SIAM J. MATRIX ANAL. APPL.
Vol. 32, No. 4, pp. 1095–1124
© 2011 Society for Industrial and Applied Mathematics

SHIFTED POWER METHOD FOR COMPUTING TENSOR EIGENPAIRS*

TAMARA G. KOLDA† AND JACKSON R. MAYO†

Abstract. Recent work on eigenvalues and eigenvectors for tensors of order $m \geq 3$ has been motivated by applications in blind source separation, magnetic resonance imaging, molecular conformation, and more. In this paper, we consider methods for computing real symmetric-tensor eigenpairs of the form $\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathbf{x}$ subject to $\|\mathbf{x}\| = 1$, which is closely related to optimal rank-1 approximation of a symmetric tensor. Our contribution is a shifted symmetric higher-order power method (SS-HOPM), which we show is guaranteed to converge to a tensor eigenpair. SS-HOPM can be viewed as a generalization of the power iteration method for matrices or of the symmetric higher-order power method. Additionally, using fixed point analysis, we can characterize exactly which eigenpairs can and cannot be found by the method. Numerical examples are presented, including examples from an extension of the method to finding complex eigenpairs.

Key words. tensor eigenvalues, E-eigenpairs, Z-eigenpairs, l^2-eigenpairs, rank-1 approximation, symmetric higher-order power method (S-HOPM), shifted symmetric higher-order power method (SS-HOPM)

AMS subject classifications. 15A18, 15A69

DOI. 10.1137/100801482

1. Introduction.

Tensor eigenvalues and eigenvectors have received much attention lately in the literature [13, 18, 20, 19, 4, 15, 27]. The tensor eigenproblem is important because it has applications in blind source separation [11], magnetic resonance imaging [25, 22], molecular conformation [7], etc. There is more than one possible definition for a tensor eigenpair [18]; in this paper, we specifically use the following definition.

DEFINITION 1.1. *Assume that $\mathcal{A}$ is a symmetric mth-order n-dimensional real-valued tensor. For any n-dimensional vector $\mathbf{x}$, define*

$$\left(\mathcal{A}\mathbf{x}^{m-1}\right)_{i_1} \equiv \sum_{i_2=1}^{n} \cdots \sum_{i_m=1}^{n} a_{i_1 i_2 \cdots i_m} x_{i_2} \cdots x_{i_m} \quad \text{for} \quad i_1 = 1, \ldots, n. \tag{1.1}$$

Then $\lambda \in \mathbb{R}$ is an eigenvalue *of $\mathcal{A}$ if there exists $\mathbf{x} \in \mathbb{R}^n$ such that*

$$\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathbf{x} \quad \text{and} \quad \mathbf{x}^T\mathbf{x} = 1. \tag{1.2}$$

The vector $\mathbf{x}$ is a corresponding eigenvector*, and $(\lambda, \mathbf{x})$ is called an* eigenpair*.*

Definition 1.1 is equivalent to the Z-eigenpairs defined by Qi [18, 19] and the l^2-eigenpairs defined by Lim [13]. In particular, Lim [13] observes that any eigenpair

*Received by the editors July 8, 2010; accepted for publication (in revised form) April 5, 2011; published electronically October 4, 2011. 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 U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. Copyright is owned by SIAM to the extent not limited by these rights.

http://www.siam.org/journals/simax/32-4/80148.html

†Sandia National Laboratories, Livermore, CA 94551 (tgkolda@sandia.gov, jmayo@sandia.gov).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$(\lambda, \mathbf{x})$ is a Karush–Kuhn–Tucker (KKT) point (i.e., a constrained stationary point) of the nonlinear optimization problem

(1.3)
$$\max_{\mathbf{x}\in\mathbb{R}^n} \mathcal{A}\mathbf{x}^m \quad \text{subject to} \quad \mathbf{x}^T\mathbf{x} = 1, \quad \text{where} \quad \mathcal{A}\mathbf{x}^m \equiv \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} a_{i_1\cdots i_m} x_{i_1} \cdots x_{i_m}.$$

This is equivalent to the problem of finding the best *symmetric* rank-1 approximation of a symmetric tensor [6]. We present the more general definition that incorporates complex-valued eigenpairs in section 5.

In this paper, we build upon foundational work by Kofidis and Regalia [11] for solving (1.3). Their paper is extremely important for computing tensor eigenvalues even though it predates the definition of the eigenvalue problem by three years. Kofidis and Regalia consider the higher-order power method (HOPM) [6], a well-known technique for approximation of higher-order tensors, and show that its symmetric generalization (S-HOPM) is not guaranteed to converge. They go on, however, to use convexity theory to provide theoretical results (as well as practical examples) explaining conditions under which the method is convergent for even-order tensors (i.e., m even). Further, these conditions are shown to hold for many problems of practical interest.

In the context of independent component analysis (ICA), both Regalia and Kofidis [23] and Erdogan [8] have developed shifted variants of the power method and shown that they are monotonically convergent. We present a similar method in the context of finding real-valued tensor eigenpairs, called the shifted symmetric higher-order power method (SS-HOPM), along with theory showing that it is guaranteed to converge to a constrained stationary point of (1.3). The proof is general and works for both odd- and even-order tensors (i.e., all $m \geq 3$). The effectiveness of SS-HOPM is demonstrated on several examples, including a problem noted previously [11] for which S-HOPM does not converge. We also present a version of SS-HOPM for finding complex-valued tensor eigenpairs and provide examples of its effectiveness.

As mentioned, there is more than one definition of a tensor eigenpair. In the case of the *l^m-eigenpair* (we use m for the tensor order instead of k as in some references) or *H-eigenpair*, the eigenvalue equation becomes $\mathcal{A}\mathbf{x}^{m-1} = \lambda\mathbf{x}^{[m-1]}$, where $\mathbf{x}^{[m-1]}$ denotes the vector $\mathbf{x}$ with each element raised to the $(m-1)$st power [13, 18]. In this context, Qi, Wang, and Wang [21] propose some methods specific to third-order tensors ($m = 3$). Unlike the (l^2-)eigenvalues we consider here, it is possible to guarantee convergence to the *largest* l^m-eigenvalue for certain classes of nonnegative tensors. For example, see the power methods proposed by Ng, Qi, and Zhou [15] and Liu, Zhou, and Ibrahim [14], the latter of which also uses a shift to guarantee convergence for any irreducible nonnegative tensor.

2. Preliminaries. Throughout, let Γ and Σ denote the unit ball and sphere on $\mathbb{R}^n$, i.e.,

$$\Gamma = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq 1\} \quad \text{and} \quad \Sigma = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}.$$

Additionally, define

$$\Pi_m \equiv \text{the set of all permutations of } (1, \ldots, m).$$

Let $\mathbf{x}\perp\mathbf{y}$ denote $\mathbf{x}^T\mathbf{y} = 0$, and define $\mathbf{x}^\perp \equiv \{\mathbf{y} \in \mathbb{R}^n : \mathbf{x}\perp\mathbf{y}\}$. Let $\rho(\mathbf{A})$ denote the spectral radius of a square matrix $\mathbf{A}$, i.e., the maximum of the magnitudes of its eigenvalues.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

2.1. Tensors. A tensor is an m-way array. We let $\mathbb{R}^{[m,n]}$ denote the space of mth-order real-valued tensors with dimension n, e.g., $\mathbb{R}^{[3,2]} = \mathbb{R}^{2\times 2\times 2}$. We adopt the convention that $\mathbb{R}^{[0,n]} = \mathbb{R}$.

We formally introduce the notion of a symmetric tensor, sometimes also called supersymmetric, which is invariant under any permutation of its indices. Further, we define a generalization of the tensor-vector multiplication in (1.1) and (1.3).

DEFINITION 2.1 (symmetric tensor [5]). *A tensor* $\mathcal{A} \in \mathbb{R}^{[m,n]}$ *is* symmetric *if*

$$a_{i_{p(1)}\cdots i_{p(m)}} = a_{i_1\cdots i_m} \quad \textit{for all} \quad i_1,\ldots,i_m \in \{1,\ldots,n\} \quad \textit{and} \quad p \in \Pi_m.$$

DEFINITION 2.2 (symmetric tensor-vector multiply). *Let* $\mathcal{A} \in \mathbb{R}^{[m,n]}$ *be symmetric and* $\mathbf{x} \in \mathbb{R}^n$. *Then for* $0 \le r \le m-1$, *the* $(m-r)$-times product *of the tensor* $\mathcal{A}$ *with the vector* $\mathbf{x}$ *is denoted by* $\mathcal{A}\mathbf{x}^{m-r} \in \mathbb{R}^{[r,n]}$ *and defined by*

$$(\mathcal{A}\mathbf{x}^{m-r})_{i_1\cdots i_r} \equiv \sum_{i_{r+1},\ldots,i_m} a_{i_1\cdots i_m} x_{i_{r+1}} \cdots x_{i_m} \quad \textit{for all} \quad i_1,\ldots,i_r \in \{1,\ldots,n\}.$$

Example 2.3. The identity matrix plays an important role in matrix analysis. This notion can be extended in a sense to the domain of tensors. We may define an identity tensor as a symmetric tensor $\mathcal{E} \in \mathbb{R}^{[m,n]}$ such that

$$\mathcal{E}\mathbf{x}^{m-1} = \mathbf{x} \quad \text{for all} \quad \mathbf{x} \in \Sigma.$$

We restrict $\mathbf{x} \in \Sigma$ since it is not possible to have a tensor with $m > 2$ such that the above equation holds for all $\mathbf{x} \in \mathbb{R}^n$. For any $\mathbf{x} \notin \Sigma$, the above equation implies

$$\mathcal{E}\mathbf{x}^{m-1} = \|\mathbf{x}\|^{m-1}\mathcal{E}(\mathbf{x}/\|\mathbf{x}\|)^{m-1} = \|\mathbf{x}\|^{m-1}(\mathbf{x}/\|\mathbf{x}\|) = \|\mathbf{x}\|^{m-2}\mathbf{x}.$$

Consider the case of $m = 4$ and $n = 2$. The system of equations that must be satisfied for all $\mathbf{x} \in \Sigma$ is

$$\begin{aligned} e_{1111}x_1^3 + 3e_{1112}x_1^2x_2 + 3e_{1122}x_1x_2^2 + e_{1222}x_2^3 &= x_1,\\ e_{1112}x_1^3 + 3e_{1122}x_1^2x_2 + 3e_{1222}x_1x_2^2 + e_{2222}x_2^3 &= x_2. \end{aligned}$$

Consider $\mathbf{x} = \begin{bmatrix}1 & 0\end{bmatrix}^T$. This yields $e_{1111} = 1$ and $e_{1112} = 0$. Similarly, $\mathbf{x} = \begin{bmatrix}0 & 1\end{bmatrix}^T$ yields $e_{2222} = 1$ and $e_{1222} = 0$. The only remaining unknown is e_{1122}, and choosing, e.g., $\mathbf{x} = \begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}^T$ yields $e_{1122} = 1/3$. In summary, the identity tensor for $m = 4$ and $n = 2$ is

$$e_{ijkl} = \begin{cases} 1 & \text{if } i=j=k=l,\\ 1/3 & \text{if } i=j\neq k=l,\\ 1/3 & \text{if } i=k\neq j=l,\\ 1/3 & \text{if } i=l\neq j=k,\\ 0 & \text{otherwise.}\end{cases}$$

We generalize this idea in the next property. □

PROPERTY 2.4. *For* m *even, the identity tensor* $\mathcal{E} \in \mathbb{R}^{[m,n]}$ *satisfying* $\mathcal{E}\mathbf{x}^{m-1} = \mathbf{x}$ *for all* $\mathbf{x} \in \Sigma$ *is given by*

$$e_{i_1\cdots i_m} = \frac{1}{m!}\sum_{p\in\Pi_m} \delta_{i_{p(1)}i_{p(2)}}\delta_{i_{p(3)}i_{p(4)}}\cdots\delta_{i_{p(m-1)}i_{p(m)}} \tag{2.1}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

for $i_1, \ldots, i_m \in \{1, \ldots, n\}$, where δ is the standard Kronecker delta, i.e.,

$$\delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

This identity tensor appears in a previous work [18], where it is denoted by I_E and used to define a generalization of the characteristic polynomial for symmetric even-order tensors.

Example 2.5. There is no identity tensor for m odd. This is seen because if $\mathcal{E}\mathbf{x}^{m-1} = \mathbf{x}$ for some odd m and some $\mathbf{x} \in \Sigma$, then we would have $-\mathbf{x} \in \Sigma$ but $\mathcal{E}(-\mathbf{x})^{m-1} = \mathbf{x} \neq -\mathbf{x}$. $\square$

For any even-order tensor (i.e., m even), observe that if $(\lambda, \mathbf{x})$ is an eigenpair, then $(\lambda, -\mathbf{x})$ is also an eigenpair since

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

Likewise, for any odd-order tensor (i.e., m odd), $(-\lambda, -\mathbf{x})$ is also an eigenpair since

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

These are *not* considered to be distinct eigenpairs.

We later present, as Theorem 5.3, a recently derived result [3] that bounds the number of real eigenpairs by $((m-1)^n - 1)/(m-2)$. We defer discussion of this result until section 5, where we discuss complex eigenpairs.

Because the tensor eigenvalue equation for $m > 2$ amounts to a system of nonlinear equations in the components of $\mathbf{x}$, a direct solution is challenging. Numerical algorithms exist for finding all solutions of a system of polynomial equations, but become computationally expensive for systems with many variables (here, large n) and with high-order polynomials (here, large m). A polynomial system solver (`NSolve`) using a Gröbner basis method is available in *Mathematica* [28] and has been employed to generate a complete list of eigenpairs for some of the examples in this paper. The solver is instructed to find all solutions $(\lambda, \mathbf{x})$ of the system (1.2). Redundant solutions with the opposite sign of $\mathbf{x}$ (for even m) or the opposite signs of $\mathbf{x}$ and λ (for odd m) are then eliminated.

2.2. Convex functions. Convexity theory plays an important role in our analysis. Here we recall two important properties of convex functions [2].

PROPERTY 2.6 (gradient of convex function). *A differentiable function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex if and only if Ω is a convex set and $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \Omega$.*

PROPERTY 2.7 (Hessian of convex function). *A twice differentiable function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex if and only if Ω is a convex set and the Hessian[1] of f is positive semidefinite on Ω, i.e., $\nabla^2 f(\mathbf{x}) \succeq 0$ for all $\mathbf{x} \in \Omega$.*

We prove an interesting fact about convex functions on vectors of unit norm that will prove useful in our later analysis. This fact is implicit in a proof given previously [11, Theorem 4] and explicit in [23, Theorem 1].

THEOREM 2.8 (Kofidis and Regalia [11, 23]). *Let f be a function that is convex and continuously differentiable on Γ. Let $\mathbf{w} \in \Sigma$ with $\nabla f(\mathbf{w}) \neq \mathbf{0}$. If $\mathbf{v} = \nabla f(\mathbf{w})/\|\nabla f(\mathbf{w})\| \neq \mathbf{w}$, then $f(\mathbf{v}) - f(\mathbf{w}) > 0$.*

[1] By ∇^2 we denote the Hessian matrix and not its trace, the Laplacian.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof. For arbitrary nonzero $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z}^T\mathbf{x}$ is strictly maximized for $\mathbf{x} \in \Sigma$ by $\mathbf{x} = \mathbf{z}/\|\mathbf{z}\|$. Substituting $\mathbf{z} = \nabla f(\mathbf{w})$, it follows that $\nabla f(\mathbf{w})^T\mathbf{v} > \nabla f(\mathbf{w})^T\mathbf{w}$, since $\mathbf{v} = \nabla f(\mathbf{w})/\|\nabla f(\mathbf{w})\| \neq \mathbf{w}$ and $\mathbf{w} \in \Sigma$. By the convexity of f on Γ and Property 2.6, we have $f(\mathbf{v}) \geq f(\mathbf{w}) + \nabla f(\mathbf{w})^T(\mathbf{v}-\mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in \Gamma$. Consequently, $f(\mathbf{v}) - f(\mathbf{w}) \geq \nabla f(\mathbf{w})^T(\mathbf{v}-\mathbf{w}) > 0$. ∎

2.3. Constrained optimization. Here we extract relevant theory from constrained optimization [17].

THEOREM 2.9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable. A point $\mathbf{x}_* \in \Sigma$ is a (constrained) stationary point of*

$$\max f(\mathbf{x}) \quad \textit{subject to} \quad \mathbf{x} \in \Sigma$$

if there exists $\mu_* \in \mathbb{R}$ such that $\nabla f(\mathbf{x}_*) + \mu_*\mathbf{x}_* = \mathbf{0}$. The point $\mathbf{x}_*$ is a (constrained) isolated local maximum if, additionally,

$$\mathbf{w}^T(\nabla^2 f(\mathbf{x}_*) + \mu_*\mathbf{I})\mathbf{w} < 0 \quad \textit{for all} \quad \mathbf{w} \in \Sigma \cap \mathbf{x}_*^{\perp}.$$

Proof. The constraint $\mathbf{x} \in \Sigma$ can be expressed as $c(\mathbf{x}) = \frac{1}{2}(\mathbf{x}^T\mathbf{x} - 1) = 0$. The Lagrangian for the constrained problem is then given by

$$\mathcal{L}(\mathbf{x}, \mu) = f(\mathbf{x}) + \mu c(\mathbf{x}).$$

Its first and second derivatives with respect to $\mathbf{x}$ are

$$\nabla\mathcal{L}(\mathbf{x}, \mu) = \nabla f(\mathbf{x}) + \mu\mathbf{x} \quad \text{and} \quad \nabla^2\mathcal{L}(\mathbf{x}, \mu) = \nabla^2 f(\mathbf{x}) + \mu\mathbf{I}.$$

By assumption, $\nabla\mathcal{L}(\mathbf{x}_*, \mu_*) = \mathbf{0}$ and $c(\mathbf{x}_*) = 0$. Therefore, the pair $(\mathbf{x}_*, \mu_*)$ satisfies the KKT conditions [17, Theorem 12.1] and so is a constrained stationary point. It is additionally a constrained isolated local maximum if it meets the second-order sufficient condition [17, Theorem 12.6]. ∎

2.4. Fixed point theory. We consider the properties of iterations of the form

$$\mathbf{x}_{k+1} = \phi(\mathbf{x}_k).$$

Under certain conditions, the iterates are guaranteed to converge to a fixed point. In particular, we are interested in "attracting" fixed points.

DEFINITION 2.10 (fixed point). *A point $\mathbf{x}_* \in \mathbb{R}^n$ is a* fixed point *of $\phi : \mathbb{R}^n \to \mathbb{R}^n$ if $\phi(\mathbf{x}_*) = \mathbf{x}_*$. Further, $\mathbf{x}_*$ is an* attracting *fixed point if there exists $\delta > 0$ such that the sequence $\{\mathbf{x}_k\}$ defined by $\mathbf{x}_{k+1} = \phi(\mathbf{x}_k)$ converges to $\mathbf{x}_*$ for any $\mathbf{x}_0$ such that $\|\mathbf{x}_0 - \mathbf{x}_*\| \leq \delta$.*

THEOREM 2.11 (see [24, Theorem 2.8]). *Let $\mathbf{x}_* \in \mathbb{R}^n$ be a fixed point of $\phi : \mathbb{R}^n \to \mathbb{R}^n$, and let $J : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be the Jacobian of ϕ. Then $\mathbf{x}_*$ is an attracting fixed point if $\sigma \equiv \rho(J(\mathbf{x}_*)) < 1$; further, if $\sigma > 0$, then the convergence of $\mathbf{x}_{k+1} = \phi(\mathbf{x}_k)$ to $\mathbf{x}_*$ is linear with rate σ.*

This condition on the Jacobian for an attracting fixed point is sufficient but not necessary. In particular, if $\sigma \equiv \rho(J(\mathbf{x}_*)) = 1$, then $\mathbf{x}_*$ may or may not be attracting, but there is no neighborhood of *linear* convergence to it. For $\sigma < 1$, the rate of linear convergence depends on σ and is slower for σ values closer to 1. On the other hand, for $\sigma > 1$, an attractor is ruled out by the following.

THEOREM 2.12 (see [26, Theorem 1.3.7]). *Let $\mathbf{x}_* \in \mathbb{R}^n$ be a fixed point of $\phi : \mathbb{R}^n \to \mathbb{R}^n$, and let $J : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be the Jacobian of ϕ. Then $\mathbf{x}_*$ is an unstable fixed point if $\sigma \equiv \rho(J(\mathbf{x}_*)) > 1$.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

3. Symmetric higher-order power method (S-HOPM). We review the symmetric higher-order power method (S-HOPM), introduced by De Lathauwer, De Moor, and Vandewalle [6] and analyzed further by Kofidis and Regalia [11]. The purpose of S-HOPM is to solve the optimization problem

$$\max_{\mathbf{x}\in\mathbb{R}^n} |\mathcal{A}\mathbf{x}^m| \quad \text{subject to} \quad \mathbf{x}\in\Sigma. \tag{3.1}$$

The solution of this problem will be a solution of either the following maximization problem (lacking the absolute value) or its opposite minimization problem:

$$\max_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{x}\in\Sigma, \quad \text{where} \quad f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m. \tag{3.2}$$

Setting $\lambda = f(\mathbf{x})$, these problems are equivalent to finding the best symmetric rank-1 approximation of a symmetric tensor $\mathcal{A}\in\mathbb{R}^{[m,n]}$, i.e.,

$$\min_{\lambda,\mathbf{x}} \|\mathcal{A}-\mathcal{B}\| \quad \text{subject to} \quad b_{i_1\ldots i_m} = \lambda x_{i_1}\cdots x_{i_m} \quad \text{and} \quad \mathbf{x}\in\Sigma. \tag{3.3}$$

Details of the connection between (3.2) and (3.3) are available elsewhere [6]. The S-HOPM algorithm is shown in Algorithm 1. We discuss its connection to the eigenvalue problem in section 3.1 and its convergence properties in section 3.2.

Algorithm 1 Symmetric higher-order power method (S-HOPM) [6, 11].

Given a symmetric tensor $\mathcal{A}\in\mathbb{R}^{[m,n]}$.

Require: $\mathbf{x}_0\in\mathbb{R}^n$ with $\|\mathbf{x}_0\|=1$. Let $\lambda_0 = \mathcal{A}\mathbf{x}_0^m$.

1: **for** $k = 0, 1, \ldots$ **do**
2: $\quad\hat{\mathbf{x}}_{k+1} \leftarrow \mathcal{A}\mathbf{x}_k^{m-1}$
3: $\quad\mathbf{x}_{k+1} \leftarrow \hat{\mathbf{x}}_{k+1}/\|\hat{\mathbf{x}}_{k+1}\|$
4: $\quad\lambda_{k+1} \leftarrow \mathcal{A}\mathbf{x}_{k+1}^m$
5: **end for**

3.1. Properties of $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$. The function $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$ plays an important role in the analysis of eigenpairs of $\mathcal{A}$ because all eigenpairs are constrained stationary points of f, as we show below.

We first need to derive the gradient of f. This result is perhaps generally well known [13, equation 4], but here we provide a proof.

LEMMA 3.1. *Let $\mathcal{A}\in\mathbb{R}^{[m,n]}$ be symmetric. The gradient of $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$ is*

$$g(\mathbf{x}) \equiv \nabla f(\mathbf{x}) = m\,\mathcal{A}\mathbf{x}^{m-1} \in \mathbb{R}^n. \tag{3.4}$$

Proof. We use the basic relation $\nabla_k x_j = \delta_{jk}$. Applying the product rule to (3.2), we find

$$\nabla_k f(\mathbf{x}) = \sum_{i_1,\ldots,i_m} \sum_{q=1}^{m} a_{i_1 i_2\cdots i_m} x_{i_1} x_{i_2}\cdots x_{i_{q-1}} \delta_{i_q k} x_{i_{q+1}}\cdots x_{i_m}.$$

Upon bringing the sum over q to the outside, we observe that for each q the dummy indices i_1 and i_q can be interchanged (without affecting the symmetric tensor $\mathcal{A}$), and

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

the result is independent of q:

$$\begin{aligned}\nabla_k f(\mathbf{x}) &= \sum_{q=1}^{m} \sum_{i_1,\ldots,i_m} a_{i_1 i_2 \cdots i_m} \delta_{i_1 k} x_{i_2} \cdots x_{i_{q-1}} x_{i_q} x_{i_{q+1}} \cdots x_{i_m} \\ &= \sum_{q=1}^{m} \sum_{i_2,\ldots,i_m} a_{k i_2 \cdots i_m} x_{i_2} \cdots x_{i_m} \\ &= m(\mathcal{A}\mathbf{x}^{m-1})_k.\end{aligned}$$

Hence, $\nabla f(\mathbf{x}) = m\,\mathcal{A}\mathbf{x}^{m-1}$. ∎

THEOREM 3.2. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric. Then $(\lambda, \mathbf{x})$ is an eigenpair of $\mathcal{A}$ if and only if $\mathbf{x}$ is a constrained stationary point of* (3.2).

Proof. By Theorem 2.9, any constrained stationary point $\mathbf{x}_*$ of (3.2) must satisfy $m\,\mathcal{A}\mathbf{x}_*^{m-1} + \mu_*\mathbf{x}_* = 0$ for some $\mu_* \in \mathbb{R}$. Thus, $\lambda_* = -\mu_*/m$ is the eigenvalue corresponding to $\mathbf{x}_*$. Conversely, any eigenpair meets the condition for being a constrained stationary point with $\mu_* = -m\lambda_*$. ∎

This is the connection between (3.2) and the eigenvalue problem. It will also be useful to consider the Hessian of f, which we present here.

LEMMA 3.3. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric. The Hessian of $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$ is*

$$H(\mathbf{x}) \equiv \nabla^2 f(\mathbf{x}) = m(m-1)\mathcal{A}\mathbf{x}^{m-2} \in \mathbb{R}^{n\times n}. \tag{3.5}$$

Proof. The (j,k) entry of $H(\mathbf{x})$ is given by the kth entry of $\nabla g_j(\mathbf{x})$. The function $g_j(\mathbf{x})$ can be rewritten as

$$g_j(\mathbf{x}) = m \sum_{i_2,\ldots,i_m} a_{j i_2 \cdots i_m} x_{i_2} \cdots x_{i_m} = m\,\mathcal{B}^{(j)}\mathbf{x}^{m-1},$$

where $\mathcal{B}^{(j)}$ is the order-$(m-1)$ symmetric tensor that is the jth subtensor of $\mathcal{A}$, defined by $b^{(j)}_{i_1\cdots i_{m-1}} = a_{j i_1 \cdots i_{m-1}}$. From Lemma 3.1, we have

$$\nabla g_j(\mathbf{x}) = m(m-1)\mathcal{B}^{(j)}\mathbf{x}^{m-2}.$$

Consequently,

$$(H(\mathbf{x}))_{jk} = m(m-1) \sum_{i_3,\ldots,i_m} a_{j k i_3 \cdots i_m} x_{i_3} \cdots x_{i_m},$$

that is, $H(\mathbf{x}) = m(m-1)\mathcal{A}\mathbf{x}^{m-2}$. ∎

From Theorem 2.9, we know that the projected Hessian of the Lagrangian plays a role in determining whether or not a fixed point is a local maximum or minimum. In our case, since $\mu_* = -m\lambda_*$, for any eigenpair $(\lambda_*, \mathbf{x}_*)$ (which must correspond to a constrained stationary point by Theorem 3.2) we have

$$\nabla^2 \mathcal{L}(\mathbf{x}_*, \lambda_*) = m(m-1)\mathcal{A}\mathbf{x}_*^{m-2} - m\lambda_*\mathbf{I}.$$

Specifically, Theorem 2.9 is concerned with the behavior of the Hessian of the Lagrangian in the subspace orthogonal to $\mathbf{x}_*$. Thus, we define the projected Hessian of the Lagrangian as

$$C(\lambda_*, \mathbf{x}_*) \equiv \mathbf{U}_*^T \left((m-1)\mathcal{A}\mathbf{x}_*^{m-2} - \lambda_*\mathbf{I}\right)\mathbf{U}_* \in \mathbb{R}^{(n-1)\times(n-1)}, \tag{3.6}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

where the columns of $\mathbf{U}_* \in \mathbb{R}^{n\times(n-1)}$ form an orthonormal basis for $\mathbf{x}_*^{\perp}$. Note that we have removed a factor of m for convenience. We now classify eigenpairs according to the spectrum of $C(\lambda_*, \mathbf{x}_*)$. The import of this classification will be made clear in section 4.2.

DEFINITION 3.4. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be a symmetric tensor. We say an eigenpair $(\lambda, \mathbf{x})$ of $\mathcal{A} \in \mathbb{R}^{[m,n]}$ is* positive stable *if $C(\lambda, \mathbf{x})$ is positive definite,* negative stable *if $C(\lambda, \mathbf{x})$ is negative definite, and* unstable *if $C(\lambda, \mathbf{x})$ is indefinite.*

These labels are not exhaustive because we do not name the cases where $C(\lambda, \mathbf{x})$ is only semidefinite, with a zero eigenvalue. Such cases do not occur for generic tensors.

If m is odd, then $(\lambda, \mathbf{x})$ is positive stable if and only if $(-\lambda, -\mathbf{x})$ is negative stable, even though these eigenpairs are in the same equivalence class. On the other hand, if m is even, then $(\lambda, \mathbf{x})$ is a positive (negative) stable eigenpair if and only if $(\lambda, -\mathbf{x})$ is also positive (negative) stable.

3.2. S-HOPM convergence analysis. S-HOPM has been deemed unreliable [6] because convergence is not guaranteed. Kofidis and Regalia [11] provide an analysis explaining that S-HOPM will converge if certain conditions are met, as well as an example where the method does not converge, which we reproduce here.

Example 3.5 (Kofidis and Regalia [11, Example 1]). Let $\mathcal{A} \in \mathbb{R}^{[4,3]}$ be the symmetric tensor defined by

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

Kofidis and Regalia [11] observed that Algorithm 1 does not converge for this tensor. Because this problem is small, all eigenpairs can be calculated by *Mathematica* as described in section 2.1. From Theorem 5.3, this problem has at most 13 eigenpairs; we list the 11 real eigenpairs in Table 3.1. We ran 100 trials of S-HOPM using different random starting points $\mathbf{x}_0$ chosen from a uniform distribution on $[-1, 1]^n$. For these experiments, we allow up to 1000 iterations and say that the algorithm has converged if $|\lambda_{k+1} - \lambda_k| < 10^{-16}$. In every single trial for this tensor, the algorithm failed to converge. In Figure 3.1, we show an example $\{\lambda_k\}$ sequence with $\mathbf{x}_0 = \begin{bmatrix}-0.2695 & 0.1972 & 0.3370\end{bmatrix}^T$. This coincides with the results reported previously [11]. □

TABLE 3.1
Eigenpairs for $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5.

λ	$\mathbf{x}^T$	Eigenvalues of $C(\lambda, \mathbf{x})$	Type
0.8893	[0.6672 0.2471 −0.7027]	{ −0.8857, −1.8459 }	Neg. stable
0.8169	[0.8412 −0.2635 0.4722]	{ −0.9024, −2.2580 }	Neg. stable
0.5105	[0.3598 −0.7780 0.5150]	{ 0.5940, −2.3398 }	Unstable
0.3633	[0.2676 0.6447 0.7160]	{ −1.1765, −0.5713 }	Neg. stable
0.2682	[0.6099 0.4362 0.6616]	{ 0.7852, −1.1793 }	Unstable
0.2628	[0.1318 −0.4425 −0.8870]	{ 0.6181, −2.1744 }	Unstable
0.2433	[0.9895 0.0947 −0.1088]	{ −1.1942, 1.4627 }	Unstable
0.1735	[0.3357 0.9073 0.2531]	{ −1.0966, 0.8629 }	Unstable
−0.0451	[0.7797 0.6135 0.1250]	{ 0.8209, 1.2456 }	Pos. stable
−0.5629	[0.1762 −0.1796 0.9678]	{ 1.6287, 2.3822 }	Pos. stable
−1.0954	[0.5915 −0.7467 −0.3043]	{ 1.8628, 2.7469 }	Pos. stable

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

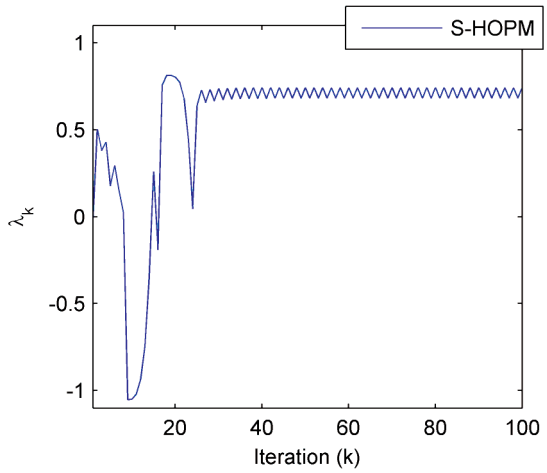


FIG. 3.1. *Example λ_k values for S-HOPM on $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example* 3.5.

Example 3.6. As a second illustrative example, we consider an odd-order tensor $\mathcal{A} \in \mathbb{R}^{[3,3]}$ defined by

$$\begin{aligned} a_{111} &= -0.1281, & a_{112} &= 0.0516, & a_{113} &= -0.0954, & a_{122} &= -0.1958, \\ a_{123} &= -0.1790, & a_{133} &= -0.2676, & a_{222} &= 0.3251, & a_{223} &= 0.2513, \\ a_{233} &= 0.1773, & a_{333} &= 0.0338. \end{aligned}$$

From Theorem 5.3, $\mathcal{A}$ has at most 7 eigenpairs; in this case we achieve that bound and the eigenpairs are listed in Table 3.2. We ran 100 trials of S-HOPM as described for Example 3.5. Every trial converged to either $\lambda = 0.8730$ or $\lambda = 0.4306$, as summarized in Table 3.3. Therefore, S-HOPM finds 2 of the 7 possible eigenvalues. □

TABLE 3.2
Eigenpairs for $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6.

λ	$\mathbf{x}^T$	Eigenvalues of $C(\lambda, \mathbf{x})$	Type
0.8730	[−0.3922 0.7249 0.5664]	{ −1.1293, −0.8807 }	Neg. stable
0.4306	[−0.7187 −0.1245 −0.6840]	{ −0.4420, −0.8275 }	Neg. stable
0.2294	[−0.8446 0.4386 −0.3070]	{ −0.2641, 0.7151 }	Unstable
0.0180	[0.7132 0.5093 −0.4817]	{ −0.4021, −0.1320 }	Neg. stable
0.0033	[0.4477 0.7740 −0.4478]	{ −0.1011, 0.2461 }	Unstable
0.0018	[0.3305 0.6314 −0.7015]	{ 0.1592, −0.1241 }	Unstable
0.0006	[0.2907 0.7359 −0.6115]	{ 0.1405, 0.0968 }	Pos. stable

TABLE 3.3
Eigenpairs for $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 *computed by S-HOPM with* 100 *random starts.*

# Occurrences	λ	$\mathbf{x}$	Median its.
62	0.8730	[−0.3922 0.7249 0.5664]	19
38	0.4306	[−0.7187 −0.1245 −0.6840]	184

In their analysis, Kofidis and Regalia [11] proved that the sequence $\{\lambda_k\}$ in Algorithm 1 converges if $\mathcal{A} \in \mathbb{R}^{[m,n]}$ is even-order and the function $f(\mathbf{x})$ is convex or

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

concave on $\mathbb{R}^n$. Since $m = 2\ell$ (because m is even), f can be expressed as

$$f(\mathbf{x}) = (\underbrace{\mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{\ell \text{ times}})^T \mathbf{A} (\underbrace{\mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{\ell \text{ times}}),$$

where $\mathbf{A} \in \mathbb{R}^{n^\ell \times n^\ell}$ is an unfolded version of the tensor $\mathcal{A}$.[2] Since $\mathcal{A}$ is symmetric, it follows that $\mathbf{A}$ is symmetric. The condition that f is convex (concave) is satisfied if the Hessian

$$\nabla^2 f(\mathbf{x}) = m(m-1)(\mathbf{I} \otimes \underbrace{\mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{\ell - 1 \text{ times}})^T \mathbf{A} (\mathbf{I} \otimes \underbrace{\mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{\ell - 1 \text{ times}})$$

is positive (negative) semidefinite for all $\mathbf{x} \in \mathbb{R}^n$.

We make a few notes regarding these results. First, even though f is convex, its restriction to the nonconvex set Σ is not. Second, $\{\lambda_k\}$ is increasing if f is convex and decreasing if f is concave. Third, only $\{\lambda_k\}$ is proved to converge for S-HOPM [11, Theorem 4]; the iterates $\{\mathbf{x}_k\}$ may not. In particular, it is easy to observe that the sign of $\mathbf{x}_k$ may flip back and forth if the concave case is not handled correctly.

4. Shifted symmetric higher-order power method (SS-HOPM). In this section, we show that S-HOPM can be modified by adding a "shift" that guarantees that the method will always converge to an eigenpair. In the context of ICA, this idea has also been proposed by Regalia and Kofidis [23] and Erdogan [8]. Based on the observation that S-HOPM is guaranteed to converge if the underlying function is convex or concave on $\mathbb{R}^n$, our method works with a suitably modified function

$$\hat{f}(\mathbf{x}) \equiv f(\mathbf{x}) + \alpha(\mathbf{x}^T\mathbf{x})^{m/2}. \tag{4.1}$$

Maximizing $\hat{f}$ on Σ is the same as maximizing f plus a constant, yet the properties of the modified function force convexity or concavity and consequently guarantee convergence to a KKT point (not necessarily the *global* maximum or minimum). Note that previous papers [23, 8] have proposed similar shifted functions that are essentially of the form $\hat{f}(\mathbf{x}) \equiv f(\mathbf{x}) + \alpha\mathbf{x}^T\mathbf{x}$, differing only in the exponent.

An advantage of our choice of $\hat{f}$ in (4.1) is that, for even m, it can be interpreted as

$$\hat{f}(\mathbf{x}) = \hat{\mathcal{A}}\mathbf{x}^m \equiv (\mathcal{A} + \alpha\mathcal{E})\mathbf{x}^m,$$

where $\mathcal{E}$ is the identity tensor as defined in (2.1). Thus, for even m, our proposed method can be interpreted as S-HOPM applied to a modified tensor that directly satisfies the convexity properties to guarantee convergence [11]. Because $\mathcal{E}\mathbf{x}^{m-1} = \mathbf{x}$ for $\mathbf{x} \in \Sigma$, the eigenvectors of $\hat{\mathcal{A}}$ are the same as those of $\mathcal{A}$ and the eigenvalues are shifted by α. Our results, however, are for both odd- and even-order tensors.

Algorithm 2 presents the shifted symmetric higher-order power method (SS-HOPM). Without loss of generality, we assume that a positive shift ($\alpha \geq 0$) is used to make the modified function in (4.1) convex and a negative shift ($\alpha < 0$) to make it concave. We have two key results. Theorem 4.4 shows that for any starting point

[2] Specifically, $\mathbf{A} \equiv \mathbf{A}_{(\mathcal{R}\times\mathcal{C})}$ with $\mathcal{R} = \{1, \ldots, \ell\}$ and $\mathcal{C} = \{\ell+1, \ldots, m\}$ in matricization notation [12].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Algorithm 2 Shifted symmetric higher-order power method (SS-HOPM).

Given a tensor $\mathcal{A} \in \mathbb{R}^{[m,n]}$.

Require: $\mathbf{x}_0 \in \mathbb{R}^n$ with $\|\mathbf{x}_0\| = 1$. Let $\lambda_0 = \mathcal{A}\mathbf{x}_0^m$.
Require: $\alpha \in \mathbb{R}$
1: **for** $k = 0, 1, \ldots$ **do**
2: **if** $\alpha \geq 0$ **then**
3: $\hat{\mathbf{x}}_{k+1} \leftarrow \mathcal{A}\mathbf{x}_k^{m-1} + \alpha \mathbf{x}_k$ ▷ Assumed Convex
4: **else**
5: $\hat{\mathbf{x}}_{k+1} \leftarrow -(\mathcal{A}\mathbf{x}_k^{m-1} + \alpha \mathbf{x}_k)$ ▷ Assumed Concave
6: **end if**
7: $\mathbf{x}_{k+1} \leftarrow \hat{\mathbf{x}}_{k+1}/\|\hat{\mathbf{x}}_{k+1}\|$
8: $\lambda_{k+1} \leftarrow \mathcal{A}\mathbf{x}_{k+1}^m$
9: **end for**

$\mathbf{x}_0 \in \Sigma$, the sequence $\{\lambda_k\}$ produced by Algorithm 2 is guaranteed to converge to an eigenvalue in the convex case if

$$\alpha > \beta(\mathcal{A}) \equiv (m-1) \cdot \max_{\mathbf{x} \in \Sigma} \rho(\mathcal{A}\mathbf{x}^{m-2}). \tag{4.2}$$

Corollary 4.6 handles the concave case where we require $\alpha < -\beta(\mathcal{A})$. Theorem 4.8 further shows that Algorithm 2 in the convex case will generically converge to an eigenpair $(\lambda, \mathbf{x})$ that is negative stable. Corollary 4.9 proves that Algorithm 2 in the concave case will generically converge to an eigenpair that is positive stable. Generally, neither version will converge to an eigenpair that is unstable.

4.1. SS-HOPM convergence analysis. We first establish a few key lemmas that guide the choice of the shift $\alpha > \beta(\mathcal{A})$ in SS-HOPM.

LEMMA 4.1. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric and let $\beta(\mathcal{A})$ be as defined in* (4.2). *Then $\beta(\mathcal{A}) \leq (m-1)\sum_{i_1,\ldots,i_m} |a_{i_1 \ldots i_m}|$.*

Proof. For all $\mathbf{x}, \mathbf{y} \in \Sigma$, we obtain $|\mathbf{y}^T(\mathcal{A}\mathbf{x}^{m-2})\mathbf{y}| \leq \sum_{i_1,\ldots,i_m} |a_{i_1 \ldots i_m}|$ by applying the triangle inequality to the sum of n^m terms. Thus $\rho(\mathcal{A}\mathbf{x}^{m-2}) \leq \sum_{i_1,\ldots,i_m} |a_{i_1 \ldots i_m}|$ for all $\mathbf{x} \in \Sigma$, and the result follows. ∎

LEMMA 4.2. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric, let $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$, and let $\beta(\mathcal{A})$ be as defined in* (4.2). *Then $|f(\mathbf{x})| \leq \beta(\mathcal{A})/(m-1)$ for all $\mathbf{x} \in \Sigma$.*

Proof. We have $|\mathcal{A}\mathbf{x}^m| = |\mathbf{x}^T(\mathcal{A}\mathbf{x}^{m-2})\mathbf{x}| \leq \rho(\mathcal{A}\mathbf{x}^{m-2}) \leq \beta(\mathcal{A})/(m-1)$. ∎

The preceding lemma upper bounds the magnitude of any eigenvalue of $\mathcal{A}$ by $\beta(\mathcal{A})/(m-1)$ since any eigenpair $(\lambda, \mathbf{x})$ satisfies $\lambda = f(\mathbf{x})$. Thus, choosing $\alpha > \beta(\mathcal{A})$ implies that α is greater than the magnitude of any eigenvalue of $\mathcal{A}$.

LEMMA 4.3. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric and let $H(\mathbf{x})$ and $\beta(\mathcal{A})$ be as defined in* (3.5) *and* (4.2). *Then $\rho(H(\mathbf{x})) \leq m\beta(\mathcal{A})$ for all $\mathbf{x} \in \Sigma$.*

Proof. This follows directly from (3.5) and (4.2). ∎

The following theorem proves that Algorithm 2 will always converge. Choosing $\alpha > (m-1)\sum_{i_1,\ldots,i_m} |a_{i_1 \ldots i_m}|$ is a conservative choice that is guaranteed to work by Lemma 4.1, but this may slow down convergence considerably, as we show in subsequent analysis and examples.

THEOREM 4.4. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric. For $\alpha > \beta(\mathcal{A})$, where $\beta(\mathcal{A})$ is defined in* (4.2), *the iterates $\{\lambda_k, \mathbf{x}_k\}$ produced by Algorithm* 2 *satisfy the following properties.* (a) *The sequence $\{\lambda_k\}$ is nondecreasing, and there exists λ_* such that $\lambda_k \to \lambda_*$.* (b) *The sequence $\{\mathbf{x}_k\}$ has an accumulation point.* (c) *For every such*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

accumulation point $\mathbf{x}_*$*, the pair* $(\lambda_*, \mathbf{x}_*)$ *is an eigenpair of* $\mathcal{A}$*.* (d) *If* $\mathcal{A}$ *has finitely many real eigenvectors, then there exists* $\mathbf{x}_*$ *such that* $\mathbf{x}_k \to \mathbf{x}_*$*.*

Proof. Our analysis depends on the modified function $\hat{f}$ defined in (4.1). Its gradient and Hessian for $\mathbf{x} \neq \mathbf{0}$ are

$$\hat{g}(\mathbf{x}) \equiv \nabla \hat{f}(\mathbf{x}) = g(\mathbf{x}) + m\alpha(\mathbf{x}^T\mathbf{x})^{m/2-1}\mathbf{x}, \tag{4.3}$$

$$\hat{H}(\mathbf{x}) \equiv \nabla^2 \hat{f}(\mathbf{x}) = H(\mathbf{x}) + m\alpha(\mathbf{x}^T\mathbf{x})^{m/2-1}\mathbf{I} + m(m-2)\alpha(\mathbf{x}^T\mathbf{x})^{m/2-2}\mathbf{x}\mathbf{x}^T, \tag{4.4}$$

where g and H are the gradient and Hessian of f from Lemma 3.1 and Lemma 3.3, respectively. And because $\hat{f}(\mathbf{x}) = O(\|\mathbf{x}\|^m)$, as $\mathbf{x} \to \mathbf{0}$, it follows that $\hat{f}(\mathbf{x})$ is of third or higher order in $\mathbf{x}$ for $m \geq 3$; thus $\hat{g}(\mathbf{0}) = \mathbf{0}$ and $\hat{H}(\mathbf{0}) = \mathbf{0}$.

Because it is important for the entire proof, we first show that $\hat{f}$ is convex on $\mathbb{R}^n$ for $\alpha > \beta(\mathcal{A})$. As noted, if $\mathbf{x} = \mathbf{0}$, we have $\hat{H}(\mathbf{x}) = \mathbf{0}$ for $m \geq 3$. Consider nonzero $\mathbf{x} \in \mathbb{R}^n$ and define $\bar{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\| \in \Sigma$; then $\hat{H}(\mathbf{x})$ is positive semidefinite (in fact, positive definite) by Lemma 4.3 since

$$\begin{aligned}\mathbf{y}^T\hat{H}(\mathbf{x})\mathbf{y} &= \|\mathbf{x}\|^{m-2}\left(\mathbf{y}^T H(\bar{\mathbf{x}})\mathbf{y} + m\alpha + m(m-1)\alpha(\bar{\mathbf{x}}^T\mathbf{y})^2\right) \\ &\geq \|\mathbf{x}\|^{m-2}\left(-m\beta(\mathcal{A}) + m\alpha + 0\right) > 0 \quad \text{for all} \quad \mathbf{y} \in \Sigma.\end{aligned}$$

By Property 2.7, $\hat{f}$ is convex on $\mathbb{R}^n$ because its Hessian is positive semidefinite.

We also note that $-\alpha$ must be an eigenvalue of $\mathcal{A}$ if $\hat{g}(\mathbf{x}) = \mathbf{0}$ for some $\mathbf{x} \in \Sigma$, since

$$\hat{g}(\mathbf{x}) = \mathbf{0} \quad \text{implies} \quad \mathcal{A}\mathbf{x}^{m-1} + \alpha\mathbf{x} = \mathbf{0}.$$

By Lemma 4.2, choosing $\alpha > \beta(\mathcal{A})$ ensures that α is greater than the magnitude of any eigenvalue, and so $\hat{g}(\mathbf{x}) \neq \mathbf{0}$ for all $\mathbf{x} \in \Sigma$. This ensures that the update in Algorithm 2, which reduces to

$$\mathbf{x}_{k+1} = \frac{\hat{g}(\mathbf{x}_k)}{\|\hat{g}(\mathbf{x}_k)\|} \tag{4.5}$$

in the convex case, is always well defined.

(a) Since $\hat{f}$ is convex on Γ and $\mathbf{x}_{k+1}, \mathbf{x}_k \in \Sigma$ and $\mathbf{x}_{k+1} = \nabla\hat{f}(\mathbf{x}_k)/\|\nabla\hat{f}(\mathbf{x}_k)\|$, Theorem 2.8 yields

$$\lambda_{k+1} - \lambda_k = \hat{f}(\mathbf{x}_{k+1}) - \hat{f}(\mathbf{x}_k) \geq 0,$$

where the nonstrict inequality covers the possibility that $\mathbf{x}_{k+1} = \mathbf{x}_k$. Thus, $\{\lambda_k\}$ is a nondecreasing sequence. By Lemma 4.2, $\lambda_k = f(\mathbf{x}_k)$ is bounded, so the sequence must converge to a limit point λ_*.[3]

(b) Since $\{\mathbf{x}_k\}$ is an infinite sequence on a compact set Σ, it must have an accumulation point $\mathbf{x}_* \in \Sigma$ by the Bolzano–Weierstrass theorem. Note also that continuity of f implies that $\lambda_* = \mathcal{A}\mathbf{x}_*^m$.

(c) By part (a) of the proof, convexity of $\hat{f}$, and Property 2.6, we have

$$\hat{f}(\mathbf{x}_{k+1}) - \hat{f}(\mathbf{x}_k) \to 0$$

[3]Note that the similar approach proposed for ICA [23, Theorem 2] allows the shift α to vary at each iteration as long as the underlying function remains convex.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

and thus

$$\hat{g}(\mathbf{x}_k)^T(\mathbf{x}_{k+1} - \mathbf{x}_k) \to 0.$$

Using (4.5), we can rewrite the above formula as

$$\|\hat{g}(\mathbf{x}_k)\| - \hat{g}(\mathbf{x}_k)^T\mathbf{x}_k \to 0. \tag{4.6}$$

By continuity of $\hat{g}$, an accumulation point $\mathbf{x}_*$ must satisfy

$$\|\hat{g}(\mathbf{x}_*)\| - \hat{g}(\mathbf{x}_*)^T\mathbf{x}_* = 0, \tag{4.7}$$

which implies

$$\|\hat{g}(\mathbf{x}_*)\| = \hat{g}(\mathbf{x}_*)^T\mathbf{x}_* = (m\,\mathcal{A}\mathbf{x}_*^{m-1} + m\alpha\mathbf{x}_*)^T\mathbf{x}_* = m(\lambda_* + \alpha).$$

Because $\mathbf{x}_* \in \Sigma$, (4.7) can hold only if

$$\mathbf{x}_* = \frac{\hat{g}(\mathbf{x}_*)}{\|\hat{g}(\mathbf{x}_*)\|} = \frac{m\,\mathcal{A}\mathbf{x}_*^{m-1} + m\alpha\mathbf{x}_*}{m(\lambda_* + \alpha)},$$

that is,

$$\mathcal{A}\mathbf{x}_*^{m-1} = \lambda_*\mathbf{x}_*.$$

Hence $(\lambda_*, \mathbf{x}_*)$ is an eigenpair of $\mathcal{A}$.

(d) Equation (4.6) gives

$$\|\hat{g}(\mathbf{x}_k)\|(1 - \mathbf{x}_{k+1}^T\mathbf{x}_k) \to 0.$$

Because $\|\hat{g}(\mathbf{x}_k)\|$ is bounded away from 0 and because $\mathbf{x}_k, \mathbf{x}_{k+1} \in \Sigma$, this requires that

$$\|\mathbf{x}_k - \mathbf{x}_{k+1}\| \to 0. \tag{4.8}$$

Recall that every accumulation point of $\{\mathbf{x}_k\}$ must be a (real) eigenvector of $\mathcal{A}$. If these eigenvectors are finite in number and thus isolated, consider removing an arbitrarily small open neighborhood of each from Σ, leaving a closed and thus compact space $Y \subset \Sigma$ containing no accumulation points of $\{\mathbf{x}_k\}$. If $\{\mathbf{x}_k\}$ had infinitely many iterates in Y, it would have an accumulation point in Y by the Bolzano–Weierstrass theorem, creating a contradiction. Therefore at most finitely many iterates are in Y, and $\{\mathbf{x}_k\}$ is ultimately confined to arbitrarily small neighborhoods of the eigenvectors. By (4.8), however, $\|\mathbf{x}_k - \mathbf{x}_{k+1}\|$ eventually remains smaller than the minimum distance between any two of these neighborhoods. Consequently, the iteration ultimately cannot jump from one neighborhood to another, and so in the limit $\{\mathbf{x}_k\}$ is confined to an arbitrarily small neighborhood of a *single* eigenvector $\mathbf{x}_*$, to which it therefore converges. Hence, the proof is complete. ∎

Note that the condition of finitely many real eigenvectors in part (d) of Theorem 4.4 holds for generic tensors. We conjecture that the convergence of $\{\mathbf{x}_k\}$ is guaranteed even without this condition.

Example 4.5. Again consider $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5. We show results using a shift of $\alpha = 2$. We ran 100 trials of SS-HOPM using the experimental conditions described in Example 3.5. We found 3 real eigenpairs; the results are summarized in Table 4.1(a). Three example runs (one for each eigenvalue) are shown in Figure 4.1(a).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

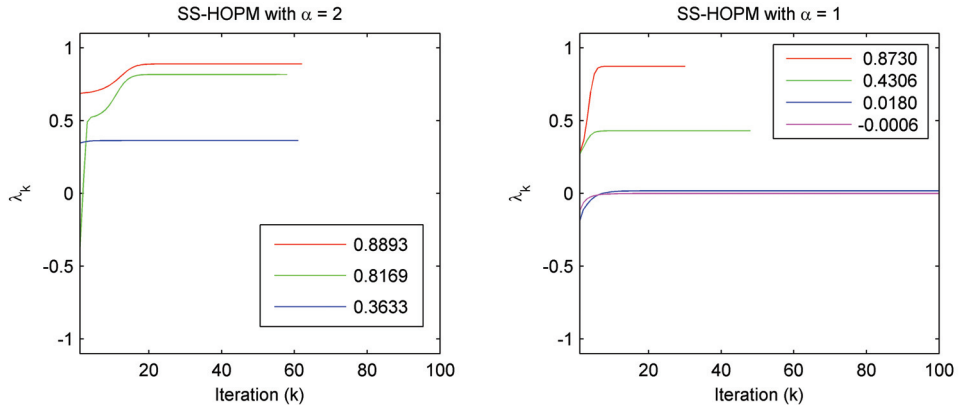
TABLE 4.1
Eigenpairs computed by SS-HOPM (convex) with 100 *random starts.*

(a) $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5 with $\alpha = 2$.

# Occurrences	λ	$\mathbf{x}$	Median its.
46	0.8893	[0.6672 0.2471 −0.7027]	63
24	0.8169	[0.8412 −0.2635 0.4722]	52
30	0.3633	[0.2676 0.6447 0.7160]	65

(b) $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 with $\alpha = 1$.

# Occurrences	λ	$\mathbf{x}$	Median its.
40	0.8730	[−0.3922 0.7249 0.5664]	32
29	0.4306	[−0.7187 −0.1245 −0.6840]	48
18	0.0180	[0.7132 0.5093 −0.4817]	116
13	−0.0006	[−0.2907 −0.7359 0.6115]	145



(a) $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5 with $\alpha = 2$. (b) $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 with $\alpha = 1$.

FIG. 4.1. *Example λ_k values for SS-HOPM (convex). One sequence is shown for each distinct eigenvalue.*

We also considered the "conservative" choice of $\alpha = (m-1)\sum_{i_1,\ldots,i_m} |a_{i_1\ldots i_m}| = 55.6620$. We ran 100 trials of SS-HOPM using the experimental conditions described in Example 3.5, except that we increased the maximum number of iterations to 10,000. Every trial converged to one of the same 3 real eigenpairs, but the number of iterations was around 1000 (versus around 60 for $\alpha = 2$); in section 4.2, we see that the rate of convergence asymptotically decreases as α increases.

Analogous results are shown for $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 with a shift of $\alpha = 1$ in Table 4.1(b) and Figure 4.1(b). Here SS-HOPM finds 2 additional eigenpairs compared to S-HOPM. In this case, we also considered $\alpha = (m-1)\sum_{i_1,\ldots,i_m} |a_{i_1\ldots i_m}| = 9.3560$, but this again increased the number of iterations up to a factor of ten.

For both tensors, $\{\lambda_k\}$ is always a nondecreasing sequence. Observe further that SS-HOPM converges only to eigenpairs that are negative stable. □

Using a large enough negative value of α makes $\hat{f}$ concave. It was observed [11] that $f(\mathbf{x}) = f(-\mathbf{x})$ for even-order tensors, and so the sequence $\{\lambda_k\}$ converges regardless of correctly handling the minus sign. The only minor problem in the concave case is that the sequence of iterates $\{\mathbf{x}_k\}$ does not converge. This is easily fixed, however, by correctly handling the sign as we do in Algorithm 2. The corresponding

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

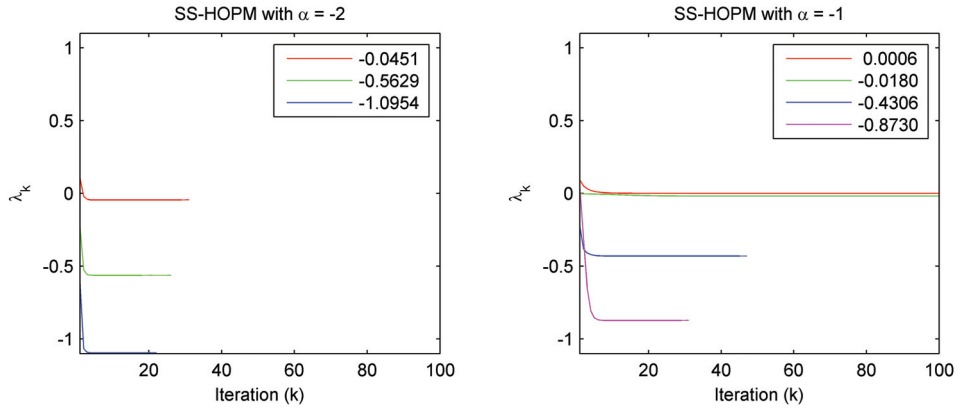
TABLE 4.2
Eigenpairs computed by SS-HOPM (concave) with 100 *random starts.*

(a) $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5 with $\alpha = -2$.

# Occurrences	λ	$\mathbf{x}$	Median its.
15	−0.0451	[−0.7797 −0.6135 −0.1250]	35
40	−0.5629	[−0.1762 0.1796 −0.9678]	23
45	−1.0954	[−0.5915 0.7467 0.3043]	23

(b) $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 with $\alpha = -1$.

# Occurrences	λ	$\mathbf{x}$	Median its.
19	0.0006	[0.2907 0.7359 −0.6115]	146
18	−0.0180	[−0.7132 −0.5093 0.4817]	117
29	−0.4306	[0.7187 0.1245 0.6840]	49
34	−0.8730	[0.3922 −0.7249 −0.5664]	33



(a) $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5 with $\alpha = 2$. (b) $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 with $\alpha = 1$.

FIG. 4.2. *Example λ_k values for SS-HOPM (concave). One sequence is shown for each distinct eigenvalue.*

theory for the concave case is presented in Corollary 4.6. In this case we choose α to be negative, i.e., the theory suggests $\alpha < -\beta(\mathcal{A})$.

COROLLARY 4.6. *Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric. For $\alpha < -\beta(\mathcal{A})$, where $\beta(\mathcal{A})$ is defined in* (4.2)*, the iterates $\{\lambda_k, \mathbf{x}_k\}$ produced by Algorithm* 2 *satisfy the following properties.* (a) *The sequence $\{\lambda_k\}$ is nonincreasing, and there exists λ_* such that $\lambda_k \to \lambda_*$.* (b) *The sequence $\{\mathbf{x}_k\}$ has an accumulation point.* (c) *For any such accumulation point $\mathbf{x}_*$, the pair $(\lambda_*, \mathbf{x}_*)$ is an eigenpair of $\mathcal{A}$.* (d) *If the eigenvalues of $\mathcal{A}$ are isolated, then $\mathbf{x}_k \to \mathbf{x}_*$.*

Proof. Apply the proof of Theorem 4.4 with $f(\mathbf{x}) = -\mathcal{A}\mathbf{x}^m$. ∎

Example 4.7. Revisiting $\mathcal{A} \in \mathbb{R}^{[4,3]}$ in Example 3.5 again, we run another 100 trials using $\alpha = -2$. We find 3 (new) real eigenpairs; the results are summarized in Table 4.2(a). Three example runs (one for each eigenvalue) are shown in Figure 4.2(a).

We also revisit $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 and use $\alpha = -1$. In this case, we find the opposites, i.e., $(-\lambda, -\mathbf{x})$, of the eigenpairs found with $\alpha = 1$, as shown in Table 4.2(b). This is to be expected for odd-order tensors since there is symmetry, i.e., $f(\mathbf{x}) = -f(-\mathbf{x})$, $C(\lambda, \mathbf{x}) = -C(-\lambda, -\mathbf{x})$, etc. Observe that the median number of iterations is nearly unchanged; this is explained in the following subsection, where

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

we discuss the rate of convergence. Four example runs (one per eigenvalue) are shown in Figure 4.2(b).

The sequence $\{\lambda_k\}$ is nonincreasing in every case. Each of the eigenpairs found in the concave case is positive stable. $\square$

4.2. SS-HOPM fixed point analysis. In this section, we show that fixed point analysis allows us to easily characterize convergence to eigenpairs according to whether they are positive stable, negative stable, or unstable. The convex version of SS-HOPM will generically converge to eigenpairs that are negative stable; the concave version will generically converge to eigenpairs that are positive stable.

To justify these conclusions, we consider Algorithm 2 in the convex case as a fixed point iteration $\mathbf{x}_{k+1} = \phi(\mathbf{x}_k; \alpha)$, where ϕ is defined as

$$\phi(\mathbf{x};\alpha) = \phi_1(\phi_2(\mathbf{x};\alpha)) \text{ with } \phi_1(\mathbf{x}) = \frac{\mathbf{x}}{(\mathbf{x}^T\mathbf{x})^{\frac{1}{2}}} \text{ and } \phi_2(\mathbf{x};\alpha) = \mathcal{A}\mathbf{x}^{m-1} + \alpha\mathbf{x}. \tag{4.9}$$

Note that an eigenpair $(\lambda, \mathbf{x})$ is a fixed point if and only if $\lambda + \alpha > 0$, which is always true for $\alpha > \beta(\mathcal{A})$.

From [9], the Jacobian of the operator ϕ is

$$J(\mathbf{x};\alpha) = \phi_1'(\phi_2(\mathbf{x};\alpha))\phi_2'(\mathbf{x};\alpha),$$

where derivatives are taken with respect to $\mathbf{x}$ and

$$\phi_1'(\mathbf{x}) = \frac{(\mathbf{x}^T\mathbf{x})\mathbf{I} - \mathbf{x}\mathbf{x}^T}{(\mathbf{x}^T\mathbf{x})^{\frac{3}{2}}} \quad \text{and} \quad \phi_2'(\mathbf{x};\alpha) = (m-1)\mathcal{A}\mathbf{x}^{m-2} + \alpha\mathbf{I}.$$

At any eigenpair $(\lambda, \mathbf{x})$, we have

$$\phi_2(\mathbf{x};\alpha) = (\lambda+\alpha)\mathbf{x}, \quad \phi_1'(\phi_2(\mathbf{x};\alpha)) = \frac{(\mathbf{I} - \mathbf{x}\mathbf{x}^T)}{\lambda+\alpha},$$
$$\text{and} \quad \phi_2'(\mathbf{x};\alpha) = (m-1)\mathcal{A}\mathbf{x}^{m-2} + \alpha\mathbf{I}.$$

Thus, the Jacobian at $\mathbf{x}$ is

$$J(\mathbf{x};\alpha) = \frac{(m-1)(\mathcal{A}\mathbf{x}^{m-2} - \lambda\mathbf{x}\mathbf{x}^T) + \alpha(\mathbf{I} - \mathbf{x}\mathbf{x}^T)}{\lambda+\alpha}. \tag{4.10}$$

Observe that the Jacobian is symmetric.

THEOREM 4.8. *Let $(\lambda, \mathbf{x})$ be an eigenpair of a symmetric tensor $\mathcal{A} \in \mathbb{R}^{[m,n]}$. Assume $\alpha \in \mathbb{R}$ such that $\alpha > \beta(\mathcal{A})$, where $\beta(\mathcal{A})$ is as defined in* (4.2). *Let $\phi(\mathbf{x})$ be given by* (4.9). *Then $(\lambda, \mathbf{x})$ is negative stable if and only if $\mathbf{x}$ is a linearly attracting fixed point of ϕ.*

Proof. Assume that $(\lambda, \mathbf{x})$ is negative stable. The Jacobian $J(\mathbf{x};\alpha)$ is given by (4.10). By Theorem 2.11, we need to show that $\rho(J(\mathbf{x};\alpha)) < 1$ or, equivalently since $J(\mathbf{x};\alpha)$ is symmetric, $|\mathbf{y}^T J(\mathbf{x};\alpha)\mathbf{y}| < 1$ for all $\mathbf{y} \in \Sigma$. We restrict our attention to $\mathbf{y} \perp \mathbf{x}$ since $J(\mathbf{x};\alpha)\mathbf{x} = \mathbf{0}$.

Let $\mathbf{y} \in \Sigma$ with $\mathbf{y} \perp \mathbf{x}$. Then

$$|\mathbf{y}^T J(\mathbf{x};\alpha)\mathbf{y}| = \left| \frac{\mathbf{y}^T\left((m-1)\mathcal{A}\mathbf{x}^{m-2}\right)\mathbf{y} + \alpha}{\lambda+\alpha} \right|.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The assumption that $(\lambda, \mathbf{x})$ is negative stable means that $C(\lambda, \mathbf{x})$ is negative definite; therefore, $\mathbf{y}^T \left((m-1)\mathcal{A}\mathbf{x}^{m-2}\right)\mathbf{y} < \lambda$. On the other hand, by the definition of β,

$$\rho\left((m-1)\mathcal{A}\mathbf{x}^{m-2}\right) \leq \beta(\mathcal{A}).$$

Thus, using the fact that $\lambda + \alpha$ is positive, we have

$$0 < \frac{-\beta(\mathcal{A}) + \alpha}{\lambda + \alpha} \leq \frac{\mathbf{y}^T \left((m-1)\mathcal{A}\mathbf{x}^{m-2}\right)\mathbf{y} + \alpha}{\lambda + \alpha} < \frac{\lambda + \alpha}{\lambda + \alpha} = 1.$$

Hence, $\rho(J(\mathbf{x}; \alpha)) < 1$, and $\mathbf{x}$ is a linearly attracting fixed point.

On the other hand, if $(\lambda, \mathbf{x})$ is not negative stable, then there exists $\mathbf{w} \in \Sigma$ such that $\mathbf{w} \perp \mathbf{x}$ and $\mathbf{w}^T \left((m-1)\mathcal{A}\mathbf{x}^{m-2}\right)\mathbf{w} \geq \lambda$. Thus,

$$\mathbf{w}^T J(\mathbf{x}; \alpha)\mathbf{w} = \frac{\mathbf{w}^T \left((m-1)\mathcal{A}\mathbf{x}^{m-2}\right)\mathbf{w} + \alpha}{\lambda + \alpha} \geq \frac{\lambda + \alpha}{\lambda + \alpha} = 1.$$

Consequently, $\rho(J(\mathbf{x}; \alpha)) \geq 1$, and $\mathbf{x}$ is not a linearly attracting fixed point by Theorems 2.11 and 2.12 ∎

In fact, we can see from the proof of Theorem 4.8 that if the eigenpair $(\lambda, \mathbf{x})$ is not negative stable, there is no choice of $\alpha \in \mathbb{R}$ that will make $\rho(J(\mathbf{x}; \alpha)) < 1$. For $\mathbf{x}$ to be a fixed point at all, we must have $\lambda + \alpha > 0$, and this is sufficient to obtain $\rho(J(\mathbf{x}; \alpha)) \geq 1$ if $(\lambda, \mathbf{x})$ is not negative stable. In other words, smaller values of α do not induce "accidental" convergence to any additional eigenpairs.

An alternative argument establishes, for $\alpha > \beta(\mathcal{A})$, the slightly broader result that any attracting fixed point, regardless of order of convergence, must be a strict constrained local maximum of $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$ on Σ. That is, the marginally attracting case corresponds to a stationary point that has degenerate $C(\lambda, \mathbf{x})$ but is still a maximum. This follows from Theorem 2.8, where the needed convexity holds for $\alpha > \beta(\mathcal{A})$, so that any vector $\mathbf{x}' \in \Sigma$ in the neighborhood of convergence of $\mathbf{x}$ must satisfy $f(\mathbf{x}') < f(\mathbf{x})$. One can convince oneself that the converse also holds for $\alpha > \beta(\mathcal{A})$, i.e., any strict local maximum corresponds to an attracting fixed point. This is because the strict monotonicity of f under iteration (other than at a fixed point) implies that the iteration acts as a contraction on the region of closed contours of f around the maximum.

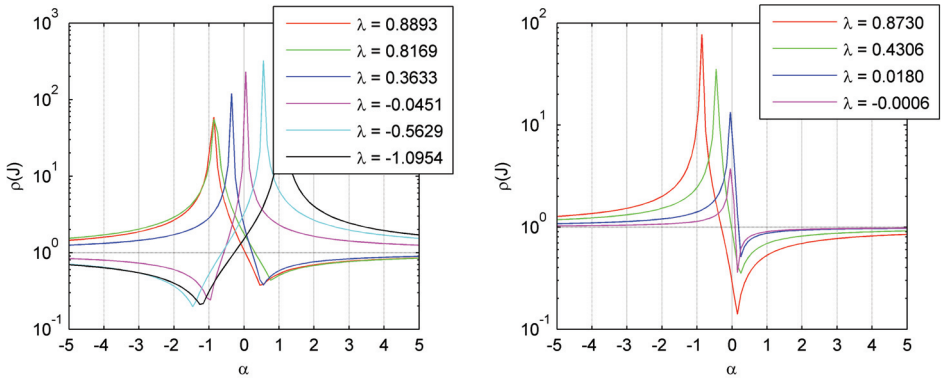
The counterpart of Theorem 4.8 for the concave case is as follows.

COROLLARY 4.9. *Let $(\lambda, \mathbf{x})$ be an eigenpair of a symmetric tensor $\mathcal{A} \in \mathbb{R}^{[m,n]}$. Assume $\alpha \in \mathbb{R}$ such that $\alpha < -\beta(\mathcal{A})$, where $\beta(\mathcal{A})$ is as defined in* (4.2). *Let $\phi(\mathbf{x})$ be given by* (4.9). *Then $(\lambda, \mathbf{x})$ is positive stable if and only if $\mathbf{x}$ is a linearly attracting fixed point of $-\phi$.*

Example 4.10. We return again to $\mathcal{A} \in \mathbb{R}^{[4,3]}$ as defined in Example 3.5. Figure 4.3(a) shows the spectral radius of the Jacobian of the fixed point iteration for varying values of α for all eigenpairs that are positive or negative stable. At $\alpha = 0$, the spectral radius is greater than 1 for every eigenvalue, and this is why S-HOPM never converges. At $\alpha = 2$, on the other hand, we see that the spectral radius is less than 1 for all of the negative stable eigenpairs. Furthermore, the spectral radius stays less than 1 as α increases. Conversely, at $\alpha = -2$, the spectral radius is less than 1 for all the eigenpairs that are positive stable.

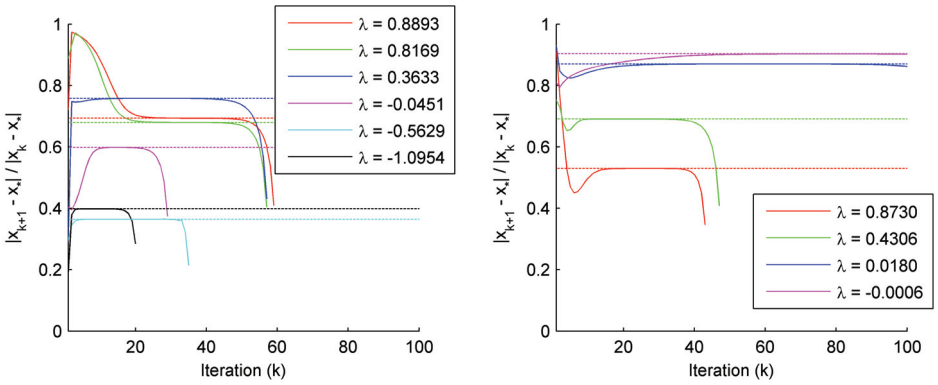
In Figure 4.4(a), we plot example iteration sequences for $\|\mathbf{x}_{k+1} - \mathbf{x}_*\|/\|\mathbf{x}_k - \mathbf{x}_*\|$ for each eigenpair, using $\alpha = 2$ for the negative stable eigenpairs and $\alpha = -2$ for the positive stable eigenpairs. We expect $\|\mathbf{x}_{k+1} - \mathbf{x}_*\| = \sigma\|\mathbf{x}_k - \mathbf{x}_*\|$, where σ is

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5. (b) $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6.

FIG. 4.3. *Spectral radii of the Jacobian $J(\mathbf{x};\alpha)$ for different eigenpairs as α varies.*



(a) $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5 using $\alpha = \pm 2$, as appropriate. (b) $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6 using $\alpha = 1$.

FIG. 4.4. *Example plots of $\|\mathbf{x}_{k+1} - \mathbf{x}_*\|/\|\mathbf{x}_k - \mathbf{x}_*\|$. The expected rate of convergence from $J(\mathbf{x}_*;\alpha)$ is shown as a dashed line.*

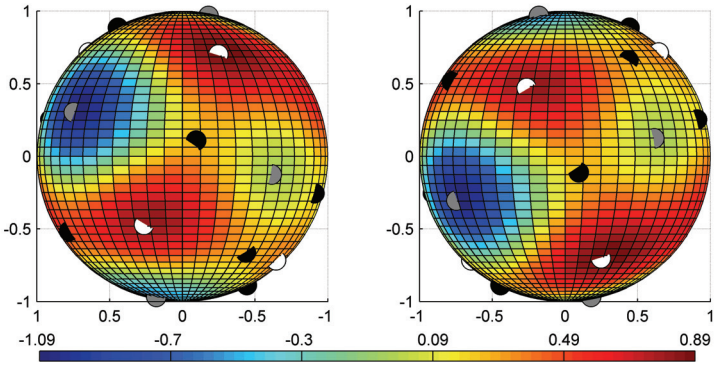
the spectral radius of the Jacobian $J(\mathbf{x};\alpha)$. For example, for $\lambda = -1.0954$, we have $\sigma = 0.4$ (shown as a dashed line), and this precisely matched the observed rate of convergence (shown as a solid line).

Figure 4.5(a) plots $f(\mathbf{x})$ on the unit sphere using color to indicate function value. We show the front and back of the sphere. Notice that the horizontal axis is from 1 to -1 in the left plot and from -1 to 1 in the right plot, as if walking around the sphere. In this image, the horizontal axis corresponds to x_2 and the vertical axis to x_3; the left image is centered at $x_1 = 1$ and the right image at $x_1 = -1$. Since m is even, the function is symmetric, i.e., $f(\mathbf{x}) = f(-\mathbf{x})$. The eigenvectors are shown as white, gray, and black circles corresponding to their classification as negative stable, positive stable, and unstable, respectively; in turn, these correspond to maxima, minima, and saddle points of $f(\mathbf{x})$.

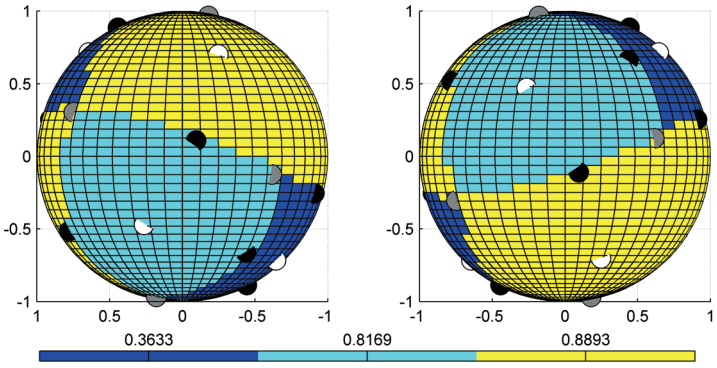
Figure 4.5(b) shows the basins of attraction for SS-HOPM with $\alpha = 2$. Every grid point on the sphere was used as a starting point for SS-HOPM, and it is colored[4]

[4]Specifically, each block on the sphere is colored according to the convergence of its lower left point.

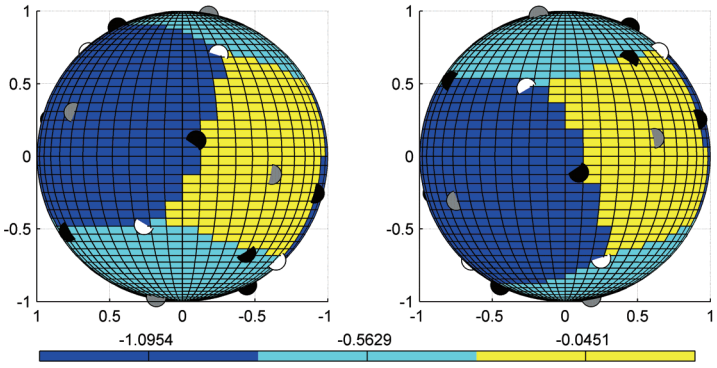
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) Function values for $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$.



(b) SS-HOPM basins of attraction using $\alpha = 2$.



(c) SS-HOPM basins of attraction using $\alpha = -2$.

FIG. 4.5. *Illustrations for* $\mathcal{A} \in \mathbb{R}^{[4,3]}$ *from Example* 3.5. *The horizontal axis corresponds to* x_2 *and the vertical axis to* x_3*; the left image is centered at* $x_1 = 1$ *and the right at* $x_1 = -1$*. White, gray, and black dots indicate eigenvectors that are negative stable, positive stable, and unstable, respectively.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

according to which eigenvalue it converged to. In this case, every run converges to a negative stable eigenpair (labeled with a white circle). Recall that SS-HOPM must converge to some eigenpair per Theorem 4.4, and Theorem 4.8 says that it is generically a negative stable eigenpair. Thus, the nonattracting points lie on the boundaries of the domains of attraction.

Figure 4.5(c) shows the basins of attraction for SS-HOPM with $\alpha = -2$. In this case, every starting point converges to an eigenpair that is positive stable (shown as gray circles). □

Example 4.11. We return again to $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from Example 3.6, which is interesting because S-HOPM was able to find 2 of its eigenpairs without any shift. In Figure 4.6(a), $f(\mathbf{x})$ is plotted on the unit sphere, along with each eigenvector, colored white, gray, or black based on whether it is negative stable, positive stable, or unstable, respectively. Observe that the function is antisymmetric, i.e., $f(\mathbf{x}) = -f(-\mathbf{x})$. Figure 4.6(b) shows the basins of attraction for S-HOPM (i.e., SS-HOPM with $\alpha = 0$). Every starting point converges to one of the 2 labeled eigenpairs. This is not surprising because Figure 4.3(b) shows that there are 2 eigenvalues for which the spectral radius of the Jacobian is less than 1 ($\lambda = 0.8730$ and 0.4306). The other 2 eigenvalues are nonattracting for $\alpha = 0$. Figure 4.4(b) shows the observed rates of convergence.

Figure 4.6(c) shows the basins of attraction for SS-HOPM with $\alpha = 1$; each negative stable eigenpair (shown as a white circle) is an attracting eigenpair. The concave case is just a mirror image and is not shown. □

As the previous example reminds us, for odd order, there is no need to try both positive and negative α because the definiteness of C flips for eigenvectors of opposite sign.

Two additional examples of SS-HOPM are presented in Appendix A.

4.3. Relationship to power method for matrix eigenpairs. The power method for matrix eigenpairs is a technique for finding the largest-magnitude eigenvalue (and corresponding eigenvector) of a diagonalizable symmetric matrix [10]. Let $\mathbf{A}$ be a symmetric real-valued $n \times n$ matrix. Then the matrix power method is defined by

$$\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|}.$$

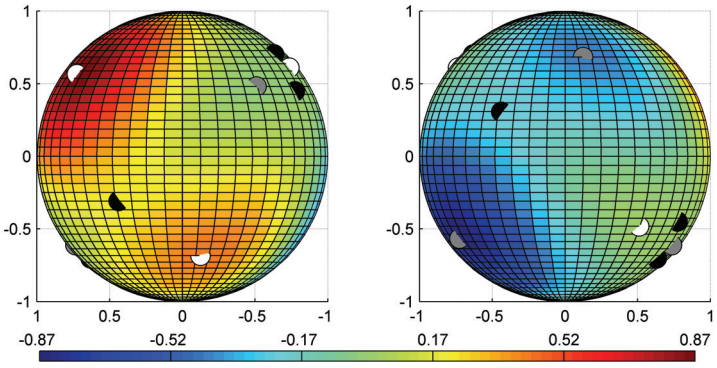
Assume that $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ is the Schur decomposition of $\mathbf{A}$ with eigenvalues satisfying $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$ (note the strict difference in the first 2 eigenvalues). The sequence $\{\mathbf{x}_k\}$ produced by the matrix power method always converges (up to sign) to the eigenvector associated with λ_1. Shifting the matrix by $\mathbf{A} \leftarrow \mathbf{A} + \alpha\mathbf{I}$ shifts the eigenvalues by $\lambda_j \leftarrow \lambda_j + \alpha$, potentially altering which eigenvalue has the largest magnitude.

In the matrix case, the eigenvalues of the Jacobian defined by (4.10) for an eigenpair $(\lambda_j, \mathbf{x}_j)$ are given by

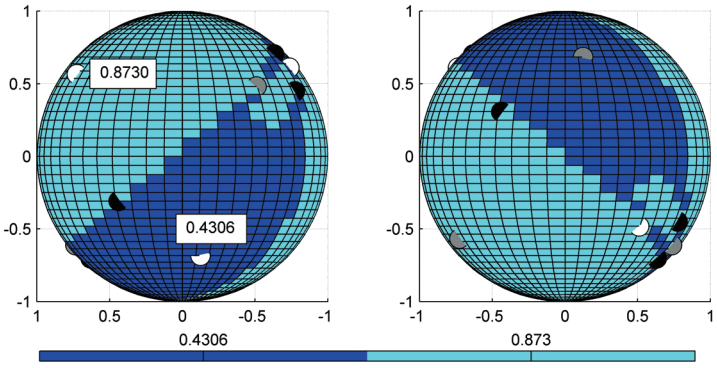
$$\{0\} \cup \left\{ \frac{\lambda_i + \alpha}{\lambda_j + \alpha} : 1 \leq i \leq n \text{ with } i \neq j \right\}.$$

Thus, the Jacobian at $\mathbf{x}_1$ is the only one such that $\rho(J(\mathbf{x}; \alpha)) < 1$; no other eigenvectors are stable fixed points of the iteration. This corresponds to Theorem 4.8 (or Corollary 4.9), since the most positive eigenvalue is negative stable, the most negative eigenvalue is positive stable, and every other eigenvalue is unstable. The eigenpair

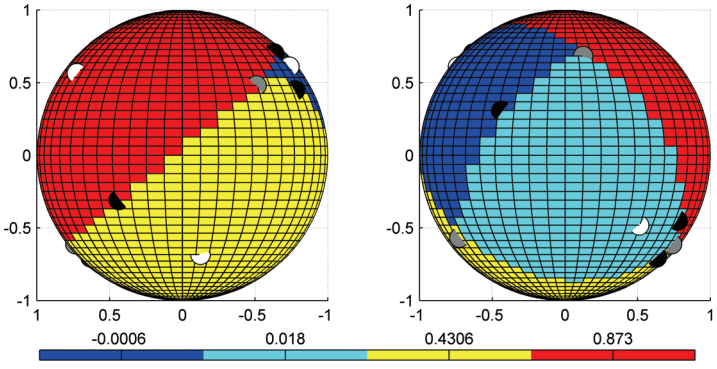
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) Function values for $f(\mathbf{x}) = \mathcal{A}\mathbf{x}^m$.



(b) SS-HOPM basins of attraction using $\alpha = 0$.



(c) SS-HOPM basins of attraction using $\alpha = 1$.

FIG. 4.6. *Illustrations for* $\mathcal{A} \in \mathbb{R}^{[3,3]}$ *from Example* 3.6. *The horizontal axis corresponds to* x_2 *and the vertical axis to* x_3*; the left image is centered at* $x_1 = 1$ *and the right at* $x_1 = -1$*. White, gray, and black dots indicate eigenvectors that are negative stable, positive stable, and unstable, respectively.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$(\lambda_1, \mathbf{x}_1)$ is an attractor for ordinary (convex) power iteration if $\lambda_1 > 0$ or for flipped (concave) power iteration if $\lambda_1 < 0$.

In contrast to the matrix power method, SS-HOPM can find multiple eigenpairs since there may be multiple positive and negative stable eigenpairs. But, as for matrices, since the most positive and most negative eigenvalues correspond to the global maximum and minimum of $f(\mathbf{x})$, they must be negative stable and positive stable, respectively. Thus, choosing α positive is necessary for finding the most positive tensor eigenvalue; conversely, choosing α negative is necessary for finding the most negative tensor eigenvalue. Unfortunately, the ability to find multiple eigenpairs means that there is no guarantee that the iterates will converge to an extremal eigenpair from every starting point. In fact, multiple starting points may be needed.

4.4. Comparison to other methods. SS-HOPM is useful for its guaranteed convergence properties and its simple implementation based on tensor-vector multiplication. For fixed m and large n, the computational complexity of each iteration of SS-HOPM is $O(n^m)$, which is the number of individual terms to be computed in $\mathcal{A}\mathbf{x}^{m-1}$. This is analogous to the $O(n^2)$ complexity of matrix-vector multiplication as used in the matrix power method. We do not yet know how the number of iterations needed for numerical convergence of SS-HOPM depends on m and n.

The convergence of SS-HOPM to only a subset of eigenvalues, which tend to be among the largest in magnitude, is beneficial when the large eigenvalues are of primary interest, as in the rank-1 approximation problem [11]. In particular, the most positive eigenvalue and most negative eigenvalue always have a region of stable convergence for a suitable choice of shift. However, the lack of stable convergence to certain other eigenvalues is a disadvantage if those eigenvalues are of interest.

One evident computational approach for finding tensor eigenpairs should be compared with SS-HOPM. This is to apply a numerical solver for nonlinear equation systems, such as Newton's method, directly to the eigenvalue equations (1.2). The computational complexity of each iteration of Newton's method for this system is that of SS-HOPM plus the construction and inversion of the $(n+1) \times (n+1)$ Jacobian for $(\lambda, \mathbf{x})$. The Jacobian construction is effectively included in SS-HOPM, since it is dominated by computing $\mathcal{A}\mathbf{x}^{m-2}$, which is a precursor of $\mathcal{A}\mathbf{x}^{m-1}$. The additional work for inversion is $O(n^3)$, and for $m \geq 3$ it does not affect the complexity scaling, which remains $O(n^m)$.

Two advantages of an approach such as Newton's method are generic locally stable convergence, which enables finding eigenpairs not found by SS-HOPM, and the quadratic order of convergence, which can be expected to require fewer iterations than the linearly convergent SS-HOPM. On the other hand, there is no known guarantee of global convergence as there is for SS-HOPM, and it is possible that many starting points fail to converge. Even those that do converge may lead to eigenpairs of less interest for a particular application. Furthermore, certain tensor structures can be more efficiently handled with SS-HOPM than with Newton's method. For example, consider a higher-order symmetric tensor expressed as a sum of terms, each of which is an outer product of matrices. The computation of $\mathcal{A}\mathbf{x}^{m-1}$ then reduces to a series of matrix-vector multiplications, which are $O(n^2)$. This compares favorably to the $O(n^3)$ of Newton's method for the same tensor. Further investigation of general nonlinear solver approaches to the tensor eigenvalue problem will be beneficial.

Finally, we consider a polynomial solver approach, such as we implemented in *Mathematica*. This can find all eigenpairs (subject to numerical conditioning issues) but becomes computationally expensive for large m and n. In part this is simply

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

because, from Theorem 5.3, the number of eigenpairs grows exponentially with n. The solver in *Mathematica* is designed to find all solutions; it is not clear whether a substantial improvement in efficiency would be possible if only one or a few solutions were required.

Nevertheless, for comparison with the iterative approaches discussed above, we have measured the computational time per eigenpair on a desktop computer for various values of m and n, as shown in Figure 4.7. The complexity of the polynomial solution, even measured per eigenpair, is seen to increase extremely rapidly (faster than exponentially) with n. Thus the polynomial solver approach is not expected to be practical for large n.

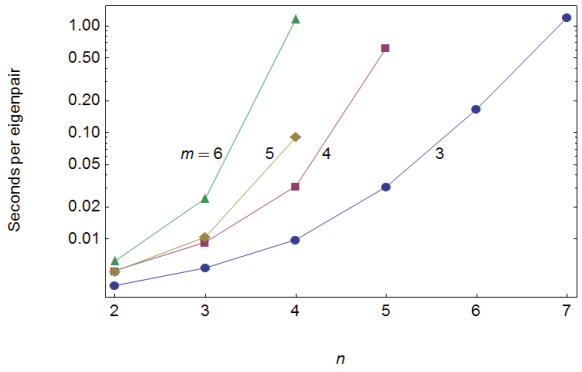


FIG. 4.7. *Average time (over* 10 *trials) required to compute all eigenpairs, divided by the number of eigenpairs, for random symmetric tensors in* $\mathbb{R}^{[m,n]}$. *Note logarithmic vertical scale. Measured using* `NSolve` *in Mathematica on a* 4 *GHz Intel Core i7.*

5. Complex case. We present the more general definition of complex eigenpairs and some related results, and propose an extension of the SS-HOPM algorithm to this case.

5.1. Eigenrings. Some of the solutions of the polynomial system that results from the eigenvalue equations may be complex; thus, the definition can be extended to the complex case as follows, where † denotes the conjugate transpose.

DEFINITION 5.1. *Assume that* $\mathcal{A}$ *is a symmetric* m*th-order* n*-dimensional real-valued tensor. Then* $\lambda \in \mathbb{C}$ *is an* eigenvalue *of* $\mathcal{A}$ *if there exists* $\mathbf{x} \in \mathbb{C}^n$ *such that*

$$\mathcal{A}\mathbf{x}^{m-1} = \lambda \mathbf{x} \quad \textit{and} \quad \mathbf{x}^\dagger \mathbf{x} = 1. \tag{5.1}$$

The vector $\mathbf{x}$ *is a corresponding* eigenvector, *and* $(\lambda, \mathbf{x})$ *is called an* eigenpair.

Definition 5.1 is closely related to the E-eigenpairs defined by Qi [18, 19] but differs in the constraint on $\mathbf{x}$.[5] It can also be considered as the obvious extension of (l^2-)eigenpairs to $\mathbb{C}$.

It has been observed [19, 3] that the complex eigenpairs of a tensor form equivalence classes under a multiplicative transformation. Specifically, if $(\lambda, \mathbf{x})$ is an eigenpair of $\mathcal{A} \in \mathbb{R}^{[m,n]}$ and $\mathbf{y} = e^{i\varphi}\mathbf{x}$ with $\varphi \in \mathbb{R}$, then $\mathbf{y}^\dagger \mathbf{y} = \mathbf{x}^\dagger \mathbf{x} = 1$ and

$$\mathcal{A}\mathbf{y}^{m-1} = e^{i(m-1)\varphi}\mathcal{A}\mathbf{x}^{m-1} = e^{i(m-1)\varphi}\lambda\mathbf{x} = e^{i(m-2)\varphi}\lambda\mathbf{y}.$$

[5] Qi [18, 19] requires $\mathbf{x}^T\mathbf{x} = 1$ rather than $\mathbf{x}^\dagger\mathbf{x} = 1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Therefore $(e^{i(m-2)\varphi}\lambda, e^{i\varphi}\mathbf{x})$ is also an eigenpair of $\mathcal{A}$ for any $\varphi \in \mathbb{R}$. Consequently, if λ is an eigenvalue, then any other $\lambda' \in \mathbb{C}$ with $|\lambda'| = |\lambda|$ is also an eigenvalue. This leads to the notion of an eigenring.

DEFINITION 5.2 (eigenring). *For any $(\lambda, \mathbf{x}) \in \mathbb{C} \times \mathbb{C}^n$ that is an eigenpair of $\mathcal{A} \in \mathbb{R}^{[m,n]}$, we define a corresponding equivalence class of (vector-normalized) eigenpairs*

$$\mathcal{P}(\lambda, \mathbf{x}) = \{(\lambda', \mathbf{x}') : \lambda' = e^{i(m-2)\varphi}\lambda, \mathbf{x}' = e^{i\varphi}\mathbf{x}, \varphi \in \mathbb{R}\},$$

as well as a corresponding eigenring

$$\mathcal{R}(\lambda) = \{\lambda' \in \mathbb{C} : |\lambda'| = |\lambda|\}.$$

Thus, any eigenring contains either 1 or 2 real eigenvalues. The special case of real eigenpairs occurs whenever the corresponding eigenvector for one of those real eigenvalues is also real.

Since we assume that $\mathcal{A}$ is real-valued, any nonreal eigenpairs must come in sets of 2 related by complex conjugation, because taking the conjugate of the eigenvalue equation does not change it. Such conjugate eigenpairs are not members of the same equivalence class unless they are equivalent to a real eigenpair.

An elegant result has recently been derived for the number of distinct (nonequivalent) eigenvalues of a symmetric tensor. The result was first derived for even-order tensors by Ni et al. [16] and later generalized by Cartwright and Sturmfels [3] for all cases. The case of $m = 2$ requires application of l'Hôpital's rule to see that there are n eigenvalues.

THEOREM 5.3 (Cartwright and Sturmfels [3]). *A generic symmetric tensor $\mathcal{A} \in \mathbb{R}^{[m,n]}$ has $((m-1)^n - 1)/(m-2)$ distinct eigenvalue equivalence classes.*

These papers [16, 3] use the condition $\mathbf{x}^T\mathbf{x} = 1$ to normalize eigenpairs, but in the generic case the result is the same for our condition $\mathbf{x}^\dagger\mathbf{x} = 1$. This is because the eigenpairs with $\mathbf{x}^\dagger\mathbf{x} = 1$ generically consist of isolated equivalence classes that have $\mathbf{x}^T\mathbf{x} \neq 0$ and thus can be rescaled to satisfy $\mathbf{x}^T\mathbf{x} = 1$, giving a one-to-one correspondence between the distinct eigenvalues in the two normalizations. In special cases, however, the condition $\mathbf{x}^\dagger\mathbf{x} = 1$ admits additional eigenpairs with $\mathbf{x}^T\mathbf{x} = 0$. Furthermore, tensors can be constructed with a continuous family of nonequivalent eigenvectors that correspond to the same eigenvalue when normalized by $\mathbf{x}^T\mathbf{x}$ but to infinitely many distinct eigenvalues when normalized by $\mathbf{x}^\dagger\mathbf{x}$ [3, Example 5.7].

The polynomial system solver using Gröbner bases mentioned earlier can also be used to find complex solutions. A complication is that our normalization condition $\mathbf{x}^\dagger\mathbf{x} = 1$ is nonpolynomial due to the complex conjugation. The system, however, becomes polynomial if the alternate normalization condition $\mathbf{x}^T\mathbf{x} = 1$ is temporarily adopted. Any such $\mathbf{x}$ can be rescaled to satisfy $\mathbf{x}^\dagger\mathbf{x} = 1$. Complex eigenvectors with $\mathbf{x}^T\mathbf{x} = 0$ will not be found, but, as just noted, these do not occur generically. Any nonreal solutions are transformed to a representative of the eigenring with positive real λ by setting

$$(\lambda, \mathbf{x}) \leftarrow \left(\frac{|\lambda|}{(\mathbf{x}^\dagger\mathbf{x})^{m/2-1}}, \left(\frac{|\lambda|}{\lambda}\right)^{1/(m-2)} \frac{\mathbf{x}}{(\mathbf{x}^\dagger\mathbf{x})^{1/2}} \right).$$

This polynomial system solution becomes prohibitively expensive for large m or n; however, for Example 3.5, the nonreal eigenpairs can be computed this way and are presented in Table 5.1. Thus, from this and Table 3.1, we have found the 13 eigenvalue equivalence classes (real and nonreal) guaranteed by Theorem 5.3.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

TABLE 5.1
Nonreal eigenpairs for $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5.

λ	$\mathbf{x}^T$
0.6694	$[\ 0.2930 + 0.0571i \quad 0.8171 - 0.0365i \quad -0.4912 - 0.0245i\ ]$
0.6694	$[\ 0.2930 - 0.0571i \quad 0.8171 + 0.0365i \quad -0.4912 + 0.0245i\ ]$

5.2. SS-HOPM for complex eigenpairs. We propose an extension of the SS-HOPM algorithm to the case of complex vectors in Algorithm 3. Observe that the division by $\lambda_k + \alpha$ keeps the phase of $\mathbf{x}_k$ from changing unintentionally. It is akin to taking the negative in the concave case in Algorithm 2. It is important to note that even if an eigenpair is real, there is no guarantee that the complex SS-HOPM will converge to the real eigenpair; instead, it will converge to some random rotation in the complex plane. We have no convergence theory in the convex case, but we present several promising numerical examples.

Algorithm 3 Complex SS-HOPM.

Given a tensor $\mathcal{A} \in \mathbb{R}^{[m,n]}$.
Require: $\mathbf{x}_0 \in \mathbb{C}^n$ with $\|\mathbf{x}_0\| = 1$. Let $\lambda_0 = \mathcal{A}\mathbf{x}_0^m$.
Require: $\alpha \in \mathbb{C}$
1: **for** $k = 0, 1, \ldots$ **do**
2: $\quad \hat{\mathbf{x}}_{k+1} \leftarrow (\mathcal{A}\mathbf{x}_k^{m-1} + \alpha\mathbf{x}_k)/(\lambda_k + \alpha)$
3: $\quad \mathbf{x}_{k+1} \leftarrow \hat{\mathbf{x}}_{k+1}/\|\hat{\mathbf{x}}_{k+1}\|$
4: $\quad \lambda_{k+1} \leftarrow \mathbf{x}_{k+1}^{\dagger}\mathcal{A}\mathbf{x}_{k+1}^{m-1}$
5: **end for**

Example 5.4. We once again revisit $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5 and test the complex version of SS-HOPM in Algorithm 3. Table 5.2(a) shows the results of 100 runs using the same experimental conditions as in Example 3.5 except with complex random starting vectors. We find 7 distinct eigenrings—the 6 stable real eigenpairs as well as a ring corresponding to the 2 complex eigenpairs. Figure 5.1(a) shows the individual λ_* values plotted on the complex plane. As mentioned above, it may converge anywhere on the eigenring, though there is clear bias toward the value of α.

To investigate this phenomenon further, we do another experiment with $\alpha = -(1+i)/\sqrt{2}$. It finds the same eigenrings as before, as shown in Table 5.2(b), but this time the λ_* values are distributed mostly in the lower left quadrant of the complex plane, as shown in Figure 5.1(b), again close to the value of α. In the case of the 2 complex eigenpairs with the same eigenring, the method finds the 2 distinct eigenvec-

TABLE 5.2
Eigenrings computed for $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example 3.5 by complex SS-HOPM with 100 random starts.

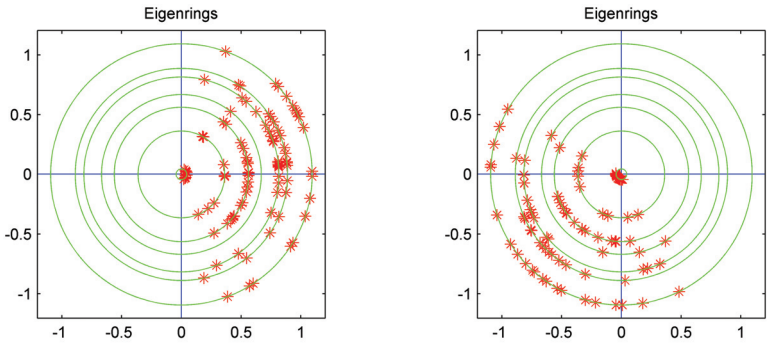
(a) $\alpha = 2$.

# Occurrences	$\lvert\lambda\rvert$
18	1.0954
18	0.8893
21	0.8169
1	0.6694
22	0.5629
8	0.3633
12	0.0451

(b) $\alpha = \sqrt{2}(1+i)$ (2 failures).

# Occurrences	$\lvert\lambda\rvert$
22	1.0954
15	0.8893
12	0.8169
4	0.6694
16	0.5629
9	0.3633
20	0.0451

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) $\alpha = 2$. (b) $\alpha = -\sqrt{2}(1+i)$ (2 failures).

FIG. 5.1. *For $\mathcal{A} \in \mathbb{R}^{[4,3]}$ from Example* 3.5*, final λ values (indicated by asterisks) for* 100 *runs of complex SS-HOPM. The circles denote the eigenrings.*

tors (i.e., defining 2 different equivalence classes) in the 4 different times it converges to that eigenvalue; this is not surprising since the complex eigenvalue has 2 different eigenvectors, as shown in Table 3.1.

We also ran an experiment with $\alpha = 0$. In this case, 95 trials converged, but to noneigenpairs (all with $|\lambda| = 0.3656$). In each case, even though $\{\lambda_k\}$ converged, we had $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \to 1.1993$, indicating that the sequence $\{\mathbf{x}_k\}$ had not converged and hence we did not obtain an $\mathbf{x}_*$ satisfying the eigenvalue equation (5.1). Although it is not shown, in all the convergent examples with the shifts mentioned above, the $\{\mathbf{x}_k\}$ sequence converged. □

6. Conclusions. We have developed the shifted symmetric higher-order power method (SS-HOPM) for finding tensor eigenvalues. The method can be considered as a higher-order analogue to the power method for matrices. Just as in the matrix case, it cannot find all possible eigenvalues, but it is guaranteed to be able to find the largest-magnitude eigenvalue. Unlike the matrix case, it can find multiple eigenvalues; multiple starting points are typically needed to find the largest eigenvalue. An implementation of SS-HOPM using a graphics processing unit (GPU) has been reported [1].

Building on previous work [11, 23, 8], we show that SS-HOPM will always converge to a real eigenpair for an appropriate choice of α. Moreover, using fixed point analysis, we characterize exactly which real eigenpairs can be found by the method, i.e., those that are positive or negative stable. Alternative methods will need to be developed for finding the unstable real eigenpairs, i.e., eigenpairs for which $C(\lambda, \mathbf{x})$ is indefinite. A topic for future investigation is that the boundaries of the basins of attraction for SS-HOPM seem to be defined by the nonattracting eigenvectors.

We present a complex version of SS-HOPM and limited experimental results indicating that it finds eigenpairs, including complex eigenpairs. Analysis of the complex version is a topic for future study.

Much is still unknown about tensor eigenpairs. For example, how do the eigenpairs change with small perturbations of the tensor entries? Is there an eigendecomposition of a tensor? Can the convergence rate of the current method be accelerated? How does one numerically compute unstable eigenpairs? For computing efficiency, what is the optimal storage for symmetric tensors? These are all potential topics of future research.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Appendix. Further examples. For additional insight, we consider two analytical examples. In the experiments in this section, each random trial used a point $\mathbf{x}_0$ chosen from a uniform distribution on $[-1,1]^n$. We allow up to 1000 iterations and say that the algorithm has converged if $|\lambda_{k+1} - \lambda_k| < 10^{-15}$.

Consider the tensor $\mathcal{A} \in \mathbb{R}^{[3,3]}$ whose entries are 0 except where the indices are all unequal, in which case the entries are 1, i.e.,

$$a_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \text{ is some permutation of } (1,2,3), \\ 0 & \text{otherwise.} \end{cases} \tag{A.1}$$

Any eigenpair $(\lambda, \mathbf{x})$ must satisfy the following system of equations:

$$2x_2x_2 = \lambda x_1, \qquad 2x_1x_3 = \lambda x_2, \qquad 2x_1x_2 = \lambda x_3, \qquad x_1^2 + x_2^2 + x_3^2 = 1.$$

The 7 real eigenpairs can be computed analytically and are listed in Table A.1(a), from which we can see that there are 4 negative stable eigenpairs that should be identifiable by SS-HOPM. Figure A.1(a) shows the spectral radius of the Jacobian as α varies; the curve is identical for all 4 negative stable eigenpairs.

Another example is given as follows. Define the tensor $\mathcal{A} \in \mathbb{R}^{[4,2]}$ by

$$a_{ijkl} = 0 \quad \text{for all} \quad i,j,k,l \quad \text{except} \quad a_{1111} = 1 \quad \text{and} \quad a_{2222} = -1. \tag{A.2}$$

The eigenpairs can be computed analytically as solutions to the following system:

$$x_1^3 = \lambda x_1, \qquad -x_2^3 = \lambda x_2, \qquad x_1^2 + x_2^2 = 1.$$

The 2 real eigenpairs are listed in Table A.1(b), from which we can see that one is negative stable and the other is positive stable. Figure A.1(a) shows the spectral radius of the Jacobian as α varies. In this case, the spectral radius of the Jacobian can be computed analytically; for $\lambda = 1$, it is $\frac{\alpha}{1+\alpha}$ and hence there is a singularity for $\alpha = -1$.

TABLE A.1
Eigenpairs for two analytical problems.

(a) Eigenpairs for $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from (A.1).

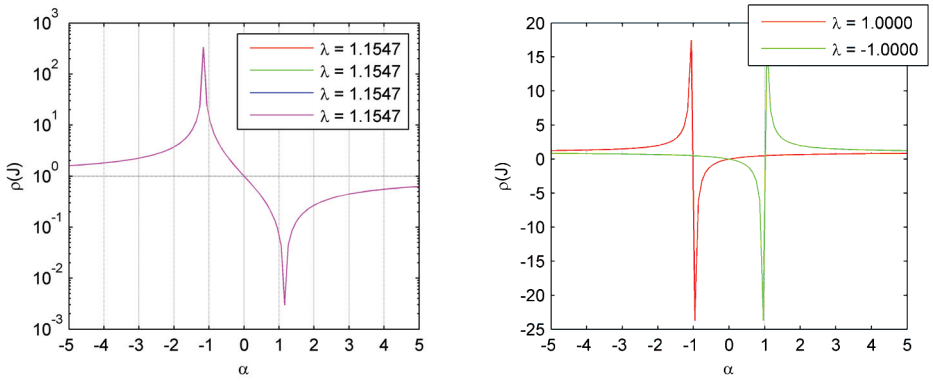
λ	$\mathbf{x}$	Eigenvalues of $C(\lambda, \mathbf{x})$	Type
0	[1 0 0]	{ −2 , 2 }	Unstable
0	[0 1 0]	{ −2 , 2 }	Unstable
0	[0 0 1]	{ −2 , 2 }	Unstable
$2/\sqrt{3}$	$[\ 1/\sqrt{3}\ \ 1/\sqrt{3}\ \ 1/\sqrt{3}\ ]$	{ −2.3094 , −2.3094 }	Neg. stable
$2/\sqrt{3}$	$[\ 1/\sqrt{3}\ -1/\sqrt{3}\ -1/\sqrt{3}\ ]$	{ −2.3094 , −2.3094 }	Neg. stable
$2/\sqrt{3}$	$[\ -1/\sqrt{3}\ \ 1/\sqrt{3}\ -1/\sqrt{3}\ ]$	{ −2.3094 , −2.3094 }	Neg. stable
$2/\sqrt{3}$	$[\ -1/\sqrt{3}\ -1/\sqrt{3}\ \ 1/\sqrt{3}\ ]$	{ −2.3094 , −2.3094 }	Neg. stable

(b) Eigenpairs for $\mathcal{A} \in \mathbb{R}^{[4,2]}$ from (A.2).

λ	$\mathbf{x}$	Eigenvalues of $C(\lambda, \mathbf{x})$	Type
1	[1 0]	{ −1 }	Neg. stable
−1	[0 1]	{ 1 }	Pos. stable

For (A.1), we ran 100 trials with $\alpha = 0$, and none converged, as expected per Figure A.1(a). The results of 100 random trials with $\alpha = 12$ (the "conservative choice") is shown in Table A.2(a), in which case every trial converged to one of

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) (A.1) (all four are identical). (b) (A.2).

FIG. A.1. *Spectral radii of the Jacobian $J(\mathbf{x};\alpha)$ for different eigenpairs as α varies.*

TABLE A.2
Eigenpairs for $\mathcal{A} \in \mathbb{R}^{[3,3]}$ from (A.1) *computed by SS-HOPM.*

(a) $\alpha = 12$.

# Occurrences	λ	$\mathbf{x}$	Median its.
22	1.1547	[−0.5774 0.5774 −0.5774]	92
18	1.1547	[0.5774 0.5774 0.5774]	90
29	1.1547	[−0.5774 −0.5774 0.5774]	91
31	1.1547	[0.5774 −0.5774 −0.5774]	94

(b) $\alpha = 1$.

# Occurrences	λ	$\mathbf{x}$	Median its.
22	1.1547	[0.5774 −0.5774 −0.5774]	9
25	1.1547	[−0.5774 0.5774 −0.5774]	9
26	1.1547	[0.5774 0.5774 0.5774]	9
27	1.1547	[−0.5774 −0.5774 0.5774]	9

TABLE A.3
Eigenpairs for $\mathcal{A} \in \mathbb{R}^{[4,2]}$ from (A.2) *computed by SS-HOPM.*

(a) $\alpha = 0.5$.

# Occurrences	λ	$\mathbf{x}$	Median its.
100	1.0000	[−1.0000 0.0000]	16

(b) $\alpha = -0.5$.

# Occurrences	λ	$\mathbf{x}$	Median its.
100	−1.0000	[− 0.0000 1.0000]	16

the 4 negative stable eigenpairs. (Note that $2/\sqrt{3} \approx 1.1547$ and $1/\sqrt{3} \approx 0.5774$.) Table A.2(b) shows the results of 100 random trials with $\alpha = 1$. As expected (per Figure A.1(a)), the convergence is much faster. For (A.2), we ran 100 trials with $\alpha = 0.5$ (Table A.3(a)) and 100 trials with $\alpha = -0.5$ (Table A.3(b)). We find the negative stable and positive stable eigenvalues as expected.

Acknowledgments. We thank Fangxiang Jiao and Chris Johnson (U. Utah), David Rogers (Sandia), and Dennis Bernstein (U. Michigan) for contacting us with interesting applications and providing test data. We thank Arindam Banergee (U. Min-

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

nesota) for providing yet another motivating application. We thank Dustin Cartwright and Bernd Sturmfels (UC Berkeley) for helpful discussions, especially about the number of eigenpairs for a problem of a given size. We also thank our colleagues at Sandia for numerous helpful conversations in the course of this work, especially Grey Ballad (intern from UC Berkeley). We are grateful to the three anonymous referees for identifying important references that we missed and for constructive feedback that greatly improved the manuscript.

REFERENCES

[1] G. Ballard, T. G. Kolda, and T. Plantenga, *Efficiently computing tensor eigenvalues on a GPU*, in IPDPSW'11: Proceedings of the 2011 IEEE International Symposium on Parallel & Distributed Processing Workshops, IEEE Computer Society Press, Los Alamitos, CA, 2011, pp. 1340–1348.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[3] D. Cartwright and B. Sturmfels, *The Number of Eigenvalues of a Tensor*, preprint, 2010; arXiv:1004.4953v1.

[4] K. C. Chang, K. Pearson, and T. Zhang, *Perron-Frobenius theorem for nonnegative tensors*, Commun. Math. Sci., 6 (2008), pp. 507–520.

[5] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.

[6] L. De Lathauwer, B. De Moor, and J. Vandewalle, *On the best rank-*1 *and rank-*$(R_1, R_2, \ldots, R_N)$ *approximation of higher-order tensors*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1324–1342.

[7] R. Diamond, *A note on the rotational superposition problem*, Acta Crystallogr. Sect. A, 44 (1988), pp. 211–216.

[8] A. T. Erdogan, *On the convergence of ICA algorithms with symmetric orthogonalization*, IEEE Trans. Signal Process., 57 (2009), pp. 2209–2221.

[9] P. L. Fackler, *Notes on Matrix Calculus*, 2005; available online from http://www4.ncsu.edu/∼pfackler/MatCalc.pdf, Sept. 2005.

[10] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.

[11] E. Kofidis and P. A. Regalia, *On the best rank-*1 *approximation of higher-order supersymmetric tensors*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 863–884.

[12] T. G. Kolda, *Multilinear Operators for Higher-Order Decompositions*, Tech. Report SAND2006-2081, Sandia National Laboratories, Albuquerque, NM, Livermore, CA, 2006.

[13] L.-H. 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.

[14] Y. Liu, G. Zhou, and N. F. Ibrahim, *An always convergent algorithm for the largest eigenvalue of an irreducible nonnegative tensor*, J. Comput. Appl. Math., 235 (2010), pp. 286–292.

[15] M. Ng, L. Qi, and G. Zhou, *Finding the largest eigenvalue of a nonnegative tensor*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1090–1099.

[16] G. Ni, L. Qi, F. Wang, and Y. Wang, *The degree of the E-characteristic polynomial of an even order tensor*, J. Math. Anal. Appl., 329 (2007), pp. 1218–1229.

[17] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, 1999.

[18] L. Qi, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Comput., 40 (2005), pp. 1302–1324.

[19] L. Qi, *Eigenvalues and invariants of tensors*, J. Math. Anal. Appl., 325 (2007), pp. 1363–1377.

[20] L. Qi, W. Sun, and Y. Wang, *Numerical multilinear algebra and its applications*, Front. Math. China, 2 (2007), pp. 501–526.

[21] L. Qi, F. Wang, and Y. Wang, *Z-eigenvalue methods for a global polynomial optimization problem*, Math. Program., 118 (2009), pp. 301–316.

[22] L. Qi, Y. Wang, and E. X. Wu, *D-eigenvalues of diffusion kurtosis tensors*, J. Comput. Appl. Math., 221 (2008), pp. 150–157.

[23] P. A. Regalia and E. Kofidis, *Monotonic convergence of fixed-point algorithms for ICA*, IEEE Trans. Neural Networks, 14 (2003), pp. 943–949.

[24] W. C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, SIAM, Philadelphia, 1974.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[25] T. Schultz and H.-P. Seidel, *Estimating crossing fibers: A tensor decomposition approach*, IEEE Trans. Visualization Comput. Graphics, 14 (2008), pp. 1635–1642.
[26] A. M. Stuart and A. R. Humphries, *Dynamical Systems and Numerical Analysis*, Cambridge University Press, Cambridge, UK, 1998.
[27] Y. Wang, L. Qi, and X. Zhang, *A practical method for computing the largest M-eigenvalue of a fourth-order partially symmetric tensor*, Numer. Linear Algebra Appl., 16 (2009), pp. 589–601.
[28] Wolfram Research, Inc., *Mathematica*, Version 7.0, 2008.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.