SIAM J. SCI. COMPUT.
Vol. 24, No. 4, pp. 1283–1302

NUMERICAL SOLUTION OF QUADRATIC EIGENVALUE PROBLEMS WITH STRUCTURE-PRESERVING METHODS*

TSUNG-MIN HWANG†, WEN-WEI LIN‡, AND VOLKER MEHRMANN§

Abstract. Numerical methods for the solution of large scale structured quadratic eigenvalue problems are discussed. We describe a new extraction procedure for the computation of eigenvectors and invariant subspaces of skew-Hamiltonian/Hamiltonian pencils using the recently proposed skew-Hamiltonian isotropic implicitly restarted Arnoldi method (SHIRA).

As an application we discuss damped gyroscopic systems. For this problem we first solve the eigenvalue problem for the undamped system using the structure-preserving method and then use the quadratic Jacobi–Davidson method as correction procedure. We also illustrate the properties of the new approach for several other application problems.

Key words. quadratic eigenvalue problems, skew-Hamiltonian/Hamiltonian pencils, invariant subspace, gyroscopic system, quadratic Jacobi–Davidson method, nonequivalence deflation technique

AMS subject classifications. 15A18, 47A15, 47J10

PII. S106482750139220X

1. Introduction.

In this paper we study the numerical methods for computing eigenpairs or invariant subspaces of the quadratic eigenvalue problem

$$(\lambda^2 M + \lambda(G + \varepsilon D) + K)x = 0, \tag{1}$$

where M, C, K, D are square $n \times n$ real matrices with $M = M^T$, $G = -G^T$, $D = D^T$, and $K = K^T$. Typically, the matrices M and K or $-K$ represent the mass matrix and the stiffness matrix, respectively; G and εD represent gyroscopic forces and the damping, respectively, of the system. We not only concentrate on the case in which ε is a small parameter, but also discuss in detail the undamped case that $\varepsilon = 0$.

Quadratic eigenvalue problems (1) arise in the solution of initial or boundary value problems for second order systems of the form

$$M\ddot{x} + (G + \varepsilon D)\dot{x} + Kx = f \tag{2}$$

and in numerous other applications. These include finite element discretization in structural analysis [34], in acoustic simulation of poro-elastic materials [22, 32, 35], or in the elastic deformation of anisotropic materials [2, 19, 33]. See also [42] for a recent survey.

The classical approach in solving the quadratic eigenvalue problems is to turn them into linear eigenvalue problems by introducing a new vector $y = \lambda x$. In the case

*Received by the editors July 10, 2001; accepted for publication (in revised form) August 19, 2002; published electronically February 6, 2003. This research was partially supported by the National Center for Theoretical Science, National Tsing Hua University, Hsinchu, Taiwan.
http://www.siam.org/journals/sisc/24-4/39220.html

†Department of Mathematics, National Taiwan Normal University, Taipei 116, Taiwan, Republic of China (min@math.ntnu.edu.tw).

‡Department of Mathematics, National Tsing Hua University, Hsinchu 300, Taiwan, Republic of China (wwlin@am.nthu.edu.tw).

§Institut für Mathematik, MA 4-5, Technische Universität Berlin, Str. des 17. Juni 136, D-10623 Germany (mehrmann@math.tu-berlin.de). This author was also partially supported by Deutsche Forschungsgemeinschaft research grant Me 790/11-3.

of (1) this leads to the *linearized* generalized eigenvalue problem

$$\left(\lambda \begin{bmatrix} M & G+\varepsilon D \\ 0 & I \end{bmatrix} - \begin{bmatrix} 0 & -K \\ I & 0 \end{bmatrix}\right) \begin{bmatrix} y \\ x \end{bmatrix} = 0. \tag{3}$$

If $(\lambda, \left[\begin{smallmatrix} y \\ x \end{smallmatrix}\right])$ is an eigenpair of (3), then x is an eigenvector of (1) associated with the eigenvalue λ. The approach (3) allows us to determine eigenpairs numerically, since for the generalized eigenvalue problem like (3) the mathematical theory, numerical methods as well as the perturbation theory are well established [1, 11, 39]. However, a difficulty of the linearization approach is that due to the embedding into the problem of double size, the condition number, i.e., the sensitivity of the eigenvalues and eigenvectors with respect to perturbations in the data matrices M, G, D, K, may increase. This is because the set of admissible perturbations for (3) is larger than for (1). If the perturbations would, however, respect the specific structure of the blocks in (3), i.e., the zero, the identity blocks, and the symmetries, then the perturbation results would be the same. Furthermore, there are many different ways to do the linearization with different conditioning, as has been demonstrated in [40]. In view of these remarks it would be ideal to have a numerical method that works directly with the original data of the quadratic eigenvalue problem and that avoids the problem of the increased condition numbers. It is difficult to develop such a method, and there are only few methods that partially fulfill these requirements, such as the quadratic or polynomial Jacobi–Davidson method [14, 36, 37].

In many applications the original quadratic eigenvalue problem has some extra structures that should be reflected in its linearization. A typical structure is the Hamiltonian eigenvalue symmetry, i.e., that the spectrum is symmetric with respect to the real and imaginary axes. This symmetry occurs if $\varepsilon = 0$ in (1), which is the case, for example, in the gyroscopic system [18] and in the elasticity problems [2, 26]; see [42] for a recent survey. A similar eigenvalue symmetry arises in the optimal control problems for which the variational principle leads to a skew-Hamiltonian/Hamiltonian (SHH) pencil that has similar properties as (3) in the undamped case; see [5, 25, 26]. In these applications the generalized eigenvalue problem arises directly and is not from any linearization of a quadratic eigenvalue problem.

Another structure that should be reflected in a proper method is preservation of the sparsity structure inherited from finite element or finite difference discretizations [3, 26, 34].

In this paper we discuss first the solution of large sparse generalized eigenvalue problems that have a Hamiltonian eigenvalue symmetry and then allow perturbations of this structure. The use of structure-preserving linearizations for problem (1) in the undamped case has first been suggested and successfully employed in [10, 26] to design structure-preserving methods for large sparse linearized quadratic eigenvalues problems with a Hamiltonian eigenvalue symmetry. We make use of these techniques, in particular, the skew-Hamiltonian isotropic implicitly-restarted Arnoldi algorithm (SHIRA) proposed in [26]. We will briefly recall this algorithm in section 2.

The key feature of the SHIRA algorithm is that it inherits the convergence properties, the implicit restart, and the reorthogonalization techniques of the standard shift-and-invert Arnoldi algorithm [20, 31, 38], while being, in general, more efficient. Furthermore, it generates isotropic Krylov subspaces to guarantee that the computed spectrum has the correct eigenvalue symmetry. In [41] the stability properties of the methods that preserve the Hamiltonian structure (although not for this algorithm) are analyzed.

A disadvantage of the SHIRA algorithm is that it does not directly generate the eigenvectors or invariant subspaces of the linearization. Instead, they are determined in [26] using a few steps of inverse or subspace iteration [11, 44], respectively. For these iterations one has to solve in every step a linear system, and hence extra sparse matrix factorizations are needed. This results in a bottleneck in the computation and also limits the possible system size. Even though, in practice, only a few iterations are needed, these extra factorizations add to the method substantial computational overhead.

In this paper we develop a new extraction method for invariant subspaces in the SHIRA algorithm that avoids extra sparse matrix factorizations. This method is a modification of an idea suggested in [45, 46] for the computation of the stable invariant subspace of a Hamiltonian matrix. We compare the performance of this extraction procedure with that of classical inverse or subspace iteration in section 5. As an application we then discuss numerical methods for problems like (1) that have a damping term with $\varepsilon \neq 0$. In this problem the Hamiltonian eigenvalue symmetry is destroyed and therefore structure-preserving methods cannot be used. On the other hand, if the damping is small, then we can expect that the damped system has a spectrum and also invariant subspaces close to those of the structured system. We discuss this topic in section 4 and introduce a method that first computes the eigenvalues, eigenvectors, or invariant subspaces of the structured system and then use them as initial vectors in the quadratic Jacobi–Davidson procedure [3, 36] to compute the eigenvalues, eigenvectors, or invariant subspaces of the damped system. A similar approach for problems with small damping in structural mechanics was suggested in [27]. In order to treat eigenvalue problems with clusters of eigenvalues, a nonequivalence low-rank transformation technique [9, 12] is used to deflate the computed eigenpairs. We compare this approach with the classical shift-and-invert subspace iteration (SISI) for several application problems in section 5.

2. Structure-preserving algorithms. In this section we briefly recall the SHIRA algorithm of [26], which is designed for the computation of a small number of specified eigenvalues of real large scale generalized eigenvalue problems for SHH pencils of the form

$$\alpha\mathcal{N} - \beta\mathcal{H} = \alpha \begin{bmatrix} F_1 & G_1 \\ H_1 & F_1^T \end{bmatrix} - \beta \begin{bmatrix} F_2 & G_2 \\ H_2 & -F_2^T \end{bmatrix}, \tag{4}$$

where $G_1 = -G_1^T,\ H_1 = -H_1^T,\ G_2 = G_2^T$, and $H_2 = H_2^T$.

Matrices $\mathcal{N}$ and $\mathcal{H}$ in (4) are called skew-Hamiltonian and Hamiltonian, respectively. If $J = \left[\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right]$ with I_n being the $n \times n$ identity matrix, then skew-Hamiltonian matrices satisfy $(\mathcal{N}J)^T = -\mathcal{N}J$ and Hamiltonian matrices satisfy $(\mathcal{H}J)^T = \mathcal{H}J$. The special structure of SHH pencils ensures [23, 24, 25] that the eigenvalues occur in quadruples $\lambda,\ \overline{\lambda},\ -\lambda,\ -\overline{\lambda}$, if λ is complex, or pairs $\lambda, -\lambda$ if λ is real. Moreover, if x is a right eigenvector associated with an eigenvalue λ, then Jx is a left eigenvector associated with $-\overline{\lambda}$.

If the skew-Hamiltonian matrix $\mathcal{N}$ in the SHH pencil $\alpha\mathcal{N} - \beta\mathcal{H}$ in (4) is invertible and given in the factored form $\mathcal{N} = \mathcal{Z}_1\mathcal{Z}_2$ with $\mathcal{Z}_2^T J = \pm J\mathcal{Z}_1$, then the pencil is equivalent to a pencil of the form $\alpha I - \beta\mathcal{W}$, where

$$\mathcal{W} = \pm\mathcal{Z}_1^{-1}\mathcal{H}\mathcal{Z}_2^{-1}$$

is Hamiltonian. The factorization of $\mathcal{N}$ can be easily determined via a Cholesky-like decomposition for skew-symmetric matrices developed in [7]; see also [4].

Suppose that one wishes to compute the eigenvalues of $\mathcal{W}$ nearest to some target value λ_0. Then typically one would use shift-and-invert of the form $(\mathcal{W} - \lambda_0 I)^{-1}$ to get the eigenvalues near λ_0 via a Krylov subspace method. But this matrix is not Hamiltonian any longer. In order to preserve the structure, [26] suggests a rational transformation with four shifts $(\lambda_0, \overline{\lambda_0}, -\lambda_0, -\overline{\lambda_0})$ given by

$$R_1(\lambda_0, \mathcal{W}) = (\mathcal{W} - \lambda_0 I)^{-1}(\mathcal{W} + \lambda_0 I)^{-1}(\mathcal{W} - \overline{\lambda_0} I)^{-1}(\mathcal{W} + \overline{\lambda_0} I)^{-1}. \tag{5}$$

If the target λ_0 is either real or purely imaginary, one may use the simpler transformation

$$R_2(\lambda_0, \mathcal{W}) = (\mathcal{W} - \lambda_0 I)^{-1}(\mathcal{W} + \lambda_0 I)^{-1}. \tag{6}$$

If the matrix $\mathcal{W}$ is real and Hamiltonian, then both matrices R_1 and R_2 are real and skew-Hamiltonian, and hence the eigenvalues have algebraic multiplicity of at least two, if they are not purely imaginary. For a real skew-Hamiltonian matrix $\mathcal{K}$ and a given vector q_1, an Arnoldi iteration applied to $\mathcal{K}$ would generate the Krylov space

$$\mathcal{V}(\mathcal{K}, q_1, k) = \text{span}\{[q_1, \mathcal{K}q_1, \ldots, \mathcal{K}^{k-1}q_1]\}.$$

Using an appropriate orthogonal basis of this space given by the columns of an orthogonal matrix Q_k, one produces a "Ritz"-projection

$$K_k = Q_k^T \mathcal{K} Q_k.$$

The "Ritz"-values, i.e., the eigenvalues of K_k are then used as eigenvalue approximations.

In order to obtain a structure-preserving method, one needs a "Ritz"-projection that is again skew-Hamiltonian. For this one would need an isotropic subspace $\mathcal{V}$, i.e., a subspace for which $x^T J y = 0$ for all $x, y \in \mathcal{V}$; see [26]. Let $V_k \in \mathbb{R}^{2n\times k}$ and let

$$Q^T V_k = \begin{bmatrix} R_k \\ 0 \\ T_k \\ 0 \end{bmatrix} \tag{7}$$

be the symplectic QR-factorization [8] of V_k, where $Q \in \mathbb{R}^{2n\times 2n}$ is orthogonal and symplectic and where R_k and $T_k \in \mathbb{R}^{k\times k}$ are upper and strictly upper triangular, respectively. If V_k is isotropic and of full column rank, then it is shown in [8] that R_k^{-1} exists and that $T_k = 0$. This means, in particular, that if V_n is of full column rank, then there exists an orthogonal symplectic matrix $\mathcal{Q} \equiv [Q_n \,|\, JQ_n]$ with $Q_n \in \mathbb{R}^{2n\times n}$, such that

$$\begin{bmatrix} Q_n^T \\ Q_n^T J^T \end{bmatrix} \mathcal{K} \, [Q_n \; JQ_n] = \begin{bmatrix} H_n & K_n \\ 0 & H_n^T \end{bmatrix}. \tag{8}$$

Here $H_n = Q_n^T \mathcal{K} Q_n$ is an upper Hessenberg matrix and $K_n = Q_n^T \mathcal{K} J Q_n$ is skew-symmetric. This idea is the basis for the structure-preserving Arnoldi algorithm SHIRA introduced in [26], which is given by the recursion

$$\mathcal{K}Q_k = Q_k H_k + q_{k+1} h_{k+1,k} e_k^T, \tag{9}$$

where H_k is the leading $k \times k$ principal submatrix of H_n and $e_k = [0, \ldots, 0, 1]^T \in \mathbb{R}^{k\times 1}$. Since Q_k is orthogonal and $Q_k^T J Q_k = 0$ for $k < n$, it is easily seen that $[Q_k | JQ_k] \in \mathbb{R}^{2n\times k}$ is orthogonal. This implies that

$$\begin{bmatrix} Q_k^T \\ Q_k^T J^T \end{bmatrix} \mathcal{K} \, [Q_k \mid JQ_k] = \begin{bmatrix} H_k & G_k \\ 0 & H_k^T \end{bmatrix} \in \mathbb{R}^{2k\times 2k} \tag{10}$$

is skew-Hamiltonian and $G_k = Q_k^T \mathcal{K} J Q_k$ is skew-symmetric.

Let (θ_i, v_i) be an eigenpair of H_k, i.e., $H_k v_i = \theta_i v_i$, and let $x_i = Q_k v_i$ be a "Ritz"-vector of the eigenvalue problem $\mathcal{K}x = \mu x$ corresponding to the "Ritz" value θ_i. Then we have the following residual bound, see [31], for the "Ritz"-pair (θ_i, x_i):

$$\begin{aligned} &\| \mathcal{K}x_i - \theta_i x_i \|_2 = \| \mathcal{K}Q_k v_i - \theta_i Q_k v_i \|_2 \\ &= \| \left(Q_k H_k + q_{k+1} h_{k+1,k} e_k^T\right) v_i - \theta_i Q_k v_i \|_2 \\ &= \| q_{k+1} h_{k+1,k} (e_k^T v_i) \|_2 = |h_{k+1,k}| |e_k^T v_i|. \end{aligned}$$

Remark 1.

(a) The transformations $R_1(\lambda_0, \mathcal{W})$ and $R_2(\lambda_0, \mathcal{W})$ can be viewed as a structure-preserving shift-and-invert strategy to accelerate the convergence of the desired eigenvalues.

(b) In exact arithmetic, the values $t_{i,k} = (Jq_i)^T \mathcal{K} q_k$ in the kth step of the SHIRA algorithm are zero, but in practice roundoff errors cause them to be nonzero. It has been shown in [26] how to design an isotropic orthogonalization scheme to ensure that the spaces $\text{span}\{q_1, \ldots, q_k\}$ are isotropic to working precision.

(c) To avoid the reorthogonalization, after a number of steps the SHIRA algorithm also uses implicit restarts as in [38].

(d) A minor disadvantage of the SHIRA method is that due to the rational transformations (5) or (6) the eigenvectors associated with pairs of eigenvalues λ and $-\lambda$ in the real case or λ and $-\bar{\lambda}$ in the complex case are both mapped to the same invariant subspace associated with $|\lambda|^2$. This has the consequence that in many applications the desired invariant subspace associated with the stable eigenvalues is not obtained directly from the method. In section 3 we discuss an extraction method that allows us to compute this subspace directly.

3. An extraction method for stable eigenspaces. In this section we discuss a method for computing specific invariant subspaces of a Hamiltonian matrix $\mathcal{W}$ directly from an isotropic invariant subspace of the skew-Hamiltonian matrix $\mathcal{W}^2$ without using inverse or subspace iteration. In many applications, we are interested in a subspace associated with eigenvalues in the left half plane. The same method can also be applied to the skew-Hamiltonian functions R_1, R_2 introduced in section 2.

The method that we describe is a modification of a technique first suggested in [45, 46]. Suppose that $\mathcal{W}$ has no eigenvalues on the imaginary axis and let Q_n be as in (8) for $\mathcal{K} = \mathcal{W}^2$. Then

$$\mathcal{W}^2 Q_n = Q_n H_n. \tag{11}$$

Let $Q_k \in \mathbb{R}^{2n\times k}$ $(k \le n)$ be a basis of an isotropic invariant subspace of $\mathcal{W}^2$ such that

$$\mathcal{W}^2 Q_k = Q_k \Omega_k, \tag{12}$$

where for the spectrum σ we have $\sigma(\Omega_k) \subseteq \sigma(H_n)$ and

$$\sigma(\Omega_k) \cap \{\sigma(H_n) \backslash \sigma(\Omega_k)\} = \emptyset,$$

i.e., the complete multiplicity of a multiple eigenvalue is included. Since we have assumed that $\mathcal{W}$ has no purely imaginary eigenvalues, it follows that there exist two bases V_k^- and V_k^+, both of dimension k, associated with the stable and unstable isotropic, invariant subspaces of $\mathcal{W}$, respectively, such that

$$\begin{aligned} \mathcal{W}V_k^- &= V_k^-\Lambda_k^-, \\ \mathcal{W}V_k^+ &= V_k^+\Lambda_k^+, \end{aligned} \tag{13}$$

where

$$\begin{aligned} \sigma(\Lambda_k^-) &= \{\lambda | \mathrm{Re}(\lambda) < 0, \lambda^2 \in \sigma(\Omega_k)\}, \\ \sigma(\Lambda_k^+) &= \{\lambda | \mathrm{Re}(\lambda) > 0, \lambda^2 \in \sigma(\Omega_k)\}. \end{aligned} \tag{14}$$

Since by assumption Ω_k has no purely imaginary eigenvalue, it follows that Ω_k has a unique positive square root X_k [16] satisfying $X_k^2 = \Omega_k$ such that all eigenvalues have positive real part. We call this root the *positive square root.* It can, for example, be computed via the MATLAB function `sqrtm` [13].

The following theorem shows how to extract the stable eigenspace $\mathrm{span}(V_k^-)$ from the isotropic invariant subspace $\mathrm{span}(Q_k)$.

THEOREM 1. *Let $\mathcal{W} \in \mathbb{R}^{2n\times 2n}$ be Hamiltonian and let $Q_k \in \mathbb{R}^{2n\times k}$ be as in* (12). *Suppose that there exist nonsingular matrices L_k^- and $L_k^+ \in \mathbb{R}^{k\times k}$ such that*

$$Q_k = V_k^+ L_k^+ + V_k^- L_k^-, \tag{15}$$

where V_k^+ and V_k^- are as in (13). *Then for X_k, the positive square root of Ω_k, we have*

$$\mathrm{span}(\mathcal{W}Q_k - Q_k X_k) = \mathrm{span}(V_k^-). \tag{16}$$

Proof. From (13) and (15) we have

$$\begin{aligned} \mathcal{W}Q_k - Q_k X_k &= \mathcal{W}(V_k^+ L_k^+ + V_k^- L_k^-) - (V_k^+ L_k^+ + V_k^- L_k^-)X_k \\ &= V_k^+ \left(\Lambda_k^+ L_k^+ - L_k^+ X_k\right) + V_k^- \left(\Lambda_k^- L_k^- - L_k^- X_k\right). \end{aligned}$$

If we can prove that $\Lambda_k^+ L_k^+ - L_k^+ X_k = 0$ and that $\Lambda_k^- L_k^- - L_k^- X_k$ is nonsingular, then assertion (16) follows. To do this, we show first that

$$X_k = (L_k^+)^{-1}\Lambda_k^+ L_k^+. \tag{17}$$

From (15) and (13) we have

$$\begin{aligned} \mathcal{W}^2 Q_k &= \mathcal{W}^2 \left(V_k^+ L_k^+ + V_k^- L_k^-\right) \\ &= V_k^+ \left(\Lambda_k^+\right)^2 L_k^+ + V_k^- \left(\Lambda_k^-\right)^2 L_k^-. \end{aligned} \tag{18}$$

On the other hand, from (12) and (15) we have

$$\mathcal{W}^2 Q_k = Q_k \Omega_k = \left(V_k^+ L_k^+ + V_k^- L_k^-\right) \Omega_k. \tag{19}$$

Subtracting (19) from (18) we obtain

$$\begin{bmatrix} V_k^+ & V_k^- \end{bmatrix} \begin{bmatrix} \left(\Lambda_k^+\right)^2 L_k^+ - L_k^+ \Omega_k \\ \left(\Lambda_k^-\right)^2 L_k^- - L_k^- \Omega_k \end{bmatrix} = 0.$$

Since $\begin{bmatrix} V_k^+ & V_k^- \end{bmatrix}$ has full rank it follows immediately that $\left(\Lambda_k^+\right)^2 L_k^+ - L_k^+ \Omega_k = 0$ and (17) follows due to the uniqueness of X_k.

We also have $\left(\Lambda_k^-\right)^2 L_k^- - L_k^- \Omega_k = 0$, and the uniqueness of the positive square root of Ω_k implies that

$$X_k = \left(L_k^-\right)^{-1} Z_k L_k^-, \tag{20}$$

where Z_k denotes the unique positive square root of $\left(\Lambda_k^-\right)^2$. Hence, $\Lambda_k^- L_k^- - L_k^- X_k = \left(\Lambda_k^- - Z_k\right) L_k^-$ is nonsingular. This completes the proof. ∎

Theorem 1 indicates a way to determine the desired stable invariant subspace. We apply SHIRA to the skew-Hamiltonian operators $R_1(\lambda_0, \mathcal{W})$ and $R_2(\lambda_0, \mathcal{W})$ as in (5) and (6), respectively, and determine the associated isotropic invariant subspace Q_k.

If the target $\lambda_0 \notin \sigma(\mathcal{W})$ is real or purely imaginary, and if the entry $h_{k+1,k}$ of H_k is negligible, then the space

$$\operatorname{span}(Q_k) = \operatorname{span}\{q_1, \ldots, q_k\}$$

generated by SHIRA is a good approximation of an isotropic invariant subspace of $R_2(\lambda_0, \mathcal{W})$. Then Q_k satisfies (approximately)

$$R_2(\lambda_0, \mathcal{W}) Q_k = (\mathcal{W}^2 - \lambda_0^2 I)^{-1} Q_k = Q_k H_k, \tag{21}$$

where H_k is an upper Hessenberg matrix. This implies that

$$\Omega_k = Q_k^T \mathcal{W}^2 Q_k = H_k^{-1} + \lambda_0^2 I. \tag{22}$$

Applying Theorem 1, we can extract the stable isotropic eigenspace V_k^- of $\mathcal{W}$ from Ω_k by computing its positive square root.

If the target $\lambda_0 \notin \sigma(\mathcal{W})$ is neither real nor purely imaginary and if $h_{k+1,k}$ is negligible, then the space $\operatorname{span}(Q_k) = \operatorname{span}\{q_1, \ldots, q_k\}$ generated by SHIRA is a good approximation to an isotropic invariant subspace of $R_1(\lambda_0, \mathcal{W})$. However, as has been shown in [26], it may happen that this subspace fails to be invariant under $\mathcal{W}^2$.

It is therefore necessary to check whether or not $\operatorname{span}(Q_k)$ is invariant under $\mathcal{W}^2$ by computing the residual

$$\| \mathcal{W}^2 Q_k - Q_k \Omega_k \|_F,$$

with $\Omega_k = Q_k^T \mathcal{W}^2 Q_k$. If this residual is small, then the procedure to compute the invariant subspace is as before.

We summarize the extraction method in the following algorithm.

ALGORITHM 1 (Extraction method (EM) for the stable invariant subspace or eigenvectors of a Hamiltonian matrix).

Input	Hamiltonian matrix $\mathcal{W}$ and a target value λ_0 with negative real part.
Output	Approximate invariant subspace V_ℓ^- of $\mathcal{W}$ associated with ℓ eigenvalues of negative real part nearest to λ_0.
(i)	**IF** $\lambda_0 = \alpha$ or $\lambda_0 = i\alpha$ for $\alpha \in \mathbb{R}$, **THEN** apply SHIRA to $R_2(\lambda_0, \mathcal{W})$, i.e.,
(a)	generate Arnoldi vectors as columns of $Q_k = [q_1, \ldots, q_k]$ and an upper Hessenberg matrix $H_k \in \mathbb{R}^{k\times k}$ such that $(\mathcal{W}^2 - \lambda_0^2 I)^{-1} Q_k = Q_k H_k$.
(b)	Compute $\Omega_k = H_k^{-1} + \lambda_0^2 I$.
(ii)	**ELSE IF** $\lambda_0 = \alpha + i\beta$, where $\alpha, \beta \in \mathbb{R}$, then apply SHIRA to $R_1(\lambda_0, \mathcal{W})$, i.e.,
(a)	generate Arnoldi vectors as columns of $Q_k = [q_1, \ldots, q_k]$ and an upper Hessenberg matrix $H_k \in \mathbb{R}^{k\times k}$ such that $R_1(\lambda_0, \mathcal{W}) Q_k = Q_k H_k$.
(b)	Compute $\Omega_k = Q_k^T \mathcal{W}^2 Q_k$.
	End
(iii)	Compute the real Schur decomposition of Ω_k as $\Omega_k = U_k T_k U_k^T$.
(iv)	Reorder the ℓ stable eigenvalues of T_k to the top of T_k using the reordering method for the Schur form of [1], i.e., $T_k = V_k \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} V_k^T$, where $T_{11} \in \mathbb{R}^{\ell\times\ell}$ has the stable eigenvalues.
(v)	Set $\tilde{Q}_\ell := Q_k U_k V_k \begin{bmatrix} I_\ell \\ 0 \end{bmatrix}$.
(vi)	Compute the unique positive square root $\sqrt{T_{11}}$ of T_{11} using the method `sqrtm` of [21].
(vii)	Compute the stable invariant subspace $V_\ell^- = \mathcal{W}\tilde{Q}_\ell - \tilde{Q}_\ell \sqrt{T_{11}}$.

The computational effort for the computation of the square root of Ω_k is of order k^3. If k is small compared to n, then this cost is negligible and it does not add much to the cost of the SHIRA iteration.

Remark 2. In a worst case analysis this procedure may suffer from a loss of accuracy due to the fact that we are computing the square root of the Schur form of a function of $\mathcal{W}$, which at least contains quadratic terms in $\mathcal{W}$. Thus, effects similar to the square reduced method for the Hamiltonian eigenvalue problem may occur, see [43], where the squaring of $\mathcal{W}$ leads to a loss of accuracy for small eigenvalues. See also the analysis in [45]. We will demonstrate this in Example 2 and also show that one step of inverse iteration fixes this accuracy loss.

Algorithm 1 can be applied directly to large sparse problems of the form (4). An important application here is the linear quadratic optimal control problem for descriptor systems, where the pencil has the form [5, 6, 25]

$$\alpha \begin{bmatrix} E & 0 \\ 0 & E^T \end{bmatrix} - \beta \begin{bmatrix} A & BB^T \\ C^TC & -A^T \end{bmatrix}. \tag{23}$$

Such problems arise in the optimal control of semidiscretized parabolic partial differential equations [15, 28, 29, 30].

Assume that the skew-Hamiltonian matrix $\mathcal{N}$ in (4) is given in factored form

$$\mathcal{N} = \mathcal{Z}_1 \mathcal{Z}_2, \quad \text{where} \quad \mathcal{Z}_2^T J = \pm J \mathcal{Z}_1. \tag{24}$$

Such a factorization (called J-Cholesky factorization) exists for all real skew-Hamiltonian matrices [5] and it can be obtained either trivially or numerically via a Cholesky-like factorization of skew-symmetric matrices [7].

If $\mathcal{N}$ is invertible, then using this factorization we can (at least formally) transform the SHH pencil $\alpha\mathcal{N} - \beta\mathcal{H}$ to a standard eigenvalue problem $\alpha I - \beta\mathcal{W}$, where $\mathcal{W} = \pm\mathcal{Z}_1^{-1}\mathcal{H}\mathcal{Z}_2^{-1}$ is Hamiltonian. If in (23) the matrix E is invertible, then the resulting Hamiltonian eigenvalue problem is

$$\alpha \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \beta \begin{bmatrix} E^{-1}A & E^{-1}BB^TE^{-T} \\ C^TC & -A^TE^{-T} \end{bmatrix}. \tag{25}$$

Numerical examples of this type are presented in section 5.

4. Damped gyroscopic systems. In this section we show how the extraction procedure given by Algorithm 1 can be used to solve quadratic eigenvalue problems for systems of the form (1). To do this we first study the undamped case ($\varepsilon = 0$), i.e.,

$$\lambda^2 Mx + \lambda Gx + Kx = 0, \tag{26}$$

with $M = M^T$ and $K = K^T$ positive definite and $G = -G^T$. Using the linearization (3) and the factorization

$$\mathcal{N} = \mathcal{Z}_1\mathcal{Z}_2 = \begin{bmatrix} I & \frac{1}{2}G \\ 0 & M \end{bmatrix} \begin{bmatrix} M & \frac{1}{2}G \\ 0 & I \end{bmatrix}$$

yields the Hamiltonian eigenvalue problem

$$(\lambda I - \mathcal{W})x = (\lambda I - \mathcal{Z}_1^{-1}\mathcal{H}\mathcal{Z}_2^{-1})x = 0. \tag{27}$$

The matrix $(\mathcal{W} - \lambda I)^{-1}$ can be factored as

$$\begin{aligned}(\mathcal{W} - \lambda I)^{-1} &\\ = \begin{bmatrix} M & \frac{1}{2}G \\ 0 & I \end{bmatrix} &\begin{bmatrix} I & \lambda I \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 & M^{-1} \\ -Q(\lambda)^{-1} & 0 \end{bmatrix} \begin{bmatrix} I & \frac{1}{2}G + \lambda M \\ 0 & M \end{bmatrix},\end{aligned} \tag{28}$$

where

$$Q(\lambda) = \lambda^2 M + \lambda G + K. \tag{29}$$

To this problem we can directly apply the extraction procedure given by Algorithm 1.

If we include the damping term in the gyroscopic system as in (1), then we cannot use the structured SHIRA method directly. But we can still linearize the system (1) and obtain the perturbed SHH system (3) or, in the different form,

$$\left(\lambda \begin{bmatrix} M & G \\ 0 & M \end{bmatrix} - \begin{bmatrix} -\varepsilon D & -K \\ M & 0 \end{bmatrix}\right) \begin{bmatrix} y \\ x \end{bmatrix} = 0, \tag{30}$$

which is now a pencil with one skew-Hamiltonian and one perturbed Hamiltonian matrix. We can still use the factorization (27) of the skew-Hamiltonian matrix and obtain the perturbed Hamiltonian eigenvalue problem

$$(\lambda I - \hat{\mathcal{W}})x = (\lambda I - \mathcal{Z}_1^{-1}\hat{\mathcal{H}}\mathcal{Z}_2^{-1})x = 0, \tag{31}$$

with

$$\hat{W} = \begin{bmatrix} -(\varepsilon D + \frac{1}{2}G)M^{-1} & -K + \frac{1}{4}GM^{-1}G + \varepsilon DM^{-1}G \\ M^{-1} & -\frac{1}{2}M^{-1}G \end{bmatrix} = W + \varepsilon \begin{bmatrix} -DM^{-1} & DM^{-1}G \\ 0 & 0 \end{bmatrix}.$$

If the perturbation ε is small, then eigenvalue/eigenvector pairs (or invariant subspaces) of the problem (1) can be regarded as small perturbations of eigenvalue/eigenvector pairs (or invariant subspaces) of problem (26). In this situation it is natural to use an eigenvalue/eigenvector pair of the undamped problem (26) computed via Algorithm 1 as start for a correction method. This method could be any method that can be used for eigenvalue, eigenvector, or subspace correction, such as subspace iteration, inverse iteration, Newton's method, or the Jacobi–Davidson method. Here we present results that are obtained with the quadratic Jacobi–Davidson method [3, 36] as correction. This method is typically very efficient and well suited for the given problem. However, if the desired eigenvalues of (1) form a cluster of nearby eigenvalues, then the quadratic Jacobi–Davidson method sometimes has difficulties in detecting and resolving such a cluster. The undesired effect is that, in this case for different starting values (eigenvalue/eigenvector pairs) obtained from the unperturbed problem, it converges to the same eigenvalue/eigenvector pair of (1). It is known that implicit deflation techniques based on Schur forms (see, e.g., section 4.7 and section 8.4 of [3]) combined with the Jacobi–Davidson method perform well for linear eigenvalue problems. However, in the quadratic eigenvalue problem, it is not clear how to incorporate an implicit deflation technique because a quadratic Schur form, in general, does not exist for quadratic matrix polynomials.

For this reason we have also analyzed the use of an explicit nonequivalence low-rank transformation deflation technique that was suggested in [9, 12] for quadratic eigenvalue problems. Let us briefly recall this technique for the polynomial

$$\mathbf{L}(\lambda) = \lambda^2 M + \lambda C + K \quad \text{with} \quad C = G + \varepsilon D. \tag{32}$$

We study the two cases of real eigenvalues or complex conjugate pairs separately. Suppose that we have computed a real eigenvalue λ_1 as well as the associated right and left eigenvectors x_1 and z_1, respectively, with $z_1^T K x_1 = 1$ such that $\mathbf{L}(\lambda_1)x_1 = 0$ and $z_1^T \mathbf{L}(\lambda_1) = 0$. Let $\theta_1 = (z_1^T M x_1)^{-1}$. We then introduce a new deflated quadratic eigenproblem as in [12] via

$$\widetilde{\mathbf{L}}(\lambda)x \equiv \left[\lambda^2 \widetilde{M} + \lambda \widetilde{C} + \widetilde{K}\right] x = 0, \tag{33}$$

where

$$\begin{aligned} \widetilde{M} &= M - \theta_1 M x_1 z_1^T M, \\ \widetilde{C} &= C + \frac{\theta_1}{\lambda_1}(M x_1 z_1^T K + K x_1 z_1^T M), \\ \widetilde{K} &= K - \frac{\theta_1}{\lambda_1^2} K x_1 z_1^T K. \end{aligned} \tag{34}$$

Suppose that we have computed an complex eigenvalue $\lambda_1 = \alpha_1 + i\beta_1$ as well as the associated right and left eigenvectors $x_1 = x_{1R} + i x_{1I}$ and $z_1 = z_{1R} + i z_{1I}$, respectively, such that $Z_1^T K X_1 = I_2$, where $X_1 = [x_{1R}, x_{1I}]$ and $Z_1 = [z_{1R}, z_{1I}]$. Let

$\Theta_1 = (Z_1^T M X_1)^{-1}$. We then introduce a new deflated quadratic eigenproblem as in [9] via (33), where

$$\begin{aligned} \widetilde{M} &= M - MX_1\Theta_1 Z_1^T M, \\ \widetilde{C} &= C + MX_1\Theta_1\Lambda_1^{-T} Z_1^T K + KX_1\Lambda_1^{-1}\Theta_1^T Z_1^T M, \\ \widetilde{K} &= K - KX_1\Lambda_1^{-1}\Theta_1\Lambda_1^{-T} Z_1^T K \end{aligned} \tag{35}$$

in which $\Lambda_1 = \left[\begin{smallmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{smallmatrix}\right]$.

Note that if the matrices M and K are symmetric, then the matrices $\widetilde{M}$ and $\widetilde{K}$ in (34) or (35) are symmetric as well. The results in [9, 12] then imply the following proposition.

PROPOSITION 1. (i) *Let λ_1 be a simple real eigenvalue of $\mathbf{L}(\lambda)$ as in* (32) *and let x_1 and z_1 with $z_1^T K x_1 = 1$ be the right and left eigenvectors, respectively. Then the spectrum of $\widetilde{\mathbf{L}}(\lambda)$ in* (33) *with coefficients as in* (34) *is given by $(\sigma(\mathbf{L}(\lambda))\setminus\{\lambda_1\}) \cup \{\infty\} = \sigma(\widetilde{\mathbf{L}}(\lambda))$ provided that $\lambda_1^2 \neq \theta_1$.*

(ii) *Let λ_1 be a simple complex eigenvalue of $\mathbf{L}(\lambda)$ as in* (32) *and let $X_1 = [x_{1R}, x_{1I}]$ and $Z_1 = [z_{1R}, z_{1I}]$ with $Z_1^T K X_1 = I_2$, where $x_1 = x_{1R} + i x_{1I}$ and $z_1 = z_{1R} + i z_{1I}$ are the associated right and left eigenvectors, respectively. Then the spectrum of $\widetilde{\mathbf{L}}(\lambda)$ in* (33) *with coefficients as in* (35) *is given by $\big(\sigma(\mathbf{L}(\lambda))\setminus\{\lambda_1, \bar{\lambda}_1\}\big)\cup\{\infty,\infty\} = \sigma(\widetilde{\mathbf{L}}(\lambda))$ provided that $\Lambda_1\Lambda_1^T \neq \Theta_1$.*

Furthermore, in both cases (i) *and* (ii)*, if $\lambda_2 \neq \lambda_1$ and (λ_2, x_2) is an eigenpair of $\mathbf{L}(\lambda)$, then the pair (λ_2, x_2) is also an eigenpair of $\widetilde{\mathbf{L}}(\lambda)$.*

Using this deflation procedure we have now presented all the ingredients for our algorithm. We solve the damped gyroscopic system (1) by the quadratic Jacobi–Davidson method combined with the explicit nonequivalence deflation technique (33), (34) or (33), (35) by using eigenpairs computed via Algorithm 1 as starting values. We summarize this approach in the following algorithm.

ALGORITHM 2 (Quadratic Jacobi–Davidson method with deflation).

Input Matrices M, G, D, and K and parameter ε as in (1). Target shift λ_0 and number ℓ of desired eigenvalue/eigenvector pairs nearest to λ_0. Tolerance *Tol* for stopping criterion.

Output The ℓ eigenpairs $\{(\lambda_j, x_j)\}_{j=1}^{\ell}$ of

$$\mathbf{L}(\lambda)x = (\lambda^2 M + \lambda(G + \varepsilon D) + K)x = 0,$$

associated with eigenvalues $\{\lambda_j\}_{j=1}^{\ell}$ that are closest to the target λ_0.

(i) Compute the ℓ eigenpairs $\{(\lambda_j^{(0)}, x_j^{(0)})\}_{j=1}^{\ell}$ of

$$(\lambda^2 M + \lambda G + K)x = 0$$

by Algorithm 1, where $\{\lambda_j^{(0)}\}_{j=1}^{\ell}$ are the closest eigenvalues to the target λ_0 and $(x_j^{(0)})^H x_j^{(0)} = 1$.

(ii) For $j = 1, \ldots, \ell$

  Compute the eigenpair $(\lambda_j, (x_j, z_j))$ of $\mathbf{L}(\lambda)$ by the quadratic Jacobi–Davidson method with target $\lambda_j^{(0)}$ and initial vector $x_j^{(0)}$, where x_j and z_j are the associated right and left eigenvectors satisfying the relations as in (i) or (ii), respectively, of Proposition 1.

  If $\|x_j - x_i\| \leq Tol$ for some $i < j$, then

    Compute the eigenpair $(\lambda_j, (x_j, z_j))$ of $\widetilde{\mathbf{L}}(\lambda)$ as in (33), (34) if λ_i is real, and as in (33), (35) if λ_i is complex, by the quadratic Jacobi–Davidson method with target $\lambda_j^{(0)}$ and initial vector $x_j^{(0)}$.

  End if

End for

Note that a recently proposed locking technique for the solution of quadratic eigenvalue problem as suggested in [22] might also be adapted instead of our non-equivalence low-rank deflation technique in Algorithm 2.

5. Numerical results. In this section we present some numerical tests for the algorithms proposed in this paper. All computations were done in MATLAB 6.0 [21] or in Fortran 90 on a Compaq DS20E workstation.

We first discuss the new extraction method and compare it with several other iterative methods. These methods are as follows:

EM:	the SHIRA algorithm combined with the extraction method given by Algorithm 1;
SISI(q):	the SHIRA algorithm followed by q steps of shift-and-invert subspace iteration;
EM_SISI:	the SHIRA algorithm combined with the extraction method EM adding one step of shift-and-invert subspace iteration;
IPI:	one step of the inverse power iteration.

Here, in the SISI the target value λ_0 with negative real part is taken as a fixed shift and the iteration starts with the subspace $\mathrm{span}(Q_k)$ generated by SHIRA, which includes the eigenvectors associated with eigenvalues near λ_0 and $-\lambda_0$. Then SISI converges to a subspace $\mathrm{span}(Q_k^-)$, associated with stable eigenvalues and the eigenpairs are computed from the "Ritz"-pairs of $(Q_k^-)^T\mathcal{W}Q_k^-$.

In the EM_SISI variant, the target value with negative real part is again taken as a fixed shift and one extra step of SISI is performed starting with $\mathrm{span}(V_k^-)$ obtained by the new extraction method EM. With the resulting subspace $\mathrm{span}(V_k)$, then the eigenpairs are computed from the "Ritz"-pairs of $V_k^T\mathcal{W}V_k$.

In the inverse power iteration (IPI) an approximate eigenvalue computed by SHIRA is taken as shift and the associated Arnoldi vector q_i is used as starting vector.

In the following we present results for three problem classes: quadratic eigenvalue problems from elasticity theory, linear quadratic optimal control problems, and damped gyroscopic systems.

5.1. Quadratic eigenproblems from elasticity theory. Consider the quadratic eigenvalue problem

$$\lambda^2 Mx + \lambda Gx + Kx = 0, \tag{36}$$

where M, G, and K are defined as follows. As in [26], let

$$B = \begin{bmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{m\times m},$$

$\tilde{M} = \frac{1}{6}(4I_m + B + B^T)$, $\tilde{G} = B - B^T$, and $\tilde{K} = -(2I_m - B - B^T)$. Define

$$\begin{aligned} M &= c_{11} I_m \otimes \tilde{M} + c_{12}\tilde{M} \otimes I_m, \\ G &= c_{21} I_m \otimes \tilde{G} + c_{22}\tilde{G} \otimes I_m, \\ K &= c_{31} I_m \otimes \tilde{K} + c_{32}\tilde{K} \otimes I_m, \end{aligned} \tag{37}$$

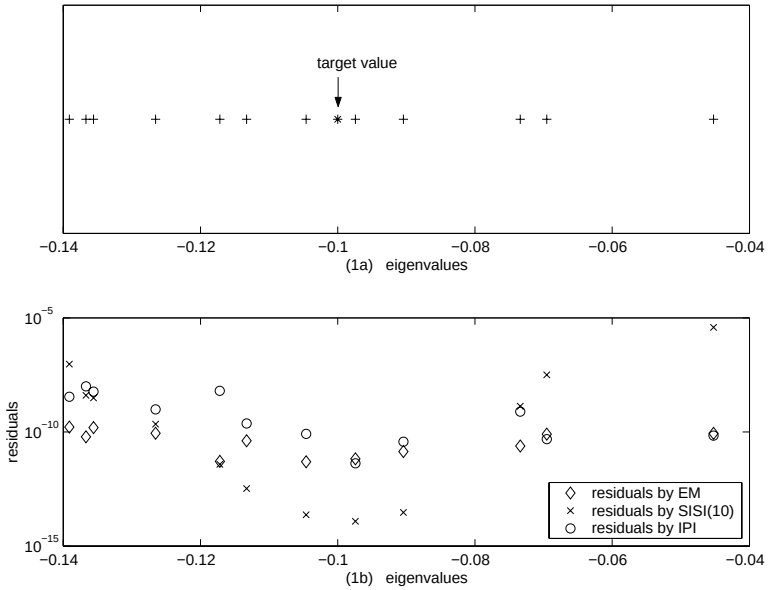

FIG. 1. *Eigenvalues and the corresponding residual.*

where c_{ij} are positive constants. Then we have $M = M^T > 0$, $G = -G^T$, and $K = K^T$.

Example 1 (see [26, Example 6.2]). We take $m = 90$, so that $\mathcal{W}$ is a 16200×16200 matrix, and

$$\begin{aligned} c_{11} &= 1.00, \qquad c_{12} = 1.30, \\ c_{21} &= 0.10, \qquad c_{22} = 1.10, \\ c_{31} &= 1.00, \qquad c_{32} = 1.20. \end{aligned} \tag{38}$$

For this example, the twelve eigenvalues computed by the SHIRA algorithm closest to the target $\lambda_0 = -0.1$ are depicted in Figure 1(a). The residuals $\| (\lambda_i^2 M + \lambda_i G + K)x_i \|_2$ for the eigenpairs (λ_i, x_i), where x_i is computed via EM, SISI(10), and IPI, respectively, are given in Figure 1(b).

We see from Figure 1(b) that the methods EM and IPI yield residuals of the same magnitude, $\approx 10^{-10}$, while SISI(10) yields smaller residual for some eigenvalues but the residuals of the eigenvalues further away from the target are much larger.

In order to evaluate the computational complexity for this problem class, we compare the major computational tasks per iteration step. One iteration of the inverse power iteration for (36) requires one forward/backward substitution for evaluating $Q(\lambda)^{-1}z$ (assuming that an LU-factorization for $Q(\lambda)$ is available). On the other hand, from (29), we see that the matrix $\mathcal{W}$ can be factored as

$$\mathcal{W} = \begin{bmatrix} I & -\frac{1}{2}GM^{-1} \\ 0 & M^{-1} \end{bmatrix} \begin{bmatrix} 0 & -K \\ I & 0 \end{bmatrix} \begin{bmatrix} I & -\frac{1}{2}G \\ 0 & I \end{bmatrix}.$$

The dominant cost of the extraction method is to determine the eigenvector in step (vii) of Algorithm 1. It consists of matrix-vector products for evaluating Kz and Gz, respectively, forward/backward substitution for evaluating $M^{-1}z$ (assuming that an

TABLE 5.1
Computational costs for EM and IPI; here $k = \ell = 12$.

EM		IPI	
0	evaluations $Q(\lambda_i)$	ℓ	evaluations $Q(\lambda_i)$
1	sparse LU factorization of M	ℓ	sparse LU factorizations of $Q(\lambda_i)$
ℓ	f/b substitutions of $M^{-1}z$	ℓ	f/b substitutions of $Q(\lambda_i)^{-1}z$
ℓ	matrix-vector products Kz	0	matrix-vector products Kz
2ℓ	matrix-vector products Gz	0	matrix-vector products Gz
$\ell(2+k)$	SAXPY operations	0	SAXPY operations

LU or Cholesky-factorization for M is available) as well as $2+k$ SAXPY operations. The cost for the evaluation of Ω_k, U_k and V_k in Algorithm 1 are negligible if $k \ll n$.

In Table 5.1 we compare the computational cost for the computation of eigenvectors using the EM and one step of the IPI, respectively. This comparison shows that the computational cost for the new extraction method compares favorably with that of one step of inverse power iteration. The major savings in computational time arise from the fact that no further factorizations of $Q(\lambda)$ are needed.

5.2. Optimal control problems. The second class of problems arises from continuous-time linear quadratic control problems of the form

$$\min_{x,u} \int_0^\infty (u^T R u + x^T C^T C x)dt \quad \text{subject to} \quad \begin{cases} \dot{x} = Ax + Bu, \\ y = Cx, \end{cases} \tag{39}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $R \in \mathbb{R}^{m\times m}$ with $p, m \ll n$ and where R is symmetric positive definite.

As test cases we consider the spatial (central difference) discretization in polar coordinates of a reaction-diffusion equation with Dirichlet boundary conditions on the two-dimensional unit disk; see [17] for details.

Performing a semidiscretization in space, one obtains a continuous-time system as in (39),

$$A = I + T, \tag{40}$$

and there exists an orthogonal matrix W, see [17], such that the matrix A in (40) can be transformed to a tridiagonal matrix $\widetilde{A} \equiv W^T A W$. Thus, the LU-factorization of $\widetilde{A}$ is easily computed in $\mathcal{O}(n)$ operations [11]. Let

$$W^T(A - \lambda_0 I)W = L_1 U_1, \qquad -W^T(A^T + \lambda_0 I)W = L_2 U_2 \tag{41}$$

be LU-factorizations of $W^T(A - \lambda_0 I)W$ and $-W^T(A^T + \lambda_0 I)W$, respectively. For this problem, $(\mathcal{W} - \lambda_0 I)^{-1}$ can be factored as

$$\begin{aligned}(\mathcal{W} - \lambda_0 I)^{-1} = &\begin{bmatrix} W & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} F^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -\tilde{B}\tilde{B}^T \\ 0 & I \end{bmatrix} \begin{bmatrix} F & 0 \\ 0 & D^{-1} \end{bmatrix} \\ &\times \begin{bmatrix} I & 0 \\ -\tilde{C}^T\tilde{C} & I \end{bmatrix} \begin{bmatrix} F^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} W^T & 0 \\ 0 & W^T \end{bmatrix},\end{aligned} \tag{42}$$

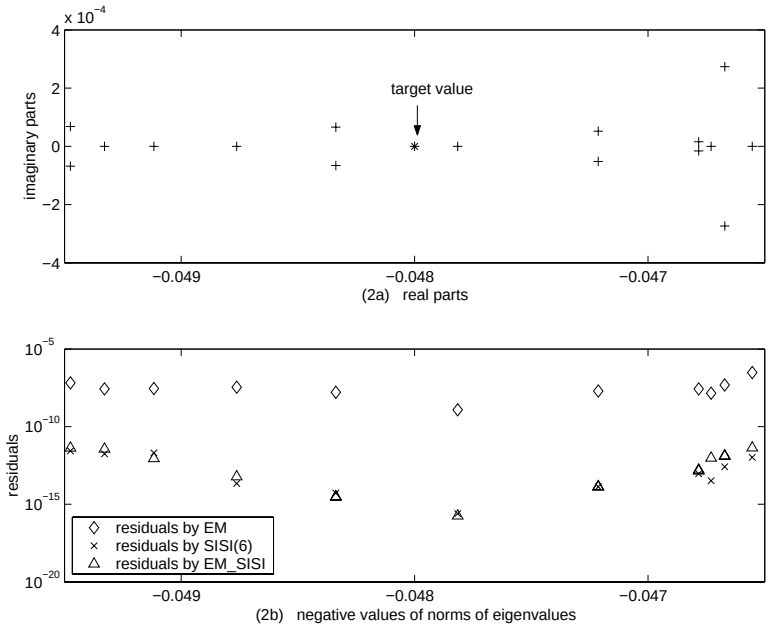

FIG. 2. *Eigenvalues and corresponding residuals.*

where $F = W^T(A - \lambda_0 I)W$, $\tilde{B} = W^T B$, $\tilde{C} = CW$, and $D = -W^T(A^T + \lambda_0 I)W - \tilde{C}^T\tilde{C}F^{-1}\tilde{B}\tilde{B}^T$. To evaluate D^{-1}, we use the LU-factorizations in (41) combined with the Sherman–Morrison–Woodbury formula [11] to obtain

$$D^{-1} = U_2^{-1}\left(I + L_2^{-1}\tilde{C}^T S^{-1}\tilde{C}U_1^{-1}L_1^{-1}\tilde{B}\tilde{B}^T U_2^{-1}\right)L_2^{-1}, \tag{43}$$

where

$$S = I - \tilde{C}U_1^{-1}L_1^{-1}\tilde{B}\tilde{B}^T U_2^{-1}L_2^{-1}\tilde{C}^T. \tag{44}$$

Example 2. With the data in [17] we get a system of size $n = 8100$. We chose $B, C \in \mathbb{R}^{n\times 2}$ randomly with uniform distribution in the interval $(0, 1)$ and used SHIRA to compute the 16 eigenvalues closest to $\lambda_0 = -0.048$. The eigenvalues are depicted in Figure 2(a). The associated eigenvectors are then computed by EM, EM_SISI, SISI(6), and IPI, respectively. The residuals for the eigenpairs computed by EM, EM_SISI, and SISI(6), respectively, are shown in Figure 2(b).

In this example the IPI had convergence problems for the vectors x_j associated with those eigenvalues with index $j = 2, 3, 5, 6, 7, 8, 10, 11, 15, 16$, because these complex conjugate pairs have relatively small imaginary parts and their use as shifts is not adequate. Here we see the loss of accuracy in the extraction method and we also observe that, as expected, this is easily compensated by adding a step of SISI.

Let us briefly compare the costs for computing an ℓ-dimensional invariant subspace associated with stable eigenvalues for optimal control problems via EM_SISI and SISI.

The matrix S in (44) is computed in SHIRA; it can be reused for SISI. It follows from the factorizations (42) and (43) that one iteration of SISI with reorthogonalization requires 5ℓ forward/backward substitutions by assuming that the

TABLE 5.2
Computational costs for EM and SISI(1); here $k = \ell = 16$.

EM		One step of SISI (SISI(1))	
0	W-transformations	4ℓ	W-transformations
0	f/b substitutions	5ℓ	f/b substitutions
ℓ	Az	0	Az
ℓ	$A^T z$	0	$A^T z$
$\ell(m+p)$	inner products	$2\ell(m+p) + \frac{1}{2}\ell(\ell+1)$	inner products
$\ell(k+m+p)$	SAXPY operations	$\ell(2m+2p-1) + \frac{1}{2}\ell(\ell+1)$	SAXPY operations

two LU-factorizations in (41) are available, as well as 4ℓ transformations with W, $2\ell(m+p) + \frac{\ell(\ell+1)}{2}$ inner products, and $\ell(2p+2m-1) + \frac{\ell(\ell+1)}{2}$ SAXPY operations.

On the other hand, the cost to compute V_ℓ^- via the Algorithm 1 consist of 2ℓ matrix-vector products plus $\ell(p+m)$ inner products and $\ell(k+p+m)$ SAXPY operations. In Table 5.2 we list these computational costs.

Again we see that the computational cost for the EM is much smaller than that of the one step of SISI.

5.3. Damped gyroscopic systems. For damped gyroscopic systems we apply the quadratic Jacobi–Davidson method as described in Algorithm 2 to compute the desired eigenpairs. The initial vectors for step (ii) of Algorithm 2 are obtained by Algorithm 1 (EM). Another way (denoted as Algorithm 2v) to start step (ii) of Algorithm 2 is to use $(\lambda_j^{(0)}, y_j)$ as the initial vector, where y_j is a randomly chosen vector and $\lambda_j^{(0)}$ is an eigenvalue of $\lambda^2 M + \lambda G + K$ obtained by SHIRA. In the following we compare these two variants of choosing initial vectors.

As a test case we consider the quadratic eigenvalue problem

$$\lambda^2 Mx + \lambda(G + \varepsilon D)x + Kx = 0, \tag{45}$$

where M, G, and K are as in (37) and $D = D^T$ is given by

$$D = c_{41} I_m \otimes \tilde{D} + c_{42} \tilde{D} \otimes I_m$$

with positive c_{41}, c_{42} and

$$\tilde{D} = \begin{bmatrix} 2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 2 \end{bmatrix}.$$

Example 3. We take $m = 90$, $c_{11}, \ldots, c_{32}$ as in (38), $c_{41} = 1.05$, and $c_{42} = 0.9$. The matrices M, G, D, and K are 8100×8100 matrices.

We have implemented both variants of choosing initial pairs as Fortran 90 subroutines in order to carry out numerical experiments. The following numerical results were performed on a Compaq DS20E workstation equipped with dual 667 MHz CPUs.

In Figure 3(a) and (b) we depict the eigenvalues of (45) with $\varepsilon = 10^{-4}$ and 10^{-3} computed by Algorithm 2 with initial eigenpairs obtained by Example 1.

In Figure 3(b), we observe three clusters of eigenvalues near to -0.14, -0.12, and -0.07. For the eigenvalues near -0.12 or -0.07 one deflation step is needed, and for those near -0.14 two deflation steps are needed. The other variant of choosing the initial vector yields results that are almost the same as those of Algorithm 2.

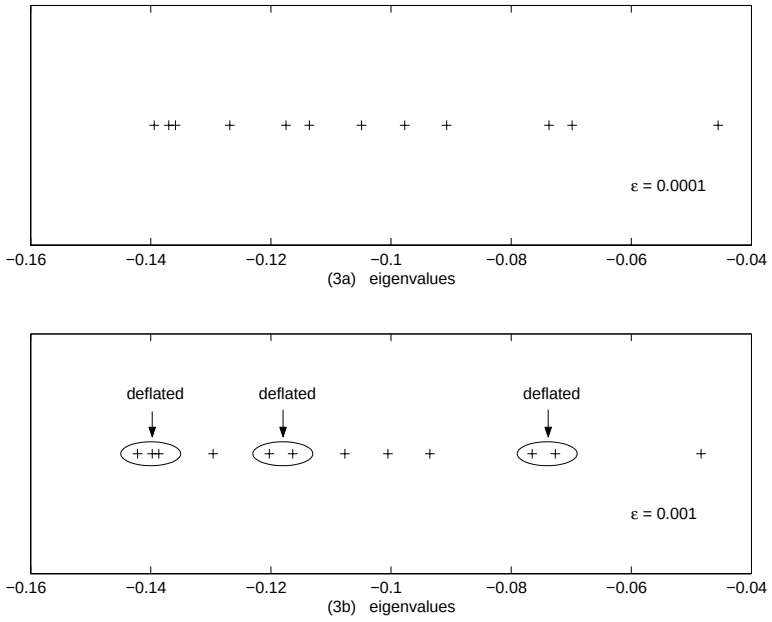

FIG. 3. *Distribution of eigenvalues closest to the target* $\lambda_0 = -0.1$.

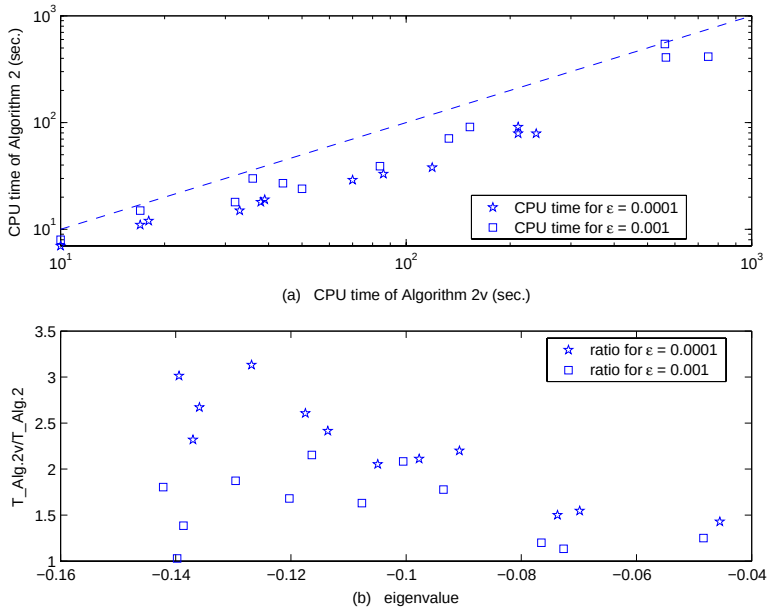

FIG. 4. *Comparison of the CPU times for Algorithm* 2 *and Algorithm* 2*v*

We also compared the CPU times for Algorithm 2 and the variant Algorithm 2v (for $\varepsilon = 10^{-4}$ and 10^{-3}). The CPU times for step (ii) of Algorithm 2, i.e., for the solution computed by quadratic Jacobi–Davidson method until convergence, are depicted in Figure 4. Here, the CPU times for EM and SHIRA are not included. The values in the x- and y-axis of Figure 4(a), denoted by T_Alg.2v and T_Alg.2,

respectively, represent the CPU times for the computation of one eigenpair of (45) by Algorithm 2 and Algorithm 2v, respectively. The ratios of T_Alg.2v and T_Alg.2 for each eigenvalue, and for $\varepsilon = 10^{-4}$ and $\varepsilon = 10^{-3}$, are given in Figure 4(b).

We conclude from Figure 4 that for small damping the approach that computes first eigenpairs of the undamped problem by EM and then uses the quadratic Jacobi–Davidson method combined with the nonequivalence deflation technique to correct these eigenpairs is an attractive approach.

6. Conclusion. In this paper we have discussed numerical methods for structured eigenvalue problems with a SHH structure as they arise from quadratic eigenvalue problems, gyroscopic systems, or optimal control problems. We have also studied problems where this structure is perturbed by a small damping term.

We have presented an extraction procedure that speeds up the computation of eigenvectors and invariant subspaces in combination with the SHIRA method considerably. We have also discussed the quadratic Jacobi–Davidson method combined with a nonequivalence deflation technique for slightly damped gyroscopic systems based on a computation of starting values for the eigenvalues and eigenvectors from the undamped problem. We have presented several numerical examples that show the efficiency of these new approaches.

REFERENCES

[1] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*, 3rd ed., SIAM, Philadelphia, 1999.

[2] T. Apel, V. Mehrmann, and D. Watkins, *Structured eigenvalue methods for the computation of corner singularities in 3D anisotropic elastic structures*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4459–4473.

[3] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templates for the solution of algebraic eigenvalue problem: A practical guide*, SIAM, Philadelphia, 2000.

[4] P. Benner, R. Byers, H. Fassbender, V. Mehrmann, and D.S. Watkins, *Cholesky-like factorizations of skew-symmetric matrices,* Electron. Trans. Numer. Anal., 11 (2000), pp. 85–93.

[5] P. Benner, R. Byers, V. Mehrmann, and H. Xu, *Numerical computation of deflating subspaces of skew Hamiltonian/Hamiltonian pencils*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 165–190.

[6] P. Benner, R. Byers, V. Mehrmann, and H. Xu, *Numerical methods for linear-quadratic and H_∞ control problems*, in Dynamical Systems, Control, Coding, Computer Vision: New Trends, Interfaces, and Interplay, Progr. Systems Control Theory 25, G. Picci and D. S. Gilliam, eds., Birkhäuser, Basel, 1999, pp. 203–222.

[7] J. R. Bunch, *A note on the stable decomposition of skew-symmetric matrices,* Math. Comp., 38 (1982), pp. 475–479.

[8] A. Bunse-Gerstner, *Matrix factorization for symplectic QR-like methods,* Linear Algebra Appl., 83 (1986), pp. 49–77.

[9] J. Carvalho, B. N. Datta, W.-W. Lin, and C.-S. Wang, *Eigenvalue Embedding in a Quadratic Pencil Using Symmetric Low Rank Updates,* Technical Report 2001-3, National Center for Theoretical Science, National Tsinghua University, Hsinchu 300, Taiwan, Republic of China, 2001.

[10] W. R. Ferng, W.-W. Lin, and C.-S. Wang, *Numerical algorithms for undamped gyroscopic systems,* Comput. Math. Appl., 37 (1999), pp. 49–66.

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

[12] J.-S. Guo, W.-W. Lin, and C.-S. Wang, *Numerical solutions for large sparse quadratic eigenvalue problems,* Linear Algebra Appl., 225 (1995), pp. 57–89.

[13] N. J. Higham, *A New* `sqrtm` *for MATLAB*, Numerical Analysis Report 336, Manchester Centre for Computational Mathematics, Manchester, England, 1999.

[14] M. E. Hochstenbach and H. A. van der Vorst, *Alternatives to the Rayleigh Quotient for the Quadratic Eigenvalue Problem*, Technical Report 1212, Department of Mathematics, Univ. Utrecht, P.O. Box 80.010, NL-3508 TA Utrecht, The Netherlands, 2001.
[15] A. S. Hodel and K. R. Poolla, *Heuristic approaches to the solution of very large sparse Lyapunov and algebraic Riccati equations*, in Proceedings of the 27th IEEE Conference on Decision and Control, Austin, TX, 1988, pp. 2217–2222.
[16] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
[17] T.-M. Hwang and W. Wang, *Analyzing and visualizing a discretized semilinear elliptic problem with Neumann boundary conditions,* Numer. Methods Partial Differential Equations, 18 (2002), pp. 261–279.
[18] P. Lancaster, *Strongly stable gyroscopic systems,* Electron. J. Linear Algebra, (1999), pp. 53–66.
[19] D. Leguillon, *Computation of* 3*d-singularities in elasticity*, in *Boundary Value Problems and Integral Equations in Nonsmooth Domains*, Lecture Notes in Pure Appl. and Math 167, M. Costabel et al., eds., Marcel Dekker, New York, 1995, pp. 161–170.
[20] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, SIAM, Philadelphia, 1998.
[21] *Manual Version* 6.0, The MathWorks, Inc., Natick, MA, 1993.
[22] K. Meerbergen, *Locking and restarting quadratic eigenvalue solvers,* SIAM J. Sci. Comput., 22 (2001), pp. 1814–1839.
[23] C. Mehl, *Compatible Lie and Jordan Algebras and Applications to Structured Matrices and Pencils*, Dissertation, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz (FRG), 1998.
[24] C. Mehl, *Condensed forms for skew-Hamiltonian/Hamiltonian pencils,* SIAM J. Matrix Anal. Appl., 21 (1999), pp. 454–476.
[25] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*, Lecture Notes in Control and Inform. Sci., 163, Springer–Verlag, Heidelberg, 1991.
[26] V. Mehrmann and D. Watkins, *Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils,* SIAM J. Sci. Comput., 22 (2001), pp. 1905–1925.
[27] A. Meyer, *Die simultane Iteration zur Eigenwertberechnung für beliebige endlichdimensionale Operatoren und die Anwendung auf Schwingungsprobleme*, Report Wissenschaftliche Informationen 36, Technische Hochschule Karl-Marx-Stadt (now Technische Universität Chemnitz), Fakultät für Mathematik, 1982.
[28] T. Penzl, *Numerische Lösung großer Lyapunov-Gleichungen.* Logos–Verlag, Berlin, 1998.
[29] T. Penzl, *Lyapack Guide*, Technical Report SFB393/00-33, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, FRG, 2000; also available online from http://www.tu-chemnitz.de/sfb393.
[30] I. G. Rosen and C. Wang, *A multilevel technique for the approximate solution of operator Lyapunov and algebraic Riccati equations*, SIAM J. Numer. Anal., 32 (1995), pp. 514–541.
[31] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, UK, 1992.
[32] R. Sandhu and K. Pister, *A variational principle for linear coupled field problems in continuum mechanics,* Internat. J. Engrg. Sci., 8 (1970), pp. 989–999.
[33] H. Schmitz, K. Volk, and W. L. Wendland, *On three-dimensional singularities of elastic fields near vertices,* Numer. Methods Partial Differential Equations, 9 (1993), pp. 323–337.
[34] H. R. Schwarz, *Methode der Finiten Elemente*, Teubner, Stuttgart, 1984.
[35] B. Simon, J. Wu, O. Zienkiewicz, and D. Paul, *Evaluation of u-w and u-p finite element methods for the dynamic response of saturated porous media using one-dimensional models,* Internat. J. Numer. Anal. Methods Geomech., 10 (1986), pp. 461–482.
[36] G. L. G. Sleijpen, G. L. Booten, D. R. Fokkema, and H. A. van der Vorst, *Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems,* BIT, 36 (1996), pp. 595–633.
[37] G. Sleijpen, H. van der Vorst, and M. van Gijzen, *Quadratic eigenproblems are no problem,* SIAM News, 29 6 (1996), pp. 8–9, 13.
[38] D. C. Sorensen, *Implicit application of polynomial filters in a k-step Arnoldi method,* SIAM J. Matrix Anal. Appl., 13 (1992), pp. 357–385.
[39] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[40] F. Tisseur, *Backward error analysis of polynomial eigenvalue problems,* Linear Algebra Appl., 309 (2000), pp. 339–361.
[41] F. Tisseur, *Stability of structured Hamiltonian eigensolvers,* SIAM J. Matrix Anal. Appl., 23 (2001), pp. 103–125.
[42] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem,* SIAM Rev., 43 (2001), pp. 235–286.
[43] C. F. Van Loan, *A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix*, Linear Algebra Appl., 61 (1984), pp. 233–251.
[44] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.
[45] H. Xu, *Solving Algebraic Riccati Equations via Skew-Hamiltonian Matrices*, Ph.D. thesis, Inst. of Math., Fudan University, Shanghai, P.R. China, 1993.
[46] H. Xu and L. Lu, *Properties of a quadratic matrix equation and the solution of the continuous-time algebraic Riccati equation,* Linear Algebra Appl., 222 (1995), pp. 127–145.