TAIWANESE JOURNAL OF MATHEMATICS
Vol. 14, No. 3A, pp. 869-883, June 2010
This paper is available online at http://www.tjm.nsysu.edu.tw/

NONLINEAR RAYLEIGH-RITZ ITERATIVE METHOD FOR SOLVING LARGE SCALE NONLINEAR EIGENVALUE PROBLEMS

Ben-Shan Liao, Zhaojun Bai, Lie-Quan Lee and Kwok Ko

Abstract. A nonlinear Rayleigh-Ritz iterative (NRRIT) method for solving nonlinear eigenvalue problems is studied in this paper. It is an extension of the nonlinear Arnoldi algorithm due to Heinrich Voss. The efficiency of the NRRIT method is demonstrated by comparing with the inverse iteration method to solve a highly nonlinear eigenvalue problem arising from finite element electromagnetic simulation in accelerator modeling.

1. INTRODUCTION

We consider the nonlinear eigenvalue problem of determining scalar values λ and nonzero vectors x for which

$$T(\lambda)x = 0, \tag{1}$$

where $T(\lambda)$ is a square matrix whose elements are analytical functions of the parameter λ. λ and x are referred to as eigenvalues and eigenvectors, respectively. The nonlinear eigenvalue problem (1) arises in a variety of applications, such as vibration simulation of fluid solid structures [20, 24], and calculation of propagation modes of circular optical fiber [16, 10]. In this paper, we will focus on solving the problem (1) derived from finite element analysis of Maxwell's equation with waveguide boundary conditions. It is of great interest in electromagnetic modeling of accelerator [7, 9, 14, 27, 15].

Inverse iteration is a simple method to find an eigenpair of the nonlinear eigenvalue problem (1) with a proper initial approximation [21]. It requires solving different linear system of equations of the same dimension as that of the original problem at each iteration. It becomes very expensive for finding multiple eigenpairs of large scale problems. Recently, a nonlinear Arnoldi algorithm is proposed [25].

Received November 25, 2008, accepted May 6, 2009.
2000 *Mathematics Subject Classification*: 65F15.
Key words and phrases: Nonlinear eigenvalue problem, Nonlinear Rayleigh-Ritz iterative method, Nonlinear Arnoldi, Accelerator modeling.

The method is a generalization of the well-known Arnoldi algorithm for solving linear problems. It uses a subspace projection to reduce the problem dimension, and is especially designed to compute multiple eigenpairs. It has been used successfully to solve large scale rational Hermitian eigenvalue problems [25, 3, 8].

In this paper, we first characterize the nonlinear Arnoldi algorithm as a generalization of the popular Rayleigh-Ritz subspace projection technique for solving large scale linear eigenvalue problems [1], and refer to it as the nonlinear Rayleigh-Ritz iterative (NRRIT) method. We then discuss a number of key issues for an efficient implementation of the NRRIT method, such as the choice of initial approximate eigenpairs, and the definition of proper projection subspaces to preserve properties of the original problem. A significant part of our effort is devoted to solving a highly nonlinear eigenvalue problem arising from electromagnetic modeling of accelerator. Numerical results are presented to demonstrate that the NRRIT method is a promising practical approach to solve this challenging problem.

The rest of the paper is organized as follows. Section 2 gives a short description of the nonlinear eigenvalue problem arising from electromagnetic modeling of accelerator. Section 3 is a review of the inverse iteration. Section 4 first presents a framework of nonlinear Rayleigh-Ritz projection technique and then discusses practical implementation issues. Numerical experiments and concluding remarks are in Sections 5 and 6, respectively.

2. Electromagnetic Simulation in Accelerator Modeling

In electromagnetic modeling of waveguide loaded accelerator cavities, one solves Maxwell's equation

$$\nabla \times (\frac{1}{\mu}\nabla \times E) - \lambda \varepsilon E = 0 \quad \text{in} \quad \Omega \tag{2}$$

with the boundary conditions

$$\vec{n} \times E = 0 \quad \text{on} \quad \Gamma_E \tag{3}$$

$$\vec{n} \times (\frac{1}{\mu}\nabla \times E) = 0 \quad \text{on} \quad \Gamma_M \tag{4}$$

$$\vec{n} \times (\frac{1}{\mu}\nabla \times E) + \mathbf{i}\sqrt{\lambda - \kappa_{c,j}^2}\, \vec{n} \times (\vec{n} \times E) = 0 \quad \text{on} \quad \Gamma_j, \;\; j = 1, 2, \ldots, p, \tag{5}$$

where E is the electric field, $\lambda = \omega^2/c^2$, ω is the wave number, c is the speed of light, ε is relative permittivity, μ is relative permeability, and Ω is the geometry of the cavity with the outward normal vector $\vec{n}$ on the boundary. Γ_E is the electric boundary with perfect conductor, Γ_M be the magnetic boundary with perfect insulator, and

Γ_j are the waveguide boundaries. $\kappa_{c,j}$ are the cutoff wave numbers of modes in the waveguide ports, or simply called "cutoff values", and p is the number of the waveguides. $\mathbf{i} = \sqrt{-1}$. We assume that the interior of the domain Ω is source-free and the dielectric materials are isotropic and homogeneous. There is only one propagating waveguide mode in each waveguide port. See [7, 9, 15] for further detail.

Using the finite element discretization, by representing of the electric field E in terms of the basis functions N_k: $E = \sum_k x_k N_k$, equations (2)–(5) lead to an algebraic nonlinear eigenvalue problem of the form

$$T(\lambda)x = 0, \tag{6}$$

where

$$T(\lambda) = K - \lambda M + \mathbf{i}\sum_{j=1}^{p} \sqrt{\lambda - \kappa_{c,j}^2}\, W_j,$$

and the entries of K, M, and W_j are given by

$$\begin{aligned} K_{k,\ell} &= \int_\Omega (\nabla \times N_k) \cdot \frac{1}{\mu}(\nabla \times N_\ell)\, d\Omega, \\ M_{k,\ell} &= \int_\Omega N_k \cdot \varepsilon N_\ell \, d\Omega, \\ (W_j)_{k,\ell} &= \int_{\Gamma_j} (\vec{n} \times N_k) \cdot (\vec{n} \times N_\ell)\, d\Gamma. \end{aligned}$$

K is the stiffness matrix, M is the mass matrix and W_j are the damping matrices. All the matrices are real symmetric. Moreover, K is positive semi-definite, and M is positive definite [14, 11].

In accelerator design, the waveguides are used to introduce damping into the cavity to suppress the so-called *Higher-Order Modes* (HOMs), which are those resonant modes forming the wakefields that can disrupt the stability of beam accelerator and transport. The effectiveness of the damping is measured by the so-called external quality factors. Specifically, let $\kappa = \sqrt{\lambda}$, the resonant frequency f and the corresponding external Q_e of the cavity are defined as

$$f(\kappa) = \frac{c}{2\pi} \cdot \mathrm{Re}(\kappa) \quad \text{and} \quad Q_e(\kappa) = \frac{1}{2} \cdot \frac{\mathrm{Re}(\kappa)}{\mathrm{Im}(\kappa)}. \tag{7}$$

The quantity Q_e measures the electromagnetic coupling between the cavity and waveguide. It characterizes the energy loss through the waveguide. With a given resonant frequency $f_0 > 0$, the cavity designers would like to seek frequencies f that are close to f_0 and $Q_e > 1$. Correspondingly, the desired eigenvalues λ of the

nonlinear eigenvalue problem (6) are the ones such that

$$(8)\quad \begin{cases} \sqrt{\lambda} \text{ is close to the value } \kappa_0 = \frac{2\pi}{c} f_0 \text{ and} \\ \lambda \in \mathcal{D} = \left\{ \lambda \;\mid\; \lambda = \kappa^2,\ \mathrm{Re}(\kappa) > \kappa_0,\ \mathrm{Im}(\kappa) > 0 \text{ and } Q_e(\kappa) > \widehat{Q}_e > 1 \right\}, \end{cases}$$

where $\widehat{Q}_e$ is a prescribed value. The shaded domain in Figure 1 is an illustration of such a region in terms of $\kappa = \sqrt{\lambda}$.

For the waveguide loaded cavity with one cutoff value ($p = 1$), as we have in Section 4, the problem (6) can be recast as a standard quadratic eigenvalue problem. There are a number of established numerical techniques for solving the quadratic eigenvalue problems [23, 2]. We have successfully solved such a nonlinear eigenvalue problem for the degrees of freedom up to 3.2 million [13, 12]. Unfortunately, in general, it is a highly nonlinear eigenvalue problem. There is no efficient transformation to convert the nonlinear eigenvalue problem (6) to the polynomial eigenvalue problems such that we can apply an established eigensolver. In the rest of this paper, we will study a nonlinear Rayleigh-Ritz iteration method to solve the problem (6).

3. Inverse Iteration

In this section, we review the simple inverse iteration, which will be used in the inner loop of the nonlinear Rayleigh-Ritz iteration method in next section. We note that the eigenvalue problem (1) is equivalent to the nonlinear system of equations

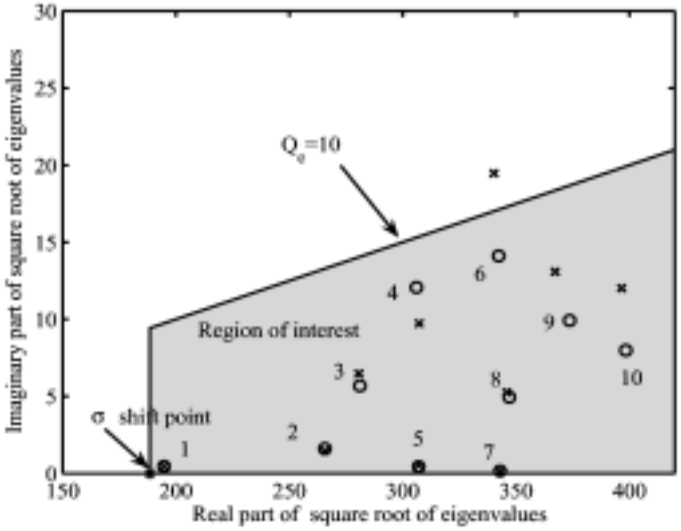

Fig. 1. The region of interest $\mathcal{D}$ (shaded area). Ten eigenvalues of the linearized pencil $\widehat{K}(\kappa_0^2) - \lambda \widehat{M}(\kappa_0^2)$ ("$\circ$") and eigenvalues λ of $T(\lambda)$ ("$\times$").

$$\begin{bmatrix} T(\lambda)x \\ w^H x - 1 \end{bmatrix} = 0, \tag{9}$$

where w is a prescribed vector for the normalization of the eigenvector. The inverse iteration stems from applying Newton's method to solve the nonlinear system (9) [21, 26]. The following is a pseudocode of the inverse iteration:

IIT (inverse iteration)

1. Choose an initial approximate eigenpair $(\lambda^{(0)}, x^{(0)})$.
2. For $j = 0, 1, \cdots$, until convergence
   (a) solve $T(\lambda^{(j)})u = T'(\lambda^{(j)})x^{(j)}$ for u.
   (b) compute the new approximate eigenpair

$$\begin{cases} \lambda^{(j+1)} &= \lambda^{(j)} - (w^H x^{(j)})/\alpha, \\ x^{(j+1)} &= u/\alpha, \end{cases}$$

   where $\alpha = w^H u$.

A common choice of the normalization vector w is $x^{(0)}/\|x^{(0)}\|_2^2$. We need to solve a linear system of the same dimension as the original problem at Step 2a. It is typically the most time consuming step. We can use a sparse linear system solver, such as UMFPACK [5] or SuperLU [6]. In practice, the convergence of the approximate eigenpair $(\lambda^{(j)}, x^{(j)})$ is declared when a properly defined residual norm is smaller than a given tolerance. One may simply use the so-called absolute residual norm $\|T(\lambda^{(\ell)})x^{(\ell)})\|$. In Section 5, we will define a relative residual norm for the nonlinear eigenvalue problem discussed in Section 2.

Under certain assumptions associated with the initial approximate eigenpair, the inverse iteration converges quadratically [21]. Variants of inverse iteration include residual inverse iteration and QR-type iteration [19, 26].

4. Nonlinear Rayleigh-Ritz Iterative Method

From a mathematical point of view, the recently proposed nonlinear Arnoldi algorithm [25], nonlinear Jacobi-Davidson method [3], and rational Krylov method [22, 8] for solving nonlinear eigenvalue problem (1) are extensions of the Rayleigh-Ritz subspace projection technique for solving linear eigenvalue problems, see [1] and reference therein. At the level of abstraction, these methods may be summarized as the following:

NRRIT (nonlinear Rayleigh-Ritz iterative procedure)

(a) Select a proper projection subspace $\mathcal{V}$.

(b) Compute approximate eigenpairs (θ, z) satisfying Galerkin condition:

$$z \in \mathcal{V} \quad \text{and} \quad T(\theta)z \perp \mathcal{V}. \tag{10}$$

(c) Evaluate residual norms for assessing the accuracy of the approximate eigenpairs (θ, z). If the approximates are satisfactory then stop. Otherwise return to Step (b) with a refined and/or augmented projection subspace $\mathcal{V}$.

We note that at Step (b), let V be an orthonormal basis of $\mathcal{V}$, then $z = Vg$ for some n-vector g, where n is the dimension of the subspace $\mathcal{V}$. Hence, Step (b) is equivalent to determining eigenpairs (θ, g) of the reduced nonlinear eigenvalue problem

$$T_V(\theta)g = 0, \tag{11}$$

where $T_V(\theta) = V^H T(\theta)V$ is a matrix of the order n. The values θ are referred to as *Ritz values* and the vectors z are the corresponding *Ritz vectors*. In a practical implementation of the NRRIT procedure, the critical issues are (a) what is a proper initial projection subspace $\mathcal{V}$ and how to compute its basis V? and (b) how to refine, expand or restart the subspace $\mathcal{V}$ when necessary? In the rest of this section, we describe how we address these issues, and present a complete implementation.

Let us begin with the choice of initial approximate eigenpairs. As we know, all nonlinear eigensolvers require initial approximate eigenpairs and it is a crucial factor deciding convergence of the procedure. A good choice is to use a subset of eigenpairs (θ_ℓ, v_ℓ) of the linear eigenvalue problem

$$\widehat{K}(\tau)v = \theta\widehat{M}(\tau)v, \tag{12}$$

where $\widehat{K}(\tau) = T(\tau) - \tau T'(\tau)$ and $\widehat{M}(\tau) = -T'(\tau)$ are extracted from the first order truncation of the Taylor series expansion of $T(\lambda)$:

$$T(\lambda) = T(\tau) + (\lambda - \tau)T'(\tau) + \mathcal{O}((\lambda - \tau)^2), \tag{13}$$

τ is a given expansion point close to the desired eigenvalues of $T(\lambda)$. The eigenvectors v_ℓ are used to form the initial projection subspace $\mathcal{V} = \text{span}\{v_\ell\}$.

Example. Consider the accelerator modeling problem (6) with a single waveguide cutoff value:

$$\left[K - \lambda M + \mathbf{i}\sqrt{\lambda - \kappa_c^2}\, W\right] x = 0. \tag{14}$$

A natural choice of the initial approximate eigenpairs (θ_ℓ, v_ℓ) is from eigenpairs of the linear term $K - \lambda M$. However, all these eigenpairs are real due to the fact that

$K - \lambda M$ is a symmetric positive definite pencil. There are no physical meaning of these eigenvalues and the corresponding external Q_e values are not defined. In Section 5, we will see that numerically, it leads to some approximate eigenpairs of (14) without practical interest. A better choice is from the linear approximate eigenvalue problem of (14) at the point $\tau = \kappa_0^2$:

$$\left[K + \mathbf{i}\left(\frac{\kappa_0^2 - 2\kappa_c^2}{2\sqrt{\kappa_0^2 - \kappa_c^2}}\right) W\right] v = \theta \left[M - \mathbf{i}\left(\frac{1}{2\sqrt{\kappa_0^2 - \kappa_c^2}}\right) W\right] v. \tag{15}$$

Let (θ_ℓ, v_ℓ) be the m eigenpairs of (15) such that θ_ℓ near to the point κ_0 and $\theta_\ell^{1/2}$ are in the region defined as (8) for $\ell = 1, 2, \ldots, m$, where the indices are labeled according to the distance between $\theta_i^{1/2}$ and κ_0:

$$|\theta_1^{1/2} - \kappa_0| \le |\theta_2^{1/2} - \kappa_0| \le \cdots \le |\theta_m^{1/2} - \kappa_0|. \tag{16}$$

Then we can use these m eigenpairs as initial approximate eigenpairs and define a projection basis V from for the initial projection subspace $\mathcal{V}$ as

$$V = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix}. \tag{17}$$

Figure 1 shows a numerical example of the problem (14) of order 10124 and the cutoff value $\kappa_c = 110.24$. The circles are ten eigenvalues θ_ℓ of the linear approximation (15) closest to the point $\kappa_0 = 188.5$ and labeled according to (16). We note that the problem (14) can be converted to the following quadratic eigenvalue problem

$$\left[\nu^2 M - \nu\,(\mathbf{i}W) + (\kappa_c^2 M - K)\right] x = 0, \tag{18}$$

where $\nu = \sqrt{\lambda - \kappa_c^2}$. Therefore, we can first extract eigenvalues ν of (18) by a quadratic eigensolver, and then compute the eigenvalues $\lambda = \nu^2 + \kappa_c^2$ of the original problem (14). The crosses in Figure 1 are the 10 eigenvalues of the original problem (14) computed via the quadratic problem (18). We can see that some of the initial approximate eigenvalues θ_ℓ are already good approximations to the eigenvalues λ, such as the ones labeled by 1, 2, 5, 7 and 8.

Now let us consider how to form a symmetry-preserving projection subspace. Note that the eigenvectors $\{v_\ell\}$ are generally complex. If V is used as a projection basis, then the matrix $T_V(\theta)$ of the reduced nonlinear eigenvalue problem will be fully complex. In the application such as the accelerator modeling problem (6) it is advantageous to preserve realness and symmetry of coefficient matrices K, M and W_j in terms of computational costs and accuracy. In this case, we can use a real orthonormal projection basis Q such that

$$Q = \mathrm{orth}(\begin{bmatrix} \mathrm{Re}(V) & \mathrm{Im}(V) \end{bmatrix}),$$

where $\mathrm{orth}(X)$ stands for an orthonormal basis for the range of X. This leads to a reduced nonlinear eigenvalue problem

$$T_Q(\theta)y = 0, \tag{19}$$

where $T_Q(\theta) = Q^T T(\theta) Q$. It is easy to see that for the problem (6), the coefficient matrices $Q^T K Q$, $Q^T M Q$ and $Q^T W_j Q$ in $T_Q(\theta)$ preserve the realness and symmetry of the original coefficient matrices K, M and W_j. In general, the computed eigenpairs will be at least as accurate as the ones computed by using the complex projection matrix V. The similar idea has been used for structure-preserving model-reduction techniques [4].

Starting with the ℓ-th initial approximate eigenpair (θ_ℓ, v_ℓ), an approximate eigenpair (θ, y) of the reduced problem (19) can be found by using a nonlinear eigensolver suitable for small dense problems, such as the IIT method with the initial approximations $(\theta_\ell, Q^T v_\ell)$. An approximate eigenpair of the original nonlinear problem (1) is then given by $(\theta, z = Qy)$. If (θ, z) is convergent, then the basis matrix V is updated by replacing the initial vector v_ℓ with the Ritz vector z. Otherwise, the projection subspace spanned by Q is augmented by the new vector

$$v = \tilde{v} - QQ^T\tilde{v}$$

where $\tilde{v} = T^{-1}(\tau)T(\theta)z$. This is one step of the residual inverse iteration [19]. Such a subspace expansion strategy can be explained by the rational Krylov method [22, 8] and is used in the nonlinear Arnoldi method [25]. In practice, to maintain the real orthonormal projection basis, Q is augmented by

$$Q := [\ Q\ \ \widetilde{Q}\ ]$$

where $\widetilde{Q} = \mathrm{orth}([\ \mathrm{Re}(v)\ \ \mathrm{Im}(v)\ ])$.

Beside the projection subspace augmentation, we also need to consder issues of purging and restarting. The NRRIT procedure with the initial approximate eigenpair (θ_ℓ, v_ℓ) may fail to converge to an eigenpair of the original problem. This is detected by the number of the NRRIT iterations exceeding a prescribed limit $i_{\max}$. If it occurs, the starting vector v_ℓ and all associated augmented vectors in Q are purged from the projection basis matrix Q. Furthermore, to control the memory and computational costs, we also need to restart the projection subspace when the number of basis vectors becomes too large. A restart is always performed after an approximate eigenpair is accepted. The restarting projection subspace is spanned by $Q = \mathrm{orth}([\ \mathrm{Re}(V)\ \ \mathrm{Im}(V)\ ])$, where V is updated with the converged Ritz vectors z and is reduced by purging those initial vectors v_ℓ, which lead to the failure of the convergence.

The following is a pseudocode for the nonlinear Rayleigh-Ritz iterative method to find k eigenvalues of the nonlinear eigenvalue problem (6) close to a prescribed

point τ. In the pseudocode, $n_{\max}$ is the maximal size of the projection subspace Q, m is the number of the converged eigenpairs, and $i_{\max}$ is the maximal number of iterations allowed.

NRRIT (nonlinear Rayleigh-Ritz iterative method to compute k eigenpairs)

1. Compute n selected eigenpairs $\{(\theta_j, v_j)\}_{j=1}^{n}$ of the matrix pair $(\widehat{K}(\tau), \widehat{M}(\tau))$ as described in (8), where $n \geq k$.
2. Order $\{(\theta_j, v_j)\}_{j=1}^{n}$ properly and let $V = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$.
3. Set $Q = \text{orth}([\text{Re}(V) \;\; \text{Im}(V)])$.
4. Set $\ell = 1$, $iters = 1$ and $m = 0$.
5. Iterate while $m \leq k$ and $\ell \leq n$:
   (a) compute an eigenpair (θ, y) of $T_Q(\theta)y = 0$ with initial $(\theta_\ell, Q^T v_\ell)$,
   (b) compute Ritz vector $z = Qy$ and residual $r = T(\theta)z$,
   (c) if the Ritz pair (θ, z) is convergent, then
      i. save (θ, z) as an approximate eigenpair,
      ii. update V by setting $v_\ell = z$,
      iii. if $|Q| > n_{\max}$, then restart with $Q = \text{orth}([\text{Re}(V) \;\; \text{Im}(V)])$, and update $T_Q(\theta)$,
      iv. reset $iters = 1$,
      v. set $\ell = \ell + 1$ and $m = m + 1$ and go to step 5 for searching next eigenpair,
   (d) solve $T(\tau)\tilde{v} = r$ for $\tilde{v}$,
   (e) orthogonalize $\tilde{v}$ against Q: $v = \tilde{v} - QQ^T\tilde{v}$,
   (f) compute $\widetilde{Q} = \text{orth}([\text{Re}(v) \;\; \text{Im}(v)])$,
   (g) augment $Q := \begin{bmatrix} Q & \widetilde{Q} \end{bmatrix}$ and update $T_Q(\theta)$,
   (h) if $iters > i_{\max}$, then
      i. flag failure with the initial approximate eigenpair (θ_ℓ, v_ℓ),
      ii. purge all associated augment vectors from Q,
      iii. update V by purging v_ℓ from V,
      iv. set $\ell = \ell + 1$ and $iters = 0$,
   (i) set $iters = iters + 1$.

A few remarks are in order. First, to solve the reduced problem at Step 5, we can use the inverse iteration discussed in Section 3. Second, at Step 5d, we need to solve a linear system $T(\tau)v = r$ of the same dimension as the original problem. It

usually takes much more time than other steps. However, we note that the coefficient matrix $T(\tau)$ of the linear system remains unchanged in the loop. Therefore, it is more efficient than a Jacobi-Davidson-based implementation, where the so-called correction equation is modified at every iteration [3, 26, 18]. Finally, in both inner and outer iterations (Steps 5 and 5c), we need to use stopping criteria. In Section 5, we will define proper stopping criteria for the accelerator problem (6).

5. Numerical Experiments

In this section, we present numerical results of using the NRRIT method to solve the nonlinear eigenvalue problem (6) arising from electromagnetic modeling of accelerator described in Section 2. We consider a model open cavity as shown in Figure 2. The cavity part is the cylindrical part. The two waveguides are coaxial waveguide at the top left and the rectangular waveguide at the bottom [17]. The underlying nonlinear eigenvalue problem is of the form

$$T(\lambda)x = \left[K - \lambda M + \mathbf{i}\sqrt{\lambda - \kappa_{c,1}^2}\, W_1 + \mathbf{i}\sqrt{\lambda - \kappa_{c,2}^2}\, W_2\right] x = 0, \tag{20}$$

where the matrices K, M, W_1 and W_2 are of order $N = 9956$. The cutoff values are $\kappa_{c,1} = 0$ and $\kappa_{c,2} = 108.8774$. The computational task is to find 10 eigenvalues λ_j close to the point $\kappa_0 = 2\pi f_0/c = 146.71$, where $f_0 = 7.0 \times 10^9$(Hz) and the corresponding external Q_e-values are greater than 10.

To test the accuracy of an approximate eigenpair (θ, z), we use the relative residual norm

$$\mathcal{E}(\theta, z) = \frac{\|T(\theta)z\|_2}{(\|K\|_1 + |\theta|\|M\|_1 + \sqrt{|\theta - \kappa_{c,1}^2|}\|W_1\|_1 + \sqrt{|\theta - \kappa_{c,2}^2|}\|_1 W_2\|_1)\|z\|_2} \tag{21}$$

where $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the 1-norm and 2-norm, respectively. All numerical experiments reported in this paper were run in MATLAB 7.0.1 on a Pentium IV PC with 2.6GHz CPU and 1GB of core memory.

Experiment 1. Let us first examine the impact of the initial approximate eigenpairs. A nature choice is to select some eigenpairs of the linear term $K - \lambda M$ of $T(\lambda)$. All these eigenvalues are real since $K - \lambda M$ is a positive definite pencil. Function `eigs` with the shift-and-invert spectral transformation took 11.8 seconds to compute 10 real eigenpairs $(\theta_j^{(\mathrm{r})}, v_j^{(\mathrm{r})})$ closest to the shift $\tau = \kappa_0^2$. Using these real initial approximations, NRRIT converges to 10 eigenpairs of (20) with all relative residual norms smaller than $\epsilon = 10^{-10}$. The left side of Table 1 reports the corresponding frequencies and Q_e-values defined as (7). In the table, we also show the results computed by the inverse iteration with the same initial approximates.

Since the computed eigenpairs by NRRIT and IIT methods are numerically the same within the relative residual errors, the same frequencies and Q_e-values are shown in the table.

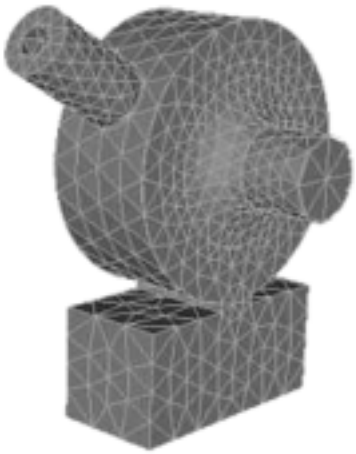

Fig. 2. A simulation model of waveguide loaded cavity with two cutoff guides.

We observe that the computed eigenvalues λ_2 and λ_3 are not physically meaningful since their corresponding external Q_e-values are less than 1. Furthermore, the computed resonant frequencies f_9 and f_{10} are far away from the prescribed frequency f_0.

An alternative choice of the initial approximate eigenpairs is from eigenvalues of linear truncation $\widehat{K}(\tau) - \lambda\widehat{M}(\tau)$ of $T(\lambda)$, as discussed in Section 4. Function `eigs` with the shift-and-invert spectral transformation took 23 seconds to compute 10 complex eigenpairs $(\theta_j^{(\mathrm{c})}, v_j^{(\mathrm{c})})$ closest to the shift $\tau = \kappa_0^2$. The right side of Table 1 reports the frequencies and Q_e-values computed by the NRRIT and IIT methods with the complex initials. As we can see that all computed resonant frequencies and Q_e-values are physically meaningful. Furthermore, the resonant frequencies f_j are closer to the prescribed frequency f_0.

Experiment 2. Let us show computational efficiency of the NRRIT method by comparing with the inverse iteration. In the inner loops, both NRRIT and IIT methods require solving linear systems of equations of the same dimension as the original problem. However, the coefficient matrix of the linear system remains unchanged in NRRIT. As a result, NRRIT is more computational efficient as shown in Table 2.

Experiment 3. In this experiment, we compare the performance of two im-

plementations of the NRRIT method, namley using real projection basis Q and complex projection basis V. Figure 3 shows the convergence history of two implementations. The horizontal axis is the number of iterations. The vertical axis is the relative residual norm of approximate eigenpairs. When an approximate eigenpair (θ, z) is declared to be convergent, it is marked by a circle ("$\circ$") for real basis Q and by an asterisk ("$*$") for complex basis V. In summary, the NRRIT method with the real projection subspace Q requires 25% fewer number of iterations, and is 20% faster in computational time.

Table 1. Results using *real* initials $(\theta_j^{(\mathrm{r})}, v_j^{(\mathrm{r})})$ (left) and *complex* initials $(\theta_j^{(\mathrm{c})}, v_j^{(\mathrm{c})})$ (right)

			IIT	NRRIT					IIT	NRRIT	
j	f_j	Q_e	iters	iters	$\|Q\|$	j	f_j	Q_e	iters	iters	$\|Q\|$
1	7.1373	34643.66	2	4	16	1	7.1373	34643.66	1	2	22
2	4.5655	0.36	6	21	58	2	9.9992	2136.73	2	10	42
3	7.2465	0.48	8	14	86	3	10.0449	12376.84	1	6	54
4	9.9992	2136.73	3	6	98	4	10.4762	1149.21	2	6	66
5	10.0449	12376.84	2	6	110	5	10.5463	7714.93	1	5	76
6	10.4762	1149.21	2	4	118	6	11.1518	118.71	2	5	86
7	10.5463	7714.93	2	5	128	7	13.1180	15.25	3	12	110
8	11.1581	118.71	3	6	140	8	13.2698	536.76	2	7	124
9	14.2617	3.17	5	14	168	9	13.5882	2500.75	2	8	140
10	15.0143	3.59	6	12	192	10	13.7688	181.23	2	7	154
Total no. of iters			39	92		Total no. of iters			18	68	
Total elapsed time (s)			178.55	53.31		Total elapsed time (s)			108.92	55.91	

Table 2. Performance of IIT and NRRIT with *complex* initials $(\theta_j^{(\mathrm{c})}, v_j^{(\mathrm{c})})$

	IIT		NRRIT	
Residual error ϵ	Iters	Elapsed-time	Iters	Elapsed-time
10^{-8}	15	94.98	46	45.95
10^{-10}	18	108.92	68	55.91
10^{-12}	22	127.47	87	75.08
10^{-14}	24	136.97	114	80.59

6. Concluding Remarks

We presented a NRRIT algorithm for solving large-scale nonlinear eigenvalue problem of the form (1). It is a generalization of the Rayleigh-Ritz subspace projection for solving linear eigenvalue problems. We showed that the choice of initial approximate eigenpairs is an important factor for the success of the method. We

also show that the method can be modified to preserve specific features of a particular application, such as using real projection subspace to preserve symmetry. It is demonstrated that the NRRIT method is practically viable method and outperforms the inverse iteration for solving the large scale nonlinear eigenvalue problems arising from finite element analysis of resonant frequencies and external Q_e values of a waveguide loaded cavity in the accelerator design. In this paper, we focus on the description of algorithm and issues in practical implementation. Theoretical analysis of the NRRIT method is an open problem.

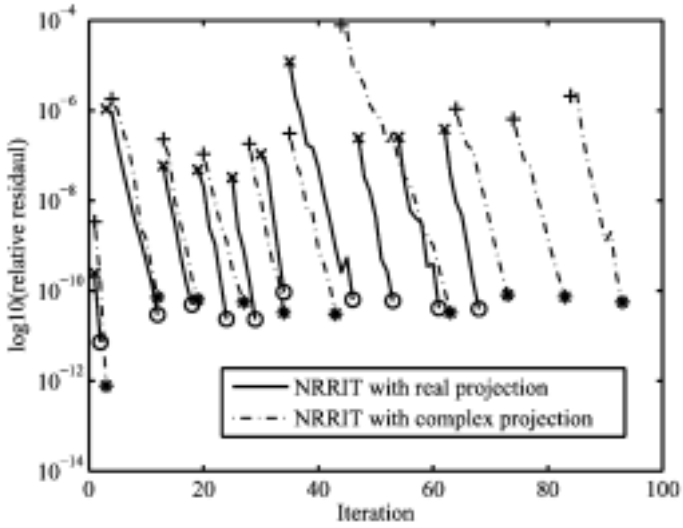

Fig. 3. Convergence history of NRRIT with real projection basis Q (solid line) and complex projection basis V (dash-dot line).

Acknowledgment

The research of Bai and Liao was supported in part by the NSF under the grant DMS-0611548 and DOE grant DE-FC0206ER25794. Ko and Lee were supported in part by the DOE under grant DE-AC02-76SF00515.

References

1. Z. Bai, J. Demmel, J. Dongarra, A. Ruhe and H. van der Vorst, eds., *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM, Philadelphia, 2000.
2. Z. Bai and Y. Su, SOAR: A second-order Arnoldi method for the solution of the quadratic eigenvalue problem, *SIAM J. Matrix Anal. Appl.*, **26(3)** (2005), 640-659.

3. T. Betcke and H. Voss, A Jacobi-Davidson-type projection method for nonlinear eigenvalue problems, *Future Generation Computer Systems*, **20(3)** (2004), 363-372.

4. D. Bindel, Z. Bai, and J. Demmel, Model reduction for RF MEMS simulation, *Lecture Notes in Computer Science*, **3732** (2006), 286-295.

5. T. Davis, Algorihm 832: UMFPACK- an unsummetric-pattern multifrontal method, *ACM Trans. on Math. Soft.*, **30** (2004), 196-199.

6. J. Demmel, S. Eisenstat, J. Gilbert, X. Li, and J. Liu, A supernodal approach to sparse partial pivoting, *SIAM J. Matrix Anal. Appl.*, **20(3)** (1999), 720-755.

7. H. Igarashi, Y. Sugawara, and T. Honma, A numerical computations of external Q of resonant cavities, *IEEE Trans. on Magnetics*, **31(3)** (1995), 1642-1645.

8. E. Jarlebring and H. Voss, Rational Krylov for nonlinear eigenproblems, an iterative projection method, *Appl. of Math.*, **50(6)** (2005), 543-554.

9. J. Jin, *The Finite Element Method in Electromagnetics*, Second Edition. Wiley, New Year, 2002.

10. L. Kaufman, Eignevalue problems in fiber optic design, *SIAM J. Matrix Anal. Appl.*, **28(1)** (2006), 105-117.

11. K. Ko, N. Folwell, L. Ge, A. Guetz, L. Lee, Z. Li, C. Ng, E. Prudencio, G. Schussman, R. Uplenchwar and L. Xiao. Advances in electromagnetic modelling through high performance computing, Technical Report SLAC-PUB-11789, Stanford Linear Accelerator Center, Menlo Park, CA, 2006.

12. L. Lee, Z. Bai, W. Gao, L. Ge, P. Husbands, M. Kowalski, X. Li, Z. Li, C. Ng, C. Yang, E. G. Ng, and K. Ko, *Solving nonlinear eigenproblems in accelerator cavity design*, Presentation at SIAM Annual Meeting, New Orleans, LA, July 2005.

13. L. Lee, L. Ge, Z. Li, C. Ng, K. Ko, B.-S. Liao, Z. Bai, W. Gao, C. Yang, P. Husbands, and E. G. Ng, *Modeling RF cavity with external coupling*, Presentation at SIAM Conference on Computational science and Engineering, Orlando, FL, Feb. 2005.

14. L. Lee, L. Ge, Z. Li, C. Ng, G. Schussman, and K. Ko, Achievements in ISICs/SAPP collaborations for electromagnetic modeling of acceleators. *J. of Physics: Conference Series*, **16** (2005), 205-209.

15. L. Lee, Z. Li, C. Ng, and K. Ko, *Omega3P: A parallel finite-element eigenmode analysis code for accelerator cavities*, Technical Report SLAC-PUB-13529, SLAC National Accelerator Laboratory, Menlo Park, CA, 2009.

16. T. A. Lenahan, Calculation of modes in an optical fiber using the finite element method and eispack, *Bell System Technical Journal*, **62** (1983), 2663-2694.

17. C. Limborg, Z. Li, L. Xiao, J. F. Schmerge, D. Dowell, S. gierman, E. Bong and S. Gilevich, *RF design of the LCLS gun*, LCLS-TN-05-3, February 2005.

18. V. Mehrmann and H. Voss, *Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods*, Technical Report 83, Arbeitsbereich Mathematik, TU Hamburg-Harburg, 2004.

19. A. Neumaier, Residual inverse iteration for the nonlinear eigenvalue problem, *SIAM J. Numer. Anal.*, **22** (1985), 914-923.

20. J. Planchard, Eigenfrequencies of a tube bundle placed in a confined fluid, *Comput. Meth. Appl. Mech. and Eng.*, **30** (1982), 75-93.

21. A. Ruhe, Algorithms for the nonlinear eigenvalue problem, *SIAM J. Numer. Anal.*, **10** (1973), 674-689.

22. A. Ruhe, A rational Krylov algorithm for nonlinear matrix eigenvalue problems, *Zapiski Nauchnyh Seminarov POMI*, **268** (2000), 176-180.

23. F. Tisseur and K. Meerbergen, The quadratic eigenvalue problem, *SIAM Rev.*, **43** (2001), 235-286.

24. H. Voss, A rational spectral problem in fluid-solid vibration, *Electr. Trans. Numer. Anal.*, **16** (2003), 94-106.

25. H. Voss, An Arnoldi method for nonlinear eigenvalue problems, *BIT Numerical Mathematics*, **44(2)** (2004), 387-401.

26. H. Voss, Numerical methods for sparse nonlinear eigenvalue problems, in: *Proc. XVth Summer School on Software and Algorithms of Numerical Mathematic*, Hejnice, Czech Republic, 2004.

27. C. Yang, Solving large-scale eigenvalue problems in SciDAC applications, *Journal of Physics: Conference Series*, **16** (2005), 425-434, (Proceedings of 2005 SciDAC Conference).

Ben-Shan Liao
Siemens PLM Software,
10824 Hope Street,
Cypress, CA 90630,
U.S.A.
E-mail: ben-shan.liao@siemens.com

Zhaojun Bai
Department of Computer Science and Department of Mathematics,
University of California,
Davis, CA 95616
U.S.A.
E-mail: bai@cs.ucdavis.edu

Lie-Quan Lee and Kwok Ko
Advanced Computations Department,
SLAC National Accelerator Laboratory,
Menlo Park, CA 94025, U.S.A.
E-mail: liequan@slac.stanford.edu
kwok@slac.stanford.edu