

AN ADAPTIVE SHIFTED POWER METHOD FOR COMPUTING GENERALIZED TENSOR EIGENPAIRS*

TAMARA G. KOLDA[†] AND JACKSON R. MAYO[†]

Abstract. Several tensor eigenpair definitions have been put forth in the past decade, but these can all be unified under generalized tensor eigenpair framework, introduced by Chang, Pearson, and Zhang (2009). Given m th-order, n -dimensional real-valued symmetric tensors \mathcal{A} and \mathcal{B} , the goal is to find $\lambda \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq 0$ such that $\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathcal{B}\mathbf{x}^{m-1}$. Different choices for \mathcal{B} yield different versions of the tensor eigenvalue problem. We present our generalized eigenproblem adaptive power (GEAP) method for solving the problem, which is an extension of the shifted symmetric higher-order power method (SS-HOPM) for finding Z-eigenpairs. A major drawback of SS-HOPM was that its performance depended in choosing an appropriate shift, but our GEAP method also includes an adaptive method for choosing the shift automatically.

Key words. tensor eigenvalues, E-eigenpairs, Z-eigenpairs, l^2 -eigenpairs, generalized tensor eigenpairs, shifted symmetric higher-order power method (SS-HOPM), generalized eigenproblem adaptive power (GEAP) method

AMS subject classifications. 15A18, 15A69

1. Introduction. Suppose \mathcal{A} is a real-valued, m th-order, n -dimensional tensors and \mathbf{x} is a real-valued n -vector. We let $\mathcal{A}\mathbf{x}^{m-1}$ denote the n -vector defined by

$$(\mathcal{A}\mathbf{x}^{m-1})_{i_1} = \sum_{i_2=1}^n \cdots \sum_{i_m=1}^n a_{i_1 \dots i_m} x_{i_2} \cdots x_{i_m} \quad \text{for } i_1 = 1, \dots, n.$$

We let $\mathcal{A}\mathbf{x}^m$ denote the scalar defined by $\mathcal{A}\mathbf{x}^m = \mathbf{x}^\top (\mathcal{A}\mathbf{x}^{m-1})$. We say the tensor \mathcal{A} is *symmetric* if its entries are invariant under permutation. We say the tensor \mathcal{A} is *positive definite* if $\mathcal{A}\mathbf{x}^m > 0$ for all $\mathbf{x} \neq 0$.

The notion of generalized eigenpairs has been defined for tensors by Chang, Pearson, and Zhang [2] as follows. Let \mathcal{A} and \mathcal{B} be real-valued, m th-order, n -dimensional symmetric tensors. Assume further that m is even and \mathcal{B} is positive definite. We say $(\lambda, \mathbf{x}) \in \mathbb{R} \times \{\mathbb{R}^n \setminus \{\mathbf{0}\}\}$ is a *generalized eigenpair* (also known as a \mathcal{B} -eigenpair) if

$$\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathcal{B}\mathbf{x}^{m-1}. \tag{1.1}$$

Taking the dot product with \mathbf{x} , it is clear that any solution satisfies

$$\lambda = \frac{\mathcal{A}\mathbf{x}^m}{\mathcal{B}\mathbf{x}^m}. \tag{1.2}$$

The advantage of the generalized eigenpair framework is that it nicely encapsulates multiple definitions of tensor eigenvalues, as follows.

- A *Z-eigenpair* [10, 7] is defined as a pair $(\lambda, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^n$ such that

$$\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathbf{x} \quad \text{and} \quad \|\mathbf{x}\| = 1. \tag{1.3}$$

*This work was funded by the applied mathematics program at the U.S. Department of Energy and by an Excellence Award from the Laboratory Directed Research & Development (LDRD) program at Sandia National Laboratories. Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

[†]Sandia National Laboratories, Livermore, CA. Email: {tgkolda,jmayo}@sandia.gov.

This is equivalent to a generalized tensor eigenpair with $\mathcal{B} = \mathcal{E}$, the identity tensor such that $\mathcal{E}\mathbf{x}^{m-1} = \|\mathbf{x}\|^{m-2}\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$ [2]. Note that, unlike ordinary tensor Z-eigenpairs, generalized tensor eigenpairs allow arbitrary rescaling of the eigenvector \mathbf{x} with no effect on the eigenvalue λ . In this way, the generalized tensor eigenvalue problem preserves the homogeneity of the corresponding matrix eigenproblem.

- An *H-eigenpair* is defined as a pair $(\lambda, \mathbf{x}) \in \mathbb{R} \times \{\mathbb{R}^n \setminus \{\mathbf{0}\}\}$ such that

$$\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathbf{x}^{[m-1]}. \quad (1.4)$$

Here $\mathbf{x}^{[m-1]}$ denotes elementwise power, i.e., $(\mathbf{x}^{[m-1]})_i \equiv \mathbf{x}_i^{m-1}$, for $i = 1, \dots, n$. This is equivalent to a generalized tensor eigenpair with $b_{i_1 i_2 \dots i_m} = \delta_{i_1 i_2 \dots i_m}$ [2].

- Let \mathbf{D} be a symmetric $n \times n$ matrix and assume $m = 4$. We say (λ, \mathbf{x}) is a *D-eigenpair* [11] if

$$\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathbf{D}\mathbf{x} \quad \text{and} \quad \mathbf{x}^\top \mathbf{D}\mathbf{x} = 1.$$

This is equivalent to a \mathcal{B} -eigenpair where \mathcal{B} is the symmetrized tensor outer product of \mathbf{D} with itself [2].

In this paper, we describe a method for computing generalized eigenpairs. Our method is a generalization of the shifted symmetric higher-order power method (SS-HOPM) that we previously introduced for computing Z-eigenvalues [6]. In addition to generalizing the method, we have also significantly improved it by adding an adaptive method for choosing the shift. To derive the method, we reformulate the generalized eigenproblem, (1.1), as a nonlinear program such that any generalized eigenpair is equivalent to a KKT point in §3. We develop an adaptive, monotonically convergent, shifted power method for solving the optimization problem in §4. We call our method the Generalized Eigenproblem Adaptive Power (GEAP) method. In §5, we show that the GEAP method is much faster than the SS-HOPM method for finding Z-eigenpairs due to its adaptive shift selection. Furthermore, the GEAP method is shown to find other types of generalized eigenpairs, by illustrating it on examples from related literature as well as a randomly generated example. This is the only known method for finding generalized eigenpairs besides direct numerical solution; we survey related work in §6.

2. Notation and preliminaries. A symmetric tensor has entries that are invariant under any permutation of its indices. More formally, a real-valued, m th-order, n -dimensional tensor \mathcal{A} is *symmetric* if

$$a_{i_{p(1)} \dots i_{p(m)}} = a_{i_1 \dots i_m} \quad \text{for all } i_1, \dots, i_m \in \{1, \dots, n\} \quad \text{and} \quad p \in \Pi_m,$$

where Π_m denotes the space of all m -permutations. We let $\mathbb{S}^{[m,n]}$ denote the space of all symmetric, real-valued, m th-order, n -dimensional tensors.

Let $\mathcal{A} \in \mathbb{S}^{[m,n]}$, then we can define the following tensor-vector products.

$$\mathcal{A}\mathbf{x}^m = \sum_{i_1=1}^n \dots \sum_{i_m=1}^n a_{i_1 i_2 \dots i_m} x_{i_1} \dots x_{i_m}, \quad (2.1)$$

$$(\mathcal{A}\mathbf{x}^{m-1})_{i_1} = \sum_{i_2=1}^n \dots \sum_{i_m=1}^n a_{i_1 i_2 \dots i_m} x_{i_2} \dots x_{i_m} \quad \text{for all } i_1 = 1, \dots, n, \quad (2.2)$$

$$(\mathcal{A}\mathbf{x}^{m-2})_{i_1 i_2} = \sum_{i_3=1}^n \dots \sum_{i_m=1}^n a_{i_1 i_2 \dots i_m} x_{i_3} \dots x_{i_m} \quad \text{for all } i_1, i_2 = 1, \dots, n. \quad (2.3)$$

Observe that the derivatives of the tensor-vector product w.r.t. \mathbf{x} are given by

$$\nabla(\mathcal{A}\mathbf{x}^m) = m\mathcal{A}\mathbf{x}^{m-1}, \quad \nabla^2(\mathcal{A}\mathbf{x}^m) = m(m-1)\mathcal{A}\mathbf{x}^{m-2}.$$

We say a tensor $\mathcal{A} \in \mathbb{S}^{[m,n]}$ is *positive definite* if

$$\mathcal{A}\mathbf{x}^m > 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq 0.$$

We let $\mathbb{S}_+^{[m,n]}$ denote the space of positive definite tensors in $\mathbb{S}^{[m,n]}$.

We use the symbol \odot to mean symmetrized outer product, i.e.,

$$\mathbf{a} \odot \mathbf{b} = \mathbf{a}\mathbf{b}^\top + \mathbf{b}\mathbf{a}^\top.$$

3. Problem reformulation. Let Σ denote the unit sphere, i.e.,

$$\Sigma = \{ \mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|^m = 1 \}.$$

Let $\mathcal{A} \in \mathbb{S}^{[m,n]}$ and $\mathcal{B} \in \mathbb{S}_+^{[m,n]}$. Then we may define the nonlinear program

$$\max f(\mathbf{x}) = \frac{\mathcal{A}\mathbf{x}^m}{\mathcal{B}\mathbf{x}^m} \|\mathbf{x}\|^m \quad \text{subject to } \mathbf{x} \in \Sigma. \quad (3.1)$$

The constraint makes the term $\|\mathbf{x}\|^m$ in $f(\mathbf{x})$ superfluous; nevertheless, we retain this form since choosing $\mathcal{B} = \mathcal{E}$ yields $f(x) = \mathcal{A}\mathbf{x}^m$, as in [6].

The details of computing the derivatives are provided in [Appendix A](#). Here we simply state the results as a theorem.

THEOREM 3.1. *Let $f(\mathbf{x})$ be as defined in (3.1). For $\mathbf{x} \in \Sigma$, the gradient is*

$$\mathbf{g}(\mathbf{x}) \equiv \nabla f(\mathbf{x}) = \frac{m}{\mathcal{B}\mathbf{x}^m} \left[(\mathcal{A}\mathbf{x}^m) \mathbf{x} + \mathcal{A}\mathbf{x}^{m-1} - \left(\frac{\mathcal{A}\mathbf{x}^m}{\mathcal{B}\mathbf{x}^m} \right) \mathcal{B}\mathbf{x}^{m-1} \right]. \quad (3.2)$$

For $\mathbf{x} \in \Sigma$, the Hessian is

$$\begin{aligned} \mathbf{H}(\mathbf{x}) \equiv \nabla^2 f(\mathbf{x}) &= \frac{m^2 \mathcal{A}\mathbf{x}^m}{(\mathcal{B}\mathbf{x}^m)^3} (\mathcal{B}\mathbf{x}^{m-1} \odot \mathcal{B}\mathbf{x}^{m-1}) \\ &+ \frac{m}{\mathcal{B}\mathbf{x}^m} \left[(m-1)\mathcal{A}\mathbf{x}^{m-2} + \mathcal{A}\mathbf{x}^m (\mathbf{I} + (m-2)\mathbf{x}\mathbf{x}^\top) + m(\mathcal{A}\mathbf{x}^{m-1} \odot \mathbf{x}) \right] \\ &- \frac{m}{(\mathcal{B}\mathbf{x}^m)^2} \left[(m-1)\mathcal{A}\mathbf{x}^m \mathcal{B}\mathbf{x}^{m-2} + m(\mathcal{A}\mathbf{x}^{m-1} \odot \mathcal{B}\mathbf{x}^{m-1}) \right. \\ &\quad \left. + m\mathcal{A}\mathbf{x}^m (\mathbf{x} \odot \mathcal{B}\mathbf{x}^{m-1}) \right]. \end{aligned} \quad (3.3)$$

These complicated derivatives reduce for $\mathcal{B} = \mathcal{E}$. In that case, we have $\mathcal{B}\mathbf{x}^m = 1$ and $\mathcal{B}\mathbf{x}^{m-1} = \mathbf{x}$, so these equations become.

$$\mathbf{g}(\mathbf{x}) = m\mathcal{A}\mathbf{x}^{m-1}, \quad \mathbf{H}(\mathbf{x}) = m(m-1)\mathcal{A}\mathbf{x}^{m-2}.$$

These match the derivatives of $f(x) = \mathcal{A}\mathbf{x}^m$, as proved in [6]. Note that we have used the fact that $(m-1)\mathcal{E}\mathbf{x}^{m-2} = \mathbf{I} + (m-2)\mathbf{x}\mathbf{x}^\top$ for all $\mathbf{x} \in \Sigma$.

We are considering the nonlinear program in (3.1) because there is a correspondence between it and the generalized tensor eigenvalue problem in (1.1). Note that the \mathbf{x} in (1.1) can be arbitrarily rescaled.

THEOREM 3.2. *Any pair (λ, \mathbf{x}) is a solution to (1.1) iff the scaled version with $\|\mathbf{x}\| = 1$ is a KKT point of (3.1) with λ as the Lagrange multiplier.*

Proof. First, assume (λ, \mathbf{x}) is a solution to (1.1). Let constraint $\mathbf{x} \in \Sigma$ be expressed as $\|\mathbf{x}\|^m = 1$. Then the Lagrangian is

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda(\|\mathbf{x}\|^m - 1).$$

Hence, using the derivatives in [Appendix A](#), we have

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda) = \frac{m}{\mathbf{Bx}^m} \left[(\mathbf{Ax}^m) \mathbf{x} + \mathbf{Ax}^{m-1} - \left(\frac{\mathbf{Ax}^m}{\mathbf{Bx}^m} \right) \mathbf{Bx}^{m-1} \right] - m\lambda \mathbf{x} = 0. \quad (3.4)$$

So, \mathbf{x} is a KKT point of (3.1) with Lagrange multiplier λ as defined in (1.2).

To prove the reverse, assume \mathbf{x} is a KKT point of (3.1) with Lagrange multiplier λ . Then, (3.4) must hold. If we multiply each term in (3.4) by \mathbf{x} , then the third and fourth terms cancel out, and we conclude that λ satisfies (1.2). Substituting that back into (3.4), we see that (1.1) is satisfied. Hence, the claim. \square

From the previous theorem, there is an equivalence between generalized tensor eigenpairs and KKT points of (3.1). Hence, solving (3.1) yields eigenpairs. An eigenpair may correspond to a local maximum, a local minimum, or a saddle point. For a given eigenpair (λ, \mathbf{x}) normalized so that $\mathbf{x} \in \Sigma$, we can categorize it by considering the projected Hessian of the Lagrangian, i.e.,

$$\mathbf{C}(\lambda, \mathbf{x}) = \mathbf{U}^T (\mathbf{H}(\mathbf{x}) - \lambda m \mathbf{I}) \mathbf{U} \in \mathbb{R}^{(n-1) \times (n-1)}, \quad (3.5)$$

where $\mathbf{U} \in \mathbb{R}^{n \times (n-1)}$ is an orthonormal basis for \mathbf{x}^\perp . We can then say the following:

$\mathbf{C}(\lambda, \mathbf{x})$ positive definite	\Rightarrow local minimum of (3.1),
$\mathbf{C}(\lambda, \mathbf{x})$ negative definite	\Rightarrow local maximum of (3.1),
$\mathbf{C}(\lambda, \mathbf{x})$ indefinite	\Rightarrow saddle point of (3.1).

The argument is very similar to that presented in [6] and so is omitted. Optimization approaches cannot easily find saddle points, but they can find local minima and maxima. We describe such an optimization approach in the next section.

4. Derivation of GEAP algorithm. We propose to use a property of convex functions of the sphere to develop a monotonically convergent method. We consider an idea originally from [5, 12]; see also [6] for a proof. We have modified the theorem here to focus on its *local* applicability by considering just an open neighborhood of \mathbf{w} rather than all of \mathbb{R}^n .

THEOREM 4.1 (Kofidis and Regalia [5, 12]). *Let $f(\mathbf{x})$ be a given function, and let $\mathbf{w} \in \Sigma$ such that $\nabla f(\mathbf{w}) \neq 0$. Let Ω be an open neighborhood of \mathbf{w} , and assume f is convex and continuously differentiable on Ω . Define $\mathbf{v} = \nabla f(\mathbf{w})/\|\nabla f(\mathbf{w})\|$. If $\mathbf{v} \in \Omega$ and $\mathbf{v} \neq \mathbf{w}$, then $f(\mathbf{v}) - f(\mathbf{w}) > 0$.*

COROLLARY 4.2. *Let $f(x)$ be a given function, and let $\mathbf{w} \in \Sigma$ such that $\nabla f(\mathbf{w}) \neq 0$. Let Ω be an open neighborhood of \mathbf{w} , and assume f is concave and continuously differentiable on Ω . Define $\mathbf{v} = -\nabla f(\mathbf{w})/\|\nabla f(\mathbf{w})\|$. If $\mathbf{v} \in \Omega$ and $\mathbf{v} \neq \mathbf{w}$, then $f(\mathbf{v}) - f(\mathbf{w}) < 0$.*

Hence, if f is locally convex, then a simple algorithm, i.e.,

$$\mathbf{x}_{\text{new}} = \mathbf{g}(\mathbf{x}) / \|\mathbf{g}(\mathbf{x})\|.$$

will yield ascent. Conversely, if f is locally concave, we can expect descent from $\mathbf{x}_{\text{new}} = -\mathbf{g}(\mathbf{x}) / \|\mathbf{g}(\mathbf{x})\|$. Unfortunately, the function $f(\mathbf{x})$ in (3.1) may not be convex or concave.

To fix this, we work with a shifted function,

$$\hat{f}(\mathbf{x}) = f(\mathbf{x}) + \alpha \|\mathbf{x}\|^m. \quad (4.1)$$

From [6], we have that for $\mathbf{x} \in \Sigma$,

$$\hat{\mathbf{g}}(\mathbf{x}) \equiv \nabla \hat{f}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) + \alpha m \mathbf{x}, \quad (4.2)$$

$$\hat{\mathbf{H}}(\mathbf{x}) \equiv \nabla^2 \hat{f}(\mathbf{x}) = \mathbf{H}(\mathbf{x}) + \alpha m \mathbf{I} + \alpha m(m-2) \mathbf{x} \mathbf{x}^\top. \quad (4.3)$$

If we choose α appropriately, then we can ensure that $\hat{\mathbf{H}}(\mathbf{x})$ is positive or negative definite, ensuring that $\hat{f}(\mathbf{x})$ is locally convex or concave. In [6] for the special case of $\mathcal{B} = \mathcal{E}$, we proposed choosing a single value for α in SS-HOPM that ensured convexity on the entire sphere. But it is difficult to choose a reasonable value in advance, and poor choices lead to either very slow convergence or a complete lack of convergence. In this work, we propose to choose α adaptively.

For an arbitrary matrix $n \times n$ symmetric matrix \mathbf{M} , the following notation denotes its eigenvalues: $\lambda_{\min}(\mathbf{M}) = \lambda_1(\mathbf{M}) \leq \lambda_2(\mathbf{M}) \leq \dots \leq \lambda_n(\mathbf{M}) = \lambda_{\max}(\mathbf{M})$.

THEOREM 4.3. *Assume $\mathbf{x} \in \Sigma$. Let $\mathbf{H}(\mathbf{x})$ and $\hat{\mathbf{H}}(\mathbf{x})$ be defined as in (3.3) and (4.3), respectively. For $\alpha \geq 0$, the eigenvalues of $\hat{\mathbf{H}}$ are bounded as*

$$\lambda_i(\mathbf{H}) + m\alpha \leq \lambda_i(\hat{\mathbf{H}}) \leq \lambda_i(\mathbf{H}) + m\alpha + m(m-2)\alpha \quad (4.4)$$

for $i = 1, \dots, n$. Likewise, for $\alpha \leq 0$, the eigenvalues of $\hat{\mathbf{H}}$ are bounded as

$$\lambda_i(\mathbf{H}) + m\alpha + m(m-2)\alpha \leq \lambda_i(\hat{\mathbf{H}}) \leq \lambda_i(\mathbf{H}) + m\alpha \quad (4.5)$$

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

Proof. The proof follows immediately from Weyl's inequality. \square

In the convex case, our goal is to choose α so that $\hat{\mathbf{H}}$ is positive semi-definite in a local neighborhood of the current iterate, \mathbf{x} . By the smoothness of $\hat{f}(\mathbf{x})$ when \mathbf{x} is away from zero, we may argue that for every $\tau > 0$, there exists and $\delta > 0$ such that $\hat{\mathbf{H}}$ is positive semi-definite for all $\|\mathbf{x} - \mathbf{x}_k\| \leq \delta$ whenever $\lambda_{\min}(\hat{\mathbf{H}}) \geq \tau$. In other words, τ is the threshold for positive definiteness.

COROLLARY 4.4. *Assume $\mathbf{x} \in \Sigma$. Let $\tau > 0$. If*

$$\alpha = \max \{ 0, (\tau - \lambda_{\min}(\mathbf{H}))/m \}, \quad (4.6)$$

then $\lambda_{\min}(\hat{\mathbf{H}}) \geq \tau$.

In the concave case, our goal is to choose α so that $\hat{\mathbf{H}}$ is negative semi-definite in a local neighborhood of the current iterate.

COROLLARY 4.5. *Assume $\mathbf{x} \in \Sigma$. Let $\tau > 0$. If*

$$\alpha = \min \{ 0, -(\tau + \lambda_{\max}(\mathbf{H}))/m \} = -\max \{ 0, \tau - \lambda_{\min}(-\mathbf{H}))/m \}, \quad (4.7)$$

then $\lambda_{\max}(\hat{\mathbf{H}}) \leq -\tau$.

From [Theorem 4.1](#), if α is selected to make $\hat{f}(\mathbf{x})$ locally convex, we have

$$\mathbf{x}_+ = \hat{\mathbf{g}}(\mathbf{x})/\|\hat{\mathbf{g}}(\mathbf{x})\| \Rightarrow \hat{f}(\mathbf{x}_+) > \hat{f}(\mathbf{x}) \Rightarrow f(\mathbf{x}_+) > f(\mathbf{x})$$

so long as $\mathbf{x}_+ \in \Omega$, the convex neighborhood of \mathbf{x} . Even though we are adaptively changing α , we see increase in the original function at each step. A similar argument applies in the concave case, with the function decreasing at each step.

The potential problem with this approach is that it may be the case that $\mathbf{x}_+ \notin \Omega$. If that happens, we may observe that the function values (i.e., λ_k) are not increasing (or decreasing) monotonically as expected. To fix this, we make a more conservative choice for τ (at least temporarily), which will in turn enforce a more conservative choice of α . If τ is large enough, then we will satisfy the lower bound on α that guarantees convergence for the shifted algorithm (this is proven for Z-eigenvalue in [\[6\]](#); the proof for the general problem is similar and so omitted). Thus far in our experiments, such contingencies have not been necessary, so we have not included the details in the algorithm.

4.1. GEAP Algorithm. The full algorithm is presented in [Algorithm 1](#).

Algorithm 1 Generalized Eigenpair Adaptive Power (GEAP) Method

Given tensors $\mathcal{A} \in \mathbb{S}^{[m,n]}$ and $\mathcal{B} \in \mathbb{S}_+^{[m,n]}$ and an initial guess $\hat{\mathbf{x}}_0$. Let $\beta = 1$ if we want to find local maxima (and the function is convex); otherwise, let $\beta = -1$, indicating that we are seeking local minima (and the function is concave). Let τ be the tolerance on being positive/negative definite.

```

1:  $\mathbf{x}_0 \leftarrow \hat{\mathbf{x}}_0/\|\hat{\mathbf{x}}_0\|$ 
2: for  $k = 0, 1, \dots$  do
3:   Precompute  $\mathcal{A}\mathbf{x}_k^{m-2}, \mathcal{B}\mathbf{x}_k^{m-2}, \mathcal{A}\mathbf{x}_k^{m-1}, \mathcal{B}\mathbf{x}_k^{m-1}, \mathcal{A}\mathbf{x}_k^m, \mathcal{B}\mathbf{x}_k^m$ 
4:    $\lambda_k \leftarrow \mathcal{A}\mathbf{x}_k^m/\mathcal{B}\mathbf{x}_k^m$ 
5:    $\mathbf{H}_k \leftarrow \mathbf{H}(\mathbf{x}_k)$ 
6:    $\alpha_k \leftarrow \beta \max\{0, (\tau - \lambda_{\min}(\beta\mathbf{H}_k))/m\}$ 
7:    $\hat{\mathbf{x}}_{k+1} \leftarrow \beta(\mathcal{A}\mathbf{x}_k^{m-1} - \lambda_k \mathcal{B}\mathbf{x}_k^{m-1} + (\alpha_k + \lambda_k) \mathcal{B}\mathbf{x}_k^m \mathbf{x}_k)$ 
8:    $\mathbf{x}_{k+1} = \hat{\mathbf{x}}_{k+1}/\|\hat{\mathbf{x}}_{k+1}\|$ 
9: end for
```

The cost per iteration of [Algorithm 1](#) is as follows. Assuming \mathcal{A} and \mathcal{B} are dense, the dominant cost is computing products with these tensors. Computing the Hessian requires six products: $\mathcal{A}\mathbf{x}^{m-2}, \mathcal{A}\mathbf{x}^{m-1}, \mathcal{A}\mathbf{x}^m, \mathcal{B}\mathbf{x}^{m-2}, \mathcal{B}\mathbf{x}^{m-1}, \mathcal{B}\mathbf{x}^m$. Recall that $\mathcal{A}\mathbf{x}^{m-2}$ is given by

$$(\mathcal{A}\mathbf{x}^{m-2})_{i_1 i_2} = \sum_{i_3=1}^n \cdots \sum_{i_m=1}^n a_{i_1 i_2 \dots i_m} x_{i_3} \cdots x_{i_m} \quad \text{for all } i_1, i_2 = 1, \dots, n.$$

The cost is $(m-2) \cdot n^{m-2}$ multiplies and n^{m-2} additions per each of n^2 entries; therefore, the total cost is $(m-1)n^m$ operations. Exploiting symmetry yields reduces the cost to $O(n^m/m!)$ [\[13\]](#). We can compute

$$\mathcal{A}\mathbf{x}^{m-1} = (\mathcal{A}\mathbf{x}^{m-2})\mathbf{x} \quad \text{and} \quad \mathcal{A}\mathbf{x}^m = (\mathcal{A}\mathbf{x}^{m-1})^\top \mathbf{x},$$

for an additional cost of $2n^2$ and $2n$ operations, respectively. (These can also be computed directly, but at a cost of $(m-1)n^m$ operations each.) These values have to

be computed for both \mathcal{A} and \mathcal{B} at every iteration for a total cost (ignoring symmetry) of $2(m-1)n^m + 4n^2 + 4n$. Once these six products are computed, the cost for computing $\mathbf{H}(\mathbf{x})$ is a series of matrix operations, for a total cost of $20n^2$. The cost of computing the eigenvalues of the symmetric matrix $\mathbf{H}(\mathbf{x})$ is $(4/3)n^3$, which is less than the cost of the products. Updating \mathbf{x} requires a 5 vector operations at a cost of n operations each. Hence, cost of the method is dominated by the computation of $\mathcal{A}\mathbf{x}^{m-2}$ and $\mathcal{B}\mathbf{x}^{m-2}$, at an expense of $O(n^m/m!)$.

4.2. Specialization of GEAP to Z-eigenpairs. In [Algorithm 2](#), we show the specialization of the method to the Z-eigenvalue problem. This is the same as SS-HOPM, except for the adaptive shift. Note that unlike [Algorithm 1](#), this algorithm can be used even when m is odd. The cost per iteration of [Algorithm 2](#) is the same order as for [Algorithm 1](#), but it does not need to do any computations with \mathcal{B} .

Algorithm 2 Z-Eigenpair Adaptive Power Method

Given tensor $\mathcal{A} \in \mathbb{S}^{[m,n]}$ and an initial guess $\hat{\mathbf{x}}_0$. Let $\beta = 1$ if we want to find local maxima (and the function is convex); otherwise, let $\beta = -1$, indicating that we are seeking local minima (and the function is concave). Let τ be the tolerance on being positive/negative definite.

```

1:  $\mathbf{x}_0 \leftarrow \hat{\mathbf{x}}_0 / \|\hat{\mathbf{x}}_0\|$ 
2: for  $k = 0, 1, \dots$  do
3:   Precompute  $\mathcal{A}\mathbf{x}_k^{m-2}, \mathcal{A}\mathbf{x}_k^{m-1}, \mathcal{A}\mathbf{x}_k^m$ 
4:    $\lambda_k \leftarrow \mathcal{A}\mathbf{x}_k^m$ 
5:    $\mathbf{H}_k \leftarrow m(m-1)\mathcal{A}\mathbf{x}_k^{m-2}$ 
6:    $\alpha_k \leftarrow \beta \max\{0, (\tau - \lambda_{\min}(\beta\mathbf{H}_k))/m\}$ 
7:    $\hat{\mathbf{x}}_{k+1} \leftarrow \beta(\mathcal{A}\mathbf{x}_k^{m-1} + \alpha_k \mathbf{x}_k)$ 
8:    $\mathbf{x}_{k+1} = \hat{\mathbf{x}}_{k+1} / \|\hat{\mathbf{x}}_{k+1}\|$ 
9: end for
```

5. Numerical experiments. All numerical tests were done using MATLAB Version R2012b and the Tensor Toolbox Version 2.5 [\[1\]](#). The experiments were performed a laptop computer with an Intel Dual-Core i7-3667UCPU (2GHz) and 8GB of RAM.

In all numerical experiments, we used the following settings. We set $\tau = 10^{-6}$, where τ is the tolerance on being positive or negative definite. We consider the iterates to be converged once $|\lambda_{k+1} - \lambda_k| \leq 10^{-15}$. The maximum iterations is 500.

5.1. Comparison to SS-HOPM for computing Z-eigenpairs. The following example is originally from [\[5\]](#) and was used in evaluating the SS-HOPM algorithm in [\[6\]](#). Our goal is to compute the Z-eigenpairs [\(1.3\)](#) using the Z-Eigenpair Adaptive Power Method in [Algorithm 2](#) and show that it is faster than SS-HOPM [\[6\]](#) using a fixed value for the shift.

EXAMPLE 5.1 (Kofidis and Regalia [\[5\]](#)). Our objective is to compute the Z-eigenpairs. Let $\mathcal{A} \in \mathbb{S}^{[4,3]}$ be the symmetric tensor given by Kofidis and Regalia [\[5, Example 1\]](#) whose entries are specified in [Appendix B](#) ([Figure B.1](#)). Since we are computing Z-eigenpairs, we have $\mathcal{B} = \mathcal{E}$. A complete list of the 11 Z-eigenpairs is provided in [Appendix C](#) ([Table C.1](#)); there are three maxima and three minima. \square

A comparison of the fixed and adaptive shift results are provided in [Table 5.1](#). There are six different experiments looking at maxima ($\beta = 1$) and minima ($\beta = -1$) and different shifts ($\alpha = 2, 10, \text{adaptive}$) in [Algorithm 2](#). Note that using a fixed shift

means that [Algorithm 2](#) is equivalent to SS-HOPM and no adaptive update of the shift is performed in Step 5.

We used 100 random starting guesses, each entry selected uniformly at random from the interval $[-1, 1]$; the same set of random starts was used for each set of experiments. For each eigenpair, the table lists the number of occurrences in the 100 experiments, the median number of iterations until convergence, the number of runs that violated monotonicity, the average error and standard deviation in the final result, and the average run time and standard deviation. The error is computed as $\|\mathcal{A}\mathbf{x}^{m-1} - \lambda\mathbf{x}\|_2$. The two monotonicity violations were both extremely small, i.e., $O(10^{-12})$. These violations indicate that a step went outside the region of local convexity.

Table 5.1: Different shifts to calculate Z-eigenpairs for $\mathcal{A} \in \mathbb{S}^{[2,3]}$ from [Example 5.1](#).

(a) α adaptive, $\beta = 1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
53	0.8893	[0.6672 0.2471 -0.7027]	30	—	9e-9 ± 3e-9	0.05 ± 0.02
29	0.8169	[0.8412 -0.2635 0.4722]	34	—	1e-8 ± 3e-9	0.04 ± 0.01
18	0.3633	[0.2676 0.6447 0.7160]	26	—	7e-9 ± 2e-9	0.03 ± 0.00

(b) $\alpha = 2$, $\beta = 1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
53	0.8893	[0.6672 0.2471 -0.7027]	49	1 (1e-15)	2e-8 ± 3e-9	0.07 ± 0.03
29	0.8169	[0.8412 -0.2635 0.4722]	45	—	2e-8 ± 3e-9	0.04 ± 0.00
18	0.3633	[0.2676 0.6447 0.7160]	57	—	2e-8 ± 2e-9	0.06 ± 0.00

(c) $\alpha = 10$, $\beta = 1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
48	0.8893	[0.6672 0.2471 -0.7027]	192	—	5e-8 ± 6e-9	0.24 ± 0.12
29	0.8169	[0.8412 -0.2635 0.4722]	185	—	5e-8 ± 5e-9	0.17 ± 0.02
18	0.3633	[0.2676 0.6447 0.7160]	261	—	5e-8 ± 2e-9	0.24 ± 0.02
5	<i>Failed to converge in 500 iterations</i>			—	N/A	0.43 ± 0.01

(d) α adaptive, $\beta = -1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
22	-0.0451	[0.7797 0.6135 0.1250]	18	—	4e-9 ± 2e-9	0.02 ± 0.00
37	-0.5629	[0.1762 -0.1796 0.9678]	17	—	6e-9 ± 2e-9	0.02 ± 0.00
41	-1.0954	[0.5915 -0.7467 -0.3043]	17	—	6e-9 ± 3e-9	0.02 ± 0.01

(e) $\alpha = -2$, $\beta = -1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
22	-0.0451	[0.7797 0.6135 0.1250]	34	—	1e-8 ± 2e-9	0.03 ± 0.00
37	-0.5629	[0.1762 -0.1796 0.9678]	20	—	7e-9 ± 2e-9	0.02 ± 0.00
41	-1.0954	[0.5915 -0.7467 -0.3043]	21	1 (1e-15)	8e-9 ± 4e-9	0.02 ± 0.00

(f) $\alpha = -10$, $\beta = -1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
22	-0.0451	[0.7797 0.6135 0.1250]	186	—	5e-8 ± 2e-9	0.16 ± 0.01
37	-0.5629	[0.1762 -0.1796 0.9678]	103	—	4e-8 ± 4e-9	0.09 ± 0.01
41	-1.0954	[0.5915 -0.7467 -0.3043]	94	—	4e-8 ± 6e-9	0.09 ± 0.01

The first three experiments use $\beta = 1$ to look for local maxima. The first experiment varies α , the second uses $\alpha = 2$ (as in [\[6\]](#)), and the third uses $\alpha = 10$. All three variations find all three local maxima. The results for $\alpha = 2$ and the adaptive method are nearly identical — they find the same local maxima with the same frequency.

The difference is that $\alpha = 2$ uses more iterations than the adaptive shift. Choosing $\alpha = 10$ is similar, except now five of the runs do not converge within the allotted 500 iterations. There was no breakdown in monotonicity, and these runs would converge eventually. If the shift is too small (e.g., $\alpha = 0$), then some or all of the runs may fail to converge [6].

The last three experiments use $\beta = -1$ to find local minima. Again, we vary α using an adaptive choice along with $\alpha = -2$ and $\alpha = -10$. The adaptive method requires the fewest number of iterations. Each experiments finds all three local minima with the exact same frequencies.

To compare the convergence in terms of the number of iterations, Figure 5.1 shows sample results for one run for computing Z-eigenpairs of \mathcal{A} from Example 5.1. The left hand plot shows the selected shift values at each iteration. The right hand plot shows the convergence of the eigenvalue. The adaptive shift is the fastest to converge.

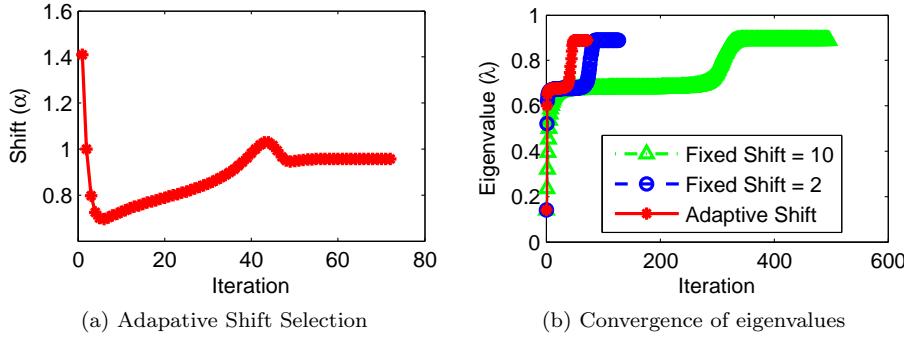


Fig. 5.1: GEAP sample results for $\mathcal{A} \in \mathbb{S}^{[4,3]}$ from Example 5.1 with $\beta = 1$ and starting point $\mathbf{x}_0 = [0.0417 \ -0.5618 \ 0.6848]$.

5.2. Numerical results for H-eigenpairs. Here we demonstrate that the GEAP method in Algorithm 1 calculates H-eigenpairs (1.4) with an appropriate choice for \mathcal{B} .

EXAMPLE 5.2. We generate a random symmetric tensor $\mathcal{A} \in \mathbb{S}^{[6,4]}$ as follows: we select random entries from $[-1, 1]$, symmetrize the result, and round to four decimal places. The tensor entries are specified in Appendix B (Figure B.2). Since we are computing H-eigenpairs, we specify \mathcal{B} as $b_{i_1 i_2 \dots i_m} = \delta_{i_1 i_2 \dots i_m}$. A complete list of the H-eigenpairs is provided in Appendix C (Table C.2); there are five maxima and five minima. \square

A summary of the results are provided in Table 5.2. There are two different experiments looking at maxima ($\beta = 1$) and minima ($\beta = -1$). We used 1000 random starting guesses, each entry selected uniformly at random from the interval $[-1, 1]$; the same set of random starts was used for each experiment. The columns are the same as for Table 5.1. The error is computed as $\|\mathcal{A}\mathbf{x}^{m-1} - \lambda\mathbf{x}^{[m-1]}\|_2$.

The first experiment uses $\beta = 1$ to look for local maxima. We find all five local maxima. There are several monotonicity violations, including at least one for $\lambda = 4.8422$ that is relatively large. These violations indicate that a step went outside the region of local convexity. Nevertheless, in all cases the algorithm is able to recover and converge, as can be seen from the small error. In general, these monotonicity

Table 5.2: H-eigenpairs for $\mathcal{A} \in \mathbb{S}^{[6,4]}$ from Example 5.2(a) $\beta = 1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
211	14.6941	[0.5426 -0.4853 0.4760 0.4936]	28	124 (2e-15)	$2e-9 \pm 2e-9$	0.08 ± 0.02
144	9.6386	[0.5342 -0.5601 0.5466 -0.3197]	110	53 (4e-15)	$9e-9 \pm 3e-9$	0.27 ± 0.04
338	8.7371	[0.4837 0.5502 0.6671 -0.1354]	100	143 (4e-15)	$1e-8 \pm 5e-9$	0.26 ± 0.06
169	5.8493	[0.6528 0.5607 -0.0627 -0.5055]	54	19 (2e-15)	$8e-9 \pm 3e-9$	0.13 ± 0.02
138	4.8422	[0.5895 -0.2640 -0.4728 0.5994]	66	13 (8e-1)	$6e-9 \pm 1e-9$	0.16 ± 0.01

(b) $\beta = -1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
130	-2.9314	[0.3161 0.5173 0.4528 -0.6537]	76	3 (2e-15)	$7e-9 \pm 1e-9$	0.18 ± 0.02
149	-3.7179	[0.6843 0.5519 0.3136 0.3589]	59	10 (1e-15)	$7e-9 \pm 2e-9$	0.15 ± 0.02
152	-4.1781	[0.4397 0.5139 -0.5444 0.4962]	99	7 (2e-15)	$5e-9 \pm 1e-9$	0.23 ± 0.03
224	-8.3200	[0.5970 -0.5816 -0.4740 -0.2842]	65	73 (2e-15)	$8e-9 \pm 3e-9$	0.16 ± 0.02
345	-10.7440	[0.4664 0.4153 -0.5880 -0.5140]	47	181 (2e-15)	$4e-9 \pm 3e-9$	0.12 ± 0.03

violations do not cause the algorithm to fail. However, such violations can be avoided by increasing τ , the tolerance on the definiteness of the Hessian matrix. Once τ is large enough, the shift will be so great that the function will be convex over the entire unit sphere. The downside of choosing a large value for τ (and the shift α_k) is that convergence will be slow.

The second experiment uses $\beta = -1$ to look for local minima. We find all five local minima. There are several monotonicity violations, but they are all small.

5.3. Numerical results for D-eigenpairs. Next we consider a different type of tensor eigenpair that also conforms to the generalized tensor eigenpair framework.

EXAMPLE 5.3 (D-eigenpairs [11]). Qi, Wang, and Wu [11] propose D-eigenpairs for diffusion kurtosis imaging (DKI). The tensors $\mathcal{A}, \mathcal{B} \in \mathbb{S}^{[4,3]}$ are specified in Appendix B (Figure B.3 and Figure B.4, respectively). We consider this example here since it can be expressed as a generalized tensor eigenproblem.¹ There are a total of 13 distinct real-valued D-eigenpairs, computed by solving the polynomial equations using Mathematica and listed in Appendix C (Table C.3); there are four maxima and three minima. \square

Table 5.3 shows the eigenpairs calculated by Algorithm 1. The error is computed as $\|\mathcal{A}\mathbf{x}^{m-1} - \lambda\mathcal{B}\mathbf{x}^{m-1}\|_2$. With 100 random starts, we find the four local maxima with $\beta = 1$. Likewise, with 100 random starts, we find the three local minima with $\beta = -1$. There are no violations to monotonicity.

5.4. Generalized eigenpairs for randomly generated \mathcal{A} and \mathcal{B} . Here we consider a randomly generated problem. We use the randomly generated $\mathcal{A} \in \mathbb{S}^{[6,4]}$ described in §5.2. However, we need a method to generate a positive definite \mathcal{B} . We use the notation $\mathcal{B} = (\mathcal{E}, \mathbf{S}, \dots, \mathbf{S}) \in \mathbb{S}^{[m,n]}$ to denote tensor-matrix multiplication in

¹Note that only four digits of precision for \mathcal{A} and \mathcal{B} are provided in [11]. We were unable to validate the solutions provided in the original paper. It is not clear if this is due to a lack of precision or a typo in the paper. Here, the problem is rescaled as well: \mathcal{D} is multiplied by 10^2 , λ is divided by 10^4 , and \mathbf{x} is divided by 10.

Table 5.3: D-eigenpairs for $\mathcal{A} \in \mathbb{S}^{[2,3]}$ and $\mathbf{D} \in \mathbb{S}^{[2,3]}$ from Example 5.3(a) $\beta = 1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
31	0.5356	[0.9227 -0.1560 -0.3526]	39	—	4e-8 ± 5e-9	0.08 ± 0.01
19	0.4359	[0.5388 0.8342 -0.1179]	48	—	3e-8 ± 4e-9	0.09 ± 0.02
25	0.2514	[0.3564 -0.8002 0.4823]	67	—	4e-8 ± 3e-9	0.13 ± 0.01
25	0.2219	[0.2184 0.3463 0.9124]	34	—	6e-8 ± 8e-9	0.07 ± 0.01

(b) $\beta = -1$

occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
39	-0.0074	[0.3669 0.5346 -0.7613]	13	—	1e-8 ± 4e-9	0.03 ± 0.01
37	-0.1242	[0.9439 0.1022 0.3141]	51	—	5e-8 ± 5e-9	0.10 ± 0.01
24	-0.3313	[0.2810 -0.9420 -0.1837]	27	—	2e-8 ± 4e-9	0.06 ± 0.01

which the tensor \mathcal{E} is multiplied by a matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$ in every mode, i.e.,

$$b_{i_1 \dots i_m} = \sum_{j_1=1}^n \dots \sum_{j_m=1}^n e_{j_1 \dots j_m} s_{i_1 j_1} \dots s_{i_m j_m}.$$

THEOREM 5.4. Let $\mathbf{S} \in \mathbb{R}^{n \times n}$ be symmetric. For m even, define $\mathcal{B} \in \mathbb{S}^{[m,n]}$ as $\mathcal{B} = (\mathcal{E}, \mathbf{S}, \dots, \mathbf{S})$. If (μ, \mathbf{x}) is a real-valued eigenpair of \mathbf{S} and $\|\mathbf{x}\| = 1$, then (λ, \mathbf{x}) is a Z-eigenpair of \mathcal{B} with $\lambda = \mu^m$. Furthermore, $\mathcal{B}\mathbf{y}^m \geq \min_i (\mu_i)^m$ for any $\mathbf{y} \in \mathbb{R}^n$ with $\|\mathbf{y}\| = 1$.

Proof. Let (μ, \mathbf{x}) be an eigenpair of \mathbf{S} such that $\|\mathbf{x}\| = 1$. Noting that $\mathbf{S}^\top = \mathbf{S}$, we have $\mathcal{B}\mathbf{x}^{m-1} = \mathbf{S}\mathcal{E}(\mathbf{S}\mathbf{x})^{m-1} = \mathbf{S}\mathcal{E}(\mu\mathbf{x})^{m-1} = \mu^{m-1}\mathbf{S}\mathcal{E}\mathbf{x}^{m-1} = \mu^{m-1}\mathbf{S}\mathbf{x} = \mu^m\mathbf{x}$. To prove the lower bound, let $\mathbf{y} \in \mathbb{R}^n$ with $\|\mathbf{y}\| = 1$. We can write \mathbf{y} as a linear combination of the eigenvectors of \mathbf{S} , i.e., $\mathbf{y} = \sum_i \nu_i \mathbf{x}_i$ and $\sum \nu_i^2 = 1$. Then

$$\begin{aligned} \mathcal{B}\mathbf{y}^m &= \mathbf{y}^\top \mathcal{B}\mathbf{y}^{(m-1)} = \mathbf{y}^\top \mathbf{S}\mathcal{E}(\mathbf{S}\mathbf{y})^{(m-1)} = \|\mathbf{S}\mathbf{y}\|^{m-1} (\mathbf{S}\mathbf{y})^\top (\mathbf{S}\mathbf{y}) = \|\mathbf{S}\mathbf{y}\|^m \\ &= \left\| \sum_i \mu_i \nu_i \mathbf{x}_i \right\|^m = \left(\sqrt{\sum_i \mu_i^2 \nu_i^2} \right)^m \geq \left(\sqrt{\min_i \mu_i^2} \right)^m = \min_i (\mu_i)^m \end{aligned}$$

Hence, the claim. \square

EXAMPLE 5.5 (Random). We use the same randomly-generated symmetric tensor $\mathcal{A} \in \mathbb{S}^{[6,4]}$ as for Example 5.2 and listed in Appendix B (Figure B.2). To generate a random positive definite symmetric tensor $\mathcal{B} \in \mathbb{S}_+^{[m,n]} = \mathbb{S}_+^{[6,4]}$, we use Theorem 5.4. (Note that this approach samples a convenient subset of $\mathbb{S}_+^{[m,n]}$ and does not draw from the entire space.) We compute a matrix $\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{U}^\top \in \mathbb{R}^{4 \times 4}$, where $\mathbf{U} \in \mathbb{R}^{4 \times 4}$ is a random orthonormal matrix and $\mathbf{D} \in \mathbb{R}^{4 \times 4}$ is a diagonal matrix with entries selected uniformly at random from $[-1, -\gamma] \cup [\gamma, 1]$ with $\gamma = \sqrt[m]{0.1} = \sqrt[6]{0.1}$. We let $\mathcal{B} = (\mathcal{E}, \mathbf{S}, \dots, \mathbf{S})$, so that \mathcal{B} has all its Z-eigenvalues in $[0, 1]$ and is positive definite. In this case, the randomly selected diagonal for \mathbf{D} is $\{-0.8620, 0.8419, 0.7979, 0.6948\}$. The random \mathcal{B} is then rounded to four decimal places, and the entries are given in Appendix B (Figure B.5). Its minimum Z-eigenvalue (computed by GEAP) is $0.1125 = 0.6948^6$, as expected. There are a total of 26 real-valued \mathcal{B} -eigenpairs of \mathcal{A} , listed in Appendix C (Table C.4). There are three maxima and four minima. \square

Table 5.4 shows the generalized eigenpairs calculated by Algorithm 1. The error is computed as $\|\mathcal{A}\mathbf{x}^{m-1} - \lambda\mathcal{B}\mathbf{x}^{m-1}\|_2$. With 1000 random starts, we find the three

local maxima with $\beta = 1$. Likewise, with 1000 random starts, we find the four local minima with $\beta = -1$. There are only small violations to monotonicity; the maximum of any violation was $O(10^{-14})$.

Table 5.4: Generalized eigenpairs for $\mathcal{A}, \mathcal{B} \in \mathbb{S}^{[6,4]}$ from Example 5.5

(a) $\beta = 1$						
occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
683	11.3476	[0.4064 0.2313 0.8810 0.0716]	59	420 (4e-15)	5e-9 ± 4e-9	0.14 ± 0.03
128	3.7394	[0.2185 -0.9142 0.2197 -0.2613]	140	11 (2e-15)	1e-8 ± 3e-9	0.30 ± 0.04
189	2.9979	[0.8224 0.4083 -0.0174 -0.3958]	23	9 (1e-15)	3e-9 ± 1e-9	0.06 ± 0.01

(b) $\beta = -1$						
occ.	λ	\mathbf{x}	its.	viol. (max.)	error	time (sec.)
151	-1.1507	[0.1935 0.5444 0.2991 -0.7594]	88	-	8e-9 ± 8e-10	0.19 ± 0.02
226	-3.2777	[0.6888 -0.6272 -0.2914 -0.2174]	33	14 (1e-15)	6e-9 ± 2e-9	0.08 ± 0.01
140	-3.5998	[0.7899 0.4554 0.2814 0.2991]	22	21 (1e-15)	2e-9 ± 1e-9	0.05 ± 0.01
483	-6.3985	[0.0733 0.1345 0.3877 0.9090]	82	73 (2e-15)	9e-9 ± 3e-9	0.17 ± 0.03

6. Related work. Like its predecessor SS-HOPM [6], the GEAP method has the desirable qualities of guaranteed convergence and simple implementation. Additionally, the adaptive choice of α in GEAP (as opposed to SS-HOPM) means that there are no parameters for the user to specify.

Also like SS-HOPM, the GEAP method can only converge to local maxima and minima of (3.1) and so will miss any saddle point solutions. Nevertheless, the largest and smallest magnitude eigenvalues can always be discovered by GEAP since they will not be saddle points.

6.1. Numerical Optimization Approaches. An alternative to GEAP is to solve (1.1) or (3.1) using a numerical nonlinear, homotopy, or optimization approach. The advantage of GEAP is that it guarantees decrease at each iteration *without* any globalization techniques (like line search or trust region) and is generally as cheap or cheaper per iteration than any competing numerical method. The disadvantage is that the rate of convergence of GEAP is only linear, as opposed to quadratic for, say, Newton's method.

Han [4] proposed an unconstrained variations principle for finding generalized eigenpairs. In the general case, the function to be optimized is

$$f(\mathbf{x}) = \frac{(\mathcal{B}\mathbf{x}^m)^2}{2m} - \beta \frac{\mathcal{A}\mathbf{x}^m}{m}. \quad (6.1)$$

The β has the same meaning as for GEAP: choosing $\beta = 1$ finds local maxima and $\beta = -1$ finds local minima. For comparison, the final solution is rescaled as $\mathbf{x} = \mathbf{x}/\sqrt[m]{\mathcal{B}\mathbf{x}^m}$, and then we calculate $\lambda = \mathcal{A}\mathbf{x}^m$ (since $\mathcal{B}\mathbf{x}^m = 1$).

The computational experiment settings are the same as specified in §5. Han used the MATLAB Optimization Toolbox, and we use Version 2.6.1. Following Han, we use the `fminunc` function and the default settings from calling `optimset('fminunc')` except that we explicitly specify

- `GradObj:on`
- `LargeScale:off`
- `TolX:1e-10`

- TolFun:1e-8
- MaxIter:10000
- Display:off

This means that the toolbox uses a quasi-Newton method with a line search which should have superlinear convergence.

The results of Han's method for [Example 5.5](#) are shown in [Table 6.1](#). For each eigenpair, the table lists the number of occurrences in the 1000 experiments, the median number of function evaluations (fevals) until convergence, the average error and standard deviation in the final result, and the average run time and standard deviation. The error is computed as $\|\mathcal{A}\mathbf{x}^{m-1} - \lambda\mathcal{B}\mathbf{x}^{m-1}\|_2$; both methods achieved comparable errors.

[Table 6.1:](#) Generalized eigenpairs from Han's method for $\mathcal{A}, \mathcal{B} \in \mathbb{S}^{[6,4]}$ from [Example 5.5](#)

(a) $\beta = 1$								
occ.	λ	\mathbf{x}			fevals	error	time (sec.)	
718	11.3476	[0.5544	0.3155	1.2018	0.0977]	45	1e-8 ± 2e-8	0.17 ± 0.06
134	3.7394	[0.2642	-1.1056	0.2657	-0.3160]	31	4e-9 ± 7e-9	0.12 ± 0.05
144	2.9979	[1.0008	0.4969	-0.0212	-0.4817]	31	4e-9 ± 5e-9	0.12 ± 0.05
4		— Failed to converge —				0.21 ± 0.10		

(b) $\beta = -1$								
occ.	λ	\mathbf{x}			fevals	error	time (sec.)	
72	-1.1507	[0.2291	0.6444	0.3540	-0.8990]	34	9e-9 ± 3e-8	0.14 ± 0.06
150	-3.2777	[0.8349	-0.7603	-0.3532	-0.2635]	33	5e-9 ± 7e-9	0.14 ± 0.07
148	-3.5998	[1.0486	0.6046	0.3736	0.3971]	41	6e-9 ± 8e-9	0.16 ± 0.08
624	-6.3985	[0.1003	0.1840	0.5305	1.2438]	48	7e-9 ± 1e-8	0.19 ± 0.08
4		— Converged to wrong solution —				0.10 ± 0.11		
2		— Failed to converge —				0.23 ± 0.02		

For $\beta = 1$, Han's method finds all three local maxima, though it fails to converge within 10,000 iterations for four starting points. There are no consistent results with respect to time. Han's method is faster than GEAP for $\lambda = 3.7394$ but slower for the other two eigenpairs. This is consistent if we compare the number of function evaluations and the number of iterations for GEAP, which are measuring comparable amounts of work.

For $\beta = -1$, Han's method finds all four local minima, but it fails to converge for two starting points and converges to wrong solutions for four starting points. In those four cases, it terminated because the gradient was small (flag = 1) for three cases and the fourth it stopped because it could no improve the function value (flag = 5). In this case, GEAP was faster on average for all eigenpairs.

In general, Han's method represents an alternative approach to solving the generalized tensor eigenpair problem. In [4], Han's method was compared to SS-HOPM with a fixed shift (for Z-eigenpairs only) and was superior. However, GEAP is usually as fast as Han's method and perhaps a little more robust in terms of its convergence behavior. The speed being similar is thanks to the adaptive shift in GEAP. It may be that Han's method could avoid problems of converging to incorrect solutions with tighter tolerances, but then the speed would be slower.

6.2. Other Related Work. Since (1.1) is a polynomial system of equations, we can also consider a polynomial solver approach. This does not scale to larger

problems and may be slow even for small problems. Nevertheless, it finds *all* solutions (even saddle points). We have used the Gröbner basis polynomial solver `NSolve` in Mathematica to compute the full set of solutions for the problems discussed in this paper.

In terms of methods specifically geared to tensor eigenvalues, most work has focused on computing the largest H-eigenvalue for a *nonnegative* tensor: [9, 8]. The method of Liu, Zhou, and Ibrahim [8] is guaranteed to always find the largest eigenvalue and also uses a “shift” approach.

7. Conclusions. The paper has proposed two improvements to the SS-HOPM method [6]. First, we have adapted the method to the *generalized* tensor eigenproblem. Second, we have proposed a method for adaptively and automatically selecting the shift, overcoming a major problem with the SS-HOPM method because choosing the shift too large dramatically slows convergence whereas choosing it too small can cause the method to fail completely.

We have tested our method numerically on several problems from the literature, including computing of Z-, H-, and D-eigenpairs. We have also proposed a novel method for generating random symmetric positive definite tensors.

As this paper was in review, a new method has been proposed to compute *all* real general eigenvalues using Jacobian semidefinite programming relaxations [3]. Comparing to this method will be a topic of future study.

Appendix A. Useful derivatives. First, we consider the gradient and Hessian of the general function

$$f(\mathbf{x}) = \frac{f_1(\mathbf{x})f_2(\mathbf{x})}{f_3(\mathbf{x})}.$$

Let $\mathbf{g}_i(\mathbf{x})$ denote $\nabla f_i(\mathbf{x})$. From matrix calculus, the gradient of $f(\mathbf{x})$ is

$$\mathbf{g} = \left(f_1\mathbf{g}_2 + f_2\mathbf{g}_1 - (f_1f_2/f_3)\mathbf{g}_3 \right) / f_3.$$

Here we have dropped the argument, \mathbf{x} , to simply the notation. Let $\mathbf{H}_i(\mathbf{x})$ denote $\nabla^2 f_i(\mathbf{x})$. The Hessian of $f(\mathbf{x})$ is

$$\begin{aligned} \mathbf{H} = & \frac{f_1f_2}{f_3^3}(\mathbf{g}_3 \odot \mathbf{g}_3) + \frac{1}{f_3} \left[f_2\mathbf{H}_1 + f_1\mathbf{H}_2 + (\mathbf{g}_1 \odot \mathbf{g}_2) \right] \\ & - \frac{1}{f_3^2} \left[f_1f_2\mathbf{H}_3 + f_2(\mathbf{g}_1 \odot \mathbf{g}_3) + f_1(\mathbf{g}_2 \odot \mathbf{g}_3) \right]. \end{aligned}$$

Now we specialize $f(\mathbf{x})$ to (3.1): let $f_1 = \mathcal{A}\mathbf{x}^m$, $f_2 = \|\mathbf{x}\|^m$, and $f_3 = \mathcal{B}\mathbf{x}^m$. The following derivatives are proved in [6]:

$$\begin{aligned} \mathbf{g}_1 &= m\mathcal{A}\mathbf{x}^{m-1}, & \mathbf{H}_1 &= m(m-1)\mathcal{A}\mathbf{x}^{m-2}, \\ \mathbf{g}_2 &= m\|\mathbf{x}\|^{m-2}\mathbf{x}, & \mathbf{H}_2 &= m\|\mathbf{x}\|^{m-2}\mathbf{I} + m(m-2)\|\mathbf{x}\|^{m-4}\mathbf{x}\mathbf{x}^\top, \\ \mathbf{g}_3 &= m\mathcal{B}\mathbf{x}^{m-1}, & \mathbf{H}_3 &= m(m-1)\mathcal{B}\mathbf{x}^{m-2}. \end{aligned}$$

We need only consider the case for $\mathbf{x} \in \Sigma$, so we may assume

$$f_2 = 1, \quad \mathbf{g}_2 = m\mathbf{x}, \quad \mathbf{H}_2 = m(\mathbf{I} + (m-2)\mathbf{x}\mathbf{x}^\top).$$

Putting everything together, we have for $\mathbf{x} \in \Sigma$,

$$\mathbf{g}(\mathbf{x}) = \frac{m}{\mathcal{B}\mathbf{x}^m} \left[(\mathcal{A}\mathbf{x}^m) \mathbf{x} + \mathcal{A}\mathbf{x}^{m-1} - \left(\frac{\mathcal{A}\mathbf{x}^m}{\mathcal{B}\mathbf{x}^m} \right) \mathcal{B}\mathbf{x}^{m-1} \right].$$

For the Hessian, assuming $\mathbf{x} \in \Sigma$, we have

$$\begin{aligned} \mathbf{H}(\mathbf{x}) &= \frac{m^2 \mathcal{A}\mathbf{x}^m}{(\mathcal{B}\mathbf{x}^m)^3} (\mathcal{B}\mathbf{x}^{m-1} \odot \mathcal{B}\mathbf{x}^{m-1}) \\ &\quad + \frac{m}{\mathcal{B}\mathbf{x}^m} \left[(m-1)\mathcal{A}\mathbf{x}^{m-2} + \mathcal{A}\mathbf{x}^m (\mathbf{I} + (m-2)\mathbf{x}\mathbf{x}^\top) + m(\mathcal{A}\mathbf{x}^{m-1} \odot \mathbf{x}) \right] \\ &\quad - \frac{m}{(\mathcal{B}\mathbf{x}^m)^2} \left[(m-1)\mathcal{A}\mathbf{x}^m \mathcal{B}\mathbf{x}^{m-2} + m(\mathcal{A}\mathbf{x}^{m-1} \odot \mathcal{B}\mathbf{x}^{m-1}) + m\mathcal{A}\mathbf{x}^m (\mathbf{x} \odot \mathcal{B}\mathbf{x}^{m-1}) \right]. \end{aligned}$$

Appendix B. Tensor specifications. The tensor for [Example 5.1](#) comes from Example 1 in Kofidis and Regalia [5] and is specified in [Figure B.1](#).

$$\begin{array}{llll} a_{1111} = 0.2883, & a_{1112} = -0.0031, & a_{1113} = 0.1973, & a_{1122} = -0.2485, \\ a_{1123} = -0.2939, & a_{1133} = 0.3847, & a_{1222} = 0.2972, & a_{1223} = 0.1862, \\ a_{1233} = 0.0919, & a_{1333} = -0.3619, & a_{2222} = 0.1241, & a_{2223} = -0.3420, \\ a_{2233} = 0.2127, & a_{2333} = 0.2727, & a_{3333} = -0.3054. \end{array}$$

Fig. B.1: \mathcal{A} from Kofidis and Regalia [5], used in [Example 5.1](#)

The tensor \mathcal{A} used for [Example 5.2](#) is randomly generated as described in [§5.2](#); its entries are specified in [Figure B.2](#).

The tensor used in [Example 5.3](#). The DKI tensor $\mathcal{A} \in \mathbb{S}^{[4,3]}$ (called \mathcal{W} in the original paper) is the symmetric tensor defined by the unique elements shown in [Figure B.3](#). The tensor \mathcal{B} is the symmetrized outer product of the matrix \mathbf{D} with itself where

$$\mathbf{D} = \begin{bmatrix} 1.755 & 0.035 & 0.132 \\ 0.035 & 1.390 & 0.017 \\ 0.132 & 0.017 & 4.006 \end{bmatrix},$$

so \mathcal{B} is the tensor whose unique elements are given in [Figure B.4](#).

The tensor \mathcal{A} used in [Example 5.5](#) is the same as is used in [Example 5.2](#) and specified in [Figure B.2](#). The tensor \mathcal{B} is a random positive definite tensor from [Example 5.5](#); its entries are specified in [Figure B.5](#).

$$\begin{aligned}
a_{111111} &= 0.2888, & a_{111112} &= -0.0013, & a_{111113} &= -0.1422, & a_{111114} &= -0.0323, \\
a_{111122} &= -0.1079, & a_{111123} &= -0.0899, & a_{111124} &= -0.2487, & a_{111133} &= 0.0231, \\
a_{111134} &= -0.0106, & a_{111144} &= 0.0740, & a_{111222} &= 0.1490, & a_{111223} &= 0.0527, \\
a_{111224} &= -0.0710, & a_{111233} &= -0.1039, & a_{111234} &= -0.0250, & a_{111244} &= 0.0169, \\
a_{111333} &= 0.2208, & a_{111334} &= 0.0662, & a_{111344} &= 0.0046, & a_{111444} &= 0.0943, \\
a_{112222} &= -0.1144, & a_{112223} &= -0.1295, & a_{112224} &= -0.0484, & a_{112233} &= 0.0238, \\
a_{112234} &= -0.0237, & a_{112244} &= 0.0308, & a_{112333} &= 0.0142, & a_{112334} &= 0.0006, \\
a_{112344} &= -0.0044, & a_{112444} &= 0.0353, & a_{113333} &= 0.0947, & a_{113334} &= -0.0610, \\
a_{113344} &= -0.0293, & a_{113444} &= 0.0638, & a_{114444} &= 0.2326, & a_{122222} &= -0.2574, \\
a_{122223} &= 0.1018, & a_{122224} &= 0.0044, & a_{122233} &= 0.0248, & a_{122234} &= 0.0562, \\
a_{122244} &= 0.0221, & a_{122333} &= 0.0612, & a_{122334} &= 0.0184, & a_{122344} &= 0.0226, \\
a_{122444} &= 0.0247, & a_{123333} &= 0.0847, & a_{123334} &= -0.0209, & a_{123344} &= -0.0795, \\
a_{123444} &= -0.0323, & a_{124444} &= -0.0819, & a_{133333} &= 0.5486, & a_{133334} &= -0.0311, \\
a_{133344} &= -0.0592, & a_{133444} &= 0.0386, & a_{134444} &= -0.0138, & a_{144444} &= 0.0246, \\
a_{222222} &= 0.9207, & a_{222223} &= -0.0908, & a_{222224} &= 0.0633, & a_{222233} &= 0.1116, \\
a_{222234} &= -0.0318, & a_{222244} &= 0.1629, & a_{222333} &= 0.1797, & a_{222334} &= -0.0348, \\
a_{222344} &= -0.0058, & a_{222444} &= 0.1359, & a_{223333} &= 0.0584, & a_{223334} &= -0.0299, \\
a_{223344} &= -0.0110, & a_{223444} &= 0.1375, & a_{224444} &= -0.1405, & a_{233333} &= 0.3613, \\
a_{233334} &= 0.0809, & a_{233344} &= 0.0205, & a_{233444} &= 0.0196, & a_{234444} &= 0.0226, \\
a_{244444} &= -0.2487, & a_{333333} &= 0.6007, & a_{333334} &= -0.0272, & a_{333344} &= -0.1343, \\
a_{333444} &= -0.0233, & a_{334444} &= -0.0227, & a_{344444} &= -0.3355, & a_{444444} &= -0.5937.
\end{aligned}$$

Fig. B.2: \mathcal{A} from Example 5.2 and Example 5.5

$$\begin{aligned}
a_{1111} &= 0.4982, & a_{1112} &= -0.0582, & a_{1113} &= -1.1719, & a_{1122} &= 0.2236, \\
a_{1123} &= -0.0171, & a_{1133} &= 0.4597, & a_{1222} &= 0.4880, & a_{1223} &= 0.1852, \\
a_{1233} &= -0.4087, & a_{1333} &= 0.7639, & a_{2222} &= 0.0000, & a_{2223} &= -0.6162, \\
a_{2233} &= 0.1519, & a_{2333} &= 0.7631, & a_{3333} &= 2.6311.
\end{aligned}$$

Fig. B.3: \mathcal{A} from Qi, Wang, and Wu [11], used in Example 5.3

$$\begin{aligned}
b_{1111} &= 3.0800, & b_{1112} &= 0.0614, & b_{1113} &= 0.2317, & b_{1122} &= 0.8140, \\
b_{1123} &= 0.0130, & b_{1133} &= 2.3551, & b_{1222} &= 0.0486, & b_{1223} &= 0.0616, \\
b_{1233} &= 0.0482, & b_{1333} &= 0.5288, & b_{2222} &= 1.9321, & b_{2223} &= 0.0236, \\
b_{2233} &= 1.8563, & b_{2333} &= 0.0681, & b_{3333} &= 16.0480.
\end{aligned}$$

Fig. B.4: \mathcal{B} from Qi, Wang, and Wu [11], used in Example 5.3

$$\begin{aligned}
b_{111111} &= 0.2678, & b_{111112} &= -0.0044, & b_{111113} &= -0.0326, & b_{111114} &= -0.0081, \\
b_{111122} &= 0.0591, & b_{111123} &= -0.0009, & b_{111124} &= -0.0045, & b_{111133} &= 0.0533, \\
b_{111134} &= -0.0059, & b_{111144} &= 0.0511, & b_{111222} &= -0.0029, & b_{111223} &= -0.0072, \\
b_{111224} &= -0.0016, & b_{111233} &= -0.0005, & b_{111234} &= 0.0007, & b_{111244} &= -0.0006, \\
b_{111333} &= -0.0185, & b_{111334} &= 0.0001, & b_{111344} &= -0.0058, & b_{111444} &= -0.0046, \\
b_{112222} &= 0.0651, & b_{112223} &= -0.0013, & b_{112224} &= -0.0050, & b_{112233} &= 0.0190, \\
b_{112234} &= -0.0023, & b_{112244} &= 0.0190, & b_{112333} &= -0.0011, & b_{112334} &= -0.0014, \\
b_{112344} &= 0.0000, & b_{112444} &= -0.0043, & b_{113333} &= 0.0498, & b_{113334} &= -0.0061, \\
b_{113344} &= 0.0169, & b_{113444} &= -0.0060, & b_{114444} &= 0.0486, & b_{122222} &= -0.0054, \\
b_{122223} &= -0.0078, & b_{122224} &= -0.0016, & b_{122233} &= -0.0006, & b_{122234} &= 0.0008, \\
b_{122244} &= -0.0006, & b_{122333} &= -0.0067, & b_{122334} &= 0.0001, & b_{122344} &= -0.0022, \\
b_{122444} &= -0.0016, & b_{123333} &= -0.0002, & b_{123334} &= 0.0006, & b_{123344} &= -0.0002, \\
b_{123444} &= 0.0006, & b_{124444} &= -0.0003, & b_{133333} &= -0.0286, & b_{133334} &= 0.0017, \\
b_{133344} &= -0.0056, & b_{133444} &= 0.0001, & b_{134444} &= -0.0051, & b_{144444} &= -0.0073, \\
b_{222222} &= 0.3585, & b_{222223} &= -0.0082, & b_{222224} &= -0.0279, & b_{222233} &= 0.0610, \\
b_{222234} &= -0.0076, & b_{222244} &= 0.0636, & b_{222333} &= -0.0042, & b_{222334} &= -0.0044, \\
b_{222344} &= -0.0002, & b_{222444} &= -0.0145, & b_{223333} &= 0.0518, & b_{223334} &= -0.0067, \\
b_{223344} &= 0.0184, & b_{223444} &= -0.0069, & b_{224444} &= 0.0549, & b_{233333} &= -0.0059, \\
b_{233334} &= -0.0034, & b_{233344} &= -0.0002, & b_{233444} &= -0.0039, & b_{234444} &= 0.0010, \\
b_{244444} &= -0.0208, & b_{333333} &= 0.2192, & b_{333334} &= -0.0294, & b_{333344} &= 0.0477, \\
b_{333444} &= -0.0181, & b_{334444} &= 0.0485, & b_{344444} &= -0.0304, & b_{444444} &= 0.2305.
\end{aligned}$$

Fig. B.5: \mathcal{B} from Example 5.5

Appendix C. Complete lists of real eigenpairs. A polynomial system solver (`NSolve`) using a Gröbner basis method is available in Mathematica and has been employed to generate a complete list of eigenpairs for the examples in this paper in Tables C.1–C.4.

Table C.1: All Z-eigenpairs for $\mathcal{A} \in \mathbb{S}^{[4,3]}$ from Example 5.1

λ	\mathbf{x}^\top	$\mathbf{C}(\lambda, \mathbf{x})$ evals.	type
-1.0954	[0.5915 -0.7467 -0.3043]	1.86, 2.75	Minima
-0.5629	[0.1762 -0.1796 0.9678]	1.63, 2.38	Minima
-0.0451	[0.7797 0.6135 0.1250]	0.82, 1.25	Minima
0.1735	[0.3357 0.9073 0.2531]	-1.10, 0.86	Saddle
0.2433	[0.9895 0.0947 -0.1088]	-1.19, 1.46	Saddle
0.2628	[0.1318 -0.4425 -0.8870]	0.62, -2.17	Saddle
0.2682	[0.6099 0.4362 0.6616]	-1.18, 0.79	Saddle
0.3633	[0.2676 0.6447 0.7160]	-1.18, -0.57	Maxima
0.5105	[0.3598 -0.7780 0.5150]	0.59, -2.34	Saddle
0.8169	[0.8412 -0.2635 0.4722]	-2.26, -0.90	Maxima
0.8893	[0.6672 0.2471 -0.7027]	-1.85, -0.89	Maxima

Table C.2: All H-eigenpairs for $\mathcal{A} \in \mathbb{S}^{[6,4]}$ from Example 5.2

λ	\mathbf{x}^\top	$\mathbf{C}(\lambda, \mathbf{x})$ evals.	type
-10.7440	[0.4664 0.4153 -0.5880 -0.5140]	75.69, 30.21, 41.28	Minima
-8.3201	[0.5970 -0.5816 -0.4740 -0.2842]	62.11, 28.56, 15.64	Minima
-4.1781	[0.4397 0.5139 -0.5444 0.4962]	5.67, 31.85, 21.21	Minima
-3.7180	[0.6843 0.5519 0.3136 0.3589]	26.89, 7.05, 12.50	Minima
-3.3137	[0.5588 0.4954 -0.6348 0.1986]	-4.83, 11.31, 17.73	Saddle
-3.0892	[0.6418 -0.2049 -0.6594 -0.3336]	-10.41, 22.10, 6.26	Saddle
-2.9314	[0.3161 0.5173 0.4528 -0.6537]	31.95, 6.88, 13.47	Minima
-2.0437	[0.6637 0.5911 -0.2205 0.4017]	15.87, -4.81, 8.30	Saddle
-1.3431	[0.0544 0.4258 0.0285 0.9027]	4.40, 2.04, -0.85	Saddle
-1.0965	[0.5156 0.3387 0.4874 0.6180]	24.09, 14.29, -13.10	Saddle
-1.0071	[0.2030 0.5656 -0.0975 -0.7933]	-3.71, 4.13, 5.35	Saddle
-0.3600	[0.6999 -0.1882 0.3292 -0.6053]	9.74, 3.89, -2.07	Saddle
-0.3428	[0.3879 -0.1700 0.5174 -0.7436]	-3.52, 6.07, 1.24	Saddle
0.0073	[0.3068 0.0539 0.3127 -0.8973]	-2.92, -1.29, 1.22	Saddle
0.1902	[0.9744 -0.0316 0.2013 -0.0952]	-1.49, 2.17, 0.65	Saddle
0.3947	[0.5416 0.4650 0.0708 0.6967]	8.59, -15.89, -3.63	Saddle
0.4679	[0.9613 0.0442 -0.2718 0.0083]	1.32, -1.33, -1.73	Saddle
0.5126	[0.4232 -0.6781 -0.2347 0.5532]	-8.44, 9.45, 7.66	Saddle
0.5236	[0.3092 0.8725 0.1389 -0.3518]	-2.58, 1.68, 3.60	Saddle
0.7573	[0.5830 -0.2565 -0.3076 -0.7069]	1.86, -5.35, -14.39	Saddle
0.8693	[0.2414 0.8332 -0.2479 -0.4313]	3.48, -3.31, -2.38	Saddle
0.9572	[0.1035 -0.9754 -0.1932 -0.0221]	-2.05, 0.83, 1.80	Saddle
1.1006	[0.2033 -0.9035 -0.1584 0.3424]	2.10, -2.38, -1.15	Saddle
2.3186	[0.1227 -0.8044 -0.0334 -0.5804]	2.50, -2.74, -10.23	Saddle
2.7045	[0.3618 -0.5607 -0.5723 0.4766]	8.78, -17.72, -21.79	Saddle
3.3889	[0.6320 0.5549 0.3596 -0.4043]	16.59, -25.41, -17.68	Saddle
3.9099	[0.6722 -0.2683 -0.1665 0.6697]	-21.17, -4.98, 5.01	Saddle
4.8422	[0.5895 -0.2640 -0.4728 0.5994]	-28.20, -6.48, -15.54	Maxima
5.1757	[0.6513 0.0021 0.7550 -0.0760]	-23.82, 3.66, -3.35	Saddle
5.8493	[0.6528 0.5607 -0.0627 -0.5055]	-34.20, -22.87, -9.58	Maxima
8.7371	[0.4837 0.5502 0.6671 -0.1354]	-7.66, -19.48, -43.93	Maxima
9.0223	[0.5927 -0.5567 0.5820 -0.0047]	-58.03, -28.84, 4.60	Saddle
9.6386	[0.5342 -0.5601 0.5466 -0.3197]	-64.78, -41.13, -9.04	Maxima
14.6941	[0.5426 -0.4853 0.4760 0.4936]	-94.14, -61.11, -54.81	Maxima

Table C.3: All D-eigenpairs for $\mathcal{A} \in \mathbb{S}^{[4,3]}$ and $\mathbf{D} \in \mathbb{S}^{[2,3]}$ from Example 5.3

λ	\mathbf{x}^\top	$\mathbf{C}(\lambda, \mathbf{x})$ evals.	type
-0.3313	[0.2309 -0.7741 -0.1509]	1.02, 2.11	Minima
-0.1242	[0.6577 0.0712 0.2189]	0.35, 1.25	Minima
-0.0074	[0.2161 0.3149 -0.4485]	0.36, 0.46	Minima
0.0611	[0.6113 -0.4573 0.1181]	-0.63, 1.14	Saddle
0.1039	[0.3314 0.5239 0.3084]	-0.46, 0.63	Saddle
0.2009	[0.2440 -0.1250 0.4601]	-0.32, 0.07	Saddle
0.2056	[0.1211 -0.2367 -0.4766]	-0.29, 0.13	Saddle
0.2219	[0.1143 0.1812 0.4773]	-0.08, -0.20	Maxima
0.2431	[0.0943 -0.6840 0.2905]	0.18, -1.11	Saddle
0.2514	[0.2485 -0.5579 0.3363]	-0.14, -0.71	Maxima
0.3827	[0.6236 0.3954 -0.1678]	-1.58, 0.32	Saddle
0.4359	[0.4336 0.6714 -0.0949]	-0.43, -1.64	Maxima
0.5356	[0.6638 -0.1123 -0.2537]	-0.48, -1.43	Maxima

Table C.4: All generalized tensor eigenpairs for $\mathcal{A} \in \mathbb{S}^{[6,4]}$ and $\mathcal{B} \in \mathbb{S}_+^{[6,4]}$ from [Example 5.5](#)

λ	\mathbf{x}^\top	$\mathbf{C}(\lambda, \mathbf{x})$ evals.	type
-6.3985	[0.0733 0.1345 0.3877 0.9090]	20.43, 4.93, 11.20	Minima
-3.5998	[0.7899 0.4554 0.2814 0.2991]	8.05, 10.39, 12.41	Minima
-3.2777	[0.6888 -0.6272 -0.2914 -0.2174]	8.27, 3.65, 5.95	Minima
-1.7537	[0.6329 -0.2966 -0.6812 -0.2180]	-4.25, 3.00, 5.56	Saddle
-1.1507	[0.1935 0.5444 0.2991 -0.7594]	0.73, 3.54, 4.20	Minima
-1.0696	[0.1372 0.5068 0.0665 -0.8485]	-1.54, 3.30, 3.64	Saddle
-1.0456	[0.2365 0.4798 -0.7212 0.4402]	-1.16, 1.54, 2.57	Saddle
-0.7842	[0.5409 0.3388 0.4698 0.6099]	16.02, 8.79, -12.47	Saddle
-0.7457	[0.6348 0.5354 -0.4388 0.3434]	2.49, 0.94, -1.59	Saddle
-0.2542	[0.3900 -0.1333 0.4946 -0.7652]	-2.51, 2.99, 0.93	Saddle
-0.2359	[0.6956 -0.1369 0.3550 -0.6094]	6.38, 2.23, -1.27	Saddle
0.0132	[0.3064 0.0541 0.3111 -0.8980]	-5.33, -2.36, 2.21	Saddle
0.1633	[0.4278 -0.6578 -0.2545 0.5652]	-2.42, 3.86, 2.36	Saddle
0.3250	[0.5265 0.4653 0.0927 0.7055]	7.50, -12.05, -3.41	Saddle
0.5206	[0.3738 -0.4806 -0.6066 0.5111]	3.19, -2.27, -1.47	Saddle
0.5463	[0.5157 -0.3055 -0.3313 -0.7287]	-9.91, -3.67, 1.37	Saddle
0.5945	[0.4015 0.8447 0.1782 -0.3058]	-3.70, 4.95, 1.87	Saddle
0.6730	[0.9634 -0.0009 0.2396 -0.1204]	-5.84, 7.88, 1.78	Saddle
0.8862	[0.3559 0.8571 -0.1675 -0.3326]	3.55, -2.24, -2.63	Saddle
1.2962	[0.9849 0.0018 -0.1681 0.0419]	2.20, -5.97, -3.18	Saddle
1.4646	[0.7396 0.4441 0.4009 -0.3083]	8.41, -2.08, -7.72	Saddle
2.9979	[0.8224 0.4083 -0.0174 -0.3958]	-4.00, -5.46, -6.56	Maxima
3.5181	[0.4494 -0.7574 0.4502 -0.1469]	-9.40, 1.89, -2.83	Saddle
3.6087	[0.0340 -0.8989 -0.0373 -0.4353]	0.87, -8.03, -5.77	Saddle
3.7394	[0.2185 -0.9142 0.2197 -0.2613]	-8.72, -0.90, -3.34	Maxima
11.3476	[0.4064 0.2313 0.8810 0.0716]	-7.20, -18.98, -21.53	Maxima

REFERENCES

- [1] BRETT W. BADER, TAMARA G. KOLDA, ET AL., *Matlab tensor toolbox version 2.5*, Jan. 2012, <http://www.sandia.gov/~tgkolda/TensorToolbox/>.
- [2] K. C. CHANG, KELLY PEARSON, AND TAN ZHANG, *On eigenvalue problems of real symmetric tensors*, Journal of Mathematical Analysis and Applications, 350 (2009), pp. 416–422, DOI:[10.1016/j.jmaa.2008.09.067](https://doi.org/10.1016/j.jmaa.2008.09.067).
- [3] CHUN-FENG CUI, YU-HONG DAI, AND JIAWANG NIE, *All real eigenvalues of symmetric tensors*, Mar. 2014, arXiv:[1403.3720](https://arxiv.org/abs/1403.3720).
- [4] LIXING HAN, *An unconstrained optimization approach for finding real eigenvalues of even order symmetric tensors*, Numerical Algebra, Control and Optimization (NACO), 3 (2012), pp. 583–599, DOI:[10.3934/naco.2013.3.583](https://doi.org/10.3934/naco.2013.3.583).
- [5] ELEFTHERIOS KOFIDIS AND PHILLIP A. REGALIA, *On the best rank-1 approximation of higher-order supersymmetric tensors*, SIAM Journal on Matrix Analysis and Applications, 23 (2002), pp. 863–884, DOI:[10.1137/S0895479801387413](https://doi.org/10.1137/S0895479801387413).
- [6] TAMARA G. KOLDA AND JACKSON R. MAYO, *Shifted power method for computing tensor eigenpairs*, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 1095–1124, DOI:[10.1137/100801482](https://doi.org/10.1137/100801482).
- [7] LEK-HENG LIM, *Singular values and eigenvalues of tensors: A variational approach*, in CAMSAP’05: Proceeding of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005, pp. 129–132, DOI:[10.1109/CAMSAP.2005.1574201](https://doi.org/10.1109/CAMSAP.2005.1574201).
- [8] YONGJUN LIU, GUANGLU ZHOU, AND NUR FADHILAH IBRAHIM, *An always convergent algorithm for the largest eigenvalue of an irreducible nonnegative tensor*, Journal of Computational and Applied Mathematics, 235 (2010), pp. 286–292, DOI:[10.1016/j.cam.2010.06.002](https://doi.org/10.1016/j.cam.2010.06.002).
- [9] MICHAEL NG, LIQUN QI, AND GUANGLU ZHOU, *Finding the largest eigenvalue of a nonnegative tensor*, SIAM Journal on Matrix Analysis and Applications, 31 (2009), pp. 1090–1099, DOI:[10.1137/09074838X](https://doi.org/10.1137/09074838X).
- [10] LIQUN QI, *Eigenvalues of a real supersymmetric tensor*, Journal of Symbolic Computation, 40 (2005), pp. 1302–1324, DOI:[10.1016/j.jsc.2005.05.007](https://doi.org/10.1016/j.jsc.2005.05.007).
- [11] LIQUN QI, YIJU WANG, AND ED X. WU, *D-eigenvalues of diffusion kurtosis tensors*, Journal of Computational and Applied Mathematics, 221 (2008), pp. 150–157, DOI:[10.1016/j.cam.2007.10.012](https://doi.org/10.1016/j.cam.2007.10.012).
- [12] PHILLIP A. REGALIA AND ELEFTHERIOS KOFIDIS, *Monotonic convergence of fixed-point algorithms for ICA*, IEEE Transactions on Neural Networks, 14 (2003), pp. 943– 949, DOI:[10.1109/TNN.2003.813843](https://doi.org/10.1109/TNN.2003.813843).
- [13] MARTIN D. SCHATZ, TZE-MENG LOW, ROBERT A. VAN DE GEIJN, AND TAMARA G. KOLDA, *Exploiting symmetry in tensors for high performance*, SIAM Journal on Scientific Computing, arXiv:[1301.7744](https://arxiv.org/abs/1301.7744). accepted 2014-03-11.