Finding the Maximum Eigenvalue of Essentially Nonnegative Symmetric Tensors via Sum of Squares Programming *

Shenglong Hu† Guoyin Li‡ Liqun Qi§ Yisheng Song¶

September 16, 2012

Abstract

Finding the maximum eigenvalue of a tensor is an important topic in tensor computation and multilinear algebra. Recently, when the tensor is non-negative in the sense that all of its entries are non-negative, efficient numerical schemes have been proposed to calculate the maximum eigenvalue based on a Perron-Frobenius type theorem for non-negative tensors. In this paper, we consider a new class of tensors called essentially non-negative tensors, which extends the non-negative tensors, and examine the maximum eigenvalue of an essentially non-negative tensor using the polynomial optimization techniques. We first establish that finding the maximum eigenvalue of an essentially non-negative symmetric tensor is equivalent to solving a sum of squares of polynomials (SOS) optimization problem, which, in turn, can be equivalently rewritten as a semi-definite programming problem. Then, using this sum of squares programming problem, we also provide upper as well as lower estimate for the maximum eigenvalue of general symmetric tensors. These upper and lower estimates can be calculated in terms of the entries of the tensor.

Key words. Symmetric Tensors, Maximum Eigenvalue, Sum of Squares of Polynomials, Semi-definite Programming Problem.

AMS Classification. 90C26, 90C22, 65K10, 65F15, 15A72

---

*Research was partially supported by a research grant form the Council of Hong Kong (Grant No. PolyU 501808, 501909, 502510 and 502111) and a research grant from Australian Research Council.

†Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: Tim.Hu@connect.polyu.hk (S. Hu)

‡(Corresponding Author) Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au (G. Li)

§Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: maqilq@polyu.edu.hk (L. Qi)

¶Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: songyisheng1@gmail.com (Y. Song)

1 Introduction

An mth-order n-dimensional tensor $\mathcal{A}$ consists of n^m entries in real number: $\mathcal{A} = (\mathcal{A}_{i_1 i_2 \cdots i_m})$ where each $\mathcal{A}_{i_1 i_2 \cdots i_m}$, $1 \leq i_1, i_2, \cdots, i_m \leq n$, is a real number. 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. Recently, to study the stability of a homogeneous polynomial dynamical system, Qi [1, 2] introduced the definition of eigenvalues of a symmetric tensor and showed that the stability of this system is tied up with the negativity of the maximum eigenvalue of the corresponding symmetric tensor. Independently, Lim [3] also gave such a definition via a variational approach and established an interesting Perron-Frobenius type theorem. Recently, the maximum eigenvalue of a symmetric tensor was shown to be ρ-semismooth with an appropriate estimate on the order ρ in [4] and numerical study on tensors also has attracted a lot of researchers due to its wide applications in polynomial optimization [2], hypergraph theory [5, 6], Higher order markov chain [7], signal processing [8] and image science [9]. 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 (cf. [2,9-14]).

If $\mathcal{A}$ is a non-negative tensor, that is, all the entries of $\mathcal{A}$ are non-negative, then various efficient methods for calculating the largest eigenvalue have been proposed recently. In particular, by using an important Perron-Frobenius theorem for non-negative tensor established in [15], Ng, Qi, and Zhou [14] proposed an iterative method for finding the maximum eigenvalue of an irreducible non-negative tensor. The NQZ method in [14] is efficient but it is not always convergent for irreducible non-negative tensors. Later on, Chang, Pearson and Zhang [16] introduced primitive tensors which is a subclass of irreducible non-negative tensors, and established the convergence of the NQZ method for primitive tensors. Moreover, Liu, Zhou, and Ibrahim [17] modified the NQZ method such that the modified algorithm is always convergent for finding the largest eigenvalue of an irreducible non-negative tensor. Recently, Zhang and Qi [18] established the linear convergence of the NQZ method for essentially positive tensors. Zhang, Qi and Xu [19] established the linear convergence of the LZI method for weakly positive tensors. Yang and Yang [20, 21] generalized the weak Perron-Frobenius theorem to general non-negative tensors. Friedland, Gaubert and Han [22] pointed out that the Perron-Frobenius theorem for non-negative tensors has a very close link with the Perron-Frobenius theorem for homogeneous monotone maps. They introduced weakly irreducible non-negative tensors and established the Perron-Frobenius theorem for such tensors. More recently, numerical method is also presented to calculate the maximum eigenvalue for non-negative tensors without the irreducible assumption by using a partion technique [23].

In this paper, we consider a new class of tensors called essentially non-negative tensors which extends the non-negative tensors, and examine the maximum eigenvalue of an essentially non-negative tensor using the polynomial optimization techniques. We establish that finding the maximum eigenvalue of an essentially non-negative tensor is equivalent to solving a sum of squares (SOS) polynomial optimization problem which is, in turn, can be equivalently rewritten as a semi-definite programming problem. Using this sum of squares programming problem, we also provide upper as well as lower estimates of the maximum eigenvalue of general tensors. These upper and lower estimates can be easily calculated in terms of the entries of the tensor. Numerical examples are also provided to illustrate the significance of the upper and lower estimates. Our approach provides the link between the maximum eigenvalue of a symmetric essentially non-negative tensor and the sum of squares programming problem, and leads to easily verifiable upper and lower estimate for the maximum eigenvalue of general tensors.

The organization of this paper is as follows. We first fix the notations and collect some basic definitions in Section 2. In Section 3, we establish that finding the maximum eigenvalue of an essentially non-negative tensor is equivalent to solving a sum of squares (SOS) polynomial optimization problem which is, in turn, can be equivalently rewritten as a semi-definite programming problem. In Section 4, we provide upper as well as lower estimates of the maximum eigenvalue of general tensors using the sum of squares (SOS) polynomial optimization techniques. Finally, we conclude our paper and present some future research topics in Section 5.

2 Preliminaries

In this section, we fix the notations and collect some basic definitions and facts which we will use later on.

We first recall some basic definitions and facts of tensor and its eigenvalues. We use $\mathbb{R}^n$ to denote the n-dimensional Euclidean space and use $\mathbb{N}$ to denote the set consisting of all the natural numbers. 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 $t \in \mathbb{R}$, $\mathcal{A} = (\mathcal{A}_{i_1,\ldots,i_m})_{1\le i_1,\ldots,i_m \le n}$ and $\mathcal{B} = (\mathcal{B}_{i_1,\ldots,i_m})_{1\le i_1,\ldots,i_m \le n}$

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

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

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

The corresponding norm is defined by $\|\mathcal{A}\|_S = (\langle \mathcal{A}, \mathcal{A} \rangle_S)^{1/2} = \big(\sum_{i_1,\ldots,i_m=1}^{n} \mathcal{A}_{i_1,\ldots,i_m}^2\big)^{1/2}$. 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.,

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

Let $\mathcal{A} \in S$. By the tensor product (cf [8]), $\mathcal{A}x^m$ is a real number defined as

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

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

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

Definition 2.1. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor. We say $\lambda \in \mathbb{R}$ is an H-eigenvalue of $\mathcal{A}$ and $x \in \mathbb{R}^n \backslash \{0\}$ is an H-eigenvector corresponding to λ iff (x, λ) satisfies*

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

To do this, we first formally define the maximum H-eigenvalue function. Since any symmetric tensor with even order always has an H-eigenvalue and the number of H-eigenvalues is finite (cf [24]), it then makes sense to define the maximum eigenvalue function $\lambda_1 : S \to \mathbb{R}$ as follows:

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

The following variational formula [4] for the maximum eigenvalue function plays an important role in our later analysis. For the convenience of the reader, we provide the proof here.

Lemma 2.1. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is even. Then, we have*

$$\lambda_1(\mathcal{A}) = \max_{x \neq 0} \frac{\mathcal{A}x^m}{\|x\|_m^m} = \max_{\|x\|_m = 1} \mathcal{A}x^m ,$$

where $\|x\|_m = (\sum_{i=1}^n |x_i|^m)^{1/m}$.

Proof. Consider the following optimization problem (P)

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

Let $f(x) := \mathcal{A}x^m$ and $g(x) := \|x\|_m^m$. 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, and so the Euler Identity implies that $\nabla g(x)^T x = m g(x)$. Thus, for any x with $g(x) = 1$, $\nabla g(x) \neq 0$. So, the standard KKT theory implies that there exists $\lambda_0 \in \mathbb{R}$ such that

$$m\mathcal{A}x_0^{m-1} - m\lambda_0 x_0^{[m-1]} = \nabla f(x_0) - \lambda_0 \nabla g(x_0) = 0$$

This implies that λ_0 is an H-eigenvalue of $\mathcal{A}$ with an H-eigenvector x_0, and so, $\lambda_0 \leq \lambda_1(\mathcal{A})$. Note that $\lambda_0 = \mathcal{A}x_0^m = v(P)$ where $v(P)$ is the optimal value of (P). It follows that $v(P) \leq \lambda_1(\mathcal{A})$, that is,

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

Finally, noting that, for any eigenvector u corresponds to $\lambda_1(\mathcal{A})$ with $\|u\|_m = 1$, we have

$$\mathcal{A}u^m = u^T(\mathcal{A}u^{m-1}) = \lambda_1(\mathcal{A})u^T u^{[m-1]} = \lambda_1(\mathcal{A})\|u\|_m^m = \lambda_1(\mathcal{A}).$$

Thus, $\lambda_1(\mathcal{A}) = \max_{\|x\|_m = 1} \mathcal{A}x^m$, and so, the conclusion follows as $\max_{\|x\|_m = 1} \mathcal{A}x^m = \max_{x \neq 0} \frac{\mathcal{A}x^m}{\|x\|_m^m}$. □

Remark 2.1. (Convexity of the maximum eigenvalue function) *From the proof of Lemma 2.1, we see that*

$$\{u : \mathcal{A}u^m = \lambda_1(\mathcal{A}),\ \|u\|_m = 1\} = \{u : (\lambda_1(\mathcal{A}), u) \text{ is a real eigenpair of } \mathcal{A},\ \|u\|_m = 1\}.$$

Since $\lambda_1(\mathcal{A}) = \max_{\|x\|_m = 1} \mathcal{A}x^m$, we have $\lambda_1(\mathcal{A}) = \max_{\mathcal{B} \in T} \langle \mathcal{B}, \mathcal{A} \rangle_S$ where $T = \{x^m : \|x\|_m = 1\}$. 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. So λ_1 is a finite valued convex function on the symmetric tensor space S.

Recall that a real polynomial f is called a sum of squares of polynomials iff there exist $r \in \mathbb{N}$ and real polynomials f_j, $j = 1, \ldots, r$, such that $f = \sum_{j=1}^{r} f_j^2$. The set of all sum of squares of real polynomials is denoted by Σ^2. Moreover, the set of all sum of squares real polynomial with degree at most d is denoted by Σ_d^2. An important property of the sum of squares of polynomials is that checking a polynomial is sum of squares or not, is equivalent to solving a semi-definite linear programming problem [25].

3 Essentially Non-negative Tensor

In this section, we show that finding the maximum eigenvalue of an essentially non-negative tensor is equivalent to solving a sum of squares programming problem. To do this, we first recall the definition of an essentially non-negative tensor.

Definition 3.1. *Let $I := \{(i, i, \ldots, i) \in \mathbb{N}^m : 1 \leq i \leq n\}$. We say an mth-order n-dimensional tensor $\mathcal{A}$ is non-negative if $\mathcal{A}_{i_1,\ldots,i_m} \geq 0$ for all $1 \leq i_j \leq n$, $j = 1, \ldots, m$. Moreover, we say an mth-order n-dimensional tensor $\mathcal{A}$ is essentially non-negative if $\mathcal{A}_{i_1,\ldots,i_m} \geq 0$ for all $\{i_1, \ldots, i_m\} \notin I$.*

The class of essentially non-negative tensor was introduced in [26]. From the definition, it is clear that any non-negative tensor is essentially non-negative while the converse may not be true in general. When the order $m = 2$, the definition collapses to the classical definition of essentially non-negative matrices.

Remark 3.1. *As pointed out by one of the referees, one of the important characterizations of the essentially non-negative matrix is the following invariant property: $e^{tA}(\mathbb{R}^n_+) \subseteq \mathbb{R}^n_+$ for all $t \geq 0$ and for all essentially non-negative $(n \times n)$ matrix A. Although some interesting log-convexity results were discussed in [26], it is not clear whether the above interesting invariant property can be extended to essentially non-negative tensors or not. One of the key difficulty is that it is not clear how to define an appropriate analog of matrix exponential for the tensor case. It seems to us that this can be tackled by using the nonlinear operator $T_{\mathcal{A}}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, defined in [16]: for an nonnegative tensor $\mathcal{A}$,*

$$T_{\mathcal{A}}(x) = (\mathcal{A}x^{m-1})^{[\frac{1}{m-1}]}.$$

Indeed, let $\lambda = \max_{1 \leq i \leq n}\{|\mathcal{A}_{i,i,\ldots,i}|\}$ and $\mathcal{I}$ be the identity tensor, i.e., $\mathcal{I}_{i_1,\ldots,i_m} = 1$ whenever $(i_1, \ldots, i_m) \in \{(i, \ldots, i) : 1 \leq i \leq n\}$ and $\mathcal{I}_{i_1,\ldots,i_m} = 0$ otherwise. Then, $\mathcal{A} + \lambda\mathcal{I}$ is a nonnegative tensor. Denote $T^k_{\mathcal{A}+\lambda\mathcal{I}} = \underbrace{T_{\mathcal{A}+\lambda\mathcal{I}} \circ \ldots \circ T_{\mathcal{A}+\lambda\mathcal{I}}}_{\text{k times}}$. One could define $e^{\mathcal{A}}$ as a nonlinear operator from $\mathbb{R}^n$ to $\mathbb{R}^n$ by

$$e^{\mathcal{A}}(x) = (\mathcal{I} + \sum_{k=1}^{\infty} \frac{T^k_{\mathcal{A}+\lambda\mathcal{I}}}{k!})(e^{-\lambda}x), \ \forall \, x \in \mathbb{R}^n.$$

As it is not the main purpose of this paper, we would like to leave the study of invariant property for essentially non-negative tensors as a future research direction and will investigate further in a next paper.

For any essentially non-negative tensor $\mathcal{A}$, we associate it a homogeneous polynomial h defined by $h(x) = -\mathcal{A}x^m$ for all $x \in \mathbb{R}^n$. Below, we present a proposition which shows that any such associated polynomial $h(x)$ is non-negative if and only if it is a sum of squares of polynomials. To do this, we first recall some definitions and a useful lemma.

Consider a homogeneous polynomial $f(x) = \sum_\alpha f_\alpha x^\alpha$ with degree m (m is an even number) where $\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{N} \cup \{0\})^n$, $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ and $|\alpha| := \sum_{i=1}^n \alpha_i = m$. Let $f_{m,i}$ be the coefficient associated with x_i^m and let

$$\Omega_f = \{\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{N} \cup \{0\})^n : f_\alpha \neq 0 \text{ and } \alpha \neq me_i,\ i = 1, \ldots, n\}, \tag{2}$$

where e_i be the vector whose ith component is one and all the other components are zero. We note that

$$f(x) = \sum_{i=1}^n f_{m,i} x_i^m + \sum_{\alpha \in \Omega_f} f_\alpha x^\alpha.$$

Recall that $2\mathbb{N}$ denotes the set consisting of all the even numbers. Define

$$\Delta_f := \{\alpha = (\alpha_1, \ldots, \alpha_n) \in \Omega_f : f_\alpha < 0 \text{ or } \alpha \notin (2\mathbb{N} \cup \{0\})^n\}. \tag{3}$$

We associate to f a new homogeneous polynomial $\tilde{f}$, defined by

$$\tilde{f}(x) = \sum_{i=1}^n f_{m,i}\, x_i^m - \sum_{\alpha \in \Delta_f} |f_\alpha| x^\alpha.$$

We now recall the following useful lemma which provides a convenient test for verifying whether f is a sum of squares of polynomials or not in terms of the non-negativity of the new homogeneous function $\tilde{f}$.

Lemma 3.1. [27, Corollary 2.8] *Let f be a homogeneous polynomial of degree m where m is an even number. If $\tilde{f}$ is a non-negative polynomial, then f is a sum of squares of polynomials.*

Proposition 3.1. *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric essentially non-negative tensor. Let $h(x) = -\mathcal{A}x^m$ for all $x \in \mathbb{R}^n$. Then h is a non-negative polynomial if and only if h is a sum of squares of polynomials.*

Proof. Note that any sum of squares of polynomials is non-negative. We only need to show the converse implication. Suppose that h is a non-negative polynomial. Note that

$$h(x) = -\sum_{i_1,\ldots,i_m=1}^n \mathcal{A}_{i_1 i_2 \cdots i_m} x_{i_1} \ldots x_{i_m} = \sum_{i=1}^n (-\mathcal{A}_{ii\cdots i}) x_i^m + \sum_{(i_1,\ldots,i_m) \notin I} (-\mathcal{A}_{i_1 i_2 \cdots i_m}) x_{i_1} \ldots x_{i_m},$$

where $I := \{(i, i, \ldots, i) \in \mathbb{N}^m : 1 \leq i \leq n\}$. As $\mathcal{A}$ is essentially non-negative, $\mathcal{A}_{i_1 i_2 \cdots i_m} \geq 0$ for all $(i_1, \ldots, i_m) \notin I$. Now, let $h(x) = \sum_{i=1}^n h_{m,i} x_i^m + \sum_{\alpha \in \Omega_h} h_\alpha x^\alpha$. Then, $h_{m,i} = -\mathcal{A}_{ii\cdots i}$ and $h_\alpha < 0$ for all $\alpha \in \Omega_h$ where

$$\Omega_h = \{\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{N} \cup \{0\})^n : h_\alpha \neq 0 \text{ and } \alpha \neq me_i,\ i = 1, \ldots, n\},$$

and e_i is the vector where its ith component is one and all the other components are zero. Recall that $\Delta_h = \{\alpha = (\alpha_1, \ldots, \alpha_n) \in \Omega_h : h_\alpha < 0 \text{ or } \alpha \notin (2\mathbb{N} \cup \{0\})^n\}$. Note that $h_\alpha < 0$ for all $\alpha \in \Omega_h$ and so, $\Delta_h = \Omega_h$. It follows that

$$\begin{aligned} \tilde{h}(x) &:= \sum_{i=1}^n h_{m,i} x_i^m - \sum_{\alpha \in \Delta_h} |h_\alpha| x^\alpha \\ &= \sum_{i=1}^n h_{m,i} x_i^m + \sum_{\alpha \in \Delta_h} h_\alpha x^\alpha \\ &= \sum_{i=1}^n h_{m,i} x_i^m + \sum_{\alpha \in \Omega_h} h_\alpha x^\alpha = h(x). \end{aligned}$$

So, $\tilde{h}$ is a non-negative polynomial. Thus the conclusion follows by Lemma 3.1. $\square$

We now show that finding the maximum eigenvalue of an essentially non-negative tensor is equivalent to solving a sum of squares programming problem.

Theorem 3.1. (Finding the Maximum Eigenvalue of Essentially Non-negative Tensors) *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric essentially non-negative tensor where m is an even number. Let $f(x) = \mathcal{A}x^m$. Consider the following sum of squares problem*

$$(P_0) \qquad \min_{\mu\in\mathbb{R}, r\in\mathbb{R}} \{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}.$$

Then,

$$\lambda_1(\mathcal{A}) = \min(P_0) = \min_{\mu\in\mathbb{R}, r\in\mathbb{R}} \{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}.$$

Proof. Consider a homogeneous polynomial optimization problem

$$(P_0') \qquad \max \quad f(x) \text{ s.t. } \|x\|_m^m = 1.$$

Denote a global maximizer for (P_0') by x^*. Clearly, $\|x^*\|_m = 1$. Define $\mu_0 = f(x^*)$. From Lemma 2.1, $\mu_0 = f(x^*) = \lambda_1(\mathcal{A})$. It follows that for all $x \in \mathbb{R}^n\backslash\{0\}$

$$\begin{aligned} -f(x) + \mu_0\|x\|_m^m &= -f(x) + f(x^*)\|x\|_m^m \\ &= \|x\|_m^m\big(-f(\frac{x}{\|x\|_m}) + f(x^*)\big) \geq 0. \end{aligned}$$

This shows that $-f(x)+\mu_0\|x\|_m^m$ is a non-negative homogeneous polynomial. Note that $-f(x)+\mu_0\|x\|_m^m = -\mathcal{C}x^m$ with

$$\mathcal{C} = \mathcal{A} - \mu_0\mathcal{I}$$

where $\mathcal{I}$ is the identity tensor, i.e., $\mathcal{I}_{i_1,\ldots,i_m} = 1$ whenever $(i_1,\ldots,i_m) \in \{(i,\ldots,i) : 1 \leq i \leq n\}$ and $\mathcal{I}_{i_1,\ldots,i_m} = 0$ otherwise. As $\mathcal{A}$ is an essentially non-negative tensor, $\mathcal{C}$ is also an essentially non-negative symmetric tensor. So, we see that $-f(x) + \mu_0\|x\|_m^m$ is a sum of squares of polynomials by Proposition 3.1. Therefore,

$$f(x^*) - f(x) + \mu_0(\|x\|_m^m - 1) = -f(x) + \mu_0\|x\|_m^m$$

is a sum of square polynomial with degree m. This implies that the optimal value of (P_0) is less or equal to $f(x^*)$. Note that, for any $r \in \mathbb{R}$ and $\mu \in \mathbb{R}$ which is feasible for (P_0), $r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2$. So, we must have

$$r \geq f(x) - \mu(\|x\|_m^m - 1) \text{ for all } x \in \mathbb{R}^n.$$

Letting $x = x^*$, we see that $r \geq f(x^*) - \mu(\|x^*\|_m^m - 1) = f(x^*)$. So, the optimal value of (P_0) is greater than $f(x^*)$. Thus, in this case, we have

$$\lambda_1(\mathcal{A}) = f(x^*) = \min_{\mu\in\mathbb{R}, r\in\mathbb{R}} \{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}.$$

This completes the proof. $\square$

Below, we present a numerical example showing that how to find the maximum eigenvalue for an essentially non-negative tensor via a sum of square programming.

Example 3.1. *Consider a 4th-order 3-dimensional symmetric tensor $\mathcal{A}$ defined by*

$$\mathcal{A}_{1111} = \mathcal{A}_{2222} = \mathcal{A}_{3333} = -4, \ \mathcal{A}_{1333} = \mathcal{A}_{3133} = \mathcal{A}_{3313} = \mathcal{A}_{3331} = 1.$$

Clearly, $\mathcal{A}$ is an essentially non-negative tensor. Let $f(x_1, x_2, x_3) = \mathcal{A}x^m$. Then, we see that $f(x_1, x_2, x_3) = -4x_1^4 - 4x_2^4 - 4x_3^4 + 4x_1x_3^3$. The corresponding equivalent sum of square programming can be written as

$$\begin{aligned} & \min_{\mu\in\mathbb{R}, r\in\mathbb{R}} \{r : r - f(x) + \mu(\|x\|_4^4 - 1) \in \Sigma_4^2\} \\ = & \min_{s_1,s_2\geq 0, r\in\mathbb{R}} \{r : r - f(x) + (s_1 - s_2)(\|x\|_4^4 - 1) \in \Sigma_4^2\} \end{aligned}$$

Solving this sum of squares programming problem via YALMIP (see [28, 29]) *gives us that $\lambda_1(\mathcal{A}) = -1.7205$*

On the other hand, direct calculation shows that for any eigenvalue λ of $\mathcal{A}$, there exists (x_1, x_2, x_3) satisfying

$$\begin{cases} -4x_1^3 + x_3^3 & = & \lambda x_1^3, \\ -4x_2^3 & = & \lambda x_2^3, \\ -4x_3^3 + 3x_1x_3^2 & = & \lambda x_3^3. \end{cases}$$

Solving this homogeneous polynomial equality system gives us that $\lambda = -4$, $\lambda = -4 \pm \sqrt[4]{27}$. So $\lambda_1(\mathcal{A}) = -4 + \sqrt[4]{27} \approx -1.7205$.

Remark 3.2. (Finding the maximum eigenvalue of an essentially non-negative tensor via semi-definite programmings) *Note that testing a polynomial is a sum of squares of polynomials or not is equivalent to solving a semi-definite programming problem. Finding the maximum eigenvalue of an essentially non-negative tensor can be converted as a semi-definite programming problem. More explicitly, let $k \in \mathbb{N}$. Let $P_k(\mathbb{R}^n)$ be the space consisting of all real polynomials on $\mathbb{R}^n$ with degree less than or equal to k and let $C(k, n)$ be the dimension of $P_k(\mathbb{R}^n)$. Write the basis of $P_k(\mathbb{R}^n)$ as*

$$z^{(k)} := [1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_2^2, \ldots, x_n^2, \ldots, x_1^k, \ldots, x_n^k]^T,$$

and let $z_\alpha^{(k)}$ be the α-th coordinate of $z^{(k)}$, $1 \leq \alpha \leq C(k, n)$. Let $f(x) = -\mathcal{A}x^m$ and $g(x) = \|x\|_m^m$. As f and g are polynomials with degree m, we can write

$$f = \sum_{1\leq\alpha\leq C(m,n)} f_\alpha z_\alpha^{(k)} \text{ and } g = \sum_{1\leq\alpha\leq C(m,n)} g_\alpha z_\alpha^{(k)}.$$

Let $p = \frac{m}{2}$. Then, the feasibility problem of the sum of square optimization problem

$$\min_{\mu\in\mathbb{R}, r\in\mathbb{R}} \{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}$$

is equivalent to finding a positive semi-definite symmetric matrix $W \in \mathbb{R}^{C(p,n)\times C(p,n)}$ and $r, \mu \in \mathbb{R}$ such that

$$(r - \mu) + \textstyle\sum_{1\leq\alpha\leq C(m,n)} (-f_\alpha + \mu g_\alpha) z_\alpha^{(m)} = z^{(p)} W z^{(p)},$$

which is in turn equivalent to finding a symmetric positive semi-definite matrix $W \in \mathbb{R}^{C(p,n)\times C(p,n)}$ *such that*

$$\begin{cases} (r-\mu) - f_1 + \mu g_1 = W_{1,1} \\ -f_\alpha + \mu g_\alpha = \sum_{1\le\beta,\gamma\le C(p,n),\beta+\gamma=\alpha} W_{\beta,\gamma} \quad (2 \le \alpha \le C(m,n)). \end{cases}$$

Therefore, the sum of square optimization problem "$\min_{\mu\in\mathbb{R},r\in\mathbb{R}}\{r : r-f(x)+\mu(\|x\|_m^m-1) \in \Sigma_m^2\}$*" is equivalent to the following semi-definite programming problem*

$$\begin{array}{lll} \min\limits_{(\mu,r)\in\mathbb{R}\times\mathbb{R}, W\in S_+^{C(p,n)}} & \mu & \\ s.t. & (r-\mu) - f_1 + \mu g_1 = W_{1,1} \\ & -f_\alpha + \mu g_\alpha = \displaystyle\sum_{1\le\beta,\gamma\le C(p,n),\beta+\gamma=\alpha} W_{\beta,\gamma} \quad (2 \le \alpha \le C(m,n)). \end{array}$$

where $S_+^{C(p,n)}$ *is the space of all symmetric positive semi-definite* $C(p,n)\times C(p,n)$ *matrices. This shows that finding the maximum eigenvalue of an essentially non-negative tensor is equivalent to solving a semi-definite linear programming problem. Note that a semi-definite linear programming problem can be solved efficiently, and numerous mature softwares are available these days for solving semi-definite linear programming problems. This provides a convenient way for finding the maximum eigenvalue of an essentially non-negative tensor. For other optimization problems which can be converted to solving a linear semi-definite programming problem see [30, 31, 32].*

Remark 3.3. (Other approaches for finding the maximum eigenvalue of an essentially non-negative tensor) *We can also develop an alternative approach to find the maximum eigenvalue of an essentially non-negative tensor. Indeed, from the definition, if* $\mathcal{A}$ *is an essentially non-negative tensor, then* $\mathcal{A}+\lambda\mathcal{I}$ *is a nonnegative tensor where* $\lambda = \max_{1\le i\le n}\{|\mathcal{A}_{i,i,\ldots,i}|\}$ *and* $\mathcal{I}$ *is the identity tensor. Since* $\lambda_1(\mathcal{A}+\lambda\mathcal{I}) = \lambda_1(\mathcal{A})+\lambda$*, this suggests that one could develop power method as in [17, 18, 26] to find the maximum eigenvalue of an essentially non-negative tensor, and establish the convergence of the power method whenever the tensor is further assumed to be irreducible. On the other hand, our current approach shows that finding the maximum eigenvalue of an essentially non-negative tensor is equivalent to solving a semi-definite linear programming problem, and so, can be solved efficiently (for example by interior point method). Note that, as pointed out in [33], "most interior-point methods for linear programming have been generalized to semi-definite programs. As in linear programming, these methods have polynomial worst-case complexity and perform very well in practice". This suggests that finding the maximum eigenvalue of an essentially non-negative tensor using our approach here also has polynomial worst-case complexity. Moreover, as we will see later in Section 4, our approach also leads to useful upper estimates for a general tensor.*

Recall that an $n\times n$ matrix is called an Z-matrix (see [30, 31]) if all its off-diagonal elements are non-positive. Extending this, we say an mth-order n-dimensional tensor $\mathcal{A}$ is a Z-tensor if $\mathcal{A}_{i_1,\ldots,i_m} \le 0$ for all $\{i_1,\ldots,i_m\} \notin I$. Clearly, a tensor $\mathcal{A}$ is a Z-tensor if and only if $-\mathcal{A}$ is essentially non-negative. Below, we show that testing whether a Z-tensor is positive definite or not can be reformulated as a sum of squares programming problem.

Corollary 3.1. (Testing Positive Definiteness of Z-Tensors) *Let* $\mathcal{A}$ *be an* m*th-order* n*-dimensional symmetric* Z*-tensor where* m *is an even number. Let* $f(x) = \mathcal{A}x^m$*. Consider the* *sum of squares problem*

$$(P_2) \qquad \min_{\mu\in\mathbb{R},r\in\mathbb{R}}\{r : r+f(x)+\mu(\|x\|_m^m-1) \in \Sigma_m^2\}.$$

Then, $\mathcal{A}$ is positive definite (i.e., $\mathcal{A}x^m > 0$ for all $x \in \mathbb{R}^n\backslash\{0\}$) if and only if the optimal value of problem (P_2) is positive.

Proof. Let $\mathcal{B} = -\mathcal{A}$. So, $\mathcal{A}$ is positive definite if and only if $\lambda_1(\mathcal{B}) < 0$. Note that a tensor $\mathcal{A}$ is a Z-tensor if and only if $-\mathcal{A}$ is essentially non-negative. We see that $\mathcal{B}$ is essentially non-negative. So, the conclusion follows from the preceding theorem. □

In Theorem 3.1, we show that finding the maximum eigenvalue of an essentially non-negative tensor is equivalent to solving a sum of squares programming problem. Next, we show that testing the positivity of the maximum eigenvalue of an essentially non-negative tensor is equivalent to a simpler sum of squares programming problem.

Theorem 3.2. (Testing Positivity of Maximum Eigenvalue of Essentially Non-negative Tensors) *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric essentially non-negative tensor where m is an even number. Let $f(x) = \mathcal{A}x^m$. Consider the sum of squares problem*

$$(P_1) \qquad \min_{\mu\geq 0,r\in\mathbb{R}}\{r : r-f(x)+\mu(\|x\|_m^m-1) \in \Sigma_m^2\}.$$

Then, $\lambda_1(\mathcal{A}) > 0$ if and only if the optimal value of problem (P_1) is positive. Moreover, if $\lambda_1(\mathcal{A}) > 0$, then

$$\lambda_1(\mathcal{A}) = \min_{\mu\geq 0,r\in\mathbb{R}}\{r : r-f(x)+\mu(\|x\|_m^m-1) \in \Sigma_m^2\}.$$

Proof. We first show that $\lambda_1(\mathcal{A}) > 0$ if and only if the optimal value of problem (P) is positive.

[⇐] Suppose that $\lambda_1(\mathcal{A}) \leq 0$. Then, by Lemma 2.1, $f(x) \leq 0$ for all $\|x\|_m = 1$, and so, $f(x) \leq 0$ for all $x \in \mathbb{R}^n$ (as f is homogeneous). So, $-f$ is a non-negative polynomial. By the preceding proposition, we have $-f(x) := -\mathcal{A}x^m$ is a sum of squares of polynomials with degree m. So, $-f(x) = 0 - f(x) + 0 \cdot (\|x\|_m^m - 1) \in \Sigma_m^2$ and hence,

$$\min_{\mu\geq 0,r\in\mathbb{R}}\{r : r-f(x)+\mu(\|x\|_m^m-1) \in \Sigma_m^2\} \leq 0.$$

[⇒] Suppose that $\lambda_1(\mathcal{A}) > 0$. Consider a homogeneous polynomial optimization problem

$$(P_1') \qquad \max \quad f(x) \text{ s.t. } \|x\|_m^m \leq 1.$$

Denote a global maximizer for (P_1') by x^*. We now claim that $\|x^*\|_m = 1$. To see this, we argue by contradiction that $\|x^*\|_m < 1$. Note that $f(x^*) \geq f(0) = 0$. If $f(x^*) > 0$, then, as for all small $t > 0$, $(1+t)x^*$ is still feasible for (P_0), we see that $f\big((1+t)x^*\big) = (1+t)^m f(x^*) > f(x^*)$ which is impossible. So, $f(x^*) = 0$. This implies that $f(x) \leq f(x^*) = 0$ for all $\|x\|_m \leq 1$, and so, $f(x) \leq 0$ for all $x \in \mathbb{R}^n$ (as f is homogeneous). This implies that $\lambda_1(\mathcal{A}) \leq 0$ which is impossible. So, $\|x^*\|_m = 1$. This together with Lemma 2.1 implies that the optimal value of (P_1') equals $\lambda_1(\mathcal{A})$, i.e., $f(x^*) = \lambda_1(\mathcal{A}) > 0$.

Let $\mu_0 = f(x^*) > 0$. Then, it follows that for all $x \in \mathbb{R}^n\backslash\{0\}$

$$\begin{aligned} -f(x) + \mu_0\|x\|_m^m &= -f(x) + f(x^*)\|x\|_m^m \\ &= \|x\|_m^m\big(-f(\frac{x}{\|x\|_m}) + f(x^*)\big) \geq 0. \end{aligned}$$

This shows that $-f(x)+\mu_0\|x\|_m^m$ is a non-negative homogeneous polynomial. Let $\mathcal{C}=\mathcal{A}-\mu_0\mathcal{I}$ where $\mathcal{I}$ is the identity tensor, i.e., $\mathcal{I}_{i_1,\ldots,i_m}=1$ whenever $(i_1,\ldots,i_m)\in\{(i,\ldots,i):1\le i\le n\}$ and $\mathcal{I}_{i_1,\ldots,i_m}=0$ otherwise. Then,

$$-f(x)+\mu_0\|x\|_m^m=-\mathcal{C}x^m.$$

As $\mathcal{A}$ is essentially non-negative and symmetric, we see that $\mathcal{C}$ is also an essentially non-negative symmetric tensor. So, $-f(x)+\mu_0\|x\|_m^m$ is a sum of squares of polynomials by Proposition 3.1. Therefore,

$$f(x^*)-f(x)+\mu_0(\|x\|_m^m-1)=-f(x)+\mu_0\|x\|_m^m$$

is a sum of square polynomial with degree m. This implies that the optimal value of (P_1') is less or equal to $f(x^*)$. Note that, for any $r\in\mathbb{R}$ and $\mu\ge 0$ which is feasible for (P_1'), $r-f(x)+\mu(\|x\|_m^m-1)\in\Sigma_m^2$. So, we must have

$$r\ge f(x)-\mu(\|x\|_m^m-1) \text{ for all } x\in\mathbb{R}^n.$$

Letting $x=x^*$, we see that $r\ge f(x^*)-\mu(\|x^*\|_m^m-1)=f(x^*)$. So, the optimal value of (P_1') is greater than $f(x^*)$. Thus, in this case, we have

$$\lambda_1(\mathcal{A})=f(x^*)=\min_{\mu\ge 0,r\in\mathbb{R}}\{r:r-f(x)+\mu(\|x\|_m^m-1)\in\Sigma_m^2\}. \tag{4}$$

Finally, the second assertion for the formula calculating $\lambda_1(\mathcal{A})$ follows from (4). □

4 Estimates for the Maximum Eigenvalue of General Tensors

In this section, we provide upper as well as lower estimates of the maximum eigenvalue of general symmetric tensors via sum of squares programming problems. To do this, we need the following lemma which provides us a convenient test for determining whether a homogeneous polynomial with only one mixed term is a sum of squares of polynomials or not.

Lemma 4.1. [27, Theorem 2.3] *Let $b_1,\ldots,b_n\ge 0$ and $d\in\mathbb{N}$. Let $a_1,\ldots,a_n\in\mathbb{N}$ be such that $\sum_{i=1}^n a_i=2d$. Consider a homogeneous polynomial f defined by*

$$f(x)=b_1x_1^{2d}+\ldots+b_nx_n^{2d}-\mu x_1^{a_1}\ldots x_n^{a_n}.$$

Define

$$\mu_0=2d\prod_{a_i\neq 0,\,1\le i\le n}(\frac{b_i}{a_i})^{\frac{a_i}{2d}}.$$

Then, the following statements are equivalent:

(1) *f is a non-negative polynomial, i.e., $f(x)\ge 0$ for all $x\in\mathbb{R}^n$;*

(2) *either $|\mu|\le\mu_0$ or $\mu<\mu_0$ and all a_i are even.*

(3) *f is a sum of squares of polynomials.*

Now, we provide the first upper estimate for the maximum eigenvalue of a general symmetric tensor. To do this, for each homogeneous polynomial f with $f(x)=\sum_\alpha f_\alpha x^\alpha$, we first define a set of multi-index as below

$$E_f:=\{\alpha=(\alpha_1,\ldots,\alpha_n)\in\Omega_f:f_\alpha>0 \text{ or } \alpha\notin(2\mathbb{N}\cup\{0\})^n\}, \tag{5}$$

where Ω_f is defined as in (2).

Proposition 4.1. (Upper Estimate for the Maximum Eigenvalue: Type I) *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is an even number. Let $f(x) = \mathcal{A}x^m$. Then, we have*

$$\lambda_1(\mathcal{A}) \le \max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}\},$$

where E_f is defined as in (5).

Proof. Consider a homogeneous polynomial optimization problem

$$(P_0') \qquad \max \quad f(x) \text{ s.t. } \|x\|_m^m = 1.$$

Denote a global maximizer for (P_0') by x^*. From Lemma 2.1, $f(x^*) = \lambda_1(\mathcal{A})$. Note that, for any $r \in \mathbb{R}$ and $\mu \in \mathbb{R}$ which satisfy $r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2$, we must have

$$r \ge f(x) - \mu(\|x\|_m^m - 1) \text{ for all } x \in \mathbb{R}^n.$$

Letting $x = x^*$, we see that $r \ge f(x^*) - \mu(\|x^*\|_m^m - 1) = f(x^*)$. So, we see that

$$\lambda_1(\mathcal{A}) \le \min_{\mu\in\mathbb{R}, r\in\mathbb{R}}\{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}.$$

Let

$$(\overline{\mu}, \overline{r}) := (\max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}\}, \max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}\}).$$

To finish the proof, it suffices to show that $(\overline{\mu}, \overline{r})$ is feasible for the above SOS minimization $\min_{\mu\in\mathbb{R}, r\in\mathbb{R}}\{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}$. This is equivalent to the fact that

$$h(x) := \overline{r} - f(x) + \overline{\mu}(\|x\|_m^m - 1) = -f(x) + \max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}\}\|x\|_m^m \tag{6}$$

is a sum of squares of polynomials. To see this, we first show that, for each $\alpha \in E_f$ with $|\alpha| = m$,

$$\sum_{i=1}^n |f_\alpha|\frac{\alpha_i}{m}x_i^m - f_\alpha x^\alpha$$

is a sum of squares of polynomials. Indeed, since

$$m \prod_{\alpha_i\ne 0, 1\le i\le n} (\frac{|f_\alpha|\frac{\alpha_i}{m}}{\alpha_i})^{\frac{\alpha_i}{m}} = m \prod_{\alpha_i\ne 0, 1\le i\le n} (\frac{|f_\alpha|}{m})^{\frac{\alpha_i}{m}} = |f_\alpha|$$

the preceding lemma (Lemma 4.1) implies that $\sum_{i=1}^n |f_\alpha|\frac{\alpha_i}{m}x_i^m - f_\alpha x^\alpha$ is a sum of squares of polynomials for each $\alpha \in E_f \subseteq (\mathbb{N}\cup\{0\})^n$ with $|\alpha| = m$. This proves the claim. Adding α through E_f, we see that

$$\sum_{i=1}^n \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}x_i^m - \sum_{\alpha\in E_f} f_\alpha x^\alpha$$

is also a sum of squares of polynomials. Let $h(x) = \sum_\alpha h_\alpha x^\alpha$ with degree m (m is an even number) where $\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{N}\cup\{0\})^n$ and $|\alpha| = \sum_{i=1}^n \alpha_i = m$. Let $h_{m,i}$ be the coefficient of h associated with x_i^m and let

$$\Omega_h = \{\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{N}\cup\{0\})^n : h_\alpha \ne 0 \text{ and } \alpha \ne me_i,\ i = 1, \ldots, n\},$$

where e_i be the vector where its ith component is one and all the other components are zero. Define

$$\Delta_h := \{\alpha = (\alpha_1, \ldots, \alpha_n) \in \Omega_h : h_\alpha < 0 \text{ or } \alpha \notin (2\mathbb{N} \cup \{0\})^n\}. \tag{7}$$

From (6), it can be verified that $\Delta_h = E_f$, $h_\alpha = -f_\alpha$ for all $\alpha \in \Delta_h$ and, for each $i = 1, \ldots, n$,

$$h_{m,i} \geq \sum_{\alpha \in E_f} |f_\alpha| \frac{\alpha_i}{m}.$$

This implies that

$$\begin{aligned} h(x) &= \sum_{i=1}^n h_{m,i} x_i^m + \sum_{\alpha \in \Omega_h} h_\alpha x^\alpha \\ &= \sum_{i=1}^n h_{m,i} x_i^m + \sum_{\alpha \in \Delta_h} h_\alpha x^\alpha + \sum_{\alpha \in \Omega_h \backslash \Delta_h} h_\alpha x^\alpha \\ &= \sum_{i=1}^n \sum_{\alpha \in E_f} |f_\alpha| \frac{\alpha_i}{m} x_i^m - \sum_{\alpha \in E_f} f_\alpha x^\alpha + \sum_{i=1}^n (h_{m,i} - \sum_{\alpha \in E_f} |f_\alpha| \frac{\alpha_i}{m}) x_i^m + \sum_{\alpha \in \Omega_h \backslash \Delta_h} h_\alpha x^\alpha. \end{aligned}$$

As $\sum_{\alpha \in \Omega_h \backslash \Delta_h} h_\alpha x^\alpha$ and $\sum_{i=1}^n (h_{m,i} - \sum_{\alpha \in E_f} |f_\alpha| \frac{\alpha_i}{m}) x_i^m$ are sum of squares of polynomials, it follows that h is also a sum of squares of polynomials. □

Below, we provide a different type upper estimate for the maximum eigenvalue of a general symmetric tensor.

Proposition 4.2. (Upper Estimate for the maximum eigenvalue: Type II) *Let $\mathcal{A}$ be an mth-order n-dimensional symmetric tensor where m is an even number. Let $f(x) = \mathcal{A}x^m$. Then, we have*

$$\lambda_1(\mathcal{A}) \leq \max_{1 \leq i \leq n} \{f_{m,i} + \frac{1}{m} \sum_{\alpha \in E_f} |f_\alpha| (\alpha_1^{\alpha_1} \ldots \alpha_n^{\alpha_n})^{\frac{1}{m}}\},$$

where E_f is defined as in (5) and we use the convention that $0^0 = 1$.

Proof. As in the proof of the preceding proposition, we see that

$$\lambda_1(\mathcal{A}) \leq \min_{\mu \in \mathbb{R}, r \in \mathbb{R}} \{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}.$$

Let

$$(\overline{\mu}, \overline{r}) := (\max_{1 \leq i \leq n} \{f_{m,i} + \frac{1}{m} \sum_{\alpha \in E_f} |f_\alpha| (\alpha_1^{\alpha_1} \ldots \alpha_n^{\alpha_n})^{\frac{1}{m}}\}, \max_{1 \leq i \leq n} \{f_{m,i} + \frac{1}{m} \sum_{\alpha \in E_f} |f_\alpha| (\alpha_1^{\alpha_1} \ldots \alpha_n^{\alpha_n})^{\frac{1}{m}}\}).$$

To finish the proof, it suffices to show that $(\overline{\mu}, \overline{r})$ is feasible for the above SOS minimization $\min_{\mu \in \mathbb{R}, r \in \mathbb{R}} \{r : r - f(x) + \mu(\|x\|_m^m - 1) \in \Sigma_m^2\}$. This is equivalent to the fact that h is a sum of squares of polynomials where h is defined by

$$h(x) := \overline{r} - f(x) + \overline{\mu}(\|x\|_m^m - 1) = -f(x) + \max_{1 \leq i \leq n} \{f_{m,i} + \frac{1}{m} \sum_{\alpha \in E_f} |f_\alpha| (\alpha_1^{\alpha_1} \ldots \alpha_n^{\alpha_n})^{\frac{1}{m}}\} \|x\|_m^m. \tag{8}$$

Let $r_\alpha = \frac{1}{m}|f_\alpha|(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}}$ for each $\alpha \in (\mathbb{N}\cup\{0\})^n$ with $|\alpha| = m$. We first show that, for each $\alpha \in E_f \subseteq (\mathbb{N}\cup\{0\})^n$ with $|\alpha| = m$,

$$r_\alpha \sum_{i=1}^n x_i^m - f_\alpha x^\alpha$$

is a sum of squares of polynomials. Indeed, since

$$m \prod_{\alpha_i \neq 0, 1\le i\le n} (\frac{r_\alpha}{\alpha_i})^{\frac{\alpha_i}{m}} = m\frac{r_\alpha}{(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}}} = |f_\alpha|$$

the preceding lemma (Lemma 4.1) implies that $r_\alpha \sum_{i=1}^n x_i^m - f_\alpha x^\alpha$ is a sum of squares of polynomials for each $\alpha \in E_f \subseteq (\mathbb{N}\cup\{0\})^n$ with $|\alpha| = m$. This proves the claim. Adding α through E_f, we see that

$$\sum_{i=1}^n \sum_{\alpha\in E_f} r_\alpha x_i^m - \sum_{\alpha\in E_f} f_\alpha x^\alpha$$

is also a sum of squares of polynomials. Let Δ_h be defined as in (7). From (8), it can be verified that $\Delta_h = E_f$, $h_\alpha = -f_\alpha$ for all $\alpha \in \Delta_h$ and

$$h_{m,i} \ge \frac{1}{m}\sum_{\alpha\in E_f} |f_\alpha|(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}} = \sum_{\alpha\in E_f} r_\alpha.$$

This implies that

$$\begin{aligned} h(x) &= \sum_{i=1}^n h_{m,i}x_i^m + \sum_{\alpha\in\Omega_h} h_\alpha x^\alpha \\ &= \sum_{i=1}^n h_{m,i}x_i^m + \sum_{\alpha\in\Delta_h} h_\alpha x^\alpha + \sum_{\alpha\in\Omega_h\setminus\Delta_h} h_\alpha x^\alpha \\ &= \sum_{i=1}^n h_{m,i}x_i^m - \sum_{\alpha\in E_f} f_\alpha x^\alpha + \sum_{\alpha\in\Omega_h\setminus\Delta_h} h_\alpha x^\alpha \\ &= \sum_{i=1}^n \sum_{\alpha\in E_f} r_\alpha x_i^m - \sum_{\alpha\in E_f} f_\alpha x^\alpha + \sum_{i=1}^n (h_{m,i} - \sum_{\alpha\in E_f} r_\alpha)x_i^m + \sum_{\alpha\in\Omega_h\setminus\Delta_h} h_\alpha x^\alpha. \end{aligned}$$

As $\sum_{\alpha\in\Omega_h\setminus\Delta_h} h_\alpha x^\alpha$ and $\sum_{i=1}^n (h_{m,i} - \sum_{\alpha\in E_f} r_\alpha)x_i^m$ are sum of squares of polynomialss, it follows that h is also a sum of squares of polynomials. □

Next, we provide two examples showing that the upper estimates for the maximum eigenvalue provided in the above two propositions are not comparable, in general.

Example 4.1. *Consider a* 4*th-order* 3*-dimensional symmetric tensor* $\mathcal{A}$ *defined by*

$$\mathcal{A}_{1111} = \mathcal{A}_{2222} = \mathcal{A}_{3333} = 1,\ \mathcal{A}_{1333} = \mathcal{A}_{3133} = \mathcal{A}_{3313} = \mathcal{A}_{3331} = -1.$$

Let $f(x_1,x_2,x_3) = \mathcal{A}x^m$. *Then, we see that* $f(x_1,x_2,x_3) = x_1^4 + x_2^4 + x_3^4 - 4x_1x_3^3$ *and* $E_f = \{(1,0,3)\}$. *Then, Proposition 4.1 implies that*

$$\lambda_1(\mathcal{A}) \le \max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}\} = 4.$$

On the other hand, Proposition 4.2 shows that

$$\lambda_1(\mathcal{A}) \le \max_{1\le i\le n}\{f_{m,i} + \frac{1}{m}\sum_{\alpha\in E_f}|f_\alpha|(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}}\} = 1+\sqrt[4]{27}.$$

So, in this case, Proposition 4.2 gives a better estimate for $\lambda_1(\mathcal{A})$*. In this case, we can verify that* $\lambda_1(\mathcal{A}) = 1+\sqrt[4]{27}$*, and so, the upper estimate for* $\lambda_1(\mathcal{A})$ *in Proposition 4.2 is sharp. Indeed, for any eigenvalue* λ *of* $\mathcal{A}$*, there exists* (x_1, x_2, x_3) *satisfying*

$$\begin{cases} x_1^3 - x_3^3 & = & \lambda x_1^3, \\ x_2^3 & = & \lambda x_2^3, \\ x_3^3 - 3x_1x_3^2 & = & \lambda x_3^3. \end{cases}$$

Solving this homogeneous polynomial equality system gives us that $\lambda = 1$*,* $\lambda = 1 \pm \sqrt[4]{27}$*. So* $\lambda_1(\mathcal{A}) = 1+\sqrt[4]{27}$*.*

Example 4.2. *Consider a* 4*th-order* 3*-dimensional symmetric tensor* $\mathcal{A}$ *defined by*

$$\mathcal{A}_{1111} = \mathcal{A}_{2222} = \mathcal{A}_{3333} = 1,$$

and $\mathcal{A}_{1233} = \mathcal{A}_{2133} = \mathcal{A}_{2313} = \mathcal{A}_{2331} = \mathcal{A}_{1323} = \mathcal{A}_{1332} = \mathcal{A}_{3123} = \mathcal{A}_{3213} = \mathcal{A}_{3312} = \mathcal{A}_{3321} = \mathcal{A}_{3231} = \mathcal{A}_{3132} = -\frac{\sqrt{2}}{6}$*. Let* $f(x_1, x_2, x_3) = \mathcal{A}x^m$*. Then, we see that* $f(x_1, x_2, x_3) = x_1^4 + 4x_2^4 + x_3^4 - \sqrt{8}x_1x_2x_3^2$ *and* $E_f = \{(1, 1, 2)\}$*. Then, Proposition 4.1 implies that*

$$\lambda_1(\mathcal{A}) \le \max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f}|f_\alpha|\frac{\alpha_i}{m}\} = 4+\frac{\sqrt{2}}{2}.$$

On the other hand, Proposition 4.2 shows that

$$\lambda_1(\mathcal{A}) \le \max_{1\le i\le n}\{f_{m,i} + \frac{1}{m}\sum_{\alpha\in E_f}|f_\alpha|(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}}\} = 5.$$

So, in this case, Proposition 4.1 gives a better estimate for $\lambda_1(\mathcal{A})$*.*

Theorem 4.1. (Upper/Lower Estimates for the Maximum Eigenvalue) *Let* $\mathcal{A}$ *be an* m*th-order* n*-dimensional symmetric tensor where* m *is an even number. Let* $f(x) = \mathcal{A}x^m$*. Then, we have*

$$\max_{1\le i\le n} f_{m,i} \le \lambda_1(\mathcal{A}) \le \min\{\max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f}|f_\alpha|\frac{\alpha_i}{m}\}, \max_{1\le i\le n}\{f_{m,i} + \frac{1}{m}\sum_{\alpha\in E_f}|f_\alpha|(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}}\}\}.$$

Proof. We first note that $\lambda_1(\mathcal{A})$ is the optimal value of

$$(P_0') \qquad \max \quad f(x) \text{ s.t. } \|x\|_m^m = 1.$$

So, for each $1 \le i \le n$, $\lambda_1(\mathcal{A}) \ge f(e_i) = f_{m,i}$ where e_i denotes the vector whose ith component is one and all the other components are zero. Thus, $\max_{1\le i\le n} f_{m,i} \le \lambda_1(\mathcal{A})$. Therefore, the conclusion now follows from the preceding two propositions. □

To end this section, we provide two examples. The first example shows that the upper estimate we provided in Theorem 4.1 may not be sharp in general, and has room to improve. On the other hand, the second example shows that there do exist some cases such that the upper estimate is indeed equal to the maximum eigenvalue of the corresponding tensor.

Example 4.3. *Let $\mathcal{A}$ be a $6th$ order 3-dimensional real symmetric tensor such that*

$$f(x) = \mathcal{A}x^6 = -x_3^6 - x_1^2 x_2^4 - x_1^4 x_2^2 + 3x_1^2 x_2^2 x_3^2$$

The polynomial $-f$ is known as the homogeneous Motzkin polynomial (cf. [34]) *which is a polynomial with non-negative values but is not a sum of squares of polynomials (here we consider negative of the homogeneous Motzkin polynomial as calculating $\lambda_1(\mathcal{A})$ is a maximization problem instead of a minimization problem). It can be easily verified that $m = 6$, $n = 3$, and the set $E_f = \{(2, 2, 2)\}$. Direct calculation gives us that*

$$\max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}\} = 1 \text{ and } \max_{1\le i\le n}\{f_{m,i} + \frac{1}{m}\sum_{\alpha\in E_f} |f_\alpha|(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}}\} = 1.$$

So, our preceding upper estimate gives that $\lambda_1(\mathcal{A}) \le 1$. On the other hand, it can be verified that $\lambda_1(\mathcal{A}) = 0$. Therefore, our upper estimate is not sharp in this case.

Example 4.4. *Let $\mathcal{A}$ be a $4th$ order 4-dimensional real symmetric tensor such that*

$$f(x) = \mathcal{A}x^4 = x_4^4 + x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 - 4x_1 x_2 x_3 x_4$$

It can be easily verified that $m = 4$, $n = 4$, and

$$E_f = \{(2, 2, 0, 0); (2, 0, 2, 0); (0, 2, 2, 0); (1, 1, 1, 1)\}.$$

Direct calculation gives us that

$$\max_{1\le i\le n}\{f_{m,i} + \sum_{\alpha\in E_f} |f_\alpha|\frac{\alpha_i}{m}\} = 2 \text{ and } \max_{1\le i\le n}\{f_{m,i} + \frac{1}{m}\sum_{\alpha\in E_f} |f_\alpha|(\alpha_1^{\alpha_1}\dots\alpha_n^{\alpha_n})^{\frac{1}{m}}\} = 3.5.$$

So, our preceding upper estimate gives that $\lambda_1(\mathcal{A}) \le 2$. On the other hand, it can be verified that $\lambda_1(\mathcal{A}) = 2$. Therefore, our upper estimate is indeed equal to the maximum eigenvalue of the corresponding tensor in this case.

5 Conclusion and Remarks

In this paper, we considered a new class of tensors called essentially non-negative tensors. We established that finding the largest eigenvalue of an essentially non-negative tensor is equivalent to solving a sum of squares of polynomials optimization problem, which, in turn, can be equivalently rewritten as a linear semi-definite programming problem. Using this sum of squares programming problem, we also provided upper as well as lower estimate of the maximum eigenvalue of general tensors. These upper and lower estimates can be easily calculated in terms of the coefficients of the tensor. Numerical examples are also provided to illustrate the significance of the upper and lower estimates.

Our approach shows that finding the largest eigenvalue of an essentially non-negative tensor is equivalent to solving a semi-definite linear programming problem. It would be interesting to investigate how to find the other eigenvalues of an essentially non-negative tensor besides the maximum eigenvalue and study the invariant property of essentially non-negative tensors. Moreover, it would be also useful to see how one can improve the upper and lower estimate for the maximum eigenvalue of a general symmetric tensor. These will be our future research topics.

Acknowledgment: The authors would like to express their sincere thanks to the handling editor and the referees for their constructive comments and valuable suggestions which have contributed to the revision of this paper.

References

1. Qi, L.: Eigenvalues of a real symmetric tensor, *Journal of Symbolic Computation,* 40, 1302-1324 (2005).

2. Ni, Q., Qi, L., Wang, F.: An eigenvalue method for the positive definiteness identification problem, *IEEE Transactions on Automatic Control*, 53, 1096-1107, (2008).

3. Lim, L.H.: Singular values and eigenvalues of tensors, A variational approach, *Proc. 1st IEEE International workshop on computational advances of multi-tensor adaptive processing*, 129-132, (2005).

4. Li, G., Qi, L., Yu, G.: Semismoothness of the maximum eigenvalue function of a symmetric tensor and its application, *Linear Algebra and its Applications*, http://dx.doi.org/10.1016/j.laa.2011.10.043, (2011).

5. Hu, S., Qi, L.: Algebraic connectivity of an even uniform hypergraph, *Journal of Combinatorial Optimization*, doi:10.1007/s10878-011-9407-1, (2011).

6. Li, G., Qi, L., Yu, G.: The Z-eigenvalues of a symmetric tensor and its application to spectral hypergraph theory, preprint, (2012).

7. Li, X., Ng, M., Ye, Y.: Finding stationary probability vector of a transition probability tensor arising from a higher-order Markov chain, preprint, (2011).

8. Qi, L., Teo, K.L.: Multivariate polynomial minimization and its application in signal processing, *Journal of Global Optimization*, 46, 419-433, (2003).

9. Qi, L., Yu, G., Wu, E.X.: Higher order positive semi-definite diffusion tensor imaging, *SIAM Journal on Imaging Sciences*, 3, 416-433, (2010) .

10. Kofidis, E., Regalia, Ph.: On the best rank-1 approximation of higher-order symmetric tensors, *SIAM Journal of Matrix Analysis and Applications* 23, 863-884, (2002).

11. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. *SIAM Review* 51, no. 3, 455-500, (2009).

12. De Lathauwer, L., 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, 1-15, (1998).

13. De Lathauwer, L., De Moor, B., Vandewalle, J.: On the best rank-1 and rank-$(R^1, R^2, \ldots, R^N)$ approximation of higher-order tensor, *SIAM Journal of Matrix Analysis and Applications* 21 , 1324-1342, (2000).

14. Ng, M., Qi, L., Zhou, G.: Finding the largest eigenvalue of a non-negative tensor, *SIAM Journal of Matrix Analysis and Applications*, 31, 1090-1099, (2009).

15. Chang, K.C., Pearson, K., Zhang, T.: Perron-Frobenius Theorem for non-negative tensors, *Communications in Mathematical Sciences*, 6, 507-520, (2008).

16. Chang, K.C., Pearson, K., Zhang, T.: Primitivity, the convergence of the NZQ method, and the largest eigenvalue for non-negative tensors, *SIAM Journal on Matrix Analysis and Applications*, 32, 806-819, (2011).

17. Liu, Y., Zhou, G., Ibrahim, N.F.: An always convergent algorithm for the largest eigenvalue of an irreducible non-negative tensor. *Journal of Computational Applied Mathematics* 235, no. 1, 286-292, (2010).

18. Zhang, L., Qi, L.: Linear convergence of an algorithm for computing the largest eigenvalue of a non-negative tensor, *Numerical Linear Algebra and its Applications*, doi: 10.1002/nla.822, (2011).

19. Zhang, L., Qi, L., Xu, Y.: Linear convergence of the LZI algorithm for weakly positive tensors, *Journal of Computational Mathematics*, 30, 24-33, (2012).

20. Yang, Y., Yang, Q.: Further results for Perron-Frobenius theorem for non-negative tensors. *SIAM Journal of Matrix Analysis and Applications* 31, no. 5, 2517-2530, (2010).

21. Yang, Q., Yang, Y.: Further results for Perron-Frobenius Theorem for non-negative tensors II, *SIAM Journal on Matrix Analysis and Applications*, 32, 1236-1250, (2011).

22. Friedland, S., Gaubert, S., Han, L.: Perron-Frobenius theorem for non-negative multilinear forms and extensions, *Linear Algebra and its Applications*, doi:10.1016/j.laa.2011.02.04, (2011).

23. Hu, S., Huang, Z.H., Qi, L.: Finding the spectral radius of a non-negative tensor, arXiv:1111.2138v1 [math.NA].

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

25. Parrilo, P.A. : Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming* Ser. B, 96, No.2, 293-320, (2003).

26. Zhang, L., Qi, L., Luo, Z.: The dominant eigenvalue of an essentially non-negative tensor, arXiv: 1110.6261v1 [math.NA].

27. Fidalgo, C., Kovacec, A.: Positive semidefinite diagonal minus tail forms are sums of squares, *Mathematische Zeitschrift* 269, 629-645, (2011).

28. Löfberg, J.: Pre- and post-processing sum-of-squares programs in practice, *IEEE Transactions on Automatic Control*, 54, 1007-1011, (2009).

29. Löfberg, J.: YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In Proceedings of the CACSD Conference, Taipei, Taiwan, (2004).

30. Jeyakumar, V., Lee, G.M., Li, G.: Alternative theorems for quadratic inequality systems and global quadratic optimization, *SIAM Journal on Optimization*, 20, no. 2, 983-1001, (2009).

31. Jeyakumar, V., Huy, H.Q., Li, G.: Necessary and sufficient conditions for S-lemma and nonconvex quadratic optimization, *Optimization and Engineering*, 10, 491-503, (2009).

32. Li, G.: Global quadratic minimization over bivalent constraints: necessary and sufficient global optimality condition. *Journal of Optimization Theory and its Applications* 152, no. 3, 710-726, (2012).

33. Vandenberghe, L., Boyd, S.: Semidefinite programming, *SIAM Review*, 38, 49-95, (1996).

34. Reznick, B.: Some concrete aspects of Hilbert's 17th Problem. *Real algebraic geometry and ordered structures* (Baton Rouge, LA, 1996), 251–272, Contemp. Math., 253, Amer. Math. Soc., Providence, RI, (2000).