ICIC Express Letters
Part B: Applications
Volume **7**, Number **11**, November **2016**

 ISSN 2185-2766
pp. 2425–2433

UNCONSTRAINED OPTIMIZATION MODELS FOR COMPUTING REAL GENERALIZED EIGENPAIRS OF WEAKLY SYMMETRIC POSITIVE DEFINITION TENSORS

Yonghui Tong, Guangming Zhou* and Qian Zhao

School of Mathematics and Computational Science
Xiangtan University
Yanggutang Western Suburb, Xiangtan 411105, P. R. China
*Corresponding author: zhougm@xtu.edu.cn

Received June 2016; accepted September 2016

Abstract. *This paper focuses on the tensor generalized eigenproblem $\mathcal{A}x^{m-1} = \lambda\mathcal{B}x^{m-1}$, where $\mathcal{A}$ and $\mathcal{B}$ are real weakly symmetric tensors and $\mathcal{B}$ is positive definite. Particularly, when $\mathcal{A}$ is positive definite, we introduce two unconstrained variational models and analyze these models whose global minima are precisely the $\mathcal{B}$-eigenvectors of $\mathcal{A}$ associated with the largest $\mathcal{B}$-eigenvalue of $\mathcal{A}$. The main results are natural extension of unconstrained variational principles for eigenvalues of symmetric matrices. Numerical examples are reported to show the effectiveness of these methods for finding a Z-eigenvalue and an H-eigenvalue of an even order symmetric positive definite tensor.*
Keywords: Positive definite tensors, Weakly symmetric tensors, Tensor eigenproblem, Unconstrained optimization

1. Introduction.

Tensor eigenproblem has become an important topic in numerical multilinear algebra. Eigenvalues of tensors were first introduced by Qi [1] and Lim [13] in 2005. It plays a fundamental role in many fields, such as image analysis [7], spectral graph theory [8], and automatic control. However, it is very difficult to compute eigenvalues of high order tensors, which is an NP-hard problem [9]. Sometimes we only need to calculate the largest eigenvalue of a tensor, for instance, the best rank-one approximation of a symmetric tensor [10]. Various efficient approaches have been proposed for computing eigenvalues of tensors recently, see for example, [2,6,10-12] and the references therein. In particular, Han [2] proposed two unconstrained optimization models for finding generalized eigenpairs of symmetric tensors. The models are given by

$$\min F_1(x) = \frac{1}{2m}(\mathcal{B}x^m)^2 + \frac{1}{m}\mathcal{A}x^m \tag{1}$$

and

$$\min F_2(x) = \frac{1}{2m}(\mathcal{B}x^m)^2 - \frac{1}{m}\mathcal{A}x^m \tag{2}$$

where $\mathcal{A}$ and $\mathcal{B}$ are weakly symmetric tensors. He proved that problem (1) and problem (2) can be used to find the smallest and largest $\mathcal{B}$-eigenvalue of $\mathcal{A}$, respectively. In [6], a Jacobian semi-definite relaxation approach was presented to compute all of the real eigenvalues of symmetric tenors. A shifted higher order power method was proposed for computing Z-eigenpairs in [10]. In [11], an adaptive version of higher order power method was presented for generalized eigenpairs of symmetric tensor. A subspace projection method was proposed in [12] for Z-eigenvalues of symmetric tensors.

In this paper, we mainly study how to find the largest $\mathcal{B}$-eigenvalue of weakly symmetric positive definite tensors. Our first goal is to analyze some functions whose global minima point are the eigenvectors associated with the largest $\mathcal{B}$-eigenvalue of $\mathcal{A}$. An alternative method for calculating the largest $\mathcal{B}$-eigenvalue of a symmetric tensor $\mathcal{A}$ is to solve the constrained optimization problem

$$\max \mathcal{A}x^m \text{ s.t. } \mathcal{B}x^m = 1, \tag{3}$$

and the critical points of problem (3) are $\mathcal{B}$-eigenvectors of $\mathcal{A}$. However, the nonlinearity of constraint condition makes this problem less attractive [2]. It is a more attractive approach to compute eigenvalues of even order tensor by using unconstrained optimization. We further introduce two unconstrained variational principles for calculating $\mathcal{B}$-eigenpairs of weakly symmetric positive definite tensors. One of our purposes is to compare the performance of these variational principles for finding the largest $\mathcal{B}$-eigenvalue. Some numerical experiments illustrated that our models are faster than Han's models in [2] and could reach the largest eigenpair with a higher probability.

The rest of our paper is organized as follows. In Section 2, we list some notations and preliminary results in numerical multilinear algebra and polynomial optimization. In Section 3, we propose two unconstrained optimization problems and analyze some variational characterizations for the largest $\mathcal{B}$-eigenvalues of tensor $\mathcal{A}$. In Section 4, we give some numerical results. We finish the paper with some conclusions and discussions in Section 5.

2. **Preliminaries.** Tensor is also referred to as the multidimensional array. Let $\mathbb{R}$ be the real filed, and let m and n be positive integers. A real-valued mth-order n-dimensional tensor is indexed as

$$\mathcal{A} = (\mathcal{A}_{i_1 i_2 \cdots i_m}) \in \mathbb{R}^{n\times n\times\cdots\times n}, \quad 1 \le i_1, i_2, \cdots, i_m \le n.$$

The tensor A is symmetric if each entry $\mathcal{A}_{i_1 i_2 \cdots i_m}$ is invariant under any permutation of $(i_1, i_2, \cdots, i_m)$. A real-valued mth-order n-dimensional tensor uniquely defines an m degree homogeneous polynomial function

$$\mathcal{A}x^m = \sum_{i_1, i_2, \cdots, i_m}^{n} \mathcal{A}_{i_1 i_2 \cdots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}.$$

If $\mathcal{A}x^m$ is positive (or nonnegative) for all $x \in \mathbb{R}^n \setminus \{0\}$, then we call $\mathcal{A}$ is positive definite (or positive semidefinite). The $\mathcal{A}x^{m-1}$ is a vector in $\mathbb{R}^n$ with its ith component as

$$\left(\mathcal{A}x^{m-1}\right)_i = \sum_{i_2, \cdots, i_m}^{n} \mathcal{A}_{i i_2 \cdots i_m} x_{i_2} \cdots x_{i_m}.$$

For a vector $x := (x_1, x_2, \cdots, x_n)^{\mathsf{T}} \in \mathbb{R}^n$ and a positive integer k, we denote

$$x^{[k]} := \left(x_1^k, x_2^k, \cdots, x_n^k\right)^{\mathsf{T}}.$$

If the gradient of $\mathcal{A}x^m$ satisfies

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

then $\mathcal{A}$ is called weakly symmetric. If $\mathcal{A}$ is symmetric, then it is weakly symmetric, but the converse is not true in general [5].

There are many different definitions of eigenvalues for tensors have been proposed in the literature, particularly, the real eigenvalues including H-eigenvalues [1,13], Z-eigenvalues [1,13] and D-eigenvalues [14]. In [5], the unified notion of eigenvalues was given for all mth-order n-dimensional tensors, which is called generalized eigenvalue.

Definition 2.1. ([5]) *Let $\mathcal{A}$ and $\mathcal{B}$ be the mth-order n-dimensional real weakly symmetric tensors. Assume further that m is even and $\mathcal{B}$ is positive definite. We say $(\lambda, x) \in \mathbb{R} \times \{R^n \setminus \{0\}\}$ is a generalized eigenpair of $(\mathcal{A}, \mathcal{B})$. If*

$$\mathcal{A}x^{m-1} = \lambda \mathcal{B}x^{m-1},$$

we also call λ is a $\mathcal{B}$-eigenvalue of $\mathcal{A}$ and x a $\mathcal{B}$-eigenvector with respect to λ. In [2], the $\mathcal{B}$-spectrum of $\mathcal{A}$ was given by

$$\sigma_{\mathcal{B}}(\mathcal{A}) = \{\lambda : \lambda \text{ is a } \mathcal{B}\text{-eigenvalue of } \mathcal{A}\}.$$

When $\mathcal{B}$ takes different tensors, $\mathcal{B}$-eigenvalue reduces to be different types of eigenvalues. In [5], the authors have given some special $\mathcal{B}$-eigenvalues, and we recall them as follows.

- When $\mathcal{B} = \mathcal{I}$, I is the identity tensor; when $\mathcal{B}x^m = \|x\|_m^m$ and $\mathcal{B}x^{m-1} = x^{[m-1]}$, the $\mathcal{B}$-eigenvalues are just the H-eigenvalues [1,13].
- When $\mathcal{B} = \mathcal{I}_n^{m/2}$, the tensor product of $m/2$ copies of identity matrix $\mathcal{I}_n \in \mathbb{R}^{n\times n}$. When $\mathcal{B}x^m = \|x\|_2^m$ and $\mathcal{B}x^{m-1} = \|x\|_2^{m-2}x$, the $\mathcal{B}$-eigenvalues are just the Z-eigenvalues [1,13].
- When $\mathcal{B} = D^{m/2}$, the tensor product of $m/2$ copies of the symmetric matrix $D \in \mathbb{R}^{n\times n}$. When $\mathcal{B}x^m = x^TDx$ and $\mathcal{B}x^{m-1} = (x^{\mathrm{T}}Dx)^{\frac{m-2}{2}}x$, the $\mathcal{B}$-eigenvalues are just the D-eigenvalues [14].

In [5], Chang et al. proved that any mth-order n-dimensional real weakly symmetric tensor $\mathcal{A}$ has at least n $\mathcal{B}$-eigenvalues. Han [2] proved the existence of real extreme $\mathcal{B}$-eigenvalues of even order symmetric tensors, which is summarized as follows.

Theorem 2.1. ([2]) *Assume that $\mathcal{A}$ and $\mathcal{B}$ are the mth-order n-dimensional real weakly symmetric tensors and $\mathcal{B}$ is positive definite. Then $\sigma_{\mathcal{B}}(\mathcal{A})$ is not empty. Furthermore, there exist $\lambda_{\min} \in \sigma_{\mathcal{B}}(\mathcal{A})$ and $\lambda_{\max} \in \sigma_{\mathcal{B}}(\mathcal{A})$ such that*

$$-\infty < \lambda_{\min} \le \lambda \le \lambda_{\max} < \infty, \quad \forall \lambda \in \sigma_{\mathcal{B}}(\mathcal{A}).$$

The following theorem gives an important property of weakly symmetric positive definite tensors.

Theorem 2.2. *Assume that $\mathcal{B}$ is an mth-order n-dimensional weakly symmetric positive definite tensor. Let $\mu > 0$ and $\nu > 0$ be the smallest and largest H-eigenvalue of $\mathcal{B}$, respectively. Then*

$$\mu\|x\|_m^m \le \mathcal{B}x^m \le \nu\|x\|_m^m, \quad \forall x \in \mathbb{R}^n, \tag{4}$$

where $\|x\|_m$ is the m-norm of x.

Proof: Actually, the left side of inequality (4) is the (2.2) of Theorem 3 in [2]. Similar to the proof of [2], we can prove the right side of inequality (4). More precisely, when $x = 0$, the right side of inequality (4) obviously holds. It follows from Theorem 2.2 that ν is the global maximum value of

$$\max \mathcal{B}x^m, \text{ s.t. } \|x\|_m^m = 1,$$

then, for any $x \in \mathbb{R} \setminus \{0\}$, we have

$$\mathcal{B}\left(\frac{x}{\|x\|_m}\right)^m \le \nu,$$

and thus, $\mathcal{B}x^m \le \nu\|x\|_m^m$. The proof is completed. □

In [2], the notion of coercive functions was introduced.

Definition 2.2. ([2]) *Assume that function $f : \mathbb{R}^n \to \mathbb{R}$. If $f(x)$ satisfies $\lim\limits_{\|x\|\to\infty} f(x) = +\infty$, then we call it coercive.*

The coercive functions have a fundamental property which is summarized as the following.

Theorem 2.3. ([2]) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous. If $f(x)$ is coercive, then $f(x)$ has at least one global minimizer. If, in addition, the first partial derivatives exist on $\mathbb{R}^n$, then $f(x)$ attains its global minimizers at its critical points.*

3. **Variational Principles for the Largest $\mathcal{B}$-Eigenvalue of Real Weakly Symmetric Positive Definite Tensors.** Auchmuty [3] proposed some unconstrained variational principles for real symmetric matrix eigenproblem. Particularly, he considered the unconstrained optimization problem

$$\min f_1(x) = \langle Bx, x\rangle - 2\sqrt{\langle Ax, x\rangle}, \tag{5}$$

where $A, B \in \mathbb{R}^{n\times n}$ are symmetric matrices and A is positive semidefinite. He proved that problem (5) can be used to calculate an extreme eigenpair of $Ax = \lambda Bx$. For Auchmuty's variational principles, Mongeau and Torki [4] have made a detailed numerical analysis. Furthermore, they also proposed a new variational principle for finding the largest eigenvalue of a positive definite matrix. They considered the problem

$$\min f_2(x) = \|x\|^2 - \ln \langle Ax, x\rangle, \tag{6}$$

where A is positive definite matrix. The properties of the function f_1 and f_2 were discussed in [3] and [4], respectively. Based on these, we next introduce two unconstrained optimization models for computing generalized eigenpairs of even order symmetric positive definite tensors.

3.1. **The first unconstrained variational model.** We now consider the function $S_{\mathcal{A},\mathcal{B}} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$S_{\mathcal{A},\mathcal{B}}(x) := \mathcal{B}x^m - m(\mathcal{A}x^m)^{\frac{1}{m}}, \tag{7}$$

where $\mathcal{A}$ and $\mathcal{B}$ are weakly symmetric tensors and $\mathcal{A}$ is positive semidefinite. The model (7) is a natural extension of model (5). Obviously, $S_{\mathcal{A},\mathcal{B}}$ is smooth on $\mathbb{R}^n \setminus \{0\}$. By simple calculation, it is easy to get the gradient of $S_{\mathcal{A},\mathcal{B}}$, which is given by

$$\nabla S_{\mathcal{A},\mathcal{B}}(x) = m\left(\mathcal{B}x^{m-1} - (\mathcal{A}x^m)^{\frac{1-m}{m}}\mathcal{A}x^{m-1}\right). \tag{8}$$

We summarize the properties of the function $S_{\mathcal{A},\mathcal{B}}$ as follows.

Theorem 3.1. *Assume that $\mathcal{A}$ and $\mathcal{B}$ are the mth-order n-dimensional real weakly symmetric positive definite tensors. Let $\lambda_{\max}$ be the largest $\mathcal{B}$-eigenvalue of $\mathcal{A}$, and $S_{\mathcal{A},\mathcal{B}}$ be defined in (7). Then*
(a) $S_{\mathcal{A},\mathcal{B}}$ is coercive on $\mathbb{R}^n$.
(b) The minimum of $S_{\mathcal{A},\mathcal{B}}$ is

$$\min_{x\in\mathbb{R}^n} S_{\mathcal{A},\mathcal{B}}(x) = (1-m)(\lambda_{\max})^{\frac{1}{m-1}},$$

which is attained at any $\mathcal{B}$-eigenvector x corresponding to the $\mathcal{B}$-eigenvalue $\lambda_{\max}$.
(c) The (nonzero) critical points of $S_{\mathcal{A},\mathcal{B}}$ are any $\mathcal{B}$-eigenvector x of $\mathcal{A}$ associated with a $\mathcal{B}$-eigenvalue λ of $\mathcal{A}$ satisfying $\lambda =: (\mathcal{A}x^m)^{\frac{m-1}{m}}$.

Proof: Since $\mathcal{B}$ is weakly symmetric positive definite, by Theorem 2.2, we have

$$\mathcal{B}x^m \geq \mu\|x\|_m^m, \quad \forall x \in \mathbb{R}^n,$$

where $\mu > 0$ is the smallest H-eigenvalue of tensor $\mathcal{B}$, then we have

$$S_{\mathcal{A},\mathcal{B}} \geq \mu\|x\|_m^m - m(\mathcal{A}x^m)^{\frac{1}{m}},$$

and since $\mathcal{A}$ is positive definite, for any x, we also have $\mathcal{A}x^m \leq \tau\|x\|_m^m$, where $\tau > 0$ is the largest H-eigenvalue of $\mathcal{A}$. Thus,

$$S_{\mathcal{A},\mathcal{B}} \geq \mu\|x\|_m^m - m(\tau\|x\|_m^m)^{\frac{1}{m}} \to \infty \text{ as } \|x\| \to \infty,$$

this shows the function $S_{\mathcal{A},\mathcal{B}}$ is coercive on $\mathbb{R}^n$, and then *(a)* holds.

Since the gradient $\nabla S_{\mathcal{A},\mathcal{B}} = 0$ at any critical point, namely, the critical points of $S_{\mathcal{A},\mathcal{B}}$ satisfy the equation

$$\mathcal{A}x^{m-1} = (\mathcal{A}x^m)^{\frac{m-1}{m}}\mathcal{B}x^{m-1} \tag{9}$$

This indicates that a nonzero critical point x is a $\mathcal{B}$-eigenvector corresponding to the $\mathcal{B}$-eigenvalue $\lambda =: (\mathcal{A}x^m)^{\frac{m-1}{m}}$, as claimed in *(c)*.

Taking the inner product of (9) with x, we obtain

$$\mathcal{A}x^m = (\mathcal{A}x^m)^{\frac{m-1}{m}}\mathcal{B}x^m,$$

so $\mathcal{B}x^m = (\mathcal{A}x^m)^{\frac{1}{m}}$. Moreover, the critical value of $S_{\mathcal{A},\mathcal{B}}$ at this critical point x is

$$S_{\mathcal{A},\mathcal{B}}(x) = (1-m)\lambda^{\frac{1}{m-1}}.$$

Since $S_{\mathcal{A},\mathcal{B}}$ is continuous and coercive on $\mathbb{R}^n$, by Theorem 2.3, a minimum is attained for $S_{\mathcal{A},\mathcal{B}}$. Obviously, the global minimal value of $S_{\mathcal{A},\mathcal{B}}$ is $\min\limits_{x\in\mathbb{R}^n} S_{\mathcal{A},\mathcal{B}}(x) = (1-m)(\lambda_{\max})^{\frac{1}{m-1}}$, so *(b)* holds. The proof is completed. $\square$

Remark 3.1. *It can be seen from the above process, at any critical point, the $\mathcal{B}$-eigenvector satisfies $\lambda =: (\mathcal{A}x^m)^{\frac{m-1}{m}}$ and since $\mathcal{B}x^m = (\mathcal{A}x^m)^{\frac{1}{m}}$, then the $\mathcal{B}$-eigenvector also satisfies $\lambda =: (\mathcal{B}x^m)^{m-1}$.*

3.2. **The second unconstrained variational model.** We then consider the function $L_{\mathcal{A},\mathcal{B}} : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n$ defined by

$$L_{\mathcal{A},\mathcal{B}}(x) := \mathcal{B}x^m - \ln(\mathcal{A}x^m), \tag{10}$$

where $\mathcal{A}$ and $\mathcal{B}$ are weakly symmetric tensors and $\mathcal{A}$ is positive definite. The model (10) is a natural extension of model (6). This function is continuous and smooth on $\mathbb{R}^n \setminus \{0\}$. The gradient of $L_{\mathcal{A},\mathcal{B}}$ is given by

$$\nabla L_{\mathcal{A},\mathcal{B}}(x) = m\left(\mathcal{B}x^{m-1} - \frac{\mathcal{A}x^{m-1}}{\mathcal{A}x^m}\right) \tag{11}$$

The properties of this function are summarized by the following theorem.

Theorem 3.2. *Assume that $\mathcal{A}$ and $\mathcal{B}$ are the mth-order n-dimensional real weakly symmetric positive definite tensors. Let $\lambda_{\max}$ be the largest $\mathcal{B}$-eigenvalue of $\mathcal{A}$, and $L_{\mathcal{A},\mathcal{B}}$ be defined in (10). Then*
(a) $L_{\mathcal{A},\mathcal{B}}$ is coercive on $\mathbb{R}^n \setminus \{0\}$.
(b) The critical points of $L_{\mathcal{A},\mathcal{B}}$ are any $\mathcal{B}$-eigenvector x associated with a $\mathcal{B}$-eigenvalue λ.
(c) The minimum of $L_{\mathcal{A},\mathcal{B}}$ is

$$\min_{x\in\mathbb{R}^n\setminus\{0\}} L_{\mathcal{A},\mathcal{B}}(x) = 1 - \ln(\lambda_{\max}),$$

which is attained at any $\mathcal{B}$-eigenvector x corresponding to the $\mathcal{B}$-eigenvalue $\lambda_{\max}$, satisfying $\lambda_{\max} = \mathcal{A}x^m$.

Proof: *(a)* According to Theorem 2.2, for each $x \in \mathbb{R}^n$, we have

$$\mathcal{B}x^m \geq \mu\|x\|_m^m,$$

where $\mu > 0$ is the smallest H-eigenvalue of real weakly symmetric positive definite tensor $\mathcal{B}$. Hence, we have

$$L_{\mathcal{A},\mathcal{B}} \geq \mu\|x\|_m^m - \ln(\mathcal{A}x^m),$$

and for any $x \in \mathbb{R}^n$, we also have $\mathcal{A}x^m \leq \tau\|x\|_m^m$, where $\tau > 0$ is the largest H-eigenvalue of $\mathcal{A}$. Thus,

$$L_{\mathcal{A},\mathcal{B}} \geq \mu\|x\|_m^m - \ln(\tau\|x\|_m^m) \to \infty \text{ as } \|x\| \to \infty.$$

This shows the coercivity of $L_{\mathcal{A},\mathcal{B}}$.
(b) Note that the gradient $\nabla L_{\mathcal{A},\mathcal{B}} = 0$ at any critical point of $L_{\mathcal{A},\mathcal{B}}$, that is

$$\mathcal{A}x^{m-1} = \mathcal{A}x^m\mathcal{B}x^{m-1}. \tag{12}$$

This shows that a critical point x is a $\mathcal{B}$-eigenvector corresponding to the $\mathcal{B}$-eigenvalue $\lambda =: \mathcal{A}x^m$.

(c) Note that $L_{\mathcal{A},\mathcal{B}}$ is continuous and coercive on $\mathbb{R}^n\backslash\{0\}$, by Theorem 2.3, and a minimum is attained for $L_{\mathcal{A},\mathcal{B}}$. Taking the inner product of (12) with x, we then have

$$\mathcal{A}x^m = \mathcal{A}x^m\mathcal{B}x^m,$$

so $\mathcal{B}x^m = 1$. Moreover, the critical value of $L_{\mathcal{A},\mathcal{B}}$ at critical point x is

$$L_{\mathcal{A},\mathcal{B}}(x) = 1 - \ln\lambda.$$

Clearly, the global minimal value of $L_{\mathcal{A},\mathcal{B}}$ is $\min_{x\in\mathbb{R}^n\backslash\{0\}} L_{\mathcal{A},\mathcal{B}}(x) = 1 - \ln(\lambda_{\max})$. The proof is completed. □

Remark 3.2. *Both (7) and (10) are unconstrained optimization problems; thus, any local optimization methods can be used to solve such problems. However, these methods cannot guarantee finding a global maximum; they only converge to a critical point. It follows from Theorem 3.1 and Theorem 3.2 that each critical point corresponds to a $\mathcal{B}$-eigenvalue of tensor $\mathcal{A}$. Therefore, when $\mathcal{B} = \mathcal{I}$, $\mathcal{B} = I_n^{m/2}$ and $\mathcal{B} = D^{m/2}$, we can obtain Z-eigenvalue, H-eigenvalue and D-eigenvalue of tensor $\mathcal{A}$, respectively.*

4. **Numerical Experiments.**

In this section, we present some numerical results to illustrate the effectiveness of using the unconstrained variational principles (7) and (10) to calculate the largest Z-eigenvalues and H-eigenvalues of some positive definite even order symmetric tensors. The experiments were done on a desktop computer with an Intel(R) Core(TM)2 Duo CPU E7500 @2.93 GHz and a 2GB RAM running Windows 7, using MATLAB R2013a, the Tensor Toolbox [16]. We use limited-memory quasi-Newton (L-BFGS) method [15] to solve (7) and (10) and compare the performance of new variational principles with the Han's method [2]. The parameters of L-BFGS in our test are given as follows:

$$\mathtt{m} = 3,\ \beta' = 0.01,\ \beta = 0.9,\ \mathtt{TolX} = 10^{-8},\ \mathtt{TolFun} = 10^{-8},\ \mathtt{MaxIter} = 1000.$$

Example 4.1. *(see [6]) Consider the 4th-order 3-dimensional tensor $\mathcal{A}$ such that*

$$\mathcal{A}x^4 = x_1^4 + 2x_2^4 + 3x_3^4.$$

This is an even order positive definite symmetric tensor. By [6], its all Z-eigenvalues are respectively

$$\lambda_1 = 3,\ \lambda_2 = 2,\ \lambda_3 = 1.2,\ \lambda_4 = 1,\ \lambda_5 = 0.75,\ \lambda_6 = 0.6667,\ \lambda_7 = 0.5455.$$

Example 4.2. *(see [6]) Consider the 4th-order 3-dimensional tensor $\mathcal{A}$ such that*

$$\mathcal{A}x^4 = 2x_1^4 + 3x_2^4 + 5x_3^4 + 4ax_1^2x_2x_3,$$

where a is a parameter. According to [6], when $a = 1$, it is an even order positive definite symmetric tensor. In this case, there exist six Z-eigenvalues and five H-eigenvalues which are respectively

- *Z-eigenvalue:* $\lambda_1 = 5,\ \lambda_2 = 3,\ \lambda_3 = 2,\ \lambda_4 = 1.8750,\ \lambda_5 = 1.6133,\ \lambda_6 = 0.4787;$
- *H-eigenvalue:* $\lambda_1 = 5.1812,\ \lambda_2 = 5,\ \lambda_3 = 3,\ \lambda_4 = 2,\ \lambda_5 = 1.2269.$

Example 4.3. *(see [2]) Consider the diagonal tensor $\mathcal{A}$ with diagonal elements $\mathcal{A}(i,i,i,i) = 10i$, $i = 1, 2, \cdots, n$; and $\mathcal{A}(i,j,k,l) = 0$ for all other i, j, k, l.*

Example 4.4. *(see [6]) Let A be a 4th-order 2-dimensional tensor such that*

$$\mathcal{A}x^4 = 3x_1^4 + x_2^4 + 6ax_1^2x_2^2,$$

where a is a parameter. As shown in [6], when $a = 2$, this tensor is an even order positive definite symmetric tensor. And in this case, there are three Z-eigenvalues, which are

$$\lambda_1 = 4.1250,\ \lambda_2 = 3,\ \lambda_3 = 1.$$

4.1. **Computing the largest Z-eigenpairs.** In our test, we use 100 randomly generated initial vectors $x_0 = 2 * \mathrm{rand}(n, 1) - 1$, where n is the dimension of tensor $\mathcal{A}$. For each set of experiments, the same set of random starts was used. Let x be the nonzero critical point obtained by L-BFGS at termination, and the error defined by $\epsilon = \|\mathcal{A}\tilde{x}^{m-1} - \lambda\tilde{x}\|_2$, where $\tilde{x} = x/\|x\|_2$.

We compute the Z-eigenvalues of $\mathcal{A}$ from Examples 4.1-4.4. For the largest Z-eigenvalue, we list the number of occurrences in the 100 tests. We also list the mean number of iterations and function evaluations until convergence, the average error and the average CPU time in the 100 experiments in Tables 1-4.

Table 1. Results of computing Z-eigenvalues of $\mathcal{A}$ from Example 4.1

Alg.	λ	Occ.	Error	Time (sec.)	Iter.	FuncEvals
Han's (2)	3.00	54%	1.08e-08	0.40	14.70	26.37
$S_{\mathcal{A},\mathcal{B}}$	3.00	62%	1.01e-08	0.25	12.39	15.94
$L_{\mathcal{A},\mathcal{B}}$	3.00	68%	1.19e-08	0.27	12.91	16.68

Table 2. Results of computing Z-eigenvalues of $\mathcal{A}$ from Example 4.2 ($a = 1$)

Alg.	λ	Occ.	Error	Time (sec.)	Iter.	FuncEvals
Han's (2)	5.00	58%	1.26e-08	0.43	16.00	24.54
$S_{\mathcal{A},\mathcal{B}}$	5.00	56%	3.93e-08	0.26	12.14	14.76
$L_{\mathcal{A},\mathcal{B}}$	5.00	56%	1.13e-08	0.30	12.91	16.18

Table 3. Results of computing Z-eigenvalues of $\mathcal{A}$ from Example 4.3 ($n = 5$)

Alg.	λ	Occ.	Error	Time (sec.)	Iter.	FuncEvals
Han's (2)	50	55%	4.05e-07	1.09	20.36	32.09
$S_{\mathcal{A},\mathcal{B}}$	50	55%	8.64e-08	0.61	15.73	18.55
$L_{\mathcal{A},\mathcal{B}}$	50	53%	8.86e-08	0.83	17.64	24.36

Table 4. Results of computing Z-eigenvalues of $\mathcal{A}$ from Example 4.4 ($a = 2$)

Alg.	λ	Occ.	Error	Time (sec.)	Iter.	FuncEvals
Han's (2)	4.125	100%	9.71e-09	0.38	14.35	23.35
$S_{\mathcal{A},\mathcal{B}}$	4.125	100%	1.65e-08	0.23	11.80	14.10
$L_{\mathcal{A},\mathcal{B}}$	4.125	100%	1.41e-08	0.25	12.15	14.70

From the numerical results given in Tables 1-4, we see the efficiency of the variational principles (7) and (10) for finding Z-eigenpairs of even order symmetric positive definite tensors. It is easy to see from Tables 1-4, the average number of iterations and objective function evaluations of $S_{\mathcal{A},\mathcal{B}}$ and $L_{\mathcal{A},\mathcal{B}}$ are much fewer than Han's (2). And the formulations $S_{\mathcal{A},\mathcal{B}}$ and $L_{\mathcal{A},\mathcal{B}}$ are more faster for finding the largest Z-eigenvalue than Han's method. The $S_{\mathcal{A},\mathcal{B}}$ and $L_{\mathcal{A},\mathcal{B}}$ perform better than Han's (2) for finding Z-eigenvalues. Thus, $S_{\mathcal{A},\mathcal{B}}$ and $L_{\mathcal{A},\mathcal{B}}$ may be a better choice for calculating Z-eigenvalues of an even order symmetric positive definite tensor. Furthermore, as we can see from Tables 1-4, the largest Z-eigenvalue is easy to find for some tensors; however, it is difficult to find the largest Z-eigenvalue for some other tensors.

4.2. **Computing the largest H-eigenpairs.** In this subsection, we test our variational principles by calculating the largest H-eigenvalues. We still use 100 random starting guesses $x_0 = 2 * \text{rand}(n, 1) - 1$ to compute H-eigenvalues of $\mathcal{A}$ from Examples 4.2-4.4. The same set of randomly initial vectors were used for each set of experiments. For H-eigenvalue, the error is defined by $\hat{\epsilon} = \left\| \mathcal{A}\bar{x}^{m-1} - \lambda \bar{x}^{[m-1]} \right\|_2$, where $\bar{x} = x/\|x\|_2$, and x is the nonzero critical point obtained by L-BFGS at termination. In Tables 5-7, we still list the comparison results for computing the largest H-eigenvalues of each tensor $\mathcal{A}$ in the 100 experiments.

TABLE 5. Results of computing H-eigenvalues of $\mathcal{A}$ from Example 4.2 ($a = 1$)

Alg.	λ	Occ.	Error	Time (sec.)	Iter.	FuncEvals
Han's (2)	5.1812	73%	2.59e-06	0.45	24.19	39.81
$S_{\mathcal{A},\mathcal{B}}$	5.1812	75%	1.33e-08	0.40	23.18	33.42
$L_{\mathcal{A},\mathcal{B}}$	5.1812	89%	8.32e-09	0.43	25.79	39.36

TABLE 6. Results of computing H-eigenvalues of $\mathcal{A}$ from Example 4.3 ($n = 5$)

Alg.	λ	Occ.	Error	Time (sec.)	Iter.	FuncEvals
Han's (2)	50.00	100%	1.78e-07	1.28	87.29	144.00
$S_{\mathcal{A},\mathcal{B}}$	50.00	100%	2.35e-07	1.16	87.93	129.79
$L_{\mathcal{A},\mathcal{B}}$	50.00	100%	3.33e-07	0.95	73.64	107.64

TABLE 7. Results of computing H-eigenvalues of $\mathcal{A}$ from Example 4.4 ($a = 2$)

Alg.	λ	Occ.	Error	Time (sec.)	Iter.	FuncEvals
Han's (2)	8.0828	100%	2.19e-08	0.51	36.87	52.53
$S_{\mathcal{A},\mathcal{B}}$	8.0828	100%	3.26e-08	0.40	34.53	43.00
$L_{\mathcal{A},\mathcal{B}}$	8.0828	100%	2.03e-08	0.44	35.73	44.93

From Tables 5-7, we see that the formulations $S_{\mathcal{A},\mathcal{B}}$ and $L_{\mathcal{A},\mathcal{B}}$ are competitive with the formulation Han's (2) for finding the largest H-eigenvalues. For Examples 4.2-4.4, in terms of the average number of objective function evaluations, the $S_{\mathcal{A},\mathcal{B}}$ and $L_{\mathcal{A},\mathcal{B}}$ are much less than Han's (2). Especially, for Examples 4.3 and 4.4, both of these three formulations could find the maximum H-eigenvalue in all of the 100 experiments.

5. **Conclusions.** In this paper, we mainly proposed two optimization formulations for generalized tensor eigenvalue problem and analyzed some variational characterizations for the largest $\mathcal{B}$-eigenvalue of even order weakly symmetric positive definite tensors. These methods can be used to compute a $\mathcal{B}$-eigenvalue (including Z-eigenvalue, H-eigenvalue and D-eigenvalue) of an even order weakly symmetric positive definite tensor. Some numerical experiments illustrated that our models are faster than Han's models and could reach the largest eigenpair with a higher probability. However, as previously mentioned, the local optimization methods cannot guarantee finding a global maximum; hence, it is necessary to develop a global optimization method for solving such problems in the future work.

Acknowledgment. The authors are grateful to the referees for their useful and constructive suggestions. This work is partially supported by the Science and Technology Project of Hunan Province, Grant 2014FJ3119, and the National Natural Science Foundation, Grant 61202397.

REFERENCES

[1] L. Qi, Eigenvalues of a real supersymmetric tensor, *Journal of Symbolic Computation*, vol.40, no.6, pp.1302-1324, 2005.

[2] L. Han, An unconstrained optimization approach for finding real eigenvalues of even order symmetric tensors, *Numerical Algebra, Control and Optimization*, vol.3, no.3, pp.583-599, 2013.

[3] G. Auchmuty, Unconstrained variational principles for eigenvalues of real symmetric matrices, *SIAM J. Mathematical Analysis*, vol.20, no.5, pp.1186-1207, 1989.

[4] M. Mongeau and M. Torki, Computing eigenelements of real symmetric matrices via optimization, *Computational Optimization and Applications*, vol.29, no.3, pp.263-287, 2004.

[5] K. C. Chang, K, Pearson and T. Zhang, On eigenvalue problems of real symmetric tensors, *Journal of Mathematical Analysis and Applications*, vol.350, no.1, pp.416-422, 2009.

[6] C. Cui, Y. Dai and J. Nie, All real eigenvalues of symmetric tensors, *SIAM J. Control and Optimization*, vol.35, no.4, pp.1582-1601, 2014.

[7] F. Zhang, B. Zhou and L. Peng, Gradient skewness tensors and local illumination detection for images, *Journal of Computational and Applied Mathematics*, vol.237, no.1, pp.663-671, 2013.

[8] G. Li, L. Qi and G. Yu, The Z-eigenvalues of a symmetric tensor and its application to spectral hypergraph theory, *Numerical Linear Algebra with Applications*, vol.20, no.6, pp.1001-1029, 2013.

[9] L. Lim and C. Hillar, *Most Tensor Problems Are NP Hard*, University of California, Berkeley, 2009.

[10] T. G. Kolda and J. R. Mayo, Shifted power method for computing tensor eigenpairs, *SIAM J. Matrix Analysis and Applications*, vol.32, no.4, pp.1095-1124, 2011.

[11] T. G. Kolda and J. R. Mayo, An adaptive shifted power method for computing generalized tensor eigenpairs, *SIAM J. Matrix Analysis and Applications*, vol.35, no.4, pp.1563-1581, 2014.

[12] C. Hao, C. Cui and Y. Dai, A sequential subspace projection method for extreme Z-eigenvalues of supersymmetric tensors, *Numerical Linear Algebra with Applications*, vol.22, no.2, pp.283-298, 2015.

[13] L. H. Lim, Singular values and eigenvalues of tensors: A variational approach, *Proc. of the 1st IEEE International Workshop on Computational Advances of Multi-Sensor Adaptive Processing*, pp.129-132, 2005.

[14] L. Qi, Y. Wang and E. Wu, D-eigenvalues of diffusion kurtosis tensors, *Journal of Computational and Applied Mathematics*, vol.221, no.1, pp.150-157, 2008.

[15] D. C. Liu and J. Nocedal, On the limited memory method for large scale optimization, *Mathematical Programming B*, vol.45, no.3, pp.503-528, 1989.

[16] B. W. Bader and T. G. Kolda, *Matlab Tensor Toolbox Version 2.5*, http://www.sandia.gov/tgkolda/TensorToolbox, 2012.