Journal of Computational Mathematics
Vol.28, No.2, 2010, 218–234.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2009.10-m1002

NONLINEAR RANK-ONE MODIFICATION OF THE SYMMETRIC EIGENVALUE PROBLEM*

Xin Huang
School of Mathematical Sciences, Fudan University, Shanghai 200433, China
Email: xinhuang@fudan.edu.cn
Zhaojun Bai
Department of Computer Science and Department of Mathematics, University of California, Davis, CA 95616, USA
Email: bai@cs.ucdavis.edu
Yangfeng Su
School of Mathematical Sciences, Fudan University, Shanghai 200433, China
Email: yfsu@fudan.edu.cn

Abstract

Nonlinear rank-one modification of the symmetric eigenvalue problem arises from eigenvibrations of mechanical structures with elastically attached loads and calculation of the propagation modes in optical fiber. In this paper, we first study the existence and uniqueness of eigenvalues, and then investigate three numerical algorithms, namely Picard iteration, nonlinear Rayleigh quotient iteration and successive linear approximation method (SLAM). The global convergence of the SLAM is proven under some mild assumptions. Numerical examples illustrate that the SLAM is the most robust method.

Mathematics subject classification: 65F15, 65H17, 15A18, 35P30, 65Y20.
Key words: Nonlinear eigenvalue problem, Rank-one modification, Rank-one damping, Low-rank damping, Picard, Successive linear approximation method, Nonlinear Rayleigh quotient iteration, Safeguard, Global convergence.

1. Introduction

We consider a nonlinear rank-one modification of the symmetric eigenvalue problem of the form

$$\left[A + s(\lambda)uu^T\right] x = \lambda x, \tag{1.1}$$

where A is a real symmetric matrix, u is a real column vector and $s(\lambda)$ is a real-valued continuous and differentiable function. The problem (1.1) is an extension of the well-known rank-one modification of symmetric eigenvalue problem $(A+\rho uu^T)x = \lambda x$, where ρ is a real constant [1,2]. The nonlinear rank-one modification problem (1.1) arises from the study of eigenvibrations of mechanical structures with elastically attached loads [3, 4], and calculation of the propagation modes of a circular optical fiber [5, 6].

In section 2 of this paper, we study the existence of eigenvalues of (1.1) under proper assumptions of the function $s(\lambda)$. An interlacing property between eigenvalues of (1.1) and the symmetric matrix A is given. Three numerical algorithms are presented in section 3. In particular, the global convergence of the SLAM is established. In section 4, we compare numerical

* Received March 19, 2009 / Revised version received May 11, 2009 / Accepted June 16, 2009 / Published online December 21, 2009 /

performance of three algorithms for examples ranging from model problem to applications in mechanical structure analysis and fiber optic design. These numerical examples illustrate that the SLAM is the most robust method to solve the nonlinear rank-one modification problem (1.1).

2. Existence of Eigenvalues

Let us first recall the following two well-known theorems which describe the interlacing property between the eigenvalues of the symmetric matrix A and its rank-one updating matrix $A + \rho uu^T$, where ρ is a scalar.

Theorem 2.1. ([2], [7, p.442]) *Suppose that the diagonal entries of $D = diag(d_1, d_2, \ldots, d_n)$ are distinct and ordered such that $d_1 < d_2 < \cdots < d_n$. Assume the components of the vector u are nonzero. Let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be the eigenvalues of $D + \rho uu^T$. Then if $\rho > 0$, $d_1 < \lambda_1 < d_2 < \lambda_2 < \cdots < d_n < \lambda_n$, and if $\rho < 0$, $\lambda_1 < d_1 < \lambda_2 < d_2 < \cdots < \lambda_n < d_n$.*

Theorem 2.2. ([2], [7, p.397]) *Let $\mu_1 \le \mu_2 \le \cdots \le \mu_n$ be the eigenvalues of A and $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be the eigenvalues of $A + \rho uu^T$. Then if $\rho > 0$,*

$$\mu_1 \le \lambda_1 \le \mu_2 \le \lambda_2 \le \cdots \le \mu_n \le \lambda_n, \tag{2.1}$$

and if $\rho < 0$,

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \cdots \le \lambda_n \le \mu_n. \tag{2.2}$$

The following lemma shows that eigenvalue of $A + \rho uu^T$ is an increasing function of ρ.

Lemma 2.1. *Let λ_k and θ_k be the kth eigenvalues of symmetric matrices $A + \rho uu^T$ and $A + \tau uu^T$, respectively. If $\rho \ge \tau$, then $\lambda_k \ge \theta_k$.*

Proof. Note that $A+\rho uu^T$ can be written as a symmetric rank-one modification of $A+\tau uu^T$: $A + \rho uu^T = A + \tau uu^T + E$, where $E = (\rho - \tau)uu^T$. By Weyl's monotonicity theorem, see for example [8, p.203], we have $\lambda_k \ge \theta_k + \lambda_{\min}(E) \ge \theta_k + 0 = \theta_k$. □

Let us now turn to studying the existence of eigenvalues for the nonlinear rank-one modification eigenvalue problem (1.1). We begin with a special case of the form

$$\left[D + s(\lambda)uu^T\right] x = \lambda x, \tag{2.3}$$

where $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ and $d_1 < d_2 < \cdots < d_n$. Furthermore, the components u_i of u are nonzero.

Lemma 2.2. *(a) If $s(d_i) = 0$, then d_i is an eigenvalue of* (2.3). *(b) If $s(d_i) \ne 0$, then d_i is not an eigenvalue of* (2.3).

Proof. The statement (a) is obvious. The statement (b) can be proven by contradiction. If d_i is an eigenvalue of (2.3), then we have $[D + s(d_i)uu^T]x = d_i x$. By Theorem 2.1, the eigenvalues of $D + s(d_i)uu^T$ are strictly interlaced by eigenvalues of D. This leads to the contradiction $d_i < d_i$. □

Based on Lemma 2.2, for the simplicity of exposition, we assume $s(d_i) \ne 0$ for $i = 1, 2, \ldots, n$ for the rest of the paper.

Lemma 2.3. *If λ is an eigenvalue of* (2.3)*, then λ is real.*

Proof. The result is based on the assumption that $s(\lambda)$ is a real-valued function. □

Let λ be an eigenvalue of (2.3), we know from Lemma 2.2 that $D - \lambda I$ is nonsingular. Thus

$$\begin{aligned}\det(D + s(\lambda)uu^T - \lambda I) &= \det(I + s(\lambda)uu^T(D - \lambda I)^{-1}) \cdot \det(D - \lambda I)\\ &= (1 + s(\lambda)u^T(D - \lambda I)^{-1}u) \cdot \det(D - \lambda I).\end{aligned}$$

Note that the identity $\det(I + xy^T) = 1 + y^T x$ for vectors $x, y \in \mathbb{R}^n$ is used for the second equality. Hence we see that λ is an eigenvalue of (2.3) if and only if λ is a root of the function

$$f(\lambda) = 1 + s(\lambda)u^T(D - \lambda I)^{-1}u = 1 + s(\lambda)w(\lambda),$$

where

$$w(\lambda) = u^T(D - \lambda I)^{-1}u = \sum_{j=1}^{n} \frac{u_j^2}{d_j - \lambda}.$$

By Lemmas 2.2 and 2.3, we conclude that the existence of eigenvalues of the nonlinear eigenvalue problem (2.3) is equivalent to the existence of the roots of $f(\lambda)$ on the intervals $\mathcal{I}_k = (d_k, d_{k+1})$ for $k = 0, 1, \ldots, n$, where $d_0 = -\infty$ and $d_{n+1} = +\infty$.

Theorem 2.3. *Let $s(\lambda)$ be a positive decreasing function on $\mathcal{I}_k$, i.e., $s(\lambda) > 0$ and $s'(\lambda) \le 0$ for $\lambda \in \mathcal{I}_k$, then (a) If $k = 0$, there is no eigenvalue of* (2.3) *on $\mathcal{I}_0$. (b) If $1 \le k \le n$, there is a simple eigenvalue of* (2.3) *on $\mathcal{I}_k$.*

Proof. The statement (a) is established immediately based on the fact that on the interval $\mathcal{I}_0 = (-\infty, d_1)$, $w(\lambda) > 0$ and $f(\lambda) = 1 + s(\lambda)w(\lambda) > 1$.

The proof of (b) consists of two parts. First, we consider the existence of eigenvalues in $\mathcal{I}_k$. Then we prove the uniqueness and simplicity.

Consider the interval $\mathcal{I}_k = (d_k, d_{k+1})$ where $1 \le k \le n-1$. On the left end of $\mathcal{I}_k$, $w(\lambda) \to -\infty$ as $\lambda \to d_k+$. Thus for any $\tilde{\delta} \in (d_k, d_{k+1})$, there exists $\delta_1 \in (d_k, \tilde{\delta})$ such that $w(\delta_1) < -1/s(\tilde{\delta}) < 0$ and

$$f(\delta_1) = 1 + s(\delta_1)w(\delta_1) < 1 - s(\tilde{\delta})/s(\tilde{\delta}) = 0.$$

On the right end of $\mathcal{I}_k$, $w(\lambda) \to +\infty$ as $\lambda \to d_{k+1}-$. Thus there exists $\delta_2 \in (\tilde{\delta}, d_{k+1})$ such that $w(\delta_2) > 0$ and $f(\delta_2) = 1 + s(\delta_2)w(\delta_2) > 1$. Since $f(\lambda)$ is continuous on $\mathcal{I}_k$, $f(\delta_1) < 0$ and $f(\delta_2) > 0$, we conclude that $f(\lambda)$ has at least a root in $(\delta_1, \delta_2) \subset \mathcal{I}_k$ for $1 \le k \le n-1$.

For $k = n$, on the interval $\mathcal{I}_n = (d_n, +\infty)$, we have $w(\lambda) \to -\infty$ as $\lambda \to d_n+$. Hence for any $\tilde{\delta} > d_n$, there exists $\delta_1 \in (d_n, \tilde{\delta})$ and $w(\delta_1) < -1/s(\tilde{\delta}) < 0$. Therefore,

$$f(\delta_1) = 1 + s(\delta_1)w(\delta_1) < 1 - s(\tilde{\delta})/s(\tilde{\delta}) = 0.$$

On the other hand, $w(\lambda) < 0$ for $\lambda > d_n$ and $w(\lambda) \to 0$ as $\lambda \to +\infty$, Hence for the same $\tilde{\delta}$, there exists $\delta_2 \in (\tilde{\delta}, +\infty)$ such that $-1/s(\tilde{\delta}) < w(\delta_2) < 0$. Consequently, we have

$$f(\delta_2) = 1 + s(\delta_2)w(\delta_2) > 1 - s(\tilde{\delta})/s(\tilde{\delta}) = 0.$$

Since $f(\delta_1) < 0$ $f(\delta_2) > 0$, and $f(\lambda)$ is continuous on $\mathcal{I}_n$, we conclude that $f(\lambda)$ has at least one root in $(\delta_1, \delta_2) \subset \mathcal{I}_n$. This completes the proof of the existence.

We now show that $f(\lambda)$ has only one root in the interval $\mathcal{I}_k$ if $1 \le k \le n$. Assume $f(\lambda)$ has two distinct roots $\alpha, \beta \in \mathcal{I}_k$ and $\alpha < \beta$. Let us define

$$f_\alpha(\lambda) = 1 + s(\alpha)w(\lambda) \text{ and } f_\beta(\lambda) = 1 + s(\beta)w(\lambda).$$

Since $w'(\lambda) > 0$ and $s(\alpha) \ge s(\beta) > 0$,

$$f'_\alpha(\lambda) = s(\alpha)w'(\lambda) > 0 \text{ and } f'_\beta(\lambda) = s(\beta)w'(\lambda) > 0.$$

Therefore, $f_\alpha(\lambda)$ and $f_\beta(\lambda)$ are strictly increasing, and $0 = f_\alpha(\alpha) < f_\alpha(\beta)$ and $f_\beta(\alpha) < f_\beta(\beta) = 0$. On the other hand, by $f_\alpha(\alpha) = 1 + s(\alpha)w(\alpha) = 0$, we have $w(\alpha) = -1/s(\alpha) < 0$. It implies that

$$f_\beta(\alpha) = 1 + s(\beta)w(\alpha) = 1 - \frac{s(\beta)}{s(\alpha)} \ge 0.$$

This is a contradiction. Hence, $\alpha = \beta$.

Finally, we prove the root α is simple. Note that $f'(\alpha) = s'(\alpha)w(\alpha) + s(\alpha)w'(\alpha)$. Since we have shown $w(\alpha) < 0$ and $w'(\alpha) > 0$ and s is positive decreasing, $f'(\alpha) > 0$. It implies that α is a simple root of $f(\lambda)$. □

Let us explain why we need to assume that $s(\lambda)$ is a decreasing function. The nonlinear eigenvalue problem (2.3) can be cast as a parameterized eigenvalue problem:

$$\begin{cases} (D + s(\mu)uu^T)x = \lambda(\mu)x, \\ \mu = \lambda(\mu). \end{cases} \tag{2.4}$$

Denote the kth eigenvalue of the eigenproblem in (2.4) as $\lambda_k(\mu)$ for an arbitrarily fixed μ. By Lemma 2.1, $\lambda_k(\mu)$ is a monotonic function of $s(\mu)$. Since $s(\mu)$ is a decreasing function of μ, $\lambda_k(\mu)$ is also a decreasing function of μ. Hence, there exists μ_* such that $\mu_* = \lambda_k(\mu_*)$ from the intersection of the functions $\lambda = \lambda_k(\mu)$ and $\lambda = \mu$.

By an analogous proof of Theorem 2.3, we have the following theorem for the case where $s(\lambda)$ is a negative decreasing function.

Theorem 2.4. *Let $s(\lambda)$ be a negative decreasing function on $\mathcal{I}_k$, i.e., $s(\lambda) < 0$ and $s'(\lambda) \le 0$ for $\lambda \in \mathcal{I}_k$, then (a) If $0 \le k \le n-1$, there is a simple eigenvalue of (2.3) on $\mathcal{I}_k$. (b) If $k = n$, there is no eigenvalue of (2.3) on $\mathcal{I}_n$.*

Now let us consider the general case (1.1). Let $Q^TAQ = D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$ be an eigendecomposition of A and $Q^TQ = I$ and $d_1 \le d_2 \le \cdots \le d_n$. Then the problem (1.1) can be transformed into the following nonlinear rank-one modification of the diagonal matrix D:

$$\left[D + s(\lambda)uu^T\right] x = \lambda x, \tag{2.5}$$

where $u := Q^Tu$ and $x := Q^Tx$. Similar to the standard rank-one modification of a diagonal matrix (see [2,9] or [10, p.221]), it can be shown that if $d_i = d_{i+1}$ or $u_i = 0$ for some $1 \le i \le n$, then d_i is an eigenvalue of (2.5) and (1.1). We can use a *deflation procedure* to remove these eigenvalues and reduce the problem (2.5) to the form (2.3). By combining the transformation (2.5), the deflation and Theorem 2.3, we have the following theorem about the existence of eigenvalues of the problem (1.1).

Theorem 2.5. *(a) If $s(\lambda)$ is positive decreasing on $\mathcal{I}_0 = (-\infty, d_1)$, then there is no eigenvalue of* (1.1) *on $\mathcal{I}_0$. (b) If $s(\lambda)$ is positive decreasing on $\mathcal{I}_k^c = [d_k, d_{k+1}]$, where $1 \le k \le n$, then there is at least one eigenvalue of* (1.1) *on $\mathcal{I}_k^c$. Meanwhile, there is no more than one eigenvalue of* (1.1) *on $\mathcal{I}_k = (d_k, d_{k+1})$.*

We learn from Theorem 2.5 that there are possibly more than one eigenvalue of (1.1) on the interval $\mathcal{I}_k^c = [d_k, d_{k+1}]$. For indexing these eigenvalues and deriving an interlacing property, we first give the following lemma.

Lemma 2.4. *Among the distinct eigenvalues of* (1.1) *on $\mathcal{I}_k^c$, there is only one eigenvalue, denoted as λ_*, such that λ_* is the kth eigenvalue of $A + s(\lambda_*)uu^T$.*

Proof. Let us first consider the most common case that there is an eigenvalue $\widehat{\lambda}$ of (1.1) such that $\widehat{\lambda} \in \mathcal{I}_k = (d_k, d_{k+1})$ for some k, i.e.,

$$d_k < \widehat{\lambda} < d_{k+1}. \tag{2.6}$$

Denote the eigenvalues of the matrix $A + s(\widehat{\lambda})uu^T$ as $\widehat{\lambda}_1 \le \widehat{\lambda}_2 \le \cdots \le \widehat{\lambda}_n$. By Theorem 2.2, we have

$$d_1 \le \widehat{\lambda}_1 \le \cdots \le d_k \le \widehat{\lambda}_k \le d_{k+1} \le \cdots \le d_n \le \widehat{\lambda}_n. \tag{2.7}$$

By (2.6) and (2.7), we conclude that $\widehat{\lambda} = \widehat{\lambda}_k$ is the kth eigenvalue of $A + s(\widehat{\lambda})uu^T$. Hence $\lambda_* = \widehat{\lambda}$. Moreover, by a similar proof as of Theorem 2.3, such λ_* is a simple eigenvalue of the deflated problem of the form (2.3). Hence λ_* is a simple eigenvalue of the problem (1.1).

Now let us consider the case where the deflation occurs that there is no eigenvalue of (1.1) on $\mathcal{I}_k$. Consider the case d_k is an eigenvalue of (1.1) of multiplicity 1, while d_{k-1} and d_{k+1} are not eigenvalues of (1.1). Without loss of generality, assume that $s(\lambda)$ is positive decreasing on (d_{k-1}, d_{k+1}). Then the problem (1.1) has one simple eigenvalue on $(d_{k-1}, d_k]$, denoted as $\widehat{\lambda}$. We can see that $\widehat{\lambda}$ is the $(k-1)$st eigenvalue of the matrix $A + s(\widehat{\lambda})uu^T$. Denote the $(k-1)$st eigenvalue of $A + s(d_k)uu^T$ as $\widetilde{\lambda}$. By Lemma 2.1, since $s(d_k) \le s(\widehat{\lambda})$, we have $d_{k-1} < \widetilde{\lambda} \le \widehat{\lambda} \le d_k$. Thus we can conclude that d_k is the kth eigenvalue of the matrix $A + s(d_k)uu^T$ and $\lambda_* = d_k$.

For the other possible cases when the deflation occurs, we can use similar arguments to conclude that either d_k or d_{k+1} is λ_*.

Finally, let us show that λ_* is unique. Assume there are two distinct eigenvalues λ_* and $\lambda_\sharp$ in $\mathcal{I}_k^c$ such that they are the kth eigenvalues of the matrices $A + s(\lambda_*)uu^T$ and $A + s(\lambda_\sharp)uu^T$, respectively. Let $\lambda_* < \lambda_\sharp$. Since the function $s(\lambda)$ is decreasing, $s(\lambda_*) \ge s(\lambda_\sharp)$, by Lemma 2.1, we have $\lambda_* \ge \lambda_\sharp$, which is a contradiction. Hence $\lambda_* = \lambda_\sharp$. □

Combining Theorems 2.3 and 2.5 and Lemma 2.4, we have the following theorem to describe the existence of eigenvalues of (1.1) and obtain an interlacing property. It can be viewed as a generalization of the interlacing property (2.1) in Theorem 2.2 for the rank-one modification of a symmetric matrix.

Theorem 2.6. *If $s(\lambda)$ is positive decreasing on $(-\infty, \infty)$, then the nonlinear eigenvalue problem* (1.1) *has n real eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. Furthermore, these eigenvalues satisfy the interlacing property $d_1 \le \lambda_1 \le d_2 \le \lambda_2 \le \cdots \le d_n \le \lambda_n$.*

By an analogous derivation, we have the following theorem for the case when the function $s(\lambda)$ is negative decreasing.

Theorem 2.7. *If $s(\lambda)$ is negative decreasing on $(-\infty, \infty)$, then the nonlinear eigenvalue problem* (1.1) *has n real eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. Furthermore, these eigenvalues satisfy the interlacing property $\lambda_1 \le d_1 \le \lambda_2 \le d_2 \le \cdots \le \lambda_n \le d_n$.*

3. Algorithms

In this section, we study numerical algorithms for extracting an eigenpair of the nonlinear rank-one modification eigenvalue problem (1.1). Specifically, we consider the problem of finding an eigenpair of (1.1) on the interval $\mathcal{I}_k^c = [d_k, d_{k+1}]$, where d_k and d_{k+1} are two consecutive eigenvalues of the symmetric matrix A. We learn from Lemma 2.4 that if $s(\lambda)$ is positive decreasing on $\mathcal{I}_k^c$, then among possible eigenvalues of (1.1) on $\mathcal{I}_k^c$, there is only one eigenvalue, denoted as λ_*, such that λ_* is the kth eigenvalue of the matrix $A + s(\lambda_*)uu^T$. Hence, the objective of algorithms in this section is to compute this eigenpair (λ_*, x_*).

3.1. Picard iteration

Picard iteration is a fixed-point iterative method to find a root of nonlinear equations, see for example [11]. To compute the eigenpair (λ_*, x_*) on the interval $\mathcal{I}_k^c$ by the Picard iteration, the ith iterate $(\lambda^{(i)}, x^{(i)})$ is the kth eigenpair of the linear symmetric eigenvalue problem

$$\left[A + s(\lambda^{(i-1)})uu^T\right] x = \lambda x, \tag{3.1}$$

where the initial approximate eigenpair $(\lambda^{(0)}, x^{(0)})$ is chosen such that $\lambda^{(0)} \in \mathcal{I}_k^c$.

In practice, the Picard iteration is combined with a safeguard strategy for avoiding misconvergence. Safeguard strategy is a generic technique in the root-finding methods [11–13]. To apply the safeguard strategy in the Picard iteration (3.1), let $[\lambda_l, \lambda_u]$ be the search interval for the desired eigenvalue λ_* and initially $[\lambda_l, \lambda_u] = \mathcal{I}_k^c = [d_k, d_{k+1}]$. There are two steps to update the interval $[\lambda_l, \lambda_u]$:

1. If the ith iterate $\lambda^{(i)}$ is at outside the search interval, we bracket it with the lower and upper bounds λ_l and λ_u:

$$\lambda^{(i)} = \begin{cases} \alpha\lambda_l + (1-\alpha)\lambda_u, & \text{if} \quad \lambda^{(i)} < \lambda_l, \\ (1-\alpha)\lambda_l + \alpha\lambda_u, & \text{if} \quad \lambda^{(i)} > \lambda_u, \end{cases}$$

   where $0.5 \le \alpha < 1$ is a relaxation parameter.

2. The search interval is tightened:

$$\begin{cases} \lambda_l = \lambda^{(i)}, & \text{if} \quad \lambda^{(i)} < \lambda_*, \\ \lambda_u = \lambda^{(i)}, & \text{if} \quad \lambda^{(i)} > \lambda_*. \end{cases} \tag{3.2}$$

As we can see, the key issue to implement the safeguard strategy is to compare $\lambda^{(i)}$ with the target λ_* at step 2. Let us first show a "jumping" behavior of the Picard iteration.

Lemma 3.1. *Let $\lambda^{(i)}$ and $\lambda^{(i+1)}$ be the ith and $(i+1)$st iterates of the Picard iteration* (3.1). *(a) If $\lambda^{(i)} \le \lambda_*$, then $\lambda^{(i+1)} \ge \lambda_*$. (b) If $\lambda^{(i)} \ge \lambda_*$, then $\lambda^{(i+1)} \le \lambda_*$.*

Proof. Note that $\lambda^{(i+1)}$ and λ_* are the kth eigenvalues of matrices $A + s(\lambda^{(i)})uu^T$ and $A + s(\lambda_*)uu^T$, respectively. If $\lambda^{(i)} \leq \lambda_*$, then $s(\lambda^{(i)}) \geq s(\lambda_*)$. Recall that $s(\lambda)$ is a decreasing function. By Lemma 2.1, $\lambda^{(i+1)} \geq \lambda_*$. This leads to result (a). The result (b) can be proven by a similar argument.

Lemma 3.1 implies that if $\lambda^{(i+1)} < \lambda^{(i)}$ then $\lambda^{(i)} > \lambda_*$, and if $\lambda^{(i+1)} > \lambda^{(i)}$ then $\lambda^{(i)} < \lambda_*$. By this fact, step 2 can be implemented by comparing two consecutive Picard iterates $\lambda^{(i)}$ and $\lambda^{(i+1)}$.

In summary, we have the following safeguarded Picard iteration:

Algorithm3.1 Safeguarded Picard iteration

INPUT: initial $(\lambda^{(0)}, x^{(0)})$, and $\lambda^{(0)} \in [\lambda_l, \lambda_u] = [d_k, d_{k+1}]$
OUTPUT: approximate eigenpair $(\lambda^{(i)}, x^{(i)})$

```
1.  for i = 1, 2, ... until convergence
2.      compute the kth eigenpair (λ^(i), x^(i)) of (3.1)
3.      if λ^(i) > λ^(i−1)
4.          λ_l = λ^(i−1)
5.      elseif λ^(i) < λ^(i−1)
6.          λ_u = λ^(i−1)
7.      end if
8.      if λ^(i) < λ_l
9.          λ^(i) = αλ_l + (1 − α)λ_u
10.     elseif λ^(i) > λ_u
11.         λ^(i) = (1 − α)λ_l + αλ_u
12.     end if
13. end for
```

There is no known global convergence property of the Picard iteration even with the safeguard strategy. It could stagnate such that $\lambda^{(i)} = \lambda^{(i+2)}$. If the iteration converges, the rate is typically linear. However, in the case that $\lambda_* = d_k$ or d_{k+1}, Picard could converge in one step, see section 4 for a numerical example.

3.2. Nonlinear Rayleigh quotient iteration

It is easy to see that an eigenpair (λ, x) of the eigenvalue problem (1.1) is a solution of the nonlinear equations

$$\begin{bmatrix} T(\lambda)x \\ x^T T(\lambda)x/2 \end{bmatrix} = 0, \tag{3.3}$$

where $T(\lambda) = A + s(\lambda)uu^T - \lambda I$. Note that the term $x^T T(\lambda)x = 0$ is related to as a nonlinear Rayleigh quotient [14].

By using Newton's method to solve the nonlinear equations (3.3), we derive the following iteration:

$$\begin{bmatrix} T(\lambda^{(i)}) & T'(\lambda^{(i)})x^{(i)} \\ (T(\lambda^{(i)})x^{(i)})^T & (x^{(i)})^T T'(\lambda^{(i)})x^{(i)}/2 \end{bmatrix} \begin{bmatrix} x^{(i+1)} - x^{(i)} \\ \lambda^{(i+1)} - \lambda^{(i)} \end{bmatrix} = - \begin{bmatrix} T(\lambda^{(i)})x^{(i)} \\ (x^{(i)})^T T(\lambda^{(i)})x^{(i)}/2 \end{bmatrix},$$

where $T'(\lambda) = s'(\lambda)uu^T - I$. Equivalently, the iteration can be written as

$$\begin{cases} x^{(i+1)} = \tau_i\, T(\lambda^{(i)})^{-1}T'(\lambda^{(i)})x^{(i)}, \\ \lambda^{(i+1)} = \lambda^{(i)} - (x^{(i)})^T T(\lambda^{(i)})x^{(i)}/(x^{(i)})^T T'(\lambda^{(i)})x^{(i)}, \end{cases} \tag{3.4}$$

where τ_i is a scaling factor such that $\|x^{(i+1)}\| = 1$. When the newly computed vector $x^{(i+1)}$ is used instead of $x^{(i)}$ in the second equation of (3.4), we have the following nonlinear Rayleigh quotient iteration (NRQI):

$$\begin{cases} x^{(i+1)} = \tau_i\, T(\lambda^{(i)})^{-1}T'(\lambda^{(i)})x^{(i)}, \\ \lambda^{(i+1)} = \lambda^{(i)} - (x^{(i+1)})^T T(\lambda^{(i)})x^{(i+1)}/(x^{(i+1)})^T T'(\lambda^{(i)})x^{(i+1)}. \end{cases} \tag{3.5}$$

The safeguard strategy described in section 3.1 can also be used here. The key issue is again at the step to compare the current iterate $\lambda^{(i)}$ with the target eigenvalue λ_*. This can be resolved by using the inertia of $T(\lambda^{(i)})$.

Lemma 3.2. *Let $\lambda^{(i)}$ be the ith iterate of NRQI* (3.5) *and θ_k be the kth eigenvalue of $A + s(\lambda^{(i)})uu^T$. We have (a) if $\lambda^{(i)} \le \lambda_*$, then $\theta_k \ge \lambda_*$, and (b) if $\lambda^{(i)} \ge \lambda_*$, then $\theta_k \le \lambda_*$.*

Proof. The lemma is established by observing that θ_k can be regarded as $\lambda^{(i+1)}$ in the Picard iteration (3.1) and apply the proof of Lemma 3.1. □

By the contrapositive of Lemma 3.2(a), we know that if $\theta_k < \lambda^{(i)}$, then $\lambda^{(i)} > \lambda_*$. Similarly by Lemma 3.2(b), we know that if $\theta_k > \lambda^{(i)}$, then $\lambda^{(i)} < \lambda_*$. Therefore, the question is now turned to comparing θ_k with $\lambda^{(i)}$. By the fact that θ_k is the kth eigenvalue of the matrix $A + s(\lambda^{(i)})uu^T$, we immediately know that if the kth eigenvalue of $T(\lambda^{(i)})$ is negative, then $\theta_k < \lambda^{(i)}$. If it is positive, then $\theta_k > \lambda^{(i)}$. The sign of the kth eigenvalue of $T(\lambda^{(i)})$ can be obtained by the inertia of $T(\lambda^{(i)})$. Therefore, we have the following lemma to compare $\lambda^{(i)}$ with λ_*.

Lemma 3.3. *Let the number triple (ν_-, ν_0, ν_+) be the inertia of $T(\lambda^{(i)})$, where ν_-, ν_0, and ν_+ are the numbers of negative, zero and positive eigenvalues of $T(\lambda^{(i)})$ respectively. We have (a) If $k \le \nu_-$, then $\lambda^{(i)} > \lambda_*$. (b) If $\nu_- + \nu_0 < k$, then $\lambda^{(i)} < \lambda_*$. (c) If $\nu_- < k \le \nu_- + \nu_0$, then $\lambda^{(i)} = \lambda_*$.*

Combining the NRQI iteration (3.5) and Lemma 3.3, we have the following algorithm.

Algorithm 3.2 Safeguarded nonlinear Rayleigh quotient iteration (SNRQI)

INPUT: initial approximate eigenpair $(\lambda^{(0)}, x^{(0)})$ and $\lambda^{(0)} \in [\lambda_l, \lambda_u] = [d_k, d_{k+1}]$
OUTPUT: approximate eigenpair $(\lambda^{(i)}, x^{(i)})$

1. for $i = 1, 2, \ldots$ until convergence
2. compute the LDL$^{\mathrm{T}}$ decomposition: $T(\lambda^{(i-1)}) = LDL^T$
3. count the inertia (ν_-, ν_0, ν_+) from D
4. if $\nu_- + \nu_0 < k$
5. $\lambda_l = \lambda^{(i-1)}$
6. elseif $\nu_- \geq k$
7. $\lambda_u = \lambda^{(i-1)}$
8. end if
9. compute $\lambda^{(i)}$ and $x^{(i)}$ as (3.5)
10. if $\lambda^{(i)} < \lambda_l$
11. $\lambda^{(i)} = \alpha\lambda_l + (1-\alpha)\lambda_u$
12. elseif $\lambda^{(i)} > \lambda_u$
13. $\lambda^{(i)} = (1-\alpha)\lambda_l + \alpha\lambda_u$
14. end if
15. end for

To compute the inertia (ν_-, ν_0, ν_+) of $T(\lambda^{(i)})$, we can use the diagonal pivoting method [15, 16] to compute the LDL$^{\mathrm{T}}$ decomposition $T(\lambda^{(i)})$. The inertia (ν_-, ν_0, ν_+) can be obtained from D. A 2-by-2 diagonal block in D indicates a positive-negative pair of eigenvalues.

Based on the convergence property of the Newton's method [11, p.158], it can be shown the SNRQI converges locally with asymptotical quadratic rate. In practice, the SNRQI could initially converge slowly due to the bracketing process, see section 4 for numerical examples.

3.3. Successive linear approximation method

To improve the Picard iteration (3.1), we can use the first order approximation of the function $s(\lambda)$ at a prescribed point σ and derive a linear approximation of the nonlinear eigenvalue problem (1.1):

$$\left(A + \ell(\lambda, \sigma) uu^T\right) x = \lambda x, \tag{3.6}$$

where $\ell(\lambda, \sigma) = s(\sigma) + s'(\sigma)(\lambda - \sigma)$. To compute the targeted eigenpair (λ_*, x_*), the $(i+1)$st approximate eigenpair $(\lambda^{(i+1)}, x^{(i+1)})$ is chosen as the kth eigenpair of the following linear eigenvalue problem

$$\left(A + \ell(\lambda^{(i+1)}, \lambda^{(i)}) uu^T\right) x^{(i+1)} = \lambda^{(i+1)} x^{(i+1)}, \tag{3.7}$$

which can be equivalently written as the generalized symmetric eigenvalue problem

$$K(\lambda^{(i)})x = \lambda M(\lambda^{(i)})x, \tag{3.8}$$

where $K(\sigma) = A + (s(\sigma) - s'(\sigma)\sigma)\, uu^T$ and $M(\sigma) = I - s'(\sigma)uu^T$. Note that $K(\sigma)$ and $M(\sigma)$ are symmetric. Furthermore, M is positive definite since $s'(\sigma) \leq 0$.

This approach is known as the method of successive linear problems [14] or the successive linear approximation method (SLAM) [17]. The following is the pseudo-code.

Algorithm 3.3 SLAM

INPUT: initial approximate eigenpair $(\lambda^{(0)}, x^{(0)})$ and $\lambda^{(0)} \in [d_k, d_{k+1}]$
OUTPUT: approximate eigenpair $(\lambda^{(i+1)}, x^{(i+1)})$

1. for $i = 0, 1, \ldots$ until convergence
2. compute the kth eigenpair $(\lambda^{(i+1)}, x^{(i+1)})$ of (3.8)
3. end for

Let us turn to the study of convergence property of the SLAM under the assumption that $s(\lambda)$ is positive decreasing and concave upward, i.e.,

$$s(\lambda) > 0, \quad s'(\lambda) \le 0, \quad s''(\lambda) \ge 0 \quad \text{for} \quad \lambda \in \mathcal{I}_k^c. \tag{3.9}$$

First, we have the following lemma.

Lemma 3.4. *Under the assumption* (3.9), $s(\lambda) \ge \ell(\lambda, \sigma)$ *for* $\lambda, \sigma \in \mathcal{I}_k^c$.

Proof. By the Taylor expansion of $s(\lambda)$ at σ, we have $s(\lambda) - \ell(\lambda, \sigma) = s(\lambda) - s(\sigma) - s'(\sigma)(\lambda - \sigma) = \frac{1}{2}s''(\zeta)(\lambda - \sigma)^2 \ge 0$ for some $\zeta \in \mathcal{I}_k^c$. □

The following lemma shows that the iterates $\lambda^{(i)}$ generated by the SLAM are bounded and monotonic.

Lemma 3.5. *Under the assumption* (3.9), *for* $\lambda^{(0)} \in \mathcal{I}_k^c$, *the* i*-th iterate* $\lambda^{(i)}$ *of SLAM satisfies (a)* $\lambda^{(i)} \in \mathcal{I}_k^c$, *(b)* $\lambda^{(i)} \le \lambda_*$, *and (c)* $\lambda^{(i)} \le \lambda^{(i+1)}$.

Proof. (a) By induction, we just need to show that $\lambda^{(1)} \in \mathcal{I}_k^c$. Let $\rho = \ell(\lambda^{(1)}, \lambda^{(0)})$, then we have $\rho > 0$. Otherwise, if $\rho \le 0$, then by the interlacing property (2.2), we have $d_{k-1} \le \lambda^{(1)} \le d_k \le \lambda^{(0)}$. Since $\ell(\lambda, \lambda^{(0)})$ is decreasing,

$$\rho = \ell(\lambda^{(1)}, \lambda^{(0)}) \ge \ell(\lambda^{(0)}, \lambda^{(0)}) = s(\lambda^{(0)}) > 0.$$

This is a contradiction to $\rho \le 0$. Since $\rho > 0$, by the interlacing property (2.1), the kth eigenvalue $\lambda^{(1)}$ of the matrix $A + \rho uu^T$ satisfies $d_k \le \lambda^{(1)} \le d_{k+1}$.

(b) We prove $\lambda^{(i)} \le \lambda_*$ by contradiction. If $\lambda^{(i)} > \lambda_*$, then

$$\rho = \ell(\lambda^{(i)}, \lambda^{(i-1)}) \le \ell(\lambda_*, \lambda^{(i-1)}) \le s(\lambda_*),$$

where the first inequality is by the fact $\ell(\lambda, \lambda^{(i-1)})$ is decreasing and the second inequality is by Lemma 3.4. Recall that $\lambda^{(i)}$ and λ_* are the kth eigenvalues of the matrices $A + \rho uu^T$ and $A + s(\lambda_*)uu^T$ respectively. Then by Lemma 2.1, we have $\lambda^{(i)} \le \lambda_*$. This is a contradiction.

(c) This can also be shown by contradiction. If $\lambda^{(i)} > \lambda^{(i+1)}$, then we have $\lambda^{(i+1)} < \lambda^{(i)} \le \lambda_*$, where the second inequality is from (b). Since $\ell(\lambda, \lambda^{(i)})$ and $s(\lambda)$ are decreasing functions, we have

$$\rho = \ell(\lambda^{(i+1)}, \lambda^{(i)}) \ge \ell(\lambda^{(i)}, \lambda^{(i)}) = s(\lambda^{(i)}) \ge s(\lambda_*).$$

By Lemma 2.1, it leads $\lambda^{(i+1)} \ge \lambda_*$, which is a contradiction.

The following theorem shows that the SLAM is globally convergent.

Theorem 3.1. *Under the assumption* (3.9)*, the iterates* $\lambda^{(i)}$ *of SLAM converge to* λ_* *for any initial* $\lambda^{(0)} \in \mathcal{I}_k^c$*. Furthermore, the convergence rate is asymptotically quadratic if* $\lambda_* \in \mathcal{I}_k$*.*

Proof. By Lemma 3.5, the series $\{\lambda^{(i)}\}$ is monotonically increasing and bounded. Hence, it is convergent. Denote the limit as $\lambda^\dagger$, we have $\lambda^\dagger \in \mathcal{I}_k^c$ and $\lambda^\dagger \leq \lambda_*$.

We now show that $\lambda^\dagger = \lambda_*$. Since $\lim_{i\to\infty} K(\lambda^{(i)}) = K(\lambda^\dagger)$ and $\lim_{i\to\infty} M(\lambda^{(i)}) = M(\lambda^\dagger)$, then from the continuity of the eigenproblem (3.8), we have $\lambda^\dagger$ is the kth eigenvalue of the pencil $(K(\lambda^\dagger), M(\lambda^\dagger))$. Thus $\lambda^\dagger$ is the kth eigenvalue of $A + \ell(\lambda^\dagger, \lambda^\dagger)uu^T = A + s(\lambda^\dagger)uu^T$. By the uniqueness of λ_* from Lemma 2.4, we conclude $\lambda^\dagger = \lambda_*$.

Finally, we prove that the convergence rate is asymptotically quadratic if $\lambda_* \in \mathcal{I}_k$. For the simplicity of notation, let us denote the ith and $(i+1)$st iterates as $\phi = \lambda^{(i)}$ and $\psi = \lambda^{(i+1)}$, and write the equation (3.7) as

$$(A + \ell(\psi, \phi)uu^T - \psi I)x = 0, \tag{3.10}$$

where $\ell(\psi, \phi) = s(\phi) + s'(\phi)(\psi - \phi)$. After taking into the account of the orthogonal transformation (2.5) and the deflation, we know that ψ and ϕ satisfy the secular equation

$$g(\psi, \phi) = 1 + \ell(\psi, \phi)w(\psi) = 0 \tag{3.11}$$

with $g(\lambda_*, \lambda_*) = 0$. Since $s'(\phi) \leq 0$, $\rho = \ell(\psi, \phi) > 0$, $w(\psi) = -1/\ell(\psi, \phi) < 0$ and $w'(\psi) > 0$, we have

$$\begin{aligned}
\frac{\partial g}{\partial \psi} &= \frac{\partial \ell(\psi, \phi)}{\partial \psi} w(\psi) + \ell(\psi, \phi)w'(\psi) = s'(\phi)w(\psi) + \ell(\psi, \phi)w'(\psi) > 0, \\
\frac{\partial g}{\partial \phi} &= \frac{\partial \ell(\psi, \phi)}{\partial \phi} w(\psi) = (\psi - \phi)s''(\phi)w(\psi).
\end{aligned}$$

Hence we have

$$\frac{\partial g}{\partial \psi}(\lambda_*, \lambda_*) > 0 \quad \text{and} \quad \frac{\partial g}{\partial \phi}(\lambda_*, \lambda_*) = 0.$$

By the implicit function theorem, see for example [18], for a sufficient large i and ψ and ϕ are within a neighborhood of λ_*, there exists a function $\psi(\phi)$ such that $g(\psi(\phi), \phi) = 0$ and $\psi(\lambda_*) = \lambda_*$.

Note that

$$\psi'(\lambda_*) = -(\frac{\partial g}{\partial \phi})/(\frac{\partial g}{\partial \psi}) \mid_{\phi=\lambda_*, \psi=\lambda_*} = 0. \tag{3.12}$$

Furthermore, the second derivative

$$\psi''(\lambda_*) = \frac{-s''(\lambda_*)w(\lambda_*)}{s'(\lambda_*)w(\lambda_*) + s(\lambda_*)w'(\lambda_*)}$$

is bounded since the denominator is positive. By the Taylor expansion, we have

$$\begin{aligned}
\lambda^{(i+1)} - \lambda_* &= \psi - \lambda_* = \psi(\phi) - \psi(\lambda_*) \\
&= \psi'(\lambda_*)(\phi - \lambda_*) + \frac{1}{2}\psi''(\xi)(\phi - \lambda_*)^2 \\
&= \frac{1}{2}\psi''(\xi)(\phi - \lambda_*)^2 \\
&= \frac{1}{2}\psi''(\xi)(\lambda^{(i)} - \lambda_*)^2,
\end{aligned}$$

for some $\xi \in (\phi, \lambda_*)$. Therefore, we conclude that the iterates $\lambda^{(i)}$ converge asymptotically quadratically.

When the function $s(\lambda)$ is concave downward, by an analogous argument, we can show that the iterates $\lambda^{(i)}$ are monotonically decreasing and converge to λ_*. Note that in [17], the local convergence of the SLAM is established for the general nonlinear eigenvalue problem $T(\lambda)x = 0$ under the assumptions that $T(\lambda)$ is twice continuously differentiable, $T'(\lambda_*)$ is nonsingular and zero is a simple eigenvalue of $T'(\lambda_*)^{-1}T(\lambda_*)$.

4. Numerical Examples

In this section, we provide numerical examples to demonstrate the theory and algorithms presented in sections 2 and 3. The safeguarded RQI [6] is used to solve the standard eigenvalue problem (3.1) in the Picard iteration and the generalized symmetric positive definite eigenvalue problem (3.8) in the SLAM. The bracketing parameter α of the safeguard strategy is set to be $\alpha = 0.8$. The residual norm of a computed eigenpair $(\widehat{\lambda}, \widehat{x})$ is defined as $\|r\| = \|(A + s(\widehat{\lambda})uu^T)\widehat{x} - \widehat{\lambda}\widehat{x}\|/\|\widehat{x}\|$. When $\|r\| \le \tau = 10^{-12}$, the computed eigenpair $(\widehat{\lambda}, \widehat{x})$ is declared to be convergent. All numerical data were obtained with Matlab implementations of the algorithms on an Intel Core2 Duo T7500 2.2Ghz with 2GB RAM.

Example 1. Consider an artificial nonlinear rank-one modification eigenvalue problem of a diagonal matrix:

$$\left[D + s(\lambda)uu^T\right] x = \lambda x, \tag{4.1}$$

where $D = \mathrm{diag}(1, 2, \ldots, n)$, $s(\lambda) = -\tan^{-1}(\lambda) - 3$ and $u = \frac{1}{\sqrt{n}}\mathbf{1}_n$, $\mathbf{1}_n$ is the vector of length n whose elements are all ones. Note that $s(\lambda) < 0$ and $s'(\lambda) = -1/(1 + \lambda^2) < 0$ on $(-\infty, +\infty)$. By Theorem 2.4, we conclude that there is a unique eigenvalue of (4.1) on each interval $\mathcal{I}_0 = (-\infty, 1)$ and $\mathcal{I}_k = (k, k + 1)$ for $k = 1, 2, \ldots, n - 1$. There is no eigenvalue on $(n, +\infty)$.

For numerical experiment, let us set $n = 100$ and compute the smallest eigenvalue $\lambda_1 \in (-\infty, 1)$. Let the initial guess $(\lambda^{(0)}, x^{(0)}) = (0, u)$. As shown in the first plot of Figure 4.1, the safeguarded Picard iteration converges linearly. The SNRQI takes 3 initial steps before the quadratic convergence rate occurs. The slow convergence of the SNRQI at the initial steps becomes more obvious when we try to compute the largest eigenvalue $\lambda_{100} \in (99, 100)$ with the initial guess $(\lambda^{(0)}, x^{(0)}) = (99.2, u)$. It takes more than 8 steps before the quadratic convergence rate occurs, see the second plot of Figure 4.1. We note that if there is no the safeguard strategy, it converges to the eigenvalue in the interval $(98, 99)$, not the targeted one in $(99, 100)$. By contrast, SLAM converges in only 2 and 3 steps in both cases.

The convergence of the Picard iteration strongly depends on the slope of the function $s(\lambda)$. For example, if let $s(\lambda) = -a(\tan^{-1} + 3)$, then the Picard iteration takes 41 steps to converge when $a = 5$, and stagnates at the residual norm $\|r\| = \mathcal{O}(1)$ when $a = 10$. On the contrary, the slope of the function $s(\lambda)$ does not affect on the convergence speed of the SNRQI and SLAM algorithms.

Finally, we note that the safeguarded Picard and SLAM converge in one iteration if the targeted eigenvalue λ_* is one of the end-points of the interval $\mathcal{I}_k^c = [d_k, d_{k+1}]$. For example, let the second element u_2 of the vector u be zero in (4.1). Then the target eigenvalue is $\lambda_2 = 2$ in $\mathcal{I}_1^c = [1, 2]$. With the initial approximate eigenpair $(\lambda^{(0)}, x^{(0)}) = (1.5, \mathbf{1}_n/\sqrt{n})$, both the Picard and SLAM algorithms take only one step to converge to $\lambda_2 = 2$. On the other hand, the SNRQI takes 4 iterations.

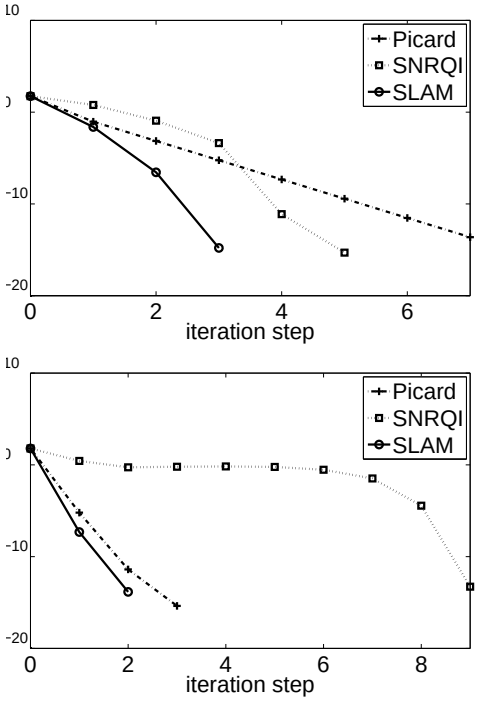

Fig. 4.1. Convergence history of the safeguarded Picard, SNRQI and SLAM. First: to compute λ_1. Second: to compute λ_{100}.

Example 2. This is an example arising from eigenvibrations of a mechanical structure with elastically attached load, such as beam, plate or shell [3, 19]. Specifically, we consider the eigenvibration of a string with a load of mass M attached by an elastic spring of stiffness K modeled by the differential eigenvalue problem

$$\begin{cases} -u''(x) = \lambda u(x), \quad x \in (0,1), \\ u(0) = 0, \quad -u'(1) = \varphi(\lambda) u(1), \end{cases} \tag{4.2}$$

where $\varphi(\lambda) = \lambda\gamma/(\lambda - \gamma)$, and $\gamma = K/M$. The eigenvalue $\lambda \in (\gamma, \infty)$ and the corresponding function $u(x)$ are the eigenfrequency and displacement of the string under natural oscillations.

Let us partition the interval $[0, 1]$ by the nodes $x_j = jh$ for $j = 0, 1, \ldots, n$, where $h = 1/n$. Then using the finite element discretization with the piecewise linear basis functions, we derive the following $n \times n$ algebraic eigenvalue problem

$$\left[A + \varphi(\lambda) e_n e_n^T\right] y = \lambda B y, \tag{4.3}$$

where A and B are tridiagonal matrices

$$A = \frac{1}{h} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix}, \quad B = \frac{h}{6} \begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix},$$

and $e_n = (0, 0, \ldots, 0, 1)^T$.

Note that $\varphi(\lambda)$ is positive decreasing and concave upward on the interval (γ, ∞), i.e., $\varphi(\lambda) > 0$, $\varphi'(\lambda) < 0$ and $\varphi''(\lambda) > 0$ for $\lambda > \gamma > 0$. Let $d_1 \le d_2 \le \cdots \le d_n$ be the eigenvalues of the matrix pair (A, B). Then we can easily verify that for $i = 1, \ldots, n$, d_i are not the eigenvalues of the problem (4.3) by the fact that

$$\begin{aligned}\det(A - d_i B + \varphi(d_i) e_n e_n^T) &= p_n(d_i) + (-1)^{n+1}\varphi(d_i) p_{n-1}(d_i) \\ &= (-1)^{n+1}\varphi(d_i) p_{n-1}(d_i) \neq 0,\end{aligned}$$

where $p_j(d_i)$ is the jth leading principal minor of the tridiagonal matrix $A - d_i B$. Therefore, if $\gamma < d_1$, by Theorem 2.3, we conclude that on the interval $(\gamma, +\infty)$, the nonlinear eigenvalue problem (4.3) has n distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ satisfying the interlacing property $d_1 < \lambda_1 < d_2 < \lambda_2 < \cdots < d_n < \lambda_n$.

Let us choose $K = M = 1$, $\gamma = 1$ and $h = 0.01$ ($n = 100$). To find the smallest eigenvalue λ_1, we compute the smallest eigenvalue of the linearized eigenvalue problem (3.8) in the iteration of the SLAM. Similarly, to find the second smallest eigenvalue λ_2, we compute the second smallest eigenvalue of the linearized eigenvalue problem (3.8). The following table shows the computed smallest four eigenvalues and the number of iterations using the same initial approximation $(\lambda^{(0)}, x^{(0)}) = (1.5, \mathbf{1}_n/\sqrt{n})$:

i	λ_i	iter	$\mathcal{I}_i = (d_i, d_{i+1})$
1	4.48217654587649	5	(2.46745, 22.2107)
2	24.2235731125643	4	(22.2107, 61.7167)
3	63.7238211419440	3	(61.7167, 121.025)
4	123.031221067616	3	(121.025, 200.193)

By the table, we see that to find $\lambda_i \in \mathcal{I}_i$, it is not necessary to compute the eigenvalues d_i and d_{i+1} of the matrix pair (A, B) explicitly. On the other hand, to use the safeguarded Picard iteration and SNRQI, we must first compute these eigenvalues d_i and d_{i+1}, and then choose an initial approximate eigenvalue $\lambda^{(0)} \in \mathcal{I}_i$.

Example 3. The nonlinear rank-one modification of the symmetric eigenvalue problem (1.1) is a computational kernel in the calculation of propagation modes in the fiber optical design [5,6]. Following the presentation of [5], we know that optical fibers are composed of fine threads of glass in layers, called the core and cladding. Consider an optical fiber which is ideally perfectly straight, circular, and uniform along its length, the cross section can be divided into core area and cladding. In the cylindrical coordinate system (r, θ, z), the Maxwell's equations on the guided wave function $f(r)$ are reduced to the scalar wave equation:

$$\left[\frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right) - \frac{m^2}{r^2} + \left(k^2(r) - k_{cl}^2\right)\right] f = (\beta^2 - k_{cl}^2) f, \tag{4.4}$$

where β is an unknown propagation constant, r is a radius in the cross section, m is an integer of the mode number (we only consider the case $m > 0$), $k(r)$ is the wave number given by $k(r) = 2\pi\eta(r, l)/l$, $\eta(r, l)$ is the refractive index at r and l is the light's vacuum wavelength. In addition, $k_{cl} = 2\pi\eta_{cl}/l$ which makes $k^2(r) - k_{cl}^2 = 0$ in the cladding, where η_{cl} denotes the refractive index in the cladding.

Assume that the cladding extends infinitely, then at a sufficient large radius r, the wave function f is of the form

$$f(r) = aK_m(\mu r),$$

where $\mu = (\beta^2 - k_{cl}^2)^{1/2}$, $K_m(z)$ is the mth order modified Bessel function of the second kind and a is an unknown scaling factor. At the center, we have the boundary condition $f(0) = 0$. On the other hand, since

$$\frac{f'(r)}{f(r)} = \mu \frac{K_m'(\mu r)}{K_m(\mu r)} \equiv G_r(\mu),$$

we have the Dirichlet-to-Neumann boundary condition at a cutoff radius R:

$$f'(R) = f(R)G_R(\mu).$$

A challenge of transmission of light in (single mode) optical fiber is chromatic dispersion [20], which leads to spreading of light pulses as they travel down the fiber (signal distortion). In the fiber optical design, when considering dispersion, derivatives of β are required. Therefore, we need to compute the propagation constant β first.

By using the finite element method to discretize the scalar wave equation (4.4), we first set $n+2$ evenly spaced sample points $r_i = i\delta$ for $i = 0, 1, \ldots, n+1$ along r with mesh size δ. The cutoff radius $R = r_n = n\delta$. The node number at the core radius is denoted as n_c such that $R_c = n_c\delta$. With the piecewise linear basis functions and proper treatment of the boundary conditions, the differential eigenvalue problem (4.4) is approximated by the following nonlinear algebraic eigenvalue problem

$$\left[A + s(\lambda)e_n e_n^T\right] x = \lambda x \quad \text{for} \quad \lambda > 0, \tag{4.5}$$

where $\lambda = \mu^2\delta^2$, A is a symmetric tridiagonal matrix with diagonal

$$a_{ii} = \begin{cases} -2 - m^2/i^2 + \delta^2(k^2(r_i) - k_{cl}^2), & 1 \le i \le n_c, \\ -2 - m^2/i^2, & n_c < i < n, \\ -1 + 1/(2n) - m^2/n^2, & i = n, \end{cases}$$

with $k(r_i) = 2\pi\eta(r_i, l)/l$ and $k_{cl} = 2\pi\eta_{cl}/l$. The sub- and super-diagonal elements of A are $a_{i+1,i} = a_{i,i+1} = (i+0.5)/\sqrt{i(i+1)}$ for $i = 1, 2, \ldots, n-1$, the function $s(\lambda)$ is of the following form:

$$s(\lambda) = \frac{n+0.5}{n}\sqrt{\lambda}\frac{K_m'(n\sqrt{\lambda})}{K_m(n\sqrt{\lambda})},$$

and the vector e_n is the last column of the identity matrix.

For the numerical experiment, we choose $m = 1$ and $l = 1.1\mu m$. Then we have

$$\eta(r_i, l) = \eta_{cl} + 1.45291C(r_i),$$

where the refractive index in the cladding is $\eta_{cl} = 1.4969$. $C(r_i)$ is a dopant concentration and has the expression

$$C(r_i) = \left(\frac{1 - 2\gamma(i/n_c)^\alpha}{1 - 2\gamma}\right)^{\frac{1}{2}} - 1 \ \ \textit{for } 0 \le i \le n_c,$$

with $\alpha = 25$ and $\gamma = 0.003$. The cutoff radius R and the core radius R_c of the fiber are set as $R = 6R_c = 24\mu m$. The mesh size is $\delta = 0.01\mu m$. As a result, we have $n_c = 400$ and the order of the problem (4.5) is $n = 6n_c = 2400$.

Note that $s(\lambda) < 0$, $s'(\lambda) < 0$ and $s''(\lambda) > 0$ for $\lambda > 0$ (see Figure 4.2). The matrix A has only one positive eigenvalue $d_n \approx 7.73 \times 10^{-7}$. Hence by Theorem 2.4, we conclude that the problem (4.5) has at most one eigenvalue on the interval $(0, d_n)$. Using the initial approximation $(\lambda^{(0)}, x^{(0)}) = (0, \mathbf{1}_n/\sqrt{n})$, the following table shows the computed eigenvalue of (4.5) on the interval $(0, d_n)$:

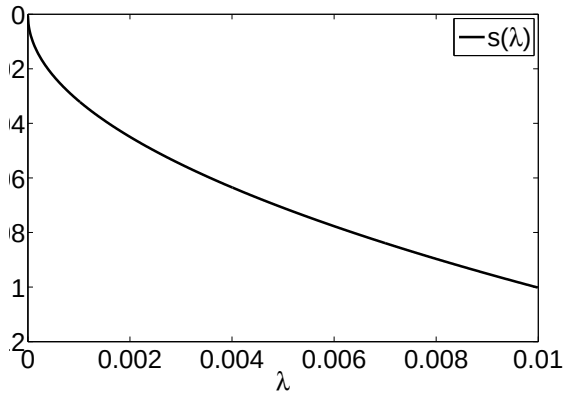

Fig. 4.2. Function $s(\lambda)$ for mode number $m = 1$ and problem size $n = 2400$.

Algorithm	iter	computed λ_*	residual
Picard	5	$7.13949431688754 \times 10^{-7}$	1.92×10^{-13}
SNRQI	4	$7.13949430586656 \times 10^{-7}$	4.95×10^{-20}
SLAM	3	$7.13949430685028 \times 10^{-7}$	4.88×10^{-16}

By the table, we observe that all three algorithms converge. The SLAM has fewest number of iterations. But the SNRQI is cheapest in terms of computational cost since at each iteration, the SNRQI only needs to solve a linear system of equations, instead of solving the linear eigenvalue problems in the Picard iteration and the SLAM.

5. Conclusion

The existence of eigenvalues for the nonlinear rank-one modification of the symmetric eigenvalue problem (1.1) is proven under the assumption that the function $s(\lambda)$ is positive decreasing. In this case, we also obtain the interlacing property (Theorem 2.6), which can be viewed as an extension of the well-known (constant) rank-one modification of the symmetric eigenvalue problem. The SLAM is shown to be the most robust method. The future work includes studying whether these results can be extended to other class of the function $s(\lambda)$. In general, it remains as a challenging problem about the existence of eigenvalues and efficient numerical methods for the general nonlinear eigenvalue problem of the form $T(\lambda)x = 0$, where the elements of the matrix $T(\lambda)$ are analytic functions of λ [21].

Acknowledgement. Huang was supported in part by China Scholarship Council while visiting University of California, Davis. Bai was supported in part by NSF grants DMS-0611548 and OCI-0749217 and DOE grant DE-FC02-06ER25794. Su was supported in part by NSF of China under the contract number 10871049 and Shanghai Down project 200601. We are grateful to the referees for valuable suggestions.

References

[1] G. Golub, Some modified matrix eigenvalue problems, *SIAM Rev.*, **15**:2 (1973), 318–334.

[2] J. Bunch, C. Nielsen and D. Sorensen, Rank-one modification of the symmetric eigenproblem, *Numer. Math.*, **31** (1978), 31–48.

[3] S. Solov'ev, Preconditioned iterative methods for a class of nonlinear eigenvalue problems, *Linear. Algebra. Appl.*, **415** (2006), 210–229.
[4] T. Betcke, N.J. Higham, V. Mehrmann, C. Schröder and F. Tisseur, NLEVP: A collection of nonlinear eigenvalue problems, `http://eprints.ma.man.ac.uk/1071/`, (2008).
[5] T. Lenahan, Calculating dispersion derivatives in fiber optic design, *Bell Syst. Tech. J.*, **62** (1983), 2663–2694.
[6] L. Kaufman, Eigenvalue problems in fiber optic design, *SIAM J. Matrix Anal. A.*, **28**:1 (2006).
[7] G. Golub and C.V. Loan, Matrix Computations (3rd ed.), The Johns Hopkins University Press, Baltimore, 1996.
[8] G. Stewart and J.G. Sun, Matrix Perturbation Theory, Academic Press, San Diego, 1990.
[9] J.M. Cuppen, A divide and conquer method for the symmetric tridiagonal eigenproblem, *Numer. Math.*, **36** (1981), 177–195.
[10] J. Demmel, Applied Numerical Linear Algebra, SIAM, Philadelphia, 1997.
[11] M. Heath, Scientific Computing — An Introductory Survey, McGraw-Hill, New York, 1997.
[12] T.. Dekker, Constructive Apects of the Fundamental Theorem of Algebra, chapter Finding a zero by means of successive linear interpolation, Wiley-Interscience, New York, 1968.
[13] R. Brent, Algorithms for Minimization Without Derivatives, Prentice-Hall, New Jersey, 1973.
[14] A. Ruhe, Algorithms for the nonlinear eigenvalue problem, *SIAM J. Numer. Anal*, **10**:4 (1973), 674–689.
[15] J. Bunch and L. Kaufman, Some stable methods for calculating inertia and solving symmetric linear systems, *Math. Comput.*, **31**:137 (1977), 163–179.
[16] C. Ashcraft, R. Grimes and J. Lewis, Accurate symmetric indefinite linear equation solvers, *SIAM J. Matrix Anal. A.*, **20** (1998), 513–561.
[17] H. Voss, Numerical methods for sparse nonlinear eigenvalue problems, Proc. XVth Summer School on Software and Algorithms of Numerical Mathematics, Hejnice, Czech Republic, 2004.
[18] W. Rudin, Principles of Mathematical Analysis, McGraw-Hill, New York, 1976.
[19] S. Solov'ev, Eigenvibrations of a plate with elastically attached loads, Preprint, 2003, `http://archiv.tu-chemnitz.de/pub/2006/0063/index_en.html.`
[20] S.V. Kartalopoulous, Introduction to DWDM Technology, IEEE Press, Piscataway New Jersey, 2000.
[21] V. Mehrmann and H. Voss, Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods, *Mitt. der Ges. f. Ang. Mathematik und Mechanik*, **27** (2005), 121–151.