A Semiorthogonal Generalized Arnoldi Method and Its Variations for Quadratic Eigenvalue Problems

Wei-Qiang Huang\textsuperscript{1}, Tiexiang Li\textsuperscript{2*}, Yung-Ta Li\textsuperscript{1} and Wen-Wei Lin\textsuperscript{3}

\textsuperscript{1}Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan
\textsuperscript{2}Department of Mathematics, Southeast University, Nanjing 211189, China
\textsuperscript{3}Department of Mathematics, National Taiwan University, Taipei 106, Taiwan

SUMMARY

In this paper, we are concerned with the computation of a few eigenpairs with smallest eigenvalues in absolute value of quadratic eigenvalue problems. We first develop a semiorthogonal generalized Arnoldi method where the name comes from the application of a pseudo inner product in the construction of a generalized Arnoldi reduction \cite{25} for a generalized eigenvalue problem. The method applies the Rayleigh-Ritz orthogonal projection technique on the quadratic eigenvalue problem. Consequently it preserves the spectral properties of the original quadratic eigenvalue problem. Furthermore, we propose a refinement scheme to improve the accuracy of the Ritz vectors for the quadratic eigenvalue problem. Given shifts, we also show how to restart the method by implicitly updating the starting vector and constructing better projection subspace. We combine the ideas of the refinement and the restart by selecting shifts upon the information of refined Ritz vectors. Finally an implicitly restarted refined semiorthogonal generalized Arnoldi method is developed. Numerical examples demonstrate that the implicitly restarted semiorthogonal generalized Arnoldi method with or without refinement has superior convergence behaviors than the implicitly restarted Anoldi method applied to the linearized quadratic eigenvalue problem.

1. INTRODUCTION

The problem of finding scalars $\lambda \in \mathbb{C}$ and nontrivial vectors $x \in \mathbb{C}^n$ such that

$$\left(\lambda^2 M + \lambda D + K\right)x = 0$$

(1)

where $M$, $D$ and $K$ are $n \times n$ large and sparse matrices, is known as the quadratic eigenvalue problem (QEP). The scalars $\lambda$ and the associated nonzero vectors $x$ are called eigenvalues and (right) eigenvectors of the QEP, respectively. Together, $(\lambda, x)$ is called an eigenpair of (1).

The QEP arises in a wide variety of applications, including electrical oscillation, vibro-acoustics, fluid mechanics, signal processing and the simulation of microelectroinical mechanical system etc. A good survey of applications, spectral theory, perturbation analysis and numerical approaches can be found in \cite[section 11.9]{5}, \cite{28} and the references therein.
In practice, some eigenvalues of a QEP near a target τ are interested. Hence we may apply the shift transformation and consider the corresponding shifted QEP

\[(\lambda^2 M + \lambda D + K)x = 0\]

where \(\lambda = \tau - \lambda, M = M, D = 2\tau M + D\) and \(K = \tau^2 M + \tau D + K\). For simplicity, we assume, without loss of generality, that \(\tau = 0\). Therefore, throughout this paper, we focus on the problem of finding eigenvalues near the zero (i.e., those small ones in modulus) under the assumption that 0 is not an eigenvalue of the QEP \((1)\) or, equivalently, that \(K\) is nonsingular.

Through the so-called “linearization” process, one may first construct a suitable matrix pair \((A, B)\) of size \(2n\) and a vector \(\varphi\) in \(\mathbb{C}^{2n}\) to rewrite the QEP \((1)\) equivalently into a generalized eigenvalue problem (GEP)

\[A\varphi = \frac{1}{\lambda}B\varphi.\]  

(2)

If \(B\) is chosen to be nonsingular, one can further reduce \((2)\) as a standard eigenvalue problem (SEP)

\[(B^{-1}A)\varphi = \frac{1}{\lambda}\varphi\]  

(3)

or

\[(AB^{-1})\psi = \frac{1}{\lambda}\psi\]  

(4)

where \(\psi = B\varphi\). We call \((3)\) and \((4)\) the left-inverted SEP (ℓ-SEP) and the right-inverted SEP (r-SEP), respectively. After transforming a QEP equivalently to a SEP, the standard Krylov subspace projection methods such as the Arnoldi algorithm can be applied to solve it [28].

The way of linearization is not unique [28]. Here, we consider the second companion form of linearization [8] for the QEP \((1)\)

\[
\begin{bmatrix}
-D & I_n \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
x \\
x \tilde{x}
\end{bmatrix}
= \frac{1}{\lambda}
\begin{bmatrix}
K & 0 \\
0 & I_n
\end{bmatrix}
\begin{bmatrix}
x \\
x \tilde{x}
\end{bmatrix}
\]

(5)

where \(\tilde{x} = -\lambda Mx\). The computational advantage of using the second companion form will be revealed in section 3.

Since \(K\) is nonsingular, from \((5)\) the corresponding ℓ-SEP and r-SEP of \((1)\) are, respectively, given by

\[
\begin{bmatrix}
-K^{-1}D & K^{-1} \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
x \\
x \tilde{x}
\end{bmatrix}
= \frac{1}{\lambda}
\begin{bmatrix}
x \\
x \tilde{x}
\end{bmatrix}
\]

(6)

and

\[
\begin{bmatrix}
-DK^{-1} & I_n \\
-MK^{-1} & 0
\end{bmatrix}
\begin{bmatrix}
x \tilde{x} \\
K\tilde{x}
\end{bmatrix}
= \frac{1}{\lambda}
\begin{bmatrix}
x \tilde{x} \\
K\tilde{x}
\end{bmatrix}
\]

(7)

In addition to solving the QEP \((1)\) by SEP \((6)\) or \((7)\), one may also work with the GEP \((5)\) to find the desired eigenpairs of \((1)\). The QZ algorithm [20] is the most popular algorithm for solving the dense GEP of the form \((2)\). This procedure reduces the matrix pair \((A, B)\) equivalently to a Hessenberg-triangular pair \((H, R)\) via unitary transformations in a finite number of steps. This truncated \(QZ\) method proposed by Sorensen [25] is one of the approaches for solving large-scale GEPs. The method generalizes the idea of the Arnoldi algorithm to construct a generalized Arnoldi reduction which is a truncation of the \(QZ\) iteration and computes the approximated eigenpairs of the original large-scale GEP from the corresponding reduced Hessenberg-triangular pair. Furthermore, in [11], the generalized \(T\)-skew-Hamiltonian implicitly restarted shift-and-invert Arnoldi (G−SHIRA) algorithm is discussed for solving the palindromic QEP arising from vibrations of fast trains. The generalized \(T\)-isotropic Arnoldi process also produces the generalized Arnoldi reduction for a GEP whose coefficient matrices are \(T\)-skew-Hamiltonian, however, a further \(T\)-bi-isotropic property is required.

However, the linearization technique will double the size of the problem and, in general, matrix structures and spectral properties of the original QEP are not preserved. More importantly, a backward stable technique for linear eigenvalue problems applied to the linearized QEP is not backward stable for the original QEP [27].
To avoid these disadvantages, numerical methods are applied to the large-scale QEP directly. In these methods, the QEP is projected onto a properly chosen low-dimensional subspace to reduce a QEP directly with matrix dimension of low order and solve the reduced QEP by a standard dense matrix approach. Methods of this type include the residual iteration method [21, 12, 19], the Jacobi–Davidson method [22, 23], a Krylov-type subspace method [17], the nonlinear Arnoldi method [29], the second-order Arnoldi (SOAR) method [1, 31, 18] and an iterated shift-and-invert Arnoldi method [32]. While these methods use a similar projection process, the main difference is the selection of projection subspaces.

In this paper, we introduce a Semiorthogonal Generalized Arnoldi (SGA) algorithm for the particular linearized problem (5) to generate a SGA decomposition. The SGA algorithm is a variation of the generalized Arnoldi reduction [25]. We then propose an orthogonal projection approach termed as the SGA method to solve the QEP (1) where the projection subspace is defined through its orthonormal basis obtained from the SGA decomposition.

For SEPs, it has been revealed that even though the approximate eigenvalues computed by orthogonal projection methods tend to converge, the corresponding approximate eigenvectors may converge very slowly and even fail to converge. To deal with this problem, Jia [13] proposed a refined Arnoldi method to compute refined approximate eigenvectors. See also [14]. We will extend this idea and use the SGA decomposition to propose a refinement scheme for QEPs.

Due to the storage requirements and computational costs, the order of the SGA decomposition can not be large and shall be limited. Therefore it is necessary to restart the SGA method. Based on the implicitly shifted QZ iterations proposed by Sorensen in [25], we develop a restart technique for the SGA method, called the Implicitly Restarted SGA (IRSGA) method. Moreover, according to the information of refined approximate eigenvectors (Ritz vectors), we will propose a procedure for selecting better shifts, termed as refined shifts, for the implicitly shifted QZ algorithm to develop an Implicitly Restarted Refined SGA (IRRSGA) method. Compared to the implicitly restarted Arnoldi method applied on the linearized problems (6) and (7), the SGA-type methods, namely IRSGA and IRRSGA, demonstrate better convergence behaviors and require less CPU time in numerical experiments.

The paper is organized as follows. In section 2, we first introduce the SGA algorithm associated with the GEP (5). In section 3, we propose an orthogonal projection method based on the orthonormal basis generated by the SGA algorithm for solving the QEP (1). In section 4, we present a refinement scheme to get better Ritz vectors by taking advantage of the SGA decomposition. In section 5, we develop a restart technique for the SGA-type methods and discuss the selection of shifts according to the information of refinement so that the faster the methods may converge. Numerical examples are presented in section 6 and the concluding remarks are given in section 7.

Throughout this paper, we use the capital letters to denote matrices and the boldface lower case letters to denote vectors. I denotes the identity matrix, e_j is the jth column of the identity matrix I, and 0 denotes zero vectors and matrices. The dimensions of these vectors and matrices are conformed with dimensions used in the context. We adopt the following MATLAB notations: v(i : j) denotes the subvector of the vector v that consists of the ith to the jth entries of v. A(i : j, k : ℓ) denotes the submatrix of the matrix A that consists of the intersection of the rows i to j and the columns k to ℓ, A(i : j,:) and A(:, k : ℓ) select the rows i to j and the columns k to ℓ, respectively, of A. We use .^T and .^H to denote the transpose and conjugate transpose. ∥ · ∥_2 and ∥ · ∥_F denotes the 2-norm and the Frobenius norm, respectively, for a vector or a matrix.

2. THE SGA DECOMPOSITION

In this section, we first give the definition of the SGA decomposition and then discuss the existence and uniqueness of the SGA decomposition in section 2.1. In section 2.2, we will propose a SGA algorithm to generate the SGA decomposition. Subsequently, we discuss the possibility of the early termination of the SGA algorithm.
Definition 2.1 (The SGA decomposition)
Given $M, D, K \in \mathbb{C}^{n \times n}$ and $m \ll n$. We define the $m$th order SGA decomposition of the QEP (1) to be the relation of the form

\begin{align}
-D & \quad I_n \\
-M & \quad 0
\end{align}

\begin{align}
- \begin{bmatrix} Q_m & P_m \end{bmatrix} & = \begin{bmatrix} V_m & U_m \end{bmatrix} \begin{bmatrix} H_m & f_m \end{bmatrix} e_m^T,
\end{align}

\begin{align}
\begin{bmatrix} K & 0 \\
0 & I_n
\end{bmatrix} & = \begin{bmatrix} V_m & U_m \end{bmatrix} R_m,
\end{align}

\begin{align}
Q_m^H Q_m = I_m, \quad V_m^H V_m = I_m \quad \text{and} \quad V_m^H g_m = 0.
\end{align}

where $Q_m, P_m, V_m, U_m \in \mathbb{C}^{n \times m}$, $g_m, f_m \in \mathbb{C}^n$, and $H_m, R_m \in \mathbb{C}^{m \times m}$ are upper Hessenberg matrix and upper triangular matrix, respectively.

Remark 2.1 (i) The orthogonality requirements in (8c) referred to as the semiorthogonality of the SGA decomposition, guarantee the linear independence of columns of $[Q_m]$ and $[V_m^H]$, respectively.

(ii) If the semiorthogonality (8c) is replaced by $Q_m^H Q_m + P_m^H P_m = I_m$, $V_m^H V_m + U_m^H U_m = I_m$ and $V_m^H g_m + U_m^H f_m = 0$, we actually obtain a generalized Arnoldi reduction [25] associated with the GEP (5). Therefore, the SGA decomposition can be also viewed as a variation of the generalized Arnoldi reduction.

2.1. Existence and uniqueness
Given an $N \times N$ matrix $C$, $N = 2n$, a nonzero vector $b \in \mathbb{C}^N$ and a positive integer $m \leq n$, the Krylov matrix of $C$ with respect to $b$ and $m$ is defined by

$$
\mathbb{K}[C, b, m] = [b \ Cb \ \cdots \ C^{m-1}b].
$$

In the following, for convenience, for a matrix $G \in \mathbb{C}^{N \times j}$ we usually partition $G$ of the form $G = [G_1, G_2]$ with $G_1 = G(1 : n, :)$ and $G_2 = G(n + 1 : 2n, :)$.

From (8), if we set

\begin{align}
A & = \begin{bmatrix} -D & I_n \\
-M & 0
\end{bmatrix}, \quad B = \begin{bmatrix} K & 0 \\
0 & I_n
\end{bmatrix},
\end{align}

\begin{align}
Z_m & = \begin{bmatrix} Q_m \\
P_m
\end{bmatrix}, \quad Y_m = \begin{bmatrix} V_m & U_m \end{bmatrix}, \quad \eta_m = \begin{bmatrix} g_m \\
f_m
\end{bmatrix},
\end{align}

then the SGA decomposition (8) can be compactly written as

\begin{align}
AZ_m & = Y_m H_m + \eta_m e_m^T,
\end{align}

\begin{align}
BZ_m & = Y_m R_m,
\end{align}

\begin{align}
Q_m^H Q_m = I_m, \quad V_m^H V_m = I_m \quad \text{and} \quad V_m^H g_m = 0.
\end{align}

Using equations (8) – (10) of the SGA decomposition (11) and based on the proof technique of Theorem 3.3 in [7], we give the following theorem.

Theorem 2.2
Given $z_1 \equiv [q_1] \in \mathbb{C}^N$ with $\|q_1\|_2 = 1$ and set $B z_1 = \rho_1 y_1 \equiv \rho_1 [y_1]$ with $\|y_1\|_2 = 1$ and $\rho_1 > 0$. Let

\begin{align}
K_\ell = \mathbb{K}[B^{-1} A, z_1, 1] & = \begin{bmatrix} K_{\ell,1} \\
K_{\ell,2}
\end{bmatrix}, \quad \text{and} \quad K_r = \mathbb{K}[A B^{-1}, y_1, 1] = \begin{bmatrix} K_{r,1} \\
K_{r,2}
\end{bmatrix}.
\end{align}

Suppose that $K_{\ell,1}$ is of full column rank and $K_{\ell,1} = Q_m R_{\ell,m}$ is the $QR$-factorization with $Q_m e_1 = q_1$ and diagonal entries of $R_{\ell,m}$ are chosen to be positive. Here and hereafter, we use the $QR_+$-factorization to indicate such a $QR$-factorization. Then
(i) $K_{r,1}$ is of full column rank and if $K_{r,1} = V_m R_{r,m}$ is the $QR_+$-factorization, then $V_m e_1 = v_1$.
(ii) Let $P_m = K_{r,2} R_{r,m}$ and $U_m = K_{r,3} R_{r,m}$. Then there exist an unreduced upper Hessenberg matrix $H_m$ with positive subdiagonal entries and an upper triangular $R_m$ with positive diagonal entries satisfying the SGA decomposition \((11)\).
(iii) The SGA decomposition \((11)\) is uniquely determined by $Z_m e_1 = [q_1]$ with $\|q_1\|_2 = 1$.

\textbf{Proof}

(i) Since
\[
\begin{bmatrix}
K_{r,1} \\
K_{r,2}
\end{bmatrix} = \mathbb{K}[AB^{-1}, y_1, m] = \begin{bmatrix}
y_1 & AB^{-1} y_1 & \cdots & (AB^{-1})^{m-1} y_1
\end{bmatrix}
= \frac{1}{\rho_1} B \begin{bmatrix}
z_1 & B^{-1} A z_1 & \cdots & (B^{-1} A)^{m-1} z_1
\end{bmatrix}
= \frac{1}{\rho_1} B K_{\ell} = \frac{1}{\rho_1} \begin{bmatrix}
K & 0 \\
I_m & K_{\ell,2}
\end{bmatrix} = \frac{1}{\rho_1} \begin{bmatrix}
K K_{\ell,1} \\
K_{\ell,2}
\end{bmatrix},
\]
the matrix $K_{r,1} = \frac{1}{\rho_1} K K_{\ell,1}$ is of full column rank and has the unique $QR_+$-factorization $K_{r,1} = V_m R_{r,m}$, with $V_m e_1 = v_1$.
(ii) By assumptions and \((10)\), we get $K_{r} = [V_m'] R_{r,m} = Z_m R_{r,m}$. From (i) and \((10)\), we also have $K_r = [U_m'] R_{r,m} = Y_m R_{r,m}$. It follows from \((12)\) that
\[
BZ_m = BK_{\ell} R_{r,m}^{-1} = \rho_1 K_{r,1} R_{r,m}^{-1} = Y_m (\rho_1 R_{r,m} R_{r,m}^{-1}) = Y_m R_{r,m}
\]
where $R_m$ is upper triangular with positive diagonal entries. On the other hand, it holds that
\[
B^{-1} A \mathbb{K}[B^{-1} A, z_1, m] = \mathbb{K}[B^{-1} A, z_1, m] H_0 + (B^{-1} A)^m z_1 e_m^\top
\]
where $H_0$ is the lower shift matrix, i.e., a matrix with ones below the main diagonal and zeros elsewhere. From \((13)\) and \((12)\) we have
\[
AZ_m = BZ_m R_{r,m} H_0 R_{r,m}^{-1} + (B^{-1} A)^m z_1 e_m^\top R_{r,m}^{-1}
= Y_m (\rho_1 R_{r,m} H_0 R_{r,m}^{-1} + \tilde{Y}_m z_m e_m^\top) + [(I - Y_m \tilde{Y}_m H_0) z_m] e_m^\top
\equiv Y_m H_m + \eta_m e_m^\top
\]
where $z_m = R_{r,m}^{-1} (m, m) B (B^{-1} A)^m z_1 \equiv [z_m^1]$ and $\tilde{Y}_m = [V_m H_0 0_m]$. Since $H_0$ is unreduced Hessenberg with subdiagonal entries “1”, $R_{r,m}$ and $R_m$ are upper triangular with positive diagonal entries, and $V_m$ is orthogonal, it is easily seen that $H_m$ is unreduced Hessenberg with positive subdiagonal entries and $V_m^H g_m = V_m^H [(I_m - V_m V_m^H) z_m, 1] = 0$.
(iii) By (i) and (ii), we know that $Y_m$, $\eta_m$, $R_m$ and $H_m$ are uniquely determined by $Z_m$ so we only need to show that $Z_m$ is unique for given $Z_m(1 :) = [q_1]$ with $\|q_1\|_2 = 1$. From \((11)\), we have $AZ_m = BZ_m R_{r,m}^{-1} H_m + \eta_m e_m^\top$. Let $Z_m = \tilde{Z}_m T_m$ be the $QR_+$-factorization of $Z_m$. Then we have the standard Arnoldi decomposition
\[
B^{-1} A Z_m = \tilde{Z}_m \tilde{H}_m + \eta_m e_m^\top
\]
where $\tilde{H}_m = (T_m R_{r,m}^{-1} H_m + \tilde{Z}_m^H B^{-1} \eta_m e_m^\top) T_m^{-1}$ and $\tilde{Y}_m = (I_m - \tilde{Z}_m \tilde{Z}_m^H) B^{-1} \eta_m e_m^\top T_m^{-1}$. Note that the standard Arnoldi decomposition \((14)\) is unreduced, it is essentially unique. It follows that $Q_m$ and $T_m^{-1}$ of the $QR_+$-factorization $\tilde{Q}_m = Q_m T_m^{-1}$ are unique, and then $P_m = \hat{P}_m T_m$ is unique. This concludes the proof.
Theorem 2.3
If the $m$th order SGA decomposition (11) exists, then
\[
K_l = \mathbb{K}[B^{-1}A, z_1, m] = Z_m[e_1 R_m^{-1}H_m e_1 \cdots (R_m^{-1}H_m)^{m-1} e_1],
\]
(15a)
\[
K_r = \mathbb{K}[AB^{-1}, y_1, m] = Y_m[e_1 H_m R_m^{-1} e_1 \cdots (H_m R_m^{-1})^{m-1} e_1].
\]
(15b)

Proof
It suffices to show
\[
K_l = [z_1 B^{-1}Az_1 \cdots (B^{-1}A)^{m-1}z_1] = Z_m[e_1 R_m^{-1}H_m e_1 \cdots (R_m^{-1}H_m)^{m-1} e_1].
\]
(16)
Since $Z_m e_1 = z_1$, we suppose $(B^{-1}A)^i z_1 = Z_m (R_m^{-1}H_m)^i e_1$, for $i < m$, and prove (16) by induction. From (11) we have
\[
(B^{-1}A)^i z_1 = (B^{-1}A)Z_m (R_m^{-1}H_m)^i e_1
\]
\[
= [Z_m (R_m^{-1}H_m) + B^{-1}\eta_m e_m^\top] (R_m^{-1}H_m)^i e_1
\]
\[
= Z_m (R_m^{-1}H_m)^i e_1 + B^{-1}\eta_m (e_m^\top (R_m^{-1}H_m)^i e_1)
\]
\[
= Z_m (R_m^{-1}H_m)^i e_1
\]
because of $e_m^\top (R_m^{-1}H_m)^i e_1 = 0$, for $i < m$. On the other hand, from (11) follows
\[
AB^{-1}Y_m = Y_m (H_m R_m^{-1}) + \tilde{\eta}_m e_m^\top,
\]
(18)
where $\tilde{\eta}_m = R_m (m, m)^{-1} \eta_m$. Similar to (17), (15b) follows from (18) immediately.

Remark 2.2
Theorem 2.2 shows that $K_{l,1}$ has the QR$_+^m$-factorization, $K_{l,1} = Q_m R_{l,m}$, then the SGA decomposition (11) exists and unique up to $Y_m e_1 = y_1$. Theorem 2.3 shows that if the SGA decomposition (11) exists, then $K_l$ and $K_r$ have the QR$_+^m$-factorizations (15a) and (15b), respectively.

2.2. The SGA algorithm
We now derive an algorithm termed as the SGA algorithm for the computation of the SGA decomposition (9). Given $q_1, p_1 \in \mathbb{C}^n$ with $\|q\|_1 = 1$, let
\[
R_1 = \|Kq_1\|_2 \neq 0, \quad v_1 = Kq_1/R_1, \quad u_1 = p_1/R_1,
\]
\[
H_1 = v_1^H (-Dq_1 + p_1), \quad g_1 = -Dq_1 + p_1 - v_1 H_1 \quad \text{and} \quad f_1 = -Mq_1 - u_1 H_1,
\]
then $q_1, p_1, v_1, u_1, g_1, f_1, R_1$ and $H_1$ satisfy the SGA decomposition (8) with $m = 1$.

Suppose that we have computed the $j$th order ($j < m$) SGA decomposition
\[
\begin{bmatrix}
-D & I_n \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
Q_j \\
P_j
\end{bmatrix}
= \begin{bmatrix}
V_j \\
U_j
\end{bmatrix} R_j
\]
(19a)
\[
\begin{bmatrix}
K & 0 \\
0 & I_n
\end{bmatrix}
\begin{bmatrix}
Q_j \\
P_j
\end{bmatrix}
= \begin{bmatrix}
V_j \\
U_j
\end{bmatrix} R_j
\]
(19b)
\[
Q_j^H Q_j = I_j, \quad V_j^H V_j = I_j \quad \text{and} \quad V_j^H g_j = 0.
\]
(19c)
To expand the SGA decomposition to order $j + 1$, we first assume that the residual vector $g_j \neq 0$. The case $g_j = 0$ will be discussed later. Our goal is to find suitable updating vectors and scalars satisfying the SGA decomposition of order $j + 1$
\[
\begin{bmatrix}
-D & I_n \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
Q_j \\
P_j
\end{bmatrix}
= \begin{bmatrix}
V_j & v \\
U_j & u
\end{bmatrix}
\begin{bmatrix}
H_j & h \\
e_j & \alpha
\end{bmatrix}
\]
\[
+ \begin{bmatrix}
g_{j+1} \\
f_{j+1}
\end{bmatrix}
\begin{bmatrix}
\alpha \\
e_{j+1}
\end{bmatrix}
\]
(20a)
\[
\begin{bmatrix}
K & 0 \\
0 & I_n
\end{bmatrix}
\begin{bmatrix}
Q_j \\
P_j
\end{bmatrix}
= \begin{bmatrix}
V_j & v \\
U_j & u
\end{bmatrix}
\begin{bmatrix}
R_j \\
0
\end{bmatrix}
\]
(20b)
\[
Q_j^H Q_{j+1} = I_{j+1}, \quad V_j^H V_{j+1} = I_{j+1} \quad \text{and} \quad V_j^H g_{j+1} = 0.
\]
(20c)
where \( Q_{j+1} = [Q_j \, q] \) and \( V_{j+1} = [V_j \, v] \). Comparing the leading \( j \) columns of (20a) with (19a), we get
\[
\gamma = \|g_j\|_2 \neq 0, \quad v = g_j/\gamma \neq 0 \quad \text{and} \quad u = f_j/\gamma.
\] (21)

Equating the \((j+1)\)st column on both sides of (20b) and noting (20c), the vector \( q \) must satisfy
\[
Kq = V_j r + v\rho \quad \text{and} \quad Q_j^H q = 0
\] (22)

Premultiplying (22) by \( Q_j^H K^{-1} \) and applying the relation \( KQ_j = V_j R_j \) gives
\[
0 = Q_j^H K^{-1} V_j r + Q_j^H K^{-1} v\rho = R_j^{-1} r + Q_j^H K^{-1} v\rho
\]
and it follows that
\[
r = -R_j Q_j^H K^{-1} v\rho.
\] (23)

Substituting (23) into (22), we have
\[
q = K^{-1} V_j r + K^{-1} v\rho
= (Q_j R_j^{-1})(-R_j Q_j^H K^{-1} v\rho) + K^{-1} v\rho
= (I_j - Q_j Q_j^H) K^{-1} v\rho
\]
where \( \rho \equiv \| (I_j - Q_j Q_j^H) K^{-1} v\|_2 \) so that \( Q_j^H v = 0 \) and \( \| q \|_2 = 1 \). Note that \( \rho \) is well-defined, otherwise, \( \| (I_j - Q_j Q_j^H) K^{-1} v\|_2 = 0 \) implies \( K^{-1} v \in \text{span}\{Q_j\} \) and hence \( v = K Q_j c = V_j R_j c \) for some constant vector \( c \). However, \( V_j^H v = 0 \) implies \( v = 0 \) which contradicts the fact (21). After determining \( u, r \) and \( \rho \), (20b) shows that \( p \) can be directly computed by
\[
p = U_j r + u\rho.
\]

Equating the \((j+1)\)st column on both sides of (20a), we know that if we take
\[
\begin{bmatrix}
\mathbf{h} \\
\alpha
\end{bmatrix}
= \begin{bmatrix}
V_j^H (-Dq + p) \\
\mathbf{v}^H (-Dq + p)
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
g_{j+1} \\
f_{j+1}
\end{bmatrix}
= \begin{bmatrix}
-D & I_n \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
\mathbf{q} \\
p
\end{bmatrix}
- \begin{bmatrix}
V_j & \mathbf{v}
\end{bmatrix}
\begin{bmatrix}
\mathbf{h} \\
\alpha
\end{bmatrix}
\] (24)

then \( V_j^H g_{j+1} = 0 \) and this completes the \((j+1)\)st expanding of the SGA decomposition.

**Breakdown and deflation.** As we encounter \( g_j = 0 \), there are two possibilities, which are called breakdown and deflation. A breakdown occurs if the vector sequence \( \{[\mathbf{v}_1], \ldots, [\mathbf{v}_r], [\mathbf{0}]\} \) is linearly dependent. In this case, both \( K_j(B^{-1} A, q_1) \) and \( K_j(AB^{-1}, v_1) \) are invariant subspaces simultaneously and hence the expanding process terminates. On the other hand, it may happen that \( \{[\mathbf{u}_1], \ldots, [\mathbf{u}_r], [\mathbf{0}]\} \) is linearly independent. This situation is called deflation and the expanding process of the SGA decomposition should continue with modified orthogonality requirements.

When a deflation is detected at step \( j \), we assign \( \gamma \) any nonzero number (say \( \gamma = 1 \), \( \mathbf{v} = g_j = 0 \) and \( \mathbf{u} = f_j/\gamma \neq 0 \) to start the \((j+1)\)st expanding process of the SGA decomposition. Without repeating the discussions above, it is easy to see that \( \mathbf{v}, \mathbf{u} \) and \( \gamma \) satisfy the \( j \)th column of (20a) but
\[
V_j^H V_{j+1} = [f_j \, 0].
\]
Equating the \((j+1)\)st column on both sides of (20b) shows that \( q = K^{-1} V_j r + Q_j(R_j^{-1} r) \) (since \( KQ_j = V_j R_j \)) and the orthogonality requirement \( \{q_1, \ldots, q_j, q\} \) in (20c) enforces \( r = 0 \) and \( q = 0 \). Again, by taking \( \rho \) any nonzero number (say \( \rho = 1 \)) and then setting \( p = u\rho = f_j/\gamma^{-1} \rho \), the updating \( q, p, r \) and \( \rho \) satisfy the \((j+1)\)st column of (20b) but \( Q_j^H Q_{j+1} = [f_j \, 0] \). This indicates that if the expanding process of the SGA decomposition encounters deflation at a certain step then the updating \( v \)-vector and \( q \)-vector will be zero simultaneous in the next expanding process. Therefore, the zero vectors of the \( V \)-matrix and the \( Q \)-matrix in a deflated SGA decomposition appear in the same columns.

To accomplish the \((j+1)\)st expanding process of the SGA decomposition, the equations in (24) are given by
\[
\begin{bmatrix}
\mathbf{h} \\
\alpha
\end{bmatrix}
= \begin{bmatrix}
V_j^H p \\
0
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
g_{j+1} \\
f_{j+1}
\end{bmatrix}
= \begin{bmatrix}
-D & I_n \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
0 \\
p
\end{bmatrix}
- \begin{bmatrix}
V_j & 0
\end{bmatrix}
\begin{bmatrix}
\mathbf{h} \\
\alpha
\end{bmatrix}.
\]
In summary, if deflations occur at step $1 < j_1, \ldots, j_d \leq m$, then we have the $m$th order deflated SGA decomposition

\[
\begin{bmatrix}
-D & I_n \\
-M & 0
\end{bmatrix}
\begin{bmatrix}
\hat{Q}_m \\
\hat{P}_m
\end{bmatrix}
= \begin{bmatrix}
\hat{V}_m \\
\hat{U}_m
\end{bmatrix}
\hat{H}_m + \begin{bmatrix}
g_m \\
f_m
\end{bmatrix}
\begin{bmatrix}
e^T_m \\
e_m
\end{bmatrix}
\tag{25a}
\]

\[
\begin{bmatrix}
K & 0 \\
0 & I_n
\end{bmatrix}
\begin{bmatrix}
\hat{Q}_m \\
\hat{P}_m
\end{bmatrix}
= \begin{bmatrix}
\hat{V}_m \\
\hat{U}_m
\end{bmatrix}
\hat{R}_m \quad \text{and} \quad \hat{V}_m \hat{g}_m = 0 
\tag{25b}
\]

\[
\hat{Q}_m^H \hat{Q}_m = J_m, \quad \hat{V}_m^H \hat{V}_m = J_m 
\tag{25c}
\]

where $Q_m(:,j_i) = V_m(:,j_i) = 0$, $R_m(1:j_i-1,j_i) = 0$, $H_m(j_i,j_i) = 0$, $R_m(j_i,j_i)$, $H_m(j_i,j_i-1)$ are nonzero numbers and

\[
J_m(s,t) = \begin{cases} 
1 & \text{if } s = t \neq j_i, \\
0 & \text{otherwise, } \\
i = 1, \ldots, d.
\end{cases}
\]

The following theorem distinguishes the deflation and breakdown.

**Theorem 2.4** ([1], Lemma 3.2) For a sequence of linearly independent vectors $\{y_1, \ldots, y_m\}$ with partition $y_i = [y_i]\gamma_i$, if there exists a subsequence $\{v_i_1, \ldots, v_i_r\}$ of the $v$ vectors that are linearly independent and the remaining vectors are zeros, $v_{i_1+1} = \cdots = v_{i_r} = 0$, then a vector $y = [\gamma_i]u$ is in span$\{y_1, \ldots, y_m\}$ if and only if $u$ is in span$\{u_{i_1+1}, \ldots, u_{i_r}\}$.

The pseudocode for the SGA algorithm that iteratively generates an $m$th order (deflated) SGA decomposition is listed in Algorithm 2.1.

**Algorithm 2.1 SGA Algorithm**

**Input:** $M, D, K \in \mathbb{C}^{n \times n}$, $q_1, p_1 \in \mathbb{C}^n$ with $\|q_1\|_2 = 1$ and $m \geq 1$.

**Output:** $Q_m$, $V_m$, $U_m$, $g_m := g$, $f_m := f$, $M_m$, $D_m$, upper Hessenberg $H_m \in \mathbb{C}^{m \times m}$ and upper triangular $R_m \in \mathbb{C}^{m \times m}$ satisfy the SGA decomposition (8) of order $m$.

1. $Q_1 := q_1$, $R_1 := \|Kq_1\|_2$, $V_1 := Kq_1/R_1$, $U_1 := p_1/R_1$, $M_1 := Mq_1$, $D_1 := Dq_1$.
2. $g := -D_1 + p_1$, $H_1 := V_1^Hg$, $f := -M_1 - U_1H_1$.
3. for $j = 1, 2, \ldots, m - 1$ do
4. if $g \neq 0$ then
5. $\gamma := \|g\|_2$; $v := g/\gamma$; $u := f/\gamma$; $V_{j+1} := [V_j v]$; $U_{j+1} := [U_j u]$; $H_j := \left[ \begin{array}{c} H_j \\ \gamma e_j^T \end{array} \right]$;
6. Solve $Kq = v_{j+1}$ for $q$.
7. $r := Q_j^Hq$; $q := q - Q_jr$; $\rho := \|q\|_2^{-1}$; $q := q/\rho$; $\mu := Mq$; $\delta := Dq$; $Q_{j+1} := [Q_j q]$; $M_{j+1} := [M_j \mu]$; $D_{j+1} := [D_j \delta]$; $R_{j+1} := \left[ \begin{array}{c} R_j \\ 0 \\ \rho \end{array} \right]$;
8. $g := -\delta + U_{j+1}R_{j+1}(j+1)$; $h := V_{j+1}^Hg$; $H_{j+1} := [H_j h]$; $f := -\mu - U_{j+1}h$.
9. else if $f \in \text{span}\{u_i \mid i : v_i = 0, 1 \leq i \leq j\}$ then breakdown
11. else
12. $V_{j+1} := [V_j 0]$; $U_{j+1} := [U_j f]$; $Q_{j+1} := [Q_j 0]$; $M_{j+1} := [M_j 0]$; $D_{j+1} := [D_j 0]$;
14. $h := V_{j+1}^Hf$; $H_{j+1} := \left[ \begin{array}{c} H_j \\ e_j^T \end{array} \right]$; $R_{j+1} := \left[ \begin{array}{c} R_j \\ 0 \\ 1 \end{array} \right]$; $g := f - V_{j+1}h$; $f := -U_{j+1}h$.
16. end if
17. end if
18. end if
19. end for
Remark 2.3

The following remarks give some detailed explanations of the SGA algorithm.

(i) At each expanding process of the SGA decomposition, we need to solve a linear system (see line 6 of the SGA algorithm). To make the computation more efficient, a factorization of $K$, such as the $LU$ factorization, should be made available outside of the first for-loop of the SGA algorithm.

(ii) At lines 8 and 15 of the SGA algorithm, we additionally store the vectors $Dq_j$ and $Mq_j$ at each expanding step and output two $n \times m$ matrices

$$
\mathbb{D}_m := [Dq_1 \cdots Dq_m] = DQ_m \quad \text{and} \quad \mathbb{M}_m := [Mq_1 \cdots Mq_m] = MQ_m.
$$

The pre-stored matrices $\mathbb{D}_m$ and $\mathbb{M}_m$ save computational costs in the subsequent projection process for solving the QEP.

(iii) From (8b), we know that $P_m$ can be completely determined by $U_m$, that is, for $j = 1, \ldots, m$, $p_j$ can be replaced by the relation $p_j = U_m(:, 1 : j)R_m(1 : j, j)$. See line 9 of the SGA algorithm. Hence we only need to evaluate and store $Q_m$, $V_m$, $g_m$, $U_m$, $f_m$, $H_m$ and $R_m$ as we implement the SGA algorithm.

(iv) At line 12 of the SGA algorithm, we decide whether the expanding process encounters a deflation or a breakdown. In practice, we use the modified Gram--Schmidt procedure to check it as suggested in [1].

3. THE SGA METHOD FOR SOLVING QEPS

In this section, we use the unitary matrix $Q_m$ produced by the SGA algorithm to develop an orthogonal projection technique to solve the QEP. For simplicity, we assume that the deflation does not occur and hence $Q_m^H Q_m = I_m$. When the deflation occurs, the same orthogonal projection technique is applied with the modification of replacing $Q_m$ with the nonzero columns of $Q_m$ shown in (25).

3.1. The SGA method

The SGA method applies the Rayleigh-Ritz subspace projection technique on the subspace $Q_m = \text{span}\{Q_m\}$ with the Galerkin condition:

$$
(\theta^2 M + \theta D + K)\psi \perp Q_m,
$$

that is, we seek an approximate eigenpair $(\theta, \psi)$ with $\theta \in \mathbb{C}$, $\psi \in Q_m$ such that

$$
\omega^*(\theta^2 M + \theta D + K)\psi = 0 \quad \text{for all } \omega \in Q_m
$$

(26)

where $\cdot^*$ denotes the transpose $\cdot^T$ when $M, D, K$ are real or complex symmetric, otherwise, $\cdot^*$ denotes the conjugate transpose $\cdot^H$ of matrices. Since $\psi \in Q_m$, it can be written as $\psi = Q_m\xi$ and (26) implies that $\theta$ and $\xi$ must satisfy the reduced QEP:

$$
(\theta^2 M_m + \theta D_m + K_m)\xi = 0
$$

(27)

where

$$
M_m = Q_m^* M Q_m, \quad D_m = Q_m^* D Q_m, \quad K_m = Q_m^* K Q_m.
$$

(28)

The eigenpair $(\theta, \xi)$ of the small-scale QEP (27) defines a Ritz pair $(\theta, Q_m\xi)$ of the QEP (1) whose accuracy is measured by the norm of the residual vector $r_{\theta, \xi} = (\theta^2 M + \theta D + K)Q_m\xi$.

Note that by explicitly formulating the matrices $M_m$, $D_m$, and $K_m$, essential structures of $M$, $D$, and $K$ are preserved. For example, if $M$ is symmetric positive definite, so is $M_m$. As a result, essential spectral properties of the QEP will be preserved. For example, if the QEP is a gyroscopic dynamical system in which $M$ and $K$ are symmetric, one of them is positive definite, and $D$ is...
skew-symmetric, then the reduced QEP is also a gyroscopic system. It is known that in this case, the eigenvalues are symmetrically placed with respect to both the real and imaginary axes [16]. Such a spectral property will be preserved in the reduced QEP.

Before we present the SGA method for solving the QEPs, we discuss how to take advantage of the SGA algorithm to efficiently generate the coefficient matrices \((M, D, K)\) of the projected QEP (28). As we describe in Remark 2.3(ii), the resultant matrices \(M_m := MQ_m\) and \(D_m := DQ_m\) produced from the SGA algorithm provide us the necessary multiplications of \(M, D\) with \(Q_m\). For the projected matrix \(K_m\), even if the SGA algorithm does not exactly perform the matrix-vector product of \(K\) and \(q_j\) at each step, \(j = 1, \ldots, m\), we can use the equality \(KQ_m = V_mR_m\) in (8b) to reduce the computational costs. The product of \(V_mR_m\) needs about \(2nm^2\) flops, but the product of \(KQ_m\) needs about \(2n^2m\) flops. Therefore, the small-scale matrices \(M_m\) and \(D_m\) can be respectively generated by

\[
M_m = Q_m^*M_m, \quad D_m = Q_m^*D_m \quad \text{and} \quad K_m = Q_m^*V_mR_m. \tag{29}
\]

Totally, (28) needs about \(6n^2m + 6nm^2\) flops to generate the coefficient matrices of the projected QEP (27), however, the matrix products (29) only need \(8nm^2\) flops. Also note that if we consider the first companion form linearization of the QEP (1), there is no such an advantage. That is, (28) is the only way to generate the coefficient matrices of the reduced QEP (27).

**Algorithm 3.1** The SGA method

**Input:** \(M, D, K \in \mathbb{C}^{n \times n}, q_1, p_1 \in \mathbb{C}^n\) with \(\|q_1\|_2 = 1\) and \(m \geq k \geq 1\).

**Output:** \(m\) Ritz pairs and their relative residuals.

1. Run the SGA algorithm (Algorithm 2.1) to generate an \(m\)th order SGA decomposition (8).
2. Compute \(M_m, D_m\) and \(K_m\) via (29).
3. Solve the reduced QEP (27) for \((\theta_i, \xi_i)\) with \(\|\xi_i\|_2 = 1, i = 1, \ldots, 2m\) and sorting Ritz values so that \(\{(\theta_1, Q_m\xi_1), \ldots, (\theta_k, Q_m\xi_k)\}\) are wanted Ritz pairs.
4. Test the accuracy of Ritz pairs \((\theta_i, v_i)\), \(v_i = Q_m\xi_i\), \(i = 1, \ldots, k\) as approximate eigenvalues and eigenvectors of the QEP (1) by the relative norms of residual vectors:

\[
\frac{\|(\theta_i^2 M + \theta_i D + K)v_i\|_2}{|\theta_i^2|\|M\|_F + |\theta_i|\|D\|_F + \|K\|_F}, \quad i = 1, \ldots, k. \tag{30}
\]

### 3.2. Projection subspace

In this subsection we explain the motivation of choosing the projection subspace \(Q_m = \text{span}\{Q_m\}\) where \(Q_m\) is generated from the SGA algorithm. We first recall a lemma in the SOAR method [1].

**Lemma 3.1** ([1], Lemma 2.2)

Let \(C\) be an arbitrary \(N \times N\) matrix. Let \(W_{m+1} = [W_m, w_{m+1}]\) be an \(N \times (m + 1)\) rectangular matrix that satisfies \(CW_m = W_{m+1}H_m\) for an \((m + 1) \times m\) upper Hessenberg matrix \(H_m\). Then there is an upper triangular matrix \(T_m\) such that

\[
W_m T_m = [w_1\ Cw_1\ \cdots\ C^{m-1}w_1].
\]

Furthermore, if the first \(m - 1\) subdiagonal elements of \(H_m\) are nonzero, then \(T_m\) is nonsingular and

\[
\text{span}\{W_m\} = \mathcal{K}_m(C, w_1).
\]

Next, we consider a Krylov subspace associated with the linearized eigenvalue problem (3) and show that it is embedded into a larger subspace spanned by some column vectors in the SGA decomposition (8).
Theorem 3.1
Consider the SGA decomposition (8) of order \( m \). Let
\[
\hat{Q}_m = \frac{n}{n} \begin{bmatrix} Q_m & m & 1 \\ 0 & -MQ_m & p_1 \end{bmatrix} \in \mathbb{C}^{2n \times (2m+1)}.
\] (31)

Then, for \( A \) and \( B \) defined in (9), we have \( K_m(B^{-1}A, [q^i_{p_1}]) \subseteq \text{span}\{\hat{Q}_m\} \).

Proof
From (11b), we have \( [V_m] = \begin{bmatrix} K & 0 \\ 0 & I_n \end{bmatrix} [P_m] R_m^{-1} \). Substituting it into the equation (11a) and then premultiplying it by \( \begin{bmatrix} K^{-1} & 0 \\ 0 & I_n \end{bmatrix} \), we get
\[
\begin{bmatrix} -K^{-1}D & K^{-1} \\ -M & 0 \end{bmatrix} [Q_m] = \begin{bmatrix} Q_m & q^t_m & e^t_m \end{bmatrix}
\] (32)
where \( H_m = R_m^{-1}H_m \) is an unreduced upper Hessenberg matrix, \( q^t_m = K^{-1}g_m \) and \( p^t_m = f_m \). By (32) and Lemma 3.1, we know that
\[
K_m \left( B^{-1}A, \begin{bmatrix} q^i \\ p_1 \end{bmatrix} \right) = K_m \left( \begin{bmatrix} -K^{-1}D & K^{-1} \\ -M & 0 \end{bmatrix}, \begin{bmatrix} q^i \\ p_1 \end{bmatrix} \right) = \text{span} \left\{ \begin{bmatrix} Q_m \\ p^t_m \end{bmatrix} \right\}
\] (33)
and the set \( \{[q^i_{p_1}], \ldots, [q^m_{p_1}]\} \) is a non-orthonormal basis of the above Krylov subspace (33). Next, we show that
\[
[q^i_{p_1}] \in \text{span}\{\hat{Q}_m\} \quad \text{for } i = 1, \ldots, m,
\] (34)
and the conclusion of Theorem 3.1 follows directly from (33) and (34).

To prove (34), it suffices to show that \( p_i \in \text{span}\{ -MQ_m, p_1 \}, 1 \leq i \leq m \). We prove this by induction. Clearly, \( p_1 \in \text{span}\{ -MQ_m, p_1 \} \). Suppose that \( p_1, \ldots, p_i \in \text{span}\{ -MQ_m, p_1 \} \) for \( 1 < i < m \). From the equality (32), we have \( -MQ_m = P_m H_m^\ell + p^t_m e^t_m \). Thus,
\[
-Q_m q^t_i = P_m H_m^\ell \left( i, i \right) = P_i H_m^\ell \left( 1 : i, i \right) + p_{i+1} H_m^\ell \left( i+1, i \right)
\]
and it follows that
\[
p_{i+1} = H_m^\ell \left( i+1, i \right)^{-1} \left(-Q_m q^t_i - P_i H_m^\ell \left( 1 : i, i \right) \right) \in \text{span}\{ -MQ_m, p_1 \}.
\]
We complete the proof. \( \square \)

Instead of using the Krylov subspace \( K_m(B^{-1}A, [q^i_{p_1}]) \), we choose the larger subspace \( K_m(B^{-1}A, [q^i_{p_1}]) \) to extract approximations of eigenpairs. To project the coefficient matrices of the GEP (5) onto the subspace \( \text{span}\{\hat{Q}_m\} \), we get
\[
\hat{Q}_m^* \begin{bmatrix} -D & I_n \\ -M & 0 \end{bmatrix} \hat{Q}_m = m \begin{bmatrix} -D_m & m & 1 \\ -N_m & -M_m & 0 \\ -p^*_m MQ_m & 0 & 0 \end{bmatrix} \equiv \hat{A}
\] (35a)
\[
\hat{Q}_m^* \begin{bmatrix} K & 0 \\ 0 & I_n \end{bmatrix} \hat{Q}_m = m \begin{bmatrix} K_m & m & 1 \\ 0 & -N_m & 0 \\ 0 & -p^*_m MQ_m & p^*_m p_1 \end{bmatrix} \equiv \hat{B}
\] (35b)
where \( M_m, D_m, K_m \) are defined in (28) and \( N_m = Q^*_m M^* MQ_m \). Therefore, the GEP (5) is reduced to the problem
\[
\hat{A}s = \nu \hat{B}s
\] (36)
with \( \tilde{A} \) and \( \tilde{B} \) defined in (35). Observe that if we premultiply (36) by the nonsingular matrix

\[
L \equiv \begin{bmatrix}
I_m & 0 & 0 \\
0 & I_m & 0 \\
0 & p_1^*M_QmN_m^{-1} & 1
\end{bmatrix}
\]

then the coefficient matrices of the resultant GEP \((L\tilde{A})s = \mu(L\tilde{B})s\) are respectively of the forms

\[
L\tilde{A} \equiv \begin{bmatrix}
-D_m & -M_m & Q_m^*p_1 \\
N_m & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}, \quad L\tilde{B} \equiv \begin{bmatrix}
K_m & 0 & 0 \\
0 & N_m & -Q_m^*M_p^*p_1 \\
0 & 0 & c
\end{bmatrix}
\]

where \( c = p_1^*(I_m - MQ_mN_m^{-1}Q_m^*M^*)p_1 \). The pencil obtained from the last component of both matrices in (37) either provides the zero eigenvalue or be a singular pencil. In both cases, the eigenvalues computed from this pencil are not wanted. Therefore, we can simply drop the last column and row of both matrices in (37) to consider the leading \( 2m \times 2m \) submatrices, which is just the first companion form linearization [8] of the reduced GEP (27), for solving QEPs.

4. REFINED SGA METHOD

As we obtain a Ritz pair \((\theta, \nu_0)\) by the SGA method, a refinement strategy for QEP is to seek a unit vector \( \nu_+ \in Q_m = \text{span}\{Q_m\} \) satisfying

\[
\nu_+ = \arg \min_{\nu \in Q_m, \|\nu\|_2 = 1} \| (\theta^2M + \theta D + K)\nu \|_2.
\]

Here we call \( \nu_+ \) the refined Ritz vector corresponding to the Ritz value \( \theta \). We next turn to propose a novel refinement scheme by taking advantage of the SGA decomposition for computing refined Ritz vectors. An other refinement scheme for QEPs we refer to [15].

Let \((\theta, \xi_0)\) be an eigenpair obtained from the small-scale QEP (27), then \((\theta, \nu_0) = (\theta, Q_m\xi_0)\) is a Ritz pair the QEP (1). To solve the optimization problem (38), we find that

\[
(\theta^2M + \theta D + K)Q_m = \theta^2(-U_mH_m - f_me_m^\top) + \theta(P_m - V_mH_m - g_me_m^\top) + V_mR_m
\]

\[
= V_m(-\theta H_m + R_m) + g_m(-\theta e_m) + U_m(-\theta^2H_m + \theta R_m) + f_m(-\theta^2e_m)
\]

\[
= \begin{bmatrix} V_m & g_m & U_m & f_m \end{bmatrix} \begin{bmatrix}
-\theta H_m + R_m \\
-\theta e_m \\
-\theta^2 H_m + \theta R_m \\
-\theta^2 e_m
\end{bmatrix},
\]

where we use the SGA decomposition (8) in the first two equalities. Since \( V_m \) is a column orthonormal matrix, the QR factorization of \( \begin{bmatrix} V_m & g_m & U_m & f_m \end{bmatrix} \) is of the form

\[
\begin{bmatrix} V_m & g_m & U_m & f_m \end{bmatrix} = \begin{bmatrix} \tilde{V}_m & \tilde{g}_m & \tilde{U}_m & \tilde{f}_m \end{bmatrix} \begin{bmatrix} I_m & t_{12} & T_{13} & t_{14} \\
t_{22} & t_{23} & i_{24} & T_{33} \\
t_{34} & t_{44} & \end{bmatrix}
\]

(40)

where \( \begin{bmatrix} \tilde{V}_m & \tilde{g}_m & \tilde{U}_m & \tilde{f}_m \end{bmatrix} \) is unitary. Since the vector 2-norm is invariant under unitary transformations, (39) and (40) imply

\[
\min_{\nu \in Q_m, \|\nu\|_2 = 1} \| (\theta^2M + \theta D + K)\nu \|_2 = \min_{\|\xi\|_2 = 1} \| (\theta^2M + \theta D + K)Q_m\xi \|_2 = \min_{\|\xi\|_2 = 1} \| S(m, \theta)\xi \|_2
\]
where

\[
S(m, \theta) = \begin{bmatrix}
I_m & t_{12} & T_{13} & t_{14} \\
t_{22} & t_{23} & T_{24} & t_{24} \\
T_{33} & T_{34} & T_{34} & t_{44}
\end{bmatrix}
\begin{bmatrix}
-\theta H_m + R_m \\
-\theta e_m^\top \\
-\theta^2 H_m + \theta R_m \\
-\theta^2 e_m^\top
\end{bmatrix} \in \mathbb{C}^{(2m+2) \times m}.
\]

(41)

Since the right singular vector \(V_0 e_m\) of \(S(m, \theta)\) corresponding to the smallest singular value \(s_{\theta, \text{min}}\) yields the minimum \(\|S(m, \theta) V_0 e_m\|_2 = s_{\theta, \text{min}}\), as a consequence, the unit vector \(u^+_i \equiv Q_m V_0 e_m\) is the solution to the minimization problem (38) with minimum \(s_{\theta, \text{min}}\). In summary, we have the following theorem.

**Theorem 4.1**

Let \((\theta, Q_m, \xi_i)\) be a Ritz pair the QEP (1) computed from the SGA method. Let \(S(m, \theta) = U_\theta \Sigma_\theta (V_\theta)^H\) be a singular value decomposition of \(S(m, \theta)\) defined in (41) and \(s_{\theta, \text{min}}\) be its smallest singular value. Then the vector \(u^+_i \equiv Q_m V_0 e_m\) is the solution to the optimization problem (38) with minimum \(s_{\theta, \text{min}}\).

When applying the refinement strategy for several Ritz pairs, we compute the QR factorization (40) only once and subsequently use the factorization for refining each Ritz pair. Combining the SGA method with the refinement strategy, we propose the refined SGA (RSGA) method in Algorithm 4.1.

**Algorithm 4.1 RSGA method**

**Input:** \(M, D, K \in \mathbb{C}^{n \times n}, q_1, p_1 \in \mathbb{C}^n\) with \(\|q_1\|_2 = 1\) and \(m \geq k \geq 1\).

**Output:** \(m\) refined Ritz pairs and their relative residuals.

1. Run steps 1–3 of the SGA method to obtain \(k\) wanted Ritz pairs \((\theta_i, Q_m \xi_i), i = 1, \ldots, k\).
2. Calculate a QR factorization of \([V_m g_m U_m f_m]\) where the Q-factor and R-factor are denoted as in (40).
3. for \(i = 1, \ldots, k\) do
4. Calculate the matrix \(S(m, \theta_i)\) as defined in (41).
5. Calculate a compact singular value decomposition of \(S(m, \theta_i) = U_\theta_i \Sigma_\theta_i (V_\theta_i)^H\).
6. Let \(s_{\theta_i, \text{min}}\) be the smallest singular value of \(S(m, \theta_i)\). Then the refined Ritz vector is given by \(u^+_i = Q_m V_0 e_m\) and the corresponding relative residual is given by

\[
\frac{\| (\theta_i^2 M + \theta_i D + K) u^+_i \|_F}{\| \theta_i \|_F M + \| \theta_i \|_F D + \| K \|_F} = \frac{s_{\theta_i, \text{min}}}{\| \theta_i \|_F M + \| \theta_i \|_F D + \| K \|_F}.
\]

(42)

7. end for

5. IMPLICIT RESTART OF THE SGA METHOD

Similar to the standard implicitly restarted Arnoldi (IRA) method [24] for SEPs, the SGA/RSGA method also needs restarting to control storage and orthogonalization expense. In this section, we will apply the implicitly shifted QZ iteration [25] to implicitly restart the SGA/RSGA method, namely IRSGA/IRRSSGA.

**5.1. The IRSGA method and the IRRSGA method**

In this subsection, we first briefly discuss the implicitly restarted step of the SGA algorithm based on the implicitly shifted QZ iteration [25]. For details, see [25, 26].

Suppose we have computed the \(m\)th order SGA decomposition (11). For given shifts \(\vartheta_1, \ldots, \vartheta_p\), \(p = m - k\), which are in general the unwanted approximate eigenvalues, let \(E_i\) and \(F_i\) be unitary matrices computed by the implicitly shifted QZ iteration with the single shift \(\vartheta_i\), \(i = 1, \ldots, p\). Write \(E^+ = E_1 \cdots E_p\) and \(F^+ = F_1 \cdots F_p\). Note that \(F_i\) is upper Hessenberg, \(i = 1, \ldots, p\).
Let $H^+_m = (E^+)^H H_m F^+$, $R^+_m = (E^+)^H R_m F^+$, $Z^+_m = Z_m F^+$ and $Y^+_m = Y_m E^+$. Then $H^+_m$ and $R^+_m$ are again upper Hessenberg and upper triangular, respectively. Let $Q^+_m = Q_m F^+$ and $V^+_m = V_m E^+$ then $(Q^+_m)^H Q^+_m = (V^+_m)^H V^+_m = I_m$. Postmultiplying (11a) and (11b) by $F^+$, we get

$$AZ^+_m = Y^+_m H^+_m + \eta_m e^+_m F^+,$$

$$BZ^+_m = Y^+_m R^+_m.$$  

Since $e^+_m F_1 = [0 \cdots 0 \alpha_1 \beta_1]$, by induction, we see that the first $k-1$ entries of $e^+_m F^+$ are zeros. Let $\eta \equiv h_{k+1} \epsilon_k Y_k + F^+(m,k) \eta_m$. Drop the last $m-k$ columns of (43a) and (43b), and then set $\eta_k \equiv \eta$. Then, by writing $Z^+_k = \left[ Q^+_k \right]$, $Y^+_k = \left[ V^+_k \right]$ and $\eta_k = \left[ \eta_k \right]$, we get the $k$ step SGA decomposition:

$$AZ^+_k = Y^+_k H^+_k + \eta_k^+ e^+_k,$$

$$BZ^+_k = Y^+_k R^+_k.$$  

$$Q_k^+ H_k = (V_k^+)^H V_k^+ = I_k,$$  

$$Q_k^+ g_k^+ = 0.$$  

Now, we present the IRSGA method and the IRRSGA method in the following algorithm.

**Algorithm 5.1 The IRSGA/IRRSGA method**

**Input:** $M, D, K \in \mathbb{C}^{n \times n}$ and $m \geq k \geq 1$.

**Output:** $k$ desired eigenpairs.

1. for $i = 1, 2, \ldots$ do
2. Run the SGA algorithm (Algorithm 2.1) to generate an $m$th order SGA decomposition.
3. Run the SGA method (Algorithm 3.1) or the RSGA method (Algorithm 4.1) to compute $k$ candidates of Ritz pairs and check their convergence by (30) or (42).
4. if $(\text{#convergent Ritz pairs}) \geq k$ then
5. break
6. else
7. Select $p := m - k$ shifts $\vartheta_1, \ldots, \vartheta_p$.
8. Let $\varepsilon := e^+_m$ and $\eta := \eta_m$
9. for $i = 1, \ldots, p$ do
10. Compute unitary matrices $E_i$ and $F_i$ by the implicit-$QZ$ step with a single shift $\vartheta_i$ so that $E_i^H H_m F_i$ and $E_i^H R_m F_i$ are upper Hessenberg and upper triangular, respectively.
11. Update $H_m := E_i^H H_m F_i$, $R_m := E_i^H R_m F_i$, $Z_m := Z_m F_i$, $Y_m := Y_m E_i$ and $\varepsilon := \varepsilon F_i$
12. end for
13. Set $\eta_k := H_m(k + 1, k) Y_m(:, k + 1) + \varepsilon(k + 1) \eta$
14. Set $Z_k := Z_m(:, 1 : k), Y_k := Y_m(:, 1 : k), H_k := H_m(1 : k, 1 : k), R_k := R_m(1 : k, 1 : k)$
15. end if
16. end for

**Remark 5.1**

Note that applying an implicitly restarted process on a deflated SGA decomposition (25) may not yield a deflated SGA decomposition. We know that the $Q$-matrix and $V$-matrix of the SGA decomposition must adhere to one of the two orthogonality requirements: (1) all column vectors form an orthonormal set and (2) when deflation occurs, all column vectors form an orthonormal set except zero columns. In the first case, the resultant $Q^+$-matrix and $V^+$-matrix maintain the same orthogonality requirement as in the $Q$-matrix and $V$-matrix of the SGA decomposition. In the second case, both $Q$-matrix and $V$-matrix contain some zero columns(s). Then the nonzero columns of the updated $Q^+$-matrix will be linearly dependent and the resultant decomposition is not a SGA decomposition. The same phenomenon occurs on the updated $V^+$-matrix.
To overcome this problem, we only need to perform column compression to make the updated $Q^+$-matrix and $V^+$-matrix of the forms $[\hat{Q}^+ 0]$ and $[\hat{V}^+ 0]$, simultaneously. On the other hand, it requires to update $H^+$-matrix and $R^+$-matrix by postmultiplying an upper triangular matrix as we perform the column compression. The resultant $H^+$-matrix and $R^+$-matrix are still upper Hessenberg form and upper triangular, respectively. Consequently, the column compression transforms a decomposition to a deflated SGA decomposition.

5.2. The selection of shifts

The above scheme involves selection of shifts $\vartheta_1, \ldots, \vartheta_{m-k}$. A good selection of shift is a key for success of the implicit restart technique. A popular choice of the shift values for IRA method [24] is to choose unwanted Ritz values, and are called exact shifts in [24]. When we solve the reduced QEP (27) to get $2m$ eigenvalues and select $k$ Ritz values as approximations to the desired eigenvalues, we may directly use the reciprocal values of the remaining unwanted Ritz values as shifts which we also call them exact shifts. Among the selection of $2m - k$ shift candidates, we always take the reciprocal values of the $m - k$ unwanted Ritz values which are farthest from the target as shifts. Applying implicitly shifted $QZ$ iteration with exact shifts to the SGA method, we have an implicitly restarted SGA (IRRSA) method.

For the RSGA method, we can also choose exact shifts. However, the refinement strategy cannot only improve the accuracy of the Ritz pairs but also provide more accurate approximations to some of the unwanted eigenvalues. Suppose that $((\vartheta, \omega) = (\vartheta, Q_m \zeta)$ is a Ritz pair of the QEP (1) which we are not interested in and the reciprocal of $\vartheta$ is one possible candidate of the shifts for the restarting process. Let $\omega^+ = Q_m \zeta^+$ be the refined Ritz vector corresponding the Ritz value $\vartheta$ as we discussed in section 4. Now, we illustrate how to find better shifts based on the unwanted refined Ritz vector $\omega^+$. For an approximate eigenvector $\omega$ of the QEP (1), the usual approach to deriving an approximate eigenvalue $\theta$ from $\omega$ is to impose the Galerkin condition

$$(\theta^2 M + \theta D + K)\omega = \omega$$

and this follows that $\theta = \theta(\omega)$ must be one of the two solutions to the quadratic equation [10]

$$a_2 \theta^2 + a_1 \theta + a_0 = 0$$

(45)

where $a_2 = \omega^* M \omega$, $a_1 = \omega^* D \omega$ and $a_0 = \omega^* K \omega$. Therefore, as we obtain the unwanted refined Ritz vector $\omega^+$, (45) provides us one way to compute more accurate Ritz value beyond our interests and should be filtered in the restarting process. Since $\omega^+ = Q_m \zeta^+$, the coefficients corresponding to the quadratic equation (45) would be reduced as follows

$$a_2 = (\zeta^+)^* M_m \zeta^+, \quad a_1 = (\zeta^+)^* D_m \zeta^+ \quad \text{and} \quad a_0 = (\zeta^+)^* K_m \zeta^+$$

(46)

where $M_m, D_m$ and $K_m$ is the projections of $M, D$ and $K$ onto the subspace $\text{span}\{Q_m\}$ respectively as described in (28).

Hence, if $\vartheta_1^+$ and $\vartheta_2^+$ are roots of the quadratic equation (45) with coefficients defined in (46) then their reciprocal values would be better candidates for the restarting process. Consequently, if $(\vartheta_1, Q_m \zeta_1), \ldots, (\vartheta_p, Q_m \zeta_p)$ are p Ritz pairs that are farthest from our target and if $\vartheta_1^{i,1}, \vartheta_2^{i,2}$ are the roots of the quadratic equation (45) with respect to the unwanted refined Ritz vector $\omega^+_i = Q_m \zeta^+_i$, $i = 1, \ldots, p$, then we choose the $p$ values from $\vartheta_1^+, \vartheta_2^+, \vartheta_1^{i,1}, \vartheta_2^{i,2}$ that are farthest from our target and take their reciprocal values as the shifts for the restarting process and call them the refined shifts. In our numerical examples, an implicitly restarted refined SGA (IRRSA) method is a restart version of the RSGA method with refined shifts.

6. NUMERICAL EXAMPLES

The purpose of this section is to present a few numerical experiments to validate that the IRRSGA method is viable for solving the QEP (1). In addition, the examples demonstrate the superior
properties of the IRSGA method and the IRRSGA method than the two versions of the IRA method [24] for solving the QEP where one IRA method is applied to the $\ell$-SEP (6) and the other is applied to the $r$-SEP (7), respectively. The abbreviations $\ell$-IRA and $r$-IRA are used to indicate that the IRA method is applied to $\ell$-SEP and $r$-SEP, respectively.

In our examples, the number $m$ denotes the order of the SGA/Arnoldi decomposition, $k$ denotes the number of desired eigenpairs. The starting vector of the SGA method and the standard Arnoldi method are chosen as a vector with all components equal to 1 and the stopping tolerance for relative residuals is chosen to be $\text{tol} = 10^{-14}$. The maximum number $r_{\text{max}}$ of restarting process is set to be $r_{\text{max}} = 30$.

**Example 6.1**

This example is obtained from “NLEVP: a collection of nonlinear eigenvalue problem” [3], namely “damped beam” arising from the vibration analysis of a beam simply supported at both ends and damped in the middle. In our MATLAB implementation, the command `nlevp('damped_beam',2000)` is used to construct real symmetric coefficient matrices $M, D, K$ with $M = M^T > 0$, $D = D^T \geq 0$ and $K = K^T > 0$. The matrix size is $n = 4,000$. Ten eigenvalues nearest the origin (i.e., $k = 10$) are computed by by four methods with $m = 20$. Figure 1(a) shows the maximum relative residuals of the ten desired eigenpairs computed by $\ell$-IRA, $r$-IRA, IRSGA and IRRSGA with respect to iterations 1, 2, $\ldots$, 30. We find that the maximum relative residuals computed by $\ell$-IRA and $r$-IRA stagnate and those computed by the IRSGA method oscillate between $10^{-12}$ and $10^{-13}$. All relative residuals of the desired eigenpairs computed by the IRRSGA method meet the stopping tolerance in 1 iteration. To investigate the convergence behaviors of the ten eigenpairs computed by $\ell$-IRA, $r$-IRA, IRSGA and IRRSGA, we depict relative residual norms of the 1 step iteration in Figure 1(b). Compared to the IRSGA method, the refinement strategy of the IRRSGA method significantly improves the accuracy of computed eigenpairs even up to 5 digits for the eight computed eigenpairs that do not meet the convergence criterion. We report the number of iterations and CPU times in Table I. In summary, among the four methods the IRRSGA method is the only viable approach that accurately finds desired eigenpairs within 30 iterations.

<table>
<thead>
<tr>
<th>Method</th>
<th>#Its</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ell$-IRA</td>
<td>30</td>
<td>32.8563</td>
</tr>
<tr>
<td>$r$-IRA</td>
<td>30</td>
<td>55.4423</td>
</tr>
<tr>
<td>IRSGA</td>
<td>30</td>
<td>38.3231</td>
</tr>
<tr>
<td>IRRSGA</td>
<td>1</td>
<td>7.3048</td>
</tr>
</tbody>
</table>

Table I. Iteration numbers and CPU times in Example 6.1.

**Example 6.2**

In Example 6.1, we see an amazing effect of the refinement strategy, i.e., ten wanted eigenpairs converge in one iteration before the restarting process with refined shifts in IRRSGA. This example illustrates that the refinement strategy with refined shifts introduced in section 5.2 for the IRRSGA method accelerate the convergence.

We consider the damped vibration mode of an acoustic fluid confined in a cavity with absorbing walls capable of dissipating acoustic energy [2]. The fluid domain $\Omega \subseteq \mathbb{R}^2$ is assumed to be polyhedral, and the boundary $\partial \Omega = \Gamma_A \cup \Gamma_R$, where the absorbing boundary $\Gamma_A$ is the union of all the different faces of $\Omega$ and is covered by damping material. The rigid boundary $\Gamma_R$ is the remaining part of $\partial \Omega$. Figure 2(a) gives an example of such a setup, where the top boundary is absorbing and
the remaining boundary is rigid. The equations characterizing the wave motion in $\Omega$ are

$$\begin{aligned}
\rho \frac{\partial^2 U}{\partial t^2} + \nabla P &= 0, \\
P &= -\rho c^2 \text{div} U \\
P &= (\alpha U \cdot n + \beta \frac{\partial U}{\partial t} \cdot n) \quad \text{on } \Gamma_A \\
U \cdot n &= 0 \quad \text{on } \Gamma_R,
\end{aligned}$$

where the acoustic pressure $P$ and the fluid displacement $U$ depend on space $x$ and time $t$, $\rho$ is the fluid density, $c$ is the speed of sound in air, $n$ is the unit outer normal vector along $\partial \Omega$, and $\alpha, \beta$ are coefficients related to the normal acoustic impedance. The absorbing boundary on $\Gamma_A$ indicates that the pressure is balanced by the effects of the viscous damping (the $\beta$ term) and the elastic behavior (the $\alpha$ term). The model induces the following QEP

$$(\lambda^2 M_u + (\alpha + \lambda \beta) A_u + K_u) u = 0,$$

where $M_u$ and $K_u$ are mass and stiffness matrices, respectively, and $A_u$ is used to describe the effect of the absorbing wall.

In this example, we adopt the geometry illustrated in Figure 2(a) and use the following physical data: $\rho = 1\text{kg/m}^3$, $c = 340\text{m/s}$, $\alpha = 5 \times 10^4\text{N/m}^3$, and $\beta = 200\text{Ns/m}^3$. The same values are used in [2]. The rectangular domain is uniformly partitioned into 384 by 288 rectangles and each rectangle...
is further refined into two triangles. The dimension of coefficient matrices in this problem is
\( n = (3 \times 384 - 1) \times 288 = 331,488 \). We compute ten analytic solutions of the desired eigenvalues\( \lambda_1, \ldots, \lambda_{10} \) plotted in Figure 2(b) with the lowest positive vibration frequencies satisfying \( 0 < \frac{\text{Im}(\lambda_i)}{2\pi} < 600 \text{Hz} \). The order \( m \) is set to be \( m = 20 \). The shift target is taken by \( \tau = -25 + 600\pi i, \quad i = \sqrt{-1} \).

Table II shows that compared to the IRSGA method, the refinement strategy used in the IRRSGA method reduces the number of iterations and CPU time. Moreover, the IRRSGA method calculates ten desired eigenpairs in the smallest number of iterations and the shortest CPU time among four competitive methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>#Its</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ell )-IRA</td>
<td>18</td>
<td>806.69</td>
</tr>
<tr>
<td>( r )-IRA</td>
<td>18</td>
<td>836.14</td>
</tr>
<tr>
<td>IRSGA</td>
<td>9</td>
<td>777.74</td>
</tr>
<tr>
<td>IRRSGA</td>
<td>7</td>
<td>735.38</td>
</tr>
</tbody>
</table>

Table II. Iteration numbers and CPU time in Example 6.2.

Example 6.3
This experiment consists of six benchmark examples from the NLEVP [3]. In the following, we describe each example and the choice of parameters for generating the coefficient matrices of corresponding QEPs. All numerical results show that regardless of iteration numbers or CPU time, both IRSGA and IRRSGA appear to be more efficient and more competitive than the traditional Arnoldi methods \( \ell \)-IRA and \( r \)-IRA. The standard Arnoldi methods can not calculate all desired eigenpairs in 30 iterations but our IRSGA and IRRSGA methods can effectively find all desired eigenpairs with high accuracy in less or around 10 iterations. The IRSGA and the IRRSGA methods have similar convergence behavior and the latter consumes a slightly more time than the former. This might be due to the fact that the IRSGA method converges in very few iterations. Figure 4 depicts the maximum of the \( k \) residual norms versus restarts and show the convergence processes of each
example. Correspondingly, Table III lists the iteration numbers and the CPU time of each method for each example.

(a) **Acoustic 1D.** This example arises from the finite element discretization of the time harmonic wave equation $-\Delta p - (2\pi f/c)^2 p = 0$ [4]. Here $p$ denotes the pressure, $f$ is the frequency, $c$ is the speed of sound in the medium, and $\zeta$ is the (possibly complex) impedance. On the domain $[0, 1]$ with $c = 1$, the $n \times n$ matrices $M, D$ and $K$ are defined by

\[
M = -4\pi^2 h^2 (I_n - \frac{1}{2}e_n e_n^T), \quad D = 2\pi^2 h\zeta I_n, \quad K = n \left( \text{tridiag}(-1, 2, -1) - e_n e_n^T \right).
\]

Observe that matrices $M, K$ are real symmetric and $D$ is complex symmetric. We use \texttt{nlevp('acoustic\_wave\_1d',5000,1)} to generate $M, D, K$ with size $n = 5,000$ and compute the six eigenvalues nearest origin (i.e., $k = 6$) with $m = 12$.

(b) **Acoustic 2D.** This example is a two-dimensional acoustic wave equation [4] on $[0, 1] \times [0, 1]$. The coefficient matrices $(M, D, K)$ are given by

\[
M = -4\pi^2 h^2 I_q - \frac{1}{2}e_q e_q^T, \quad D = 2\pi^2 h\zeta I_q, \quad K = I_q \otimes D_q + T_{q-1} \otimes (-I_q + \frac{1}{2}e_q e_q^T),
\]

where $h$ denotes the mesh size, $q = 1/h$, $\otimes$ denotes the Kronecker product, $\zeta$ is the (possibly complex) impedance, $D_q = \text{tridiag}(-1, 4, -1) - 2e_q e_q^T \in \mathbb{R}^{q \times q}$ and $T_{q-1} = \text{tridiag}(1, 0, 1) \in \mathbb{R}^{(q-1) \times (q-1)}$ to get the real symmetric matrices $(M, D, K)$. The matrix size is given by $n = 8,010$ and we compute ten eigenvalues nearest the origin (i.e., $k = 10$) with $m = 20$.

(c) **Concrete.** This problem arises from a model of a concrete structure supporting a machine assembly [6] and induces the QEP, $(\lambda^2 M + \lambda D + (1 + \mu I) K)x = 0$, where $M$ is real diagonal and low rank, $D$, the viscous damping matrix, is pure imaginary and diagonal, $K$ is complex symmetric, and the factor $1 + \mu I$ adds uniform hysteretic damping. We use \texttt{nlevp('concrete',0.04)} to generate the complex symmetric coefficient matrices. The matrix size $n = 2,472$ and we compute ten eigenvalues nearest the origin (i.e., $k = 10$) with $m = 20$.

(d) **Spring dashpot.** The QEP arises from a finite element model of a linear spring in parallel with Maxwell elements [9]. The mass matrix $M$ is rank deficient and symmetric, the damping matrix $D$ is rank deficient and block diagonal, and the stiffness matrix $K$ is symmetric and has arrowhead structure. Matrices $M, D, K$ are generated from \texttt{nlevp('spring\_dashpot',7850,5000,0)} with size $n = 10,000$. We compute 50 eigenvalues nearest the origin (i.e., $k = 50$) with $m = 100$.

(e) **Wiresaw1.** We use \texttt{nlevp('wiresaw1',10000,0.01)} to generate the coefficient matrices of the gyroscopic QEP arising in the vibration analysis of a wiresaw [30]. Here $M, D, K$ are $n \times n$ matrices defined by

\[
M = \frac{1}{2} I_n, \quad D = -D^T = [d_{ij}] \quad \text{and} \quad K = \text{diag} \left( \frac{i^2 \pi^2 (1 - v^2)}{2} \right)
\]

where $d_{ij} = \frac{4ij}{\pi^2} v$ if $i + j$ is odd and, otherwise, $d_{ij} = 0$. The matrix size for this problem is $n = 10,000$ and we compute 10 eigenvalues nearest origin (i.e., $k = 10$) with $m = 20$.

(f) **Wiresaw2.** When the effect of viscous damping is added to the problem in Wiresaw1, the corresponding QEP has the form [30]

\[
(\lambda^2 M + \lambda (D + \eta I_n) + K + \eta D)x = 0
\]

where $M$, $D$ and $K$ are the same as in Wiresaw1 and $\eta$ is a real nonnegative damping parameter. We take $\eta = 0.5$ and use \texttt{nlevp('wiresaw2',10000,0.01,0.5)} to generate the coefficient matrices. The matrix size is $n = 10,000$ and we compute 10 eigenvalues near the target $-0.5$ (i.e., $k = 10$ and $\tau = -0.5$) with $m = 20$. 

---

7. CONCLUSIONS

We have presented the SGA method, an orthogonal projection method, for solving QEPs based on a SGA decomposition. We have developed a practical algorithm to compute the SGA decomposition. The application of the SGA decomposition is three fold. First, we compute an orthonormal basis of the projection subspace in the SGA decomposition. Second, the SGA decomposition (8) has computational advantage for generating the coefficient matrices of reduced QEP (28). Third, we take advantage of the SGA decomposition to save some computational costs in the refinement process resulting a refined version of the SGA method abbreviated as the RSGA method for solving QEPs. After applying an implicit restart technique to SGA/RSGA methods, we have restart versions of SGA and RSGA, namely, the IRSGA/IRRSGA method. We have reported the numerical results on computation of the approximate eigenpairs with small eigenvalues in modulus. Compared to the standard IRA method, both the IRSGA method and IRRSGA method are superior in accuracy and convergence rate. We also see that the IRRSGA method may significantly improve the accuracy for obtaining the desired eigenpairs when the standard IRA method and the IRSGA method can not converge in a certain number of iterations.

ACKNOWLEDGEMENTS

This work is partially supported by the National Science Council, the Taida Institute for Mathematical Sciences, the National Center for Theoretical Sciences and the Center of Mathematical Modeling and Scientific Computing. The second author’s work was supported by the National Science Foundation of China (No.11101080) and the Specialized Research Fund for the Doctoral Program of Higher Education (No.20110092120023).

REFERENCES
A SGA METHOD AND ITS VARIATIONS FOR QEPS

Figure 4. Example 6.3: Convergence histories for $\ell$-IRA, $r$-IRA, IRSGA and IRRSGA.


