $(y, \mathcal{X}) \in \mathbb{R}^n \times S(4,3)$ solves the KKT system (4.17) if and only if $F(x) = 0$.

From the proceeding proposition, to obtain a solution of the KKT system, it suffices to solve the nonsmooth underdetermined equation $F(x) = 0$ where $F : \mathbb{R}^{n+15} \to \mathbb{R}^{n+2}$. Now, we state a nonsmooth Newton method for solving the space tensor conic linear problem (STCLP):

Algorithm-1

Step 0. Choose $(y^{(0)}, \mathcal{X}^{(0)}) \in \mathbb{R}^n \times S(4,3)$. Compute $z^{(0)} = L^{-1}(\mathcal{X}^{(0)})$ and let $x^{(0)} = (y^{(0)}, z^{(0)})$. If $F(x^{(0)}) \neq 0$, then set $k := 0$. Otherwise, output $(y^{(0)}, \mathcal{X}^{(0)})$.

Step 1. Compute a $V_k \in \mathbb{R}^{(n+2)\times(n+15)}$ such that $V_k \in J_C F(x^{(k)})$. †

Step 2. Let $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$ where $\Delta x^{(k)} = -(V_k^T V_k)^{-1} V_k^T F(x^{(k)})$.

Step 3. If $F(x^{(k+1)}) \neq 0$, then replace k by $k+1$ and go back to Step 1. Otherwise, let $x^{(k+1)} = (y^{(k+1)}, z^{(k+1)}) \in \mathbb{R}^n \times \mathbb{R}^{15}$ and output $\big(y^{(k+1)}, L^{-1}(z^{(k+1)})\big) \in \mathbb{R}^n \times S(4,3)$.

Next, we present the superlinear convergence result of Algorithm-1 with an explicit estimate of the convergence rate for a regular solution (see the definition in Appendix), by using the known convergence result for general nonsmooth Newton method (See Lemma 7.1) and the semismooth property of maximum Z-eigenvalue (Theorem 4.2).

Theorem 4.3
Let $(y^*, \mathcal{X}^*) \in \mathbb{R}^n \times S(4,3)$ be a solution of the KKT system of (STCLP). Let $x^* = \big(y^*, L^{-1}(\mathcal{X}^*)\big) \in \mathbb{R}^n \times \mathbb{R}^{15}$. Let x^* be a regular solution and let P be the regular space associated with x^*. Then there exists a neighborhood N of x^* such that the Algorithm-1 is well defined for any initial point $x^{(0)} \in N_0 \cap (x^* + P)$ and Algorithm-1 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges superlinearly to x^* with order at least $1 + \rho$ where $\rho = \frac{1}{m(3m-3)^{n-1}-1}$, i.e.,

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^{1+\rho}} < +\infty. \tag{4.19}$$

Proof
Clearly, F is locally Lipschitz. Moreover, we note that $(a,b) \mapsto \max\{a,b\}$ is strongly semismooth, any continuous differentiable function with locally Lipschitz gradient is strongly semismooth and composition of (ρth-order) semismooth function is still (ρth-order) semismooth. Then, Theorem 4.2(1) implies that $\mathcal{X} \mapsto \lambda_1^Z(\mathcal{X})$ is at least ρth-order semismooth with $\rho = \frac{1}{m(3m-3)^{n-1}-1}$. It follows that F is a vector valued function where each of its coordinate is at least ρth-order semismooth. Thus, F is also at least ρth-order semismooth, and so, the conclusion follows from Lemma 7.1. □

5. APPLICATION TO SPECTRAL HYPERGRAPH THEORY

Now, consider an (undirected) hypergraph which is a pair $G = (V, E)$, where $V = \{1, \ldots, n\}$ is a finite set of vertices and $E \subseteq 2^V$ is a set of subsets of V (each of which is called a hyperedge).

†For a 4th-order 3-dimensional tensor, one can efficiently find an eigenvector associated with its maximum eigenvalue (e.g. see [42]), and so, one can also efficiently compute a member of the (Clarke) generalized Hessian $J_C F(x^{(k)}) \subseteq \mathbb{R}^{(n+2)\times(n+15)}$.

DOI: 10.1002/nla

A hypergraph G is said to be m-uniform for an integer $m \geq 2$ if, $|e| = m$ for all $e \in E$, where $|\cdot|$ denotes the cardinality. That is, for an m-uniform hypergraph, each hyperedge has the same cardinality m. If $m = 2$, then 2-uniform graphs are typically called graphs. A finite path from vertex i to vertex j is a finite sequence of vertices with start from i and end with j such that each of its vertices and the next vertex belong to some hyperedge. Two vertices are called connected if there is a finite path between them. A connected component X of G is a subset of V such that any two vertices in X are connected and no other vertex in $V\backslash X$ is connected to any vertex in X.

Consider an m-uniform hypergraph $G = (V, E)$. Let $E = \{E_1, \ldots, E_p\}$ where each E_l, $l = 1, \ldots, p$, is a hyperedge. We define a homogeneous polynomial f associated with the hypergraph G defined by

$$f_G(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{ij}(x_i - x_j)^m,$$

where δ_{ij} is defined by

$$\delta_{ij} = \begin{cases} 1, & \text{if} \quad i < j \text{ and } \{i,j\} \subseteq E_l, \text{ for some } l = 1, \ldots, p, \\ 0, & \text{else.} \end{cases}$$

Note that any homogeneous polynomial uniquely determines a symmetric tensor. We can define the Laplacian tensor $\mathcal{L}$ of the hypergraph G by

$$\mathcal{L}x^m = f_G(x) \text{ for all } x \in \mathbb{R}^n.$$

In the special case when $m = 2$, our definition of Laplacian tensor reduces to the Laplacian matrix as the Laplacian matrix L of a graph $G = (V, E)$ satisfies the following property: for all $x \in \mathbb{R}^n$,

$$x^T L x = \sum_{i<j,\,(i,j)\in E} (x_i - x_j)^2.$$

Remark 5.1
Recently, [13] studied the algebraic connectivity of a 4-uniform hypergraph by introducing a different definition of the Laplacian tensor as follows: for a 4-uniform hypergraph $G = (V, E)$, its Laplacian tensor T corresponds to the following quartic form:

$$Tx^4 := \sum_{E_p \in E} L(E_p)x^4$$

where each E_p is an hyperedge of G and

$$\begin{aligned} L(E_p)x^4 &= \frac{1}{84}[(x_i + x_j + x_k - 3x_l)^4 + (x_i + x_j + x_l - 3x_k)^4 \\ &\quad + (x_i + x_k + x_l - 3x_j)^4 + (x_j + x_k + x_l - 3x_i)^4]. \end{aligned}$$

It can be seen that our definition of a Laplacian tensor is a direct extension of the matrix cases and works for m-uniform hypergraph with $m > 4$.

We now define the characteristic tensor $\mathcal{C}$ of an m-uniform hypergraph G by $\mathcal{C} = -\mathcal{L}$. The following theorem summarizes some basic features of the characteristic tensor.

Theorem 5.1
Let $\mathcal{C}$ be the characteristic tensor of an m-uniform hypergraph G where m is even. Then, the following statements hold:

(1) The characteristic tensor $\mathcal{C}$ is negative semi-definite in the sense that $\mathcal{C}x^m \leq 0$ for all $x \in \mathbb{R}^m$.
(2) The maximum Z-eigenvalue of the characteristic tensor $\mathcal{C}$ is 0 and $a = \frac{1}{\sqrt{n}}\mathbf{1}_n$ is a corresponding eigenvector where $\mathbf{1}_n \in \mathbb{R}^n$ is the vector whose components are all equal to one.

Prepared using nlaauth.cls
DOI: 10.1002/nla

(3) The dimension of the eigenspace of the maximum Z-eigenvalue equals the number of the connected component of the m-uniform hypergraph G.

Proof

[Proof of (1)] Clearly, $f_G(x) \geq 0$ for all $x \in \mathbb{R}^n$. So, we see that

$$\mathcal{C}x^m = -\mathcal{L}x^m = -f_G(x) \leq 0 \text{ for all } x \in \mathbb{R}^n.$$

[Proof of (2)] Consider $a = \frac{1}{\sqrt{n}}\mathbf{1}_n$. Then, $a = (a_1, \ldots, a_n)$ with $a_i = \frac{1}{\sqrt{n}}$ for each $i = 1, \ldots, n$ and $\|a\| = 1$. Note that

$$\mathcal{C}a^m = -f_G(a) = -\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij}(a_i - a_j)^m = 0.$$

So, a is a global maximizer of (P) $\max_{x \in \mathbb{R}^n}\{\mathcal{C}x^m : \|x\| = 1\}$ and the optimal value of (P) is 0. Thus, the conclusion follows from Lemma 3.1.

[Proof of (3)] Let the connected component of G be $\{V^1, \ldots, V^q\}$ where $q \in \mathbb{N}$. For each $l = 1, \ldots, q$, define $a^l = (a_1^l, \ldots, a_n^l)$ where, for each $i = 1, \ldots, n$,

$$a_i^l = \begin{cases} 1, & \text{if} \quad i \in V_l, \\ 0, & \text{else.} \end{cases}$$

Then, we see that

$$\mathcal{C}(a^l)^m = -f_G(a^l) = -\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij}(a_i^l - a_j^l)^m = 0$$

where the last equality holds as $a_i^l - a_j^l = 0$ if i, j in the same connected component and $\delta_{ij} = 0$ if i, j in different connected components. So, each a^l is an eigenvector associated with the maximum Z-eigenvalue of $\mathcal{L}$, $l = 1, \ldots, q$. Note that $\{a^1, \ldots, a^q\}$ is linearly independent. So, the dimension of the eigenspace is at least q. Now, consider any eigenvector $x = (x_1, \ldots, x_n)$ associated with the maximum Z-eigenvalue. Then

$$\mathcal{C}x^m = -f_G(x) = -\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij}(x_i - x_j)^m = 0.$$

As m is even, this implies that $\delta_{ij}(x_i - x_j) = 0$ for all $i, j = 1, \ldots, n$. Take each connected component V^l, it follows that $x_i = x_j$ if $i, j \in V^l$. Note that $\bigcup_{l=1}^{q} V^l = V$. It follows that

$$x = \alpha_1 a^1 + \ldots + \alpha_q a^q,$$

for some $(\alpha_1, \ldots, \alpha_q) \neq 0$. This shows that the dimension of the eigenspace equals q. □

The following corollary provides the link between the combinatorial structure and the analytic structure of the hypergraph.

Corollary 5.1
Let G be an m-uniform hypergraph G where m is even, and let $\mathcal{C}$ be its characteristic tensor. Then, the following statements are equivalent:

(1) G is connected.
(2) the geometric multiplicity of $\lambda_1^Z(\mathcal{C})$ is one.
(3) the maximum Z-eigenvalue function λ_1^Z is differentiable at $\mathcal{C}$.

Proof

[$(1) \Leftrightarrow (2)$] This equivalence follows from the preceding theorem by letting the number of the connected component be one.

[$(2) \Leftrightarrow (3)$] This equivalence follows by Corollary 3.2. □

DOI: 10.1002/nla

Next, we denote the second largest Z-eigenvalue of the characteristic tensor by $\lambda_2^Z(\mathcal{C})$, that is,

$$\lambda_2^Z(\mathcal{C}) := \max\{\lambda \in \mathbb{R} : \lambda \text{ is a Z-eigenvalue of } \mathcal{C} \text{ and } \lambda \neq \lambda_1^Z(\mathcal{C})\}.$$

The following proposition provides a variational characterization of the second largest Z-eigenvalue of the characteristic tensor.

Proposition 5.1
Let $\mathcal{C}$ be the characteristic tensor of an m-uniform hypergraph G where m is even. Then, we have

$$\lambda_2^Z(\mathcal{C}) = \max_{x\in\mathbb{R}^n}\{\mathcal{C}x^m : \|x\| = 1 \text{ and } x \in (\mathrm{span}E_1^Z(\mathcal{C}))^{\perp}\}.$$

Proof
Consider the maximization problem

$$(P_0)\ \max_{x\in\mathbb{R}^n}\{\mathcal{C}x^m : \|x\|^m = 1 \text{ and } x \in (\mathrm{span}E_1^Z(\mathcal{C}))^{\perp}\}.$$

To finish the proof, it suffices to show that $v(P_0) = \lambda_2^Z(\mathcal{C})$ where $v(P_0)$ is the optimal value of (P_0).

Let a be a maximizer of the problem (P_0). Then, by the KKT condition, we have $\|a\| = 1$, $a \in (\mathrm{span}E_1^Z(\mathcal{C}))^{\perp}$ and there exist $\lambda \in \mathbb{R}$ and $u \in \mathrm{span}E_1^Z(\mathcal{C})$ such that

$$-m\mathcal{C}a^{m-1} + m\lambda a + u = 0.$$

This implies that

$$-m(\mathcal{C}a^{m-1})^T u + \|u\|^2 = (-m\mathcal{C}a^{m-1} + m\lambda a + u)^T u = 0.$$

From the construction of the characteristic tensor, one can write $\mathcal{C} = \sum_{j=1}^r \gamma_j w_j^m$, where $\gamma_j \leq 0$, $w_j \in \mathbb{R}^n$, $j = 1, \ldots, r$, $r \in \mathbb{N}$ and y^m denotes the rank one tensor defined by $(y^m)_{i_1,\ldots,i_m} = y_{i_1}\ldots y_{i_m}$. So, we have $\mathcal{C}a^{m-1} = \sum_{j=1}^r \gamma_j(\langle w_j, a\rangle)^{m-1} w_j$, and hence

$$-m\sum_{j=1}^r \gamma_j(w_j^T a)^{m-1} w_j^T u + \|u\|^2 = 0.$$

As $u \in \mathrm{span}E_1^Z(\mathcal{C})$ and $E_1^Z(\mathcal{C})$ is symmetric (that is, if $v \in E_1^Z(\mathcal{C})$ then $-v \in E_1^Z(\mathcal{C})$), there exist $s \in \mathbb{N}$, $\alpha_i \geq 0$, $i = 1, \ldots, s$, such that $u = \sum_{i=1}^s \alpha_i v_i$ with $v_i \in E_1^Z(\mathcal{C})$. Note that, for each $i = 1, \ldots, s$

$$0 = \lambda_1^Z(\mathcal{C}) = \mathcal{C}v_i^m = \sum_{j=1}^r \gamma_j(w_j^T v_i)^m.$$

This together with $\gamma_j \leq 0$ and m is even implies that $\gamma_j w_j^T v_i = 0$ for all $j = 1, \ldots, r$ and $i = 1, \ldots, s$. So, $\gamma_j w_j^T u = \gamma_j w_j^T(\sum_{i=1}^s \alpha_i v_i) = 0$ for all $j = 1, \ldots, r$, and hence

$$0 = -m\sum_{j=1}^r \gamma_j(w_j^T a)^{m-1} w_j^T u + \|u\|^2 = \|u\|^2.$$

Thus, $u = 0$ and $\mathcal{C}a^{m-1} = \lambda a$. So, λ is a Z-eigenvalue of $\mathcal{C}$ with an eigenvector $a \in (\mathrm{span}E_1^Z(\mathcal{C}))^{\perp}$ with $\|a\| = 1$. Moreover, we also have $\lambda = \lambda a^T a = \mathcal{C}a^m = v(P_0)$. We now show that $\lambda \neq \lambda_1(\mathcal{A})$. To see this, we proceed by the method of contradiction and suppose that $\lambda = \lambda_1^Z(\mathcal{C})$. Then, $\mathcal{C}a^m = \lambda_1^Z(\mathcal{C})$ and $\|a\| = 1$. Then, we see that a is an eigenvector of the maximum Z-eigenvalue. This contradicts the fact that $a \in (\mathrm{span}E_1^Z(\mathcal{C}))^{\perp}$ with $\|a\| = 1$. So, by the definition of $\lambda_2^Z(\mathcal{C})$, we have

$$v(P_0) = \lambda \leq \lambda_2^Z(\mathcal{C}).$$

To see the reverse inequality, let a be an eigenvector of $\lambda_2^Z(\mathcal{C})$. Then, we can write $a = \rho x + u$ with $\rho \geq 0$, $x \in \mathrm{span}E_1^Z(\mathcal{C})$ with $\|x\| = 1$ and $u \in (\mathrm{span}E_1^Z(\mathcal{C}))^{\perp}$. It follows that $a^T x = (\rho x +$

DOI: 10.1002/nla

$u)^T x = \rho\|x\|^2 = \rho$. Note that there exists $r \in \mathbb{N}$ such that $\mathcal{C} = \sum_{j=1}^r \gamma_j w_j^m$. It follows that

$$\lambda_2^Z(\mathcal{C})a = \mathcal{C}a^{m-1} = \sum_{j=1}^r \gamma_j (w_j^T a)^{m-1} w_j$$

As $x \in \mathrm{span} E_1^Z(\mathcal{C})$, similar as before, we can show that $\gamma_j w_j^T x = 0$, $j = 1, \ldots, r$. It then follows that

$$\rho\lambda_2^Z(\mathcal{C}) = \lambda_2^Z(\mathcal{C})a^T x = \sum_{j=1}^r \gamma_j (w_j^T a)^{m-1} w_j^T x = 0.$$

As $\lambda_2^Z(\mathcal{C}) < \lambda_1^Z(\mathcal{C}) = 0$. So, $\rho = 0$. This implies that $a = u \in (\mathrm{span} E_1^Z(\mathcal{C}))^\perp$. This together with $\|a\| = 1$ gives us that a is feasible for (P_0). Thus, we see that

$$\lambda_2^Z(\mathcal{C}) \leq v(P_0).$$

This completes the proof. □

Consider a hypergraph $G = (V, E)$ where $V = \{1, \ldots, n\}$ is a finite set of vertices and $E \subseteq 2^V$ is a set consisting of all the hyperedges. The edge cut or coboundary, E_X, of the set $X \subseteq V$ is defined as the set of all hyperedges $e \in E$ such that there are two vertices $u, v \in e$ with $u \in X$ and $v \notin X$. A bisection of G is a 2-partition $\{X, Y\}$ of the vertex set $V = \{1, \ldots, n\}$ in which $|X| = |Y|$ if n is even or $|X| = |Y| - 1$ if n is odd. The bisection problem is to find a bisection for which E_X is as small as possible. The bipartition width $bw(G)$ of the hypergraph G is defined as the optimal value of the bisection problem, that is,

$$bw(G) := \min\{|E_X| : X \subseteq V; |X| = [\frac{n}{2}]\},$$

where $[a]$ denotes the integer part of the number a. Calculating the exact value of the bipartition width is, in general, a hard problem even for the graph case. Below, we shall see that one can use the second largest eigenvalue to provide a lower bound for the bipartition width of a connected hypergraph.

Proposition 5.2
Let $G = (V, E)$ be an m-uniform connected hypergraph where m is an even number and $V = \{1, \ldots, n\}$ is a finite set of vertices. Let $\mathcal{C}$ be the characteristic tensor of G and let $\lambda_2^Z(\mathcal{C})$ be the second largest eigenvalue of $\mathcal{C}$. Let X be a subset of the vertex set V. Then, we have

$$|E_X| \geq \frac{-4\lambda_2^Z(\mathcal{C})}{m^2}\left(\frac{|X||n - X|}{n}\right)^{\frac{m}{2}}.$$

Moreover, we have

$$bw(G) \geq \begin{cases} \frac{-4\lambda_2^Z(\mathcal{C})}{m^2}\left(\frac{n^2-1}{4n}\right)^{\frac{m}{2}}, & \text{if } n \text{ is odd,} \\ \frac{-4\lambda_2^Z(\mathcal{C})}{m^2}\left(\frac{n}{4}\right)^{\frac{m}{2}}, & \text{if } n \text{ is even.} \end{cases}$$

Proof
The inequality is immediate if $X = \emptyset$ and $X = V$. So, let us consider the case when X is a proper nonempty subset of V. Let $w = \sum_{i \in X} e_i$ where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^n$ is the vector whose ith component is one and all the other components are all zero. Let $w = \beta \mathbf{1}_n + u$ where $\beta = \frac{|X|}{n}$ and $u = \sum_{i \in X} e_i - \frac{|X|}{n}\mathbf{1}_n$. Then, $u^T \mathbf{1}_n = 0$. Note that $\mathbf{1}_n \in E_1^Z(\mathcal{C})$ and $\mathrm{span} E_1^z(\mathcal{C})$ is a one dimensional subspace as G is connected. So, $u \in \left(\mathrm{span} E_1^z(\mathcal{C})\right)^\perp$. It is easy to see that $\|u\| = \sqrt{\frac{|X|(n-|X|)}{n}}$. Let $\overline{u} = \frac{u}{\|u\|}$. Then, $\overline{u} \in \left(\mathrm{span} E_1^z(\mathcal{C})\right)^\perp$ and $\|\overline{u}\| = 1$. So, from the preceding proposition, we see that

$$\lambda_2^Z(\mathcal{C}) \geq \mathcal{C}\overline{u}^m = \frac{\mathcal{C}u^m}{\|u\|^m} = \mathcal{C}u^m\left(\frac{n}{|X|(n - |X|)}\right)^{\frac{m}{2}}. \tag{5.20}$$

DOI: 10.1002/nla

Now, from the construction of the characteristic tensor, one can write $\mathcal{C} = \sum_{k=1}^{r} \gamma_k v_k^m$ with $v_k^T \mathbf{1}_n = 0$ for some $r \in \mathbb{N}$ and $v_k \in \mathbb{R}^n$, $k = 1, \ldots, r$. So, we see that

$$\mathcal{C}u^m = \mathcal{C}(w - \beta \mathbf{1}_n)^m = \sum_{k=1}^{r} \gamma_k (v_k^T (w - \beta \mathbf{1}_n))^m = \sum_{k=1}^{r} \gamma_k (v_k^T w)^m = \mathcal{C}w^m = -\sum_{i,j} \delta_{ij}(w_i - w_j)^m. \tag{5.21}$$

Note that $w = \sum_{i \in X} e_i$, and hence,

$$\sum_{i,j} \delta_{ij}(w_i - w_j)^m = \sum_{i \in X, j \notin X, i<j, \{i,j\} \subseteq E} 1 \le \max_{e \in E}\{|e \cap X|(m - |e \cap X|)\}|E_X| \le \frac{m^2}{4}|E_X|,$$

where the last inequality follows as the discrete function $f : \{1, \ldots, m\} \to \mathbb{R}$ defined by $k \mapsto k(m-k)$ attains its maximum at $k = m/2$. Thus, it follows from (5.20) and (5.21) that

$$\begin{aligned} \lambda_2^Z(\mathcal{C}) \ge \mathcal{C}u^m \Big(\frac{n}{|X|(n-|X|)}\Big)^{\frac{m}{2}} &= -\sum_{i,j} \delta_{ij}(w_i - w_j)^m \Big(\frac{n}{|X|(n-|X|)}\Big)^{\frac{m}{2}} \\ &\ge -\frac{m^2}{4}|E_X|\Big(\frac{n}{|X|(n-|X|)}\Big)^{\frac{m}{2}}, \end{aligned}$$

and hence, the first assertion follows. The second assertion follows by taking $|X| = \frac{n}{2}$ if n is even and $|X| = \frac{n-1}{2}$ if n is odd in the first assertion. □

Remark 5.2
When $m = 2$, our lower bounds for the edge cut and bipartition width collapses the classical result for the connected graph cases (cf. [31, 32]). As an illustration, when $m = 2$ the lower bound for edge cuts reads

$$\begin{aligned} |E_X| &\ge -\lambda_2^Z(\mathcal{C})(\frac{|X||n-X|}{n}) \\ &= -\max\{x^T \mathcal{C} x : x^T \mathbf{1}_n = 0, \|x\| = 1\}(\frac{|X||n-X|}{n}) \\ &= \min\{x^T L x : x^T \mathbf{1}_n = 0, \|x\| = 1\}(\frac{|X||n-X|}{n}), \end{aligned}$$

where L is the Laplacian matrix of the graph G. Note that $\min\{x^T L x : x^T \mathbf{1}_n = 0, \|x\| = 1\} = \mu_1$ where μ_1 is the second smallest eigenvalue of the Laplacian matrix L. It follows that $|E_X| \ge \mu_1 \frac{|X||n-X|}{n}$ which is the classical result of the lower bound for the edge cut for a connected graph. The case for the bipartition width is also similar.

6. COMPUTATION OF THE LARGEST/2ND-LARGEST Z-EIGENVALUE

In this section, we explain how one can compute the largest/2nd-largest Z-eigenvalue of symmetric tensors, using polynomial optimization techniques and our variational formula developed in the previous sections.

Computation of the Largest Z-Eigenvalue

In this subsection, we first discuss how to compute the largest Z-eigenvalue for a general symmetric tensor. From Lemma 3.1, the largest Z-eigenvalue of an mth-order n-dimensional symmetric tensor can be computed as

$$\lambda_1^Z(\mathcal{A}) = \max_{\|x\|=1} \mathcal{A}x^m = -\min\{f(x) : h(x) = 0\},$$

DOI: 10.1002/nla

Let $f(x) = -\mathcal{A}x^m$ and $h(x) = \|x\|^2 - 1$. Then, the largest Z-eigenvalue cam be found by solving the following polynomial optimization problem.

$$(P^0) \quad \lambda_1^Z(\mathcal{A}) = -\min\{f(x) : h(x) = 0\}.$$

In general, solving (P^0) is an NP-hard problem when $m \geq 4$. So, finding the largest Z-eigenvalue problem is, in general, also an NP-hard problem. Recently, [18] proposed a shifted power method for finding an Z-eigenvalue of a general symmetric tensor. However, the Z-eigenvalue found by the method developed in [18] need not to be the largest one. Moreover, another power-type method was proposed in [33] for calculating the largest H-eigenvalue (which is a different notion of Z-eigenvalue) of a nonnegative tensor. Unfortunately, this method does not work for finding the Z-eigenvalue and heavily relies on the nonnegative assumption. To the best of our knowledge, the only exact method for finding the largest Z-eigenvalue was established in [43] which only works for the 4th order 3-dimensional symmetric tensor. Below, we introduce a new method for finding the largest Z-eigenvalue for a general symmetric tensor which utilize the polynomial optimization technique. With the help of large-scale SDP solvers, it can be used to find the largest Z-eigenvalue for a medium-size symmetric tensor.

To do this, we recall some basic facts as below. For a real polynomial (polynomial with real coefficients) f, we use $\deg f$ to denote the degree of f. We say that a real polynomial f is *sum of squares* (SOS) if there exist real polynomials $f_j, j = 1, \ldots, r$, such that $f = \sum_{j=1}^r f_j^2$. An important property of the sum of squares polynomials is that checking a polynomial is sum of squares or not, is equivalent to solving a semidefinite programming problem. For details see [19].

For each $k \in \mathbb{N}$, the kth Lasserre's relaxation for solving (P^0) is

$$\begin{array}{lll} (RP_k^0) & \max & \gamma \\ & \text{s.t.} & f(x) - \gamma = \sigma_0(x) + \phi(x)h(x) \\ & & \sigma_0 \text{ is SOS}, \deg\sigma_0 \leq k, \\ & & \phi \text{ are real polynomials}, \deg(\phi h) \leq k. \end{array}$$

As shown in [19], for each fixed $k \in \mathbb{N}$, (RP_k^0) can be equivalently rewritten as a semidefinite programming problem. We note that, a semi-definite programming problem (SDP) can be efficiently solved and has found numerous application in various areas. Moreover, it is easy to see that, for each $k \in \mathbb{N}$, $\max(RP_k^0) \leq \max(RP_{k+1}^0) \leq \min(P^0)$. We now show that the optimal value of the relaxation problem (RP_k^0) asymptotically convergent to $\min(P^0)$ using the celebrated positivity characterization from real algebraic geometry (see Appendix).

Proposition 6.1
It holds that $\lim_{k\to\infty} \max(RP_k^0) = \min(P^0)$.

Proof
Let $r := \min(P^0)$ and fix any $\epsilon > 0$. Then, we have $f(x) - r + \epsilon > 0$ for all $x \in K$ where $K = \{x : h(x) = 0\} = \{x : \pm h(x) \geq 0\}$. Define the quadratic module as follows

$$\mathbf{M}(h, -h) = \{\sigma_0 + (\sigma_1 - \sigma_2)h \mid \sigma_i \text{ is SOS}, i = 0, 1, 2, 3, 4\}.$$

It is clear that $-h \in \mathbf{M}(h, -h)$ and $\{x : -h(x) \geq 0\}$ is compact. Thus, Putinar positivstellensatz (Lemma 7.2) implies that $f - r + \epsilon \in M(h, -h)$. This shows that there exists $k \in \mathbb{N}$, $r - \epsilon \leq \max(RP_k)$. Thus, the conclusion follows. □

Numerical experience indicates that (and recently was justified theoretically in a generical sense in [36]), the relaxation is often exact for small relaxation order k (usually $k \leq 4$). Moreover, one can certify the global optimality by verifying some suitable technical condition called flat truncation condition [19]. On the other hand, the size of the equivalent SDP problem of the relaxation problem increase dramatically when the dimension/order of the tensor increases. For example, as illustrate in Table 1, for a 4th-order 60-dimensional tensor, the equivalent SDP problem for the 4th relaxation

DOI: 10.1002/nla

problem has 1830 variables and 595664 constraints. Fortunately, a robust SDP software (SDPNAL [48]) has been established very recently which enables us to solve large-scale SDP (dimension up to 2000 and number of constraint of the SDP up to 1 million). This, in turn, helps us to find the largest Z-eigenvalue for medium size problem.

Example 6.1
A 4th-order n-dimensional symmetric tensor defined by

$$\mathcal{A}_{ijkl} = \frac{1}{24}(i+j-k-l), \;\; 1 \le i < j < k < l \le n, \text{ and } \mathcal{A}_{ijkl} = \mathcal{A}_{\sigma(ijkl)}$$

where $\sigma(ijkl)$ denotes a permutation of its index $\{i,j,k,l\}$. The corresponding polynomial optimization problem (P^0), in this case, becomes

$$\begin{array}{ll} \min & f(x) := \displaystyle\sum_{1 \le i<j<k<l \le n} (-i-j+k+l) x_i x_j x_k x_l \\ \text{s.t.} & h(x) = \|x\|^2 - 1 = 0. \end{array}$$

Example 6.2
An 4th-order n-dimensional symmetric tensor defined by

$$\mathcal{A}_{ijkl} = \frac{1}{24}(-i-j-k-l), \;\; 1 \le i < j < k < l \le n, \text{ and } \mathcal{A}_{ijkl} = \mathcal{A}_{\sigma(ijkl)}$$

where $\sigma(ijkl)$ denotes a permutation of its index $\{i,j,k,l\}$. The corresponding polynomial optimization problem (P^0), in this case, becomes

$$\begin{array}{ll} \min & f(x) := \displaystyle\sum_{1 \le i<j<k<l \le n} (i+j+k+l) x_i x_j x_k x_l \\ \text{s.t.} & h(x) = \|x\|^2 - 1 = 0. \end{array}$$

Example 6.3
Let n be an even number. Let $V = \{1, \ldots, n\}$. We generate a random graphs $G = (V, E)$ with $|V| = n$ as follows. Select a random subset $M \subseteq V$ with $|M| = n/2$. The edges $e_{i,j}$, $(i,j) \nsubseteq M$, are generated with probability $1/2$. A 4th-order n-dimensional symmetric tensor defined by

$$\mathcal{A}_{iiii} = -1, i = 1, \ldots, n,$$

$$\mathcal{A}_{iijj} = -\frac{1}{3}, \;\; (i,j) \in E, \text{ and } \mathcal{A}_{ijkl} = 0 \text{ otherwise.}$$

The corresponding polynomial optimization problem (P^0), in this case, becomes

$$\begin{array}{ll} \min & f(x) := \displaystyle\sum_{1 \le i \le n} x_i^4 + 2 \sum_{(i,j) \in E} x_i^2 x_j^2 \\ \text{s.t.} & h(x) = \|x\|^2 - 1 = 0. \end{array}$$

In fact, the optimal value of this optimization problem (P^0) indeed returns the stability number of the random graph G we generated.

The following table summarizes the numerical results of Example 6.1-Example 6.3 where we compute the largest eigenvalue by first converting the 4nd-order relaxation of the equivalent polynomial optimization to a SDP problem and solving this SDP problem using SDPNAL. All numerical experiments are performed on a desktop, with 3.47 GHz quad-core Intel E5620 Xeon 64-bit CPUs and 4 GB RAM, equipped with Matlab 7.13 (R2011b). In particular, the data of the following table are explained as follows.

- m: the order of the symmetric tensor,

DOI: 10.1002/nla

- n: the dimension of the symmetric tensor,
- NV: the number of variables of the equivalent SDP problem,
- NC: the number of constraints in the equivalent SDP problem,
- $\lambda_1^Z(\mathcal{A})$: the calculated largest Z-eigenvalue,
- Global Optimality (Y/N) : whether the global optimality is certified or not,
- Time : the CPU-time measured in seconds.

Problem	m	n	NV	NC	$\lambda_1^Z(\mathcal{A})$	Global Optimality (Y/N)	Time (Sec.)
Example 6.1	4	20	210	8854	21.4745	Yes	2.20
Example 6.1	4	30	465	40919	48.3792	Yes	44.59
Example 6.1	4	40	820	123409	87.1374	Yes	246.59
Example 6.1	4	50	1275	292824	140.405	Yes	1159.5
Example 6.2	4	20	210	8854	46.0150	Yes	14.41
Example 6.2	4	30	465	40919	88.8139	Yes	49.72
Example 6.2	4	40	820	123409	136.4154	Yes	290.20
Example 6.2	4	50	1275	292824	187.6926	Yes	1494.64
Example 6.3	4	40	820	123409	15.9999	Yes	159.75
Example 6.3	4	50	1275	292824	18.9999	Yes	806.18
Example 6.3	4	60	1830	595664	24.0001	Yes	3387.17

Table 1

We observe that, for all the above numerical examples, the largest Z-eigenvalues can be found successfully for medium size tensors (dimension ranges from 20 to 60). On the other hand, in general, to find the largest eigenvalue of a large-scale tensor, one has to exploit the structure (e.g. sparsity structure) of the underlying tensor. It would be interesting to see how one could exploit the structure a tensor to reduce the corresponding computation cost in the above algorithm.

Computation of the Second Largest Eigenvalue for Connected Hypergraph

In this subsection, we discuss the computation of the second largest eigenvalue for the characteristic tensor of a hypergraph. For the characteristic tensor $\mathcal{C}$ of a connected hypergraph G, we have seen that the largest eigenvalue of it is zero, and the second largest eigenvalue of it provides a lower bound for the bipartition width. Thus, it is important to compute or estimate the second largest eigenvalue of the characteristic tensor. From our variational formula for second largest eigenvalue (Proposition 5.1) and noting that G is connected, Theorem 5.1 implies that

$$\lambda_2^Z(\mathcal{C}) = \max\{\mathcal{C}x^m : \mathbf{1}_n^T x = 0, \|x\| = 1\}.$$

Let $f(x) = -\mathcal{C}x^m$, $h_1(x) = \mathbf{1}_n^T x$ and $h_2(x) = \|x\|^2 - 1$. Then, the second largest eigenvalue for laplacian tensor for a connected hypergraph G equals the negative optimal value of the following polynomial optimization problem

$$(P) \quad \min\{f(x) : h_1(x) = 0, h_2(x) = 0\}.$$

For each $k \in \mathbb{N}$, the kth Lasserre's relaxation for solving (P) is

$$\begin{array}{lll} (RP_k) & \max & \gamma \\ & \text{s.t.} & f(x) - \gamma = \sigma_0(x) + \phi_l(x)h_1(x) + \phi_2(x)h_2(x) \\ & & \sigma_0 \text{ is SOS}, \deg\sigma_0 \le k, \\ & & \phi_l \text{ are real polynomials}, \deg(\phi_l h_l) \le k,\ l = 1, 2, \end{array}$$

Similarly, we can show that the optimal value of (P) can be approximated by a sequence of semi-definite programming problem.

Proposition 6.2
It holds that $\lim_{k\to\infty}\max(RP_k) = \min(P)$.

Proof
Let $r := \min(P)$ and fix any $\epsilon > 0$. Then, we have $f(x) - r + \epsilon > 0$ for all $x \in K$ where $K = \{x : h_1(x) = 0, h_2(x) = 0\} = \{x : \pm h_1(x) \geq 0, \pm h_2(x) \geq 0\}$. Define the quadratic module as follows

$$M(h_1, -h_1, h_2, -h_2) = \{\sigma_0 + (\sigma_1 - \sigma_2)h_1 + (\sigma_3 - \sigma_4)h_2 \mid \sigma_i \text{ is SOS}, i = 0, 1, 2, 3, 4\}.$$

It is clear that $-h_2 \in M(h_1, -h_1, h_2, -h_2)$ and $\{x : -h_2(x) \geq 0\}$ is compact. Thus, Putinar positivstellensatz (Lemma 7.2) implies that $f - r + \epsilon \in M(h_1, -h_1, h_2, -h_2)$. This shows that there exists $k \in \mathbb{N}$, $r - \epsilon \leq \max(RP_k)$. Thus, the conclusion follows. □

Below, as an illustration, we use two simple numerical examples to explain how one can compute/approximate the second largest eigenvalue for the characteristic tensor of a hypergraph (and so, the bipartition width) using the common global polynomial optimization software Gloptipoly3 [11]. An important feature of the software Gloptipoly3 is that it can certify the global optimality using tools from real algebraic geometry (see [11]).

Example 6.4
Consider the following connected uniform 4-hypergraph $G = (V, E)$ where $V = \{1, 2, 3, 4, 5, 6\}$ and $E = \{(1, 2, 3, 4), (1, 2, 5, 6), (3, 4, 5, 6)\}$. It can be verified that

$$\begin{aligned}\lambda_2^Z(\mathcal{C}) &= \max\{\mathcal{C}x^m : \mathbf{1}_n^T x = 0, \|x\| = 1\}\\ &= -\min\{f(x) : \sum_{i=1}^{6} x_i = 0, \sum_{i=1}^{6} x_i^2 - 1 = 0\},\end{aligned}$$

where

$$\begin{aligned}f(x) = \; & (x_1 - x_2)^4 + (x_1 - x_3)^4 + (x_1 - x_4)^4 + (x_1 - x_5)^4 + (x_1 - x_6)^4\\ & + (x_2 - x_3)^4 + (x_2 - x_4)^4 + (x_2 - x_5)^4 + (x_2 - x_6)^4\\ & + (x_3 - x_4)^4 + (x_3 - x_5)^4 + (x_3 - x_6)^4 + (x_4 - x_5)^4\\ & + (x_4 - x_6)^4 + (x_5 - x_6)^4.\end{aligned}$$

As we explained before, the problem (P) $\min\{f(x) : \sum_{i=1}^{6} x_i = 0, \sum_{i=1}^{6} x_i^2 - 1 = 0\}$ can be solved by using Gloptipoly3. Indeed, by running the following simple code using Gloptipoly3, the software indicates that the 4th order relaxation (RP_4) gives the global minimum 4 of (P) and returns a global minimizer $[-0.4082, 0.4082, -0.4082, 0.4082, 0.4082, -0.4082]^T$.

Thus, we see that $\lambda_2^Z(\mathcal{C}) = -4$. Letting $m = 4$ and $n = 6$, our estimate of bipartition width gives us that $bw(G) \geq \frac{-4\lambda_2^Z(\mathcal{C})}{m^2}(\frac{n}{4})^{\frac{m}{2}} = \frac{9}{4}$. Note that $bw(G)$ must be an integer. So, $bw(G) = 3$. On the other hand, it is easy to verify directly from the definition of bipartition width that $bw(G) = 3$ in this case.

Example 6.5
Consider the following connected uniform 4-hypergraph $G = (V, E)$ where $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and

$$E = \{(1, 2, 3, 4), (3, 4, 5, 6), (5, 6, 7, 8), (6, 7, 8, 9), (7, 8, 9, 10), (1, 2, 9, 10)\}.$$

It can be verified that

$$\begin{aligned}\lambda_2^Z(\mathcal{C}) &= \max\{\mathcal{C}x^m : \mathbf{1}_n^T x = 0, \|x\| = 1\}\\ &= -\min\{f(x) : \sum_{i=1}^{10} x_i = 0, \sum_{i=1}^{10} x_i^2 - 1 = 0\},\end{aligned}$$

DOI: 10.1002/nla

where

$$\begin{aligned} f(x) &= \sum_{j\in\{2,3,4,9,10\}} (x(1)-x(j))^4 + \sum_{j\in\{3,4,9,10\}} (x(2)-x(j))^4 + \sum_{j\in\{4,5,6\}} (x(3)-x(j))^4 \\ &+ \sum_{j\in\{5,6\}} (x(4)-x(j))^4 + \sum_{j\in\{6,7,8\}} (x(5)-x(j))^4 + \sum_{j\in\{7,8,9\}} (x(6)-x(j))^4 \\ &\sum_{j\in\{8,9,10\}} (x(7)-x(j))^4 + \sum_{j\in\{9,10\}} (x(8)-x(j))^4 + (x(9)-x(10))^4; \end{aligned}$$

As we explained before, the problem (P) $\min\{f(x) : \sum_{i=1}^{10} x_i = 0, \sum_{i=1}^{10} x_i^2 - 1 = 0\}$ can be solved by using Gloptipoly3. Indeed, by running a similar simple code as in the preceding example via Gloptipoly3, the software indicates that the 3th order relaxation (RP_3) gives the global minimum 0.4 of (P) and returns a global minimizer

$$[0.4509, 0.4509, 0.2129, 0.2129, -0.1939, -0.1879, -0.4636, -0.4636, -0.0107, -0.0079]^T.$$

So, in this case, $\lambda_2(\mathcal{C}) = -0.4920$.

7. CONCLUSION AND REMARKS

In this paper, using variational analysis techniques, we examined some fundamental analytic properties of Z-eigenvalues of a symmetric tensor with even order. As applications, we introduced the characteristic tensor of a hypergraph and showed that the maximum Z-eigenvalue function of the associated characteristic tensor provides a natural link for the combinatorial structure and the analytic structure of the underlying hypergraph. We also established a variational formula for the second largest Z-eigenvalue for the characteristic tensor of a hypergraph.

Below, we present a few open questions and remarks:

- For an mth order n-dimensional tensor $\mathcal{A}$, we showed that the maximum Z-eigenvalue function is always at least ρth-order semismooth at $\mathcal{A}$ with $\rho = \frac{1}{m(3m-3)^{n-1}-1}$. In our preceding paper [23], ρth-order semismoothness of the maximum H-eigenvalue function at $\mathcal{A}$ with $\rho = \frac{1}{(2m-1)^n}$ is shown under an additional assumption that the geometric multiplicity at $\mathcal{A}$ is one. It would be interesting to see whether our method of proof here can be used to relax the geometric multiplicity assumption in our previous ρth-order semismoothness result for maximum-H eigenvalue function with some appropriate exponent ρ.
- We made use of the concept of Z-eigenvalue to study some very basic properties of hypergraphs. It would be useful to exploit more in this direction to establish further results (for example, bounds on the maximal cliques, chromatic number etc). Recently, some progresses has been made in [5] using the concept of H-eigenvalues via an algebraic approach. It would be also interesting to see how one could link these two approaches together.
- We established the variational characterization of the second largest Z-eigenvalue for the characteristic tensor of an uniform graph. It would be useful to see whether this variational characterization continues to hold for more general tensors or not. Moreover, we have explained how one can compute the largest and the second largest Z-eigenvalue for using polynomial optimization techniques together with our variational characterization. Detail study of the convergence rate and effective error estimate of this method would be interesting topics to examine.

These will be our future research topics and will be examined in a forthcoming study.

ACKNOWLEDGEMENTS

The authors are grateful to the referees and the editor for their constructive comments and helpful suggestions which have contributed to the final preparation of the paper. This work was partially supported by the Hong Kong Research Grant Council (Grant No. PolyU 501909, 502510, 502111 and 501212), the National Natural Science Foundation of China (Grant No. 11001060 and 61262026), the JGZX programm of Jiangxi Province (20112BCB23027), the science and technology programm of Jiangxi Education Committee (LDJH12088) and a research grant from Australian Research Council.

REFERENCES

1. Bonnans J.F. and Shapiro A. *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
2. Chang K.C., Pearson K. and Zhang T. Perron-Frobenius Theorem for nonnegative tensors, *Communications in Mathematical Sciences* 2008; **6**: 507–520.
3. Chang K.C., Pearson K. and Zhang T. On eigenvalue problems of real symmetric tensors, *Journal of Mathematical Analysis and Applications* 2009; **350**: 416–422.
4. Clarke F. H. *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
5. Cooper J. and Dutle A. Spectra of uniform hypergraphs, *Linear Algebra and its Applications* 2012; **436**: 3268–3292.
6. D'Acunto D. and Kurdyka K. Explicit bounds for the Łojasiewicz exponent in the gradient inequality for polynomials. *Annales Polonici Mathematici* 2005; **87**: 51–61.
7. De Lathauwer L. and De Moor B. From matrix to tensor: Multilinear algebra and signal processing. In: J. McWhirter, Editor, *Mathematics in Signal Processing IV, Selected papers presented at 4th IMA Int. Conf. on Mathematics in Signal Processing*, Oxford University Press, Oxford, United Kingdom, 1998, 1–15.
8. De Lathauwer L., De Moor B. and Vandewalle J. On the best rank-1 and rank-$(R^1, R^2, \ldots, R^N)$ approximation of higher-order tensor, *SIAM Journal on Matrix Analysis and Applications* 2000; **21**: 1324–1342.
9. Fiacco A.V. Second order sufficient conditions for weak and strict constrained minima. *SIAM Journal on Applied Mathematics* 1968; **16**: 105–108.
10. Gwoździewicz J. The Łojasiewicz exponent of an analytic function at an isolated zero, *Commentarii Mathematici Helvetici* 1999; **74**: 364–375.
11. Henrion D., Lasserre J. B. and Loefberg J. GloptiPoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software* 2009 **24**: 761–779.
12. Hiriart-Urruty J.B. and Lemaréchal C. *Convex analysis and minimization algorithms. I. Fundamentals*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 305. Springer-Verlag, Berlin, 1993.
13. Hu S. and Qi L. Algebraic connectivity of an even uniform hypergraph, *Journal of Combinatorial Optimization* 2012; **24**: 564–579.
14. Hu S., Li G., Qi L. and Song Y. Finding the maximum eigenvalue of essentially nonnegative symmetric tensors via sum of squares programming, preprint, 2012.
15. Jeyakumar V. and Li G. Necessary global optimality conditions for nonlinear programming with polynomial constraint, *Mathematical Programming* 2011; **126**(2): 393–399.
16. Kofidis E. and Regalia Ph. On the best rank-1 approximation of higher-order symmetric tensors, *SIAM Journal on Matrix Analysis and Applications* 2002; **23**: 863–884.
17. Kolda T. G. and Bader B. W. Tensor decompositions and applications. *SIAM Review* 2009; **51**(3): 455–500.
18. Kolda T. G. and Mayo J. R. Shifted Power Method for Computing Tensor Eigenpairs, *SIAM Journal on Matrix Analysis and Applications* 2011; **32**: 1095–1124.
19. Lasserre J. B. Global optimization with polynomials and the problem of moments, *SIAM Journal on Optimization* 2001; **11**: 796–817.
20. Lewis A.S. Nonsmooth analysis of eigenvalues, *Mathematical Programming* 1999; **84**: 1–24.
21. Li G. Global error bounds for piecewise convex polynomials, *Mathematical Programming* 2011; DOI: 10.1007/s10107-011-0481-z.
22. Li G., On the asymptotically well behaved functions and global error bound for convex polynomials, *SIAM Journal on Optimization* 2010; **20**(4): 1923–1943.
23. Li G., Qi L. and Yu G. Semismoothness of the maximum H-eigenvalue function of a symmetric tensor and its application, *Linear Algebra and its Applications* 2013; **438**(2): 813–833.
24. Li G. and Ng K.F. Error bounds of generalized D-gap functions for nonsmooth and nonmonotone variational inequality problems , *SIAM Journal on Optimization* 2009; **20**(2): 667–690.
25. Lim L.H. Singular values and eigenvalues of tensors, A variational approach, *Proceeding of 1st IEEE International workshop on computational advances of multi-tensor adaptive processing*, 2005, 129-132.
26. Li G., Tang C. and Wei Z. Error bound results for generalized D-gap functions of nonsmooth variational inequality problems, *Journal of Computational and Applied Mathematics* 2010; **233**(11): 2795–2806.
27. Li X., Ng M. and Ye Y., Finding stationary probability vector of a transition probability tensor arising from a higher-order Markov chain, preprint, 2011.
28. Löfberg J. Pre- and post-processing sum-of-squares programs in practice, *IEEE Transactions on Automatic Control* 2009; **54**: 1007-1011.
29. Löfberg J. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. *In Proceedings of the CACSD Conference*, Taipei, 2004.
30. Mifflin R. Semismooth and semiconvex functions in constrained optimization. *SIAM Journal on Control and Optimization* 1977; **15**(6): 959–972.
31. Mohar B., Some applications of Laplace eigenvalues of graphs, in "Graph Symmetry: Algebraic Methods and Applications," Eds. G. Hahn and G. Sabidussi, *NATO ASI Ser. C* 497, Kluwer, 225-275.
32. Mohar B. Laplace eigenvalues of graphs-a survey, *Discrete Mathematics* 1992; **109**: 171–183.

33. Ng M., Qi L. and Zhou G. Finding the largest eigenvalue of a nonnegative tensor, *SIAM Journal on Matrix Analysis and Applications* 2009; **31**: 1090–1099.
34. Ngai H.V. and Thèra M. Error bounds for systems of lower semicontinuous functions in Asplund spaces, *Mathematical Programming* 2009; **116**(1-2): 397–427,
35. NiQ., Qi L. and Wang F. An eigenvalue method for the positive definiteness identification problem, *IEEE Transactions on Automatic Control* 2008; **53**: 1096–1107.
36. Nie J.W. Optimality Conditions and Finite Convergence of Lasserre's Hierarchy, arXiv:1206.0319 [math.OC].
37. Pang J. and Qi L. Nonsmooth equations: motivation and algorithms, *SIAM Journal on Optimization* 1993; **3**(3): 443–465.
38. Qi L. Eigenvalues of a real symmetric tensor, *Journal of Symbolic Computation* 2005; **40**: 1302–1324.
39. Qi L. and Sun J. A nonsmooth version of Newton's method, *Mathematical Programming* 1993; **78**: 353–368.
40. Qi L. and Teo K.L. Multivariate polynomial minimization and its application in signal processing, *Journal of Global Optimization* 2003; **46**: 419–433.
41. Qi L. and Ye Y. Space tensor conic programming, preprint.
42. Qi L., Yu G. and Wu E.X. Higher order positive semi-definite diffusion tensor imaging, *SIAM Journal on Imaging Sciences* 2010; **3**: 416–433.
43. Qi L., Wang F. and Wang Y. Z-eigenvalue methods for a global polynomial optimization problem, *Mathematical Programming* 2009; **118**: 301–316.
44. Reznick B. Some concrete aspects of Hilbert's 17th Problem. *Real algebraic geometry and ordered structures* (Baton Rouge, LA, 1996), 251–272, Contemporary Mathematics, 253, American Mathematical Society, Providence, RI, 2000.
45. Sun D. and Sun J. Semismooth matrix-valued functions, *Mathematics of Operation Research*2002; **27**: 150–169.
46. Ye D. Sensitivity analysis of the greatest eigenvalue of a symmetric matrix via the ϵ-subdifferential of the associated convex quadratic form, *Journal of Optimization Theory and Applications* 1993; **76**(2): 287–304.
47. Zalinescu C. *Convex Analysis in General Vector Spaces*, World Scientific, 2002.
48. Zhao X., Sun D. and Toh K.C. A Newton-CG Augmented Lagrangian Method for Semidefinite Programming, *SIAM Journal on Optimization* 2010; **20**: 1737–1765.

APPENDIX

Nonsmooth Newton method for underdetermined equations

Consider a general nonsmooth equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^l$ with $m \geq l$. The general algorithm of the nonsmooth Newton method for a underdetermined equation is stated as follows:

Algorithm-0

Step 0. Choose $x^{(0)} \in \mathbb{R}^m$. If $G(x^{(0)}) \neq 0$, then set $k := 0$ and go to Step 1. Otherwise, output $x^{(0)}$.

Step 1. Compute a $V_k \in \mathbb{R}^{l\times m}$ such that $V_k \in J_C G(x^{(k)})$.

Step 2. Let $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$ where

$$\Delta x^{(k)} = -V_k^T (V_k V_k^T)^{-1}\, G(x^{(k)}).$$

Step 3. If $G(x^{(k+1)}) \neq 0$, then replace k by $k + 1$ and go back to Step 1. Otherwise, output $x^{(k+1)} \in \mathbb{R}^m$.

We now recall the definition of regular solution and the local convergence result of a nonsmooth Newton method for underdetermined equations which was presented in [23].

Definition 7.1
For a nonlinear equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^l$ with $m \geq l$. We say x^* is a regular solution of this equation if $G(x^*) = 0$ and there exists a neighborhood N of x^* such that

(a) $\mathrm{rank}(V) = l$ for all $V \in J_C G(x)$ and $x \in N \cap \{x : G(x) \neq 0\}$.
(b) $R(V^T (V V^T)^{-1}) \equiv P$ for all $V \in J_C G(x)$ and $x \in N \cap \{x : G(x) \neq 0\}$ where $R(A)$ denotes the range of a $(m \times l)$ matrix A which is defined by $R(A) = \{Ax : x \in \mathbb{R}^l\} \subseteq \mathbb{R}^m$ and P is some vector space in $\mathbb{R}^m$.

Moreover, the vector space P in (b) is called the regular space associated with x^*.

DOI: 10.1002/nla

Lemma 7.1
(cf. [23, Theorem 5.1] Consider an underdetermined equation $G(x) = 0$ where $G : \mathbb{R}^m \to \mathbb{R}^l$ with $m \geq l$. Let x^* be a regular solution and P be the regular space associated with x^*. If we further assume that G is ρth-order semismooth for some $\rho \in (0, 1]$. Then, Algorithm-0 either terminates in finitely many iterations or generates a sequence $\{x^{(k)}\}$ such that $x^{(k)}$ converges to x^* with order $(1 + \rho)$, i.e.,

$$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^{1+\rho}} < +\infty.$$

Positivity Characterization from Real Algebraic Geometry

A *quadratic module generated by polynomials* $g_1, \ldots, g_m \in \mathbb{R}[\underline{x}]$ is defined as $\mathbf{M}(g_1, \ldots, g_m) := \{\sigma_0 + \sigma_1 g_1 + \cdots + \sigma_m g_m \quad | \quad \sigma_i$ is sum of squares polynomial, $i = 0, 1, \ldots, m\}$. We say $\mathbf{M}(g_1, \ldots, g_m)$ is archimedean if there exists $p \in \mathbf{M}(g_1, \ldots, g_m)$ such that $\{x : p(x) \geq 0\}$ is compact.

We now recall the following important certificate for positivity of a polynomial over a semi-algebraic set under the assumption that the associated quadratic module is archimedean (and hence, the semi-algebraic set must be compact).

Lemma 7.2
(Putinar positivstellensatz) Let f, g_j, $j = 1, \ldots, m$, be real polynomials with $K := \{x : g_j(x) \geq 0, j = 1, \ldots, m\} \neq \emptyset$. Suppose that $f(x) > 0$ for all $x \in K$ and $M(g_1, \ldots, g_m)$ is archimedean. Then, $f \in M(g_1, \ldots, g_m)$.

DOI: 10.1002/nla