AN IMPLICITLY-RESTARTED KRYLOV SUBSPACE METHOD FOR REAL SYMMETRIC/SKEW-SYMMETRIC EIGENPROBLEMS[§]

V. MEHRMANN[†], C. SCHRÖDER[†], AND V. SIMONCINI[‡]

Abstract. A new implicitly-restarted Krylov subspace method for real symmetric/skew-symmetric generalized eigenvalue problems is presented. The new method improves and generalizes the SHIRA method of [36] to the case where the skew symmetric matrix is singular. It computes a few eigenvalues and eigenvectors of the matrix pencil close to a given target point. Several applications from control theory are presented and the properties of the new method are illustrated by benchmark examples.

1. Introduction. In this paper we discuss the computation of a few eigenvalues and eigenvectors of large sparse structured generalized eigenvalue problems of the form

$$Mx = \lambda Nx, \tag{1.1}$$

where $M \in \mathbb{R}^{n\times n}$ is symmetric and $N \in \mathbb{R}^{n\times n}$ is skew symmetric. The numerical approximation of generalized eigenvalue problems of the form (1.1) is of great importance in a variety of applications, including the solution of linear quadratic optimal control problems [7, 12, 34, 50], robust control problems [9, 31, 38, 52], passivity analysis and passivation of linear systems [1, 17, 20, 21, 24, 42, 44], model reduction [1, 23, 19, 35, 39, 46], crack following in anisotropic materials [2, 26, 36, 37] and others; see also [32] and Section 6.

Eigenvalue problems associated with symmetric/skew-symmetric matrix pencils $M - \lambda N$, where $N = -N^\top$ and $M = M^\top$ (cf. [16]) occur in different representations under the names *alternating* eigenvalue problems [36], *(generalized or extended) Hamiltonian* eigenvalue problems [38], *skew-Hamiltonian/Hamiltonian* eigenvalue problems [7] or more recently *even/odd* eigenvalue problems [32].

For small dense problems the perturbation analysis, structured normal forms, as well as the existence of structured Schur forms are well understood, see [16, 28, 29, 49]. Numerical solution methods for small dense problems have been developed and implemented into standardized software [7, 12, 10, 14, 16, 27].

In the following we assume that the problem is too large to apply a dense method that computes all the eigenvalues as well as a generalized structured Schur form, and we restrict ourselves to the case in which the structure of M and N allows the use of sparse direct LU-factorizations of $M - \sigma N$ for some shift σ. For this class of problems, or more general for polynomial eigenvalue problems with this structure, structure preserving methods based on implicitly restarted Arnoldi methods have been suggested in [2, 26, 36, 37, 51] for the case that N is invertible. Here we allow N to be singular, but the pencil $M - \lambda N$ still is a regular pencil, i. e. its determinant does not vanish identically.

It is our goal to compute a few eigenvalues and associated eigenvectors near a given shift σ. We restrict ourselves to the case that σ is either real or purely imaginary.

[†]Institut für Mathematik, Ma 4-5, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, FRG; email: {mehrmann,schroed}@math.tu-berlin.de

[‡]Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, I-40127 Bologna, Italy and CIRSA, Ravenna, Italy; email: valeria@dm.unibo.it.

[§]Research supported by *Deutsche Forschungsgemeinschaft*, via the DFG Research Center MATHEON, Mathematics for Key Technologies, in Berlin and by Berlin Mathematical School.

An extension to general complex shifts is possible, but would require either complex arithmetic or a modified approach, see Remark 2.1 below.

Throughout the paper we denote by I_m, O_m the identity and zero matrices of size m, respectively; analogously, $O_{n,m}$ is a $n \times m$ zero matrix. Moreover, $X^\top$ denotes the transpose of the possibly complex matrix X, and X^* its conjugate transpose. If two vectors x, y are linearly dependent we write $x \propto y$.

Let $\lambda E - A$ be a matrix pencil with $E, A \in \mathbb{R}^{n,n}$. Then $\lambda E - A$ is called *regular* if $\det(\lambda E - A) \neq 0$ for some $\lambda \in \mathbb{C}$. For regular pencils, *generalized eigenvalues* are the pairs $(\alpha, \beta) \in \mathbb{C}^2 \setminus \{(0,0)\}$ for which $\det(\alpha E - \beta A) = 0$. If $\beta \neq 0$, then the pair represents the finite eigenvalue $\lambda = \alpha/\beta$. If $\beta = 0$, then (α, β) represent the eigenvalue infinity.

The basic properties of regular matrix pencils can be determined from the Weierstraß canonical form.

THEOREM 1.1 (Weierstraß canonical form). *For a regular matrix pencil $\lambda E - A$, there exist matrices $W_f, V_f \in \mathbb{R}^{n,n_f}$, $W_\infty, V_\infty \in \mathbb{R}^{n,n_\infty}$ so that $W = \begin{bmatrix} W_f & W_\infty \end{bmatrix}$, $V = \begin{bmatrix} V_f & V_\infty \end{bmatrix}$ are square and invertible, with*

$$
\begin{aligned}
W^\top E V &= \begin{bmatrix} W_f^\top \\ W_\infty^\top \end{bmatrix} E \begin{bmatrix} V_f & V_\infty \end{bmatrix} = \begin{bmatrix} I_{n_f} & O_{n_f,n_\infty} \\ O_{n_\infty,n_f} & \mathcal{N}_\infty \end{bmatrix}, \\
W^\top A V &= \begin{bmatrix} W_f^\top \\ W_\infty^\top \end{bmatrix} A \begin{bmatrix} V_f & V_\infty \end{bmatrix} = \begin{bmatrix} A_f & 0 \\ 0 & I_{n_\infty} \end{bmatrix},
\end{aligned}
$$

$A_f \in \mathbb{R}^{n_f,n_f}$ is in real Jordan canonical form and $\mathcal{N}_\infty \in \mathbb{R}^{n_\infty,n_\infty}$ is a nilpotent matrix in Jordan canonical form.

We call n_f, n_∞ the number of finite or infinite eigenvalues, and the index of nilpotency of $\mathcal{N}_\infty$ the *index* of the system. A subspace $\mathcal{L} \subset \mathbb{R}^n$ is called *deflating subspace* for the pencil $\lambda E - A$ if for a matrix $X_\mathcal{L} \in \mathbb{R}^{n,k}$ with full column rank and $\mathrm{im} X_\mathcal{L} = \mathcal{L}$, there exist $Y_\mathcal{L} \in \mathbb{R}^{n,k}$, $R_\mathcal{L} \in \mathbb{R}^{k,k}$, $U_\mathcal{L} \in \mathbb{R}^{k,k}$ such that

$$
E X_\mathcal{L} = Y_\mathcal{L} R_\mathcal{L}, \quad A X_\mathcal{L} = Y_\mathcal{L} U_\mathcal{L}.
$$

2. Properties of the skew-symmetric/symmetric eigenvalue problem. As already mentioned, the eigenvalue problem in (1.1) is characterized by a rich structure, which becomes apparent by observing that the spectrum is symmetric with respect to both the real and the imaginary axis. More precisely, if λ is a real eigenvalue of (1.1) with right eigenvector x, then $-\lambda$ is also an eigenvalue of (1.1) with left eigenvector x. Analogously, if λ is a complex eigenvalue, then $-\lambda, \bar\lambda$ and $-\bar\lambda$ are also eigenvalues. In particular, this implies that the origin is always the center of the spectral region. This symmetry carries over to the canonical form under congruence [49] and also to associated staircase forms under unitary transformations [16].

In most applications one is interested in the accurate approximation of eigenvalues (pairs or quadruples) close to the origin or to the whole imaginary axis. Since Krylov subspace iterative methods tend to better approximate exterior eigenvalues, a standard strategy consists of performing a shift-and-invert transformation that maps the eigenvalues of (1.1) close to the target σ, to the largest eigenvalues of the transformed problem $(M - \sigma N)^{-1} N x = \eta x$, with $\eta = (\lambda - \sigma)^{-1}$ [3]. Although effective in general, this approach hinders the preservation of the special symmetry of our problem: quadruple eigenvalues $(\lambda, -\lambda, \bar\lambda, -\bar\lambda)$ are approximated by two pairs of unrelated complex conjugate eigenvalues. As a consequence, it is difficult to distinguish between true (matching) quadruples and close eigenvalues, unless full accuracy approximations can be obtained.

To preserve the particular structure, we therefore consider the following spectral transformation introduced in [36, 51], and its generalization to the case that N is allowed to be singular. With

$$K := (M + \sigma N)^{-1} N (M - \sigma N)^{-1} N, \tag{2.1}$$

we study the eigenvalue problem $Kx = \theta x$. It may be readily verified that if the eigenpair (λ, x) satisfies (1.1) then (θ, x) satisfies the equation $Kx = \theta x$ with $\theta = 1/(\lambda^2 - \sigma^2)$. Note also that the matrix K is real, even in case of a purely imaginary shift σ; see also Proposition 2.2.

Remark 2.1. *If the shift σ is not purely imaginary or on the real axis, we can still work with real arithmetic by replacing K in (2.1) with the real matrix*

$$(M + \sigma N)^{-1} N (M - \sigma N)^{-1} N (M + \bar{\sigma} N)^{-1} N (M - \bar{\sigma} N)^{-1} N.$$

We refrain here from discussing this case in detail.

It is important to note that in exact arithmetic, if $-\lambda$ is a paired eigenvalue of λ in (1.1) with eigenvector y, then $-\lambda$ is mapped onto the same eigenvalue θ. This fact clearly anticipates that, while the eigenvalue approximation could be obtained directly from the problem (2.1), the approximation of the eigenvectors requires additional work. We also note in passing that the property just outlined implies that eigenvalues of K are all multiple.

We next state a few obvious properties that will be used in what follows.

Proposition 2.2. *Let M be real symmetric and N be real and skew-symmetric. Let $\sigma \in \mathbb{C}$ be such that $M - \sigma N$ is nonsingular. Then*

i) $M + \sigma N = (M - \sigma N)^\top$;

ii) the matrices $(M + \sigma N)^{-1} N$ and $(M - \sigma N)^{-1} N$ commute;

iii) the matrix $K = (M + \sigma N)^{-1} N (M - \sigma N)^{-1} N$ satisfies

$$N K^\ell = -(N K^\ell)^\top, \quad \ell \in \mathbb{N}, \tag{2.2}$$

that is, NK^ℓ is skew-symmetric for any natural number ℓ. In particular, $K^\top N = NK$.

iv) If additionally $\sigma \in \imath\mathbb{R}$, then $M + \sigma N = (M + \sigma N)^*$.

These properties imply the following lemma.

Lemma 2.3. *Let M be real symmetric and N be real skew-symmetric. Let $\lambda \neq 0$ be a simple eigenvalue of (M, N) and x_+ an associated eigenvector. Let $(-\lambda, x_-)$ be the matched eigenpair of (M, N) corresponding to (λ, x_+). Then the following results hold:*

a) $x_+^\top N x_- \neq 0$;

b) Let $\mathcal{V}$ be an N-neutral subspace of $\mathbb{C}^n$, i.e., $v^\top N w = 0$ for any $v, w \in \mathcal{V}$. If $u \in span\{x_+, x_-\}$ satisfies $u \in \mathcal{V}$, then no other linearly independent vector of $span\{x_+, x_-\}$ also belongs to $\mathcal{V}$.

Proof. a) Without loss of generality let x_- and x_+ be of unit 2-norm. We have $Mx_- = (-\lambda) N x_-$ and by transposing, $x_-^\top M = \lambda x_-^\top N$, i.e., x_- is a left eigenvector for the eigenvalue λ. Hence $x_-^\top N x_+$ is the inverse eigenvalue condition number of λ, which is nonzero and finite for simple eigenvalues of $\lambda N - M$ [22, 25].

b) We scale x_+, x_- so that $x_-^\top N x_+ = 1$. Let $u_i = \alpha_i x_+ + \beta_i x_-$, $i = 1, 2$ be two vectors in $\operatorname{span}\{x_+, x_-\} \cap \mathcal{V}$. Then by the N-neutrality of $\mathcal{V}$ we have

$$0 = u_1^\top N u_2 = \alpha_1 \alpha_2 \underbrace{x_+^\top N x_+}_{0} + \alpha_1 \beta_2 \underbrace{x_+^\top N x_-}_{-1} + \alpha_2 \beta_1 \underbrace{x_-^\top N x_+}_{1} + \beta_1 \beta_2 \underbrace{x_-^\top N x_-}_{0},$$

i.e., $\beta_1\alpha_2 - \alpha_1\beta_2 = 0$. This implies that u_1 and u_2 are linearly dependent. ☐

We conclude this section with a property of polynomials in the matrix K.

LEMMA 2.4. *Under the assumptions of Proposition 2.2 let K be as in (2.1). Let $v_1 \in \mathbb{R}^n$. Let $w = p_i(K)v_1$, $z = p_j(K)v_1$, with p_i, p_j real polynomials of degree i and j, respectively. Then $w^\top N z = 0$.*

Proof. It is sufficient to prove the result for $p_k(\lambda) = \lambda^k$, $k = i, j$. Repeatedly using Proposition 2.2 iii), we have $w^\top N z = v_1^\top (K^\top)^i N K^j v_1 = v_1^\top N K^{i+j} v_1 = 0$. ☐

3. The new iterative method. The properties of the pencil (M, N) that we have discussed in the previous section suggest a way to simultaneously approximate matched eigenvalue pairs of (M, N). Indeed, it is sufficient to build a search space that is restricted to approximate only one vector from each two-dimensional deflating subspace associated with $(+\lambda, x_+)$ and $(-\lambda, x_-)$. Moreover, the eigenvalues of (M, N) may be obtained by mapping back approximate θ values to the matching pairs of the original problem as $\pm\lambda = \pm\sqrt{1/\theta + \sigma^2}$. This combined process avoids the distinct approximation of $+\lambda$ and $-\lambda$ by means of the approximation of both eigenvectors of the multiple eigenvalue θ.

We next describe the whole procedure through its three main phases.

Phase I. Generating the approximation space. Given a normalized starting vector v_1, we iteratively generate the following Krylov subspace as search space:

$$\mathcal{K}_m(K, v_1) = \mathrm{span}\{v_1, Kv_1, \ldots, K^{m-1}v_1\},$$

until a maximum size $m = m_{\max}$ is reached. An orthonormal basis $\{v_1, v_2, \ldots, v_m\}$ is constructed by means of a standard Arnoldi procedure (see, e.g., [3]). Due to the nested nature of the space, that is $\mathcal{K}_m(K, v_1) \subset \mathcal{K}_{m+1}(K, v_1)$, such a basis may be generated iteratively by adding one vector at the time.

We recall that K has only double eigenvalues, because $+\lambda$ and $-\lambda$ are mapped to the same value $\theta = ((\pm\lambda)^2 - \sigma^2)^{-1}$, and that we are interested in one instance of every θ value only. In exact arithmetic and if a Krylov subspace is used as approximation space, multiple eigenvalues will not be captured. However, computational round-off will eventually generate them. (Usually, if all eigenvalues are wanted this is advantageous, but here it does harm, because it generates spurious copies of the λ values.) So, in finite precision, extra care has to be taken to ensure that only one copy of each θ value is found. Thanks to Lemma 2.3, the restriction to a single vector for each two-dimensional deflating subspace may be enforced by requiring that the generated search space $\mathcal{K}$ is N-neutral.

Since a vector w in $\mathcal{K}_m(K, v_1)$ may be written in terms of a polynomial in K, namely $w = p_{m-1}(K)v_1$, Lemma 2.4 ensures that N-neutrality is obtained for free in exact arithmetic. Without taking further measures, however, N-neutrality of the basis is lost in finite precision arithmetic during the expansion. Therefore, we explicitly orthogonalize each new basis vector v_{m+1} not only against all previous vectors $v_1, \ldots, v_m$, but also against $Nv_1, \ldots, Nv_m$. Let $V_m = [v_1, v_2, \ldots, v_m]$. Then the Arnoldi recurrence can be compactly rewritten as

$$KV_m = V_m H_m + v_{m+1} h_{m+1,m} e_m^\top, \quad V_m^\top N V_m = O_m, \tag{3.1}$$

where e_m is the m-th vector of the Euclidean basis, whose dimension is clear from the context.

Phase II. Extraction of spectral information. To extract approximate eigenvalues, we compute the Ritz values μ of K (i.e., the eigenvalues of H_m) and we use

$\pm\sqrt{1/\theta+\sigma^2}$ as eigenvalue approximations for (M,N). Note that these values form matched pairs.

Up to this point our new method is a generalization of the SHIRA method in [36]. For the determination of approximate eigenvectors we deviate from SHIRA, which uses inverse iteration for this task.

The problem can be described as follows: if $(+\lambda, x_+)$ and $(-\lambda, x_-)$ are matched eigenpairs of (M,N), then (θ, x_+) and (θ, x_-) are eigenpairs of K to the double eigenvalue θ. So, every linear combination of x_+ and x_- is an eigenvector of K, but not of (M,N). Therefore, the search space based on K may well contain good approximations to eigenvectors of K, but not of (M,N). However, the following result shows that a Krylov-Schur decomposition can be recovered with respect to the original problem, so that the sought after approximations can be computed. To this end we introduce the following matrix

$$W_m(\alpha) = (M-\alpha N)^{-1} N V_m. \tag{3.2}$$

We explicitly keep track of the dependence on α in W_m because of later convenience. For purely imaginary α, the matrix $W_m(\alpha)$ is complex.

Proposition 3.1. *With the previous notation, and for $\alpha \in \{\sigma, -\sigma\}$ we have*

$$\begin{aligned} M[V_m, W_m(\alpha)] \begin{bmatrix} O_m & H_m \\ I_m & O_m \end{bmatrix} = N[V_m, W_m(\alpha)] &\begin{bmatrix} I_m & -\alpha H_m \\ \alpha I_m & I_m \end{bmatrix} \\ &+ [O_m, (M+\alpha N) v_{m+1} h_{m+1,m} e_m^\top]. \end{aligned} \tag{3.3}$$

Proof. For $\alpha=\sigma$ the proof follows by explicitly rewriting and collecting the two equations $W_m(\sigma) = (M-\sigma N)^{-1} N V_m$ and $(M+\sigma N)^{-1} N W_m(\sigma) = V_m H_m + h_{m+1,m} v_{m+1} e_m^\top$. Using Proposition 2.2 ii), the same technique works for $\alpha = -\sigma$. $\square$

In the remainder of this section we use $\alpha=\sigma$, whereas the choice $\alpha=-\sigma$ will be discussed in section 4.

The relation in (3.3) shows that if V_m generates an invariant subspace of K, i. e. if $h_{m+1,m}=0$, then the matrix $[V_m, W_m(\sigma)]$ generates a deflating subspace of (M,N). In addition, we notice that due to the commutativity of the factors in K (cf. Proposition 2.2), the matrix $[V_m, W_m(-\sigma)]$ also generates a deflating subspace of (M,N).

We use the eigenvalues and eigenvectors of the *reduced* problem

$$\begin{bmatrix} I_m & -\sigma H_m \\ \sigma I_m & I_m \end{bmatrix} z = \lambda \begin{bmatrix} O_m & H_m \\ I_m & O_m \end{bmatrix} z \tag{3.4}$$

to obtain spectral approximations to the original problem. If $h_{m+1,m}=0$, then $(\lambda, [V_m, W_m(\sigma)] z(\sigma))$ is an eigenpair of (M,N).

Further savings can be achieved by noticing that the eigenpairs of (3.4) need not be explicitly computed, but can be recovered from those of H_m. More precisely, let (μ, s) be an eigenpair of H_m, that is, $(\mu, V_m s)$ is a Ritz pair of K. Then it may be verified that $\hat\lambda_\pm = \pm\sqrt{\sigma^2+1/\mu}$ and $z(\sigma) = [s/(\hat\lambda_\pm - \sigma); s]$ are the eigenvalues and the corresponding eigenvectors of (3.4). Note that the transformation $\hat\lambda_\pm = \pm\sqrt{\sigma^2+1/\mu}$ is the same as that mapping the spectrum of K back to eigenvalues of (M,N). Note also that the sign of the eigenvalue $\hat{\lambda}_\pm$ only influences the factor of the first block vector of $z(\sigma)$. Finally, we have

$$\hat{z}_\pm(\sigma) := \begin{bmatrix} I_m & -\sigma H_m \\ \sigma I_m & I_m \end{bmatrix} z(\sigma) = \begin{bmatrix} \frac{1}{\hat{\lambda}_\pm+\sigma} s \\ s \end{bmatrix} \frac{\hat{\lambda}_\pm}{\hat{\lambda}_\pm - \sigma}. \tag{3.5}$$

The pairs

$$(\pm\hat{\lambda}, \frac{\hat{x}_\pm(\sigma)}{\|\hat{x}_\pm(\sigma)\|}), \quad \text{with} \quad \hat{x}_\pm(\sigma) = [V_m, W_m(\sigma)]\hat{z}_\pm(\sigma), \tag{3.6}$$

are the desired approximate eigenpairs of (M, N). In Section 4 we will see that $\hat{x}_-(\sigma)$ may not be a good approximation, so that we will propose an alternative approximate eigenvector for $-\hat{\lambda}$.

The following result holds for the associated residual.

PROPOSITION 3.2. *With the discussed notation, given an approximate eigenpair* $(\hat{\lambda}_\pm, \hat{x}_\pm(\sigma)/\|\hat{x}_\pm(\sigma)\|)$ *as in (3.6), the residual*

$$r_m := M\frac{\hat{x}_\pm(\sigma)}{\|\hat{x}_\pm(\sigma)\|} - N\frac{\hat{x}_\pm(\sigma)}{\|\hat{x}_\pm(\sigma)\|}\hat{\lambda}_\pm$$

is proportional to the vector $(M + \sigma N)v_{m+1}$ *and*

$$\|r_m\| = \|(M + \sigma N)v_{m+1}\| \, |h_{m+1,m}| \, |e_m^\top s| \frac{|\hat{\lambda}_\pm|}{\|\hat{x}_\pm(\sigma)\|}.$$

Moreover, r_m *is orthogonal to* $W_m(\sigma)$.

Proof. Using (3.3) and (3.4) in the definition of the residual, we readily obtain

$$r_m = [O_m, (M + \sigma N)v_{m+1}h_{m+1,m}e_m^\top]z(\sigma)\hat{\lambda}_\pm/\|\hat{x}_m(\sigma)\|, \tag{3.7}$$

from which the proportionality and the value of the norm follow. Moreover, using Proposition 2.2 i) and the N-neutrality of the basis, we have that $W_m(\sigma)^\top r_m = V_m^\top N^\top (M - \sigma N)^{-\top} r_m = V_m^\top N^\top (M + \sigma N)^{-1} r_m = 0$, and the orthogonality follows. ∎

Proposition 3.2 shows that all residuals of the approximate eigenpairs are collinear to the vector $(M + \sigma N)v_{m+1}$, as is typical of Galerkin-type approximations. On the other hand, the residual is not orthogonal to the whole approximation space generated by $[V_m, W_m(\sigma)]$, as it would be the case were it a true Galerkin approximation [3].

Phase III. Implicit restart. Once the search space reaches its maximal size, it is truncated. Since we maintain a standard Arnoldi factorization of K, we can use standard truncation procedures, e.g., Krylov-Schur-restarting [47, 48], without modification. This allows us to keep information on the converging eigenvectors in the approximation space, while discarding all remaining basis vectors.

These three phases form a cycle, and this cycle is repeated until enough eigenvalues have converged, or the maximum number of cycles is reached. More details on the actual implementation are given in the next section.

In exact arithmetic, the convergence of the approximate eigenvalues is the same as that of an implicitly-restarted Krylov subspace method on the matrix K [5].

4. Implementation considerations. The actual implementation requires the description of a few more details. Firstly, there are two ways to perform the extra orthogonalization against NV_m: i) In addition to V_m store an orthonormal basis of NV_m. In this case the (modified) Gram-Schmidt procedure is used for orthogonalization; ii) In addition to V_m store the matrix $Z_m = (NV_m)^\top (NV_m)$: the orthogonalization against V_m is standard and the orthogonalization against NV_m can be carried out using the projector $I - NV_m((NV_m)^\top NV_m)^{-1}(NV_m)^\top$ whose application only needs multiplications by V_m, Z_m^{-1} and N. The second alternative essentially halves the memory requirements of the method. On the other hand multiplications by N are required, and the procedure may be unstable, if N is such that NV_m is ill-conditioned. In our experiments we implemented the first approach.

Secondly, new spectral approximations are computed only after the maximum search space dimension has been reached. Moreover, to limit computational costs, the eigenvalue/eigenvector extraction discussed above is carried out on the truncated basis. More precisely, we first compute the Schur decomposition of the matrix H_m so as to have the relevant eigenvalues lying in the upper left corner of the quasi-triangular factor. We then truncate the Schur basis by only keeping the Schur vectors associated with these eigenvalues, compute the eigenpairs of the truncated matrix H_m, and then replace H_m with the reduced triangular factor.

One more implementation aspect that needs further discussion is the eigenvector extraction. Using (3.5) and the definition in (3.6) we have that

$$\begin{aligned}\hat{x}_\pm(\sigma) = [V_m, W_m(\sigma)]\hat{z}_m(\sigma) &\propto V_m s \frac{1}{\hat{\lambda}_\pm + \sigma} + W_m(\sigma)s \\ &= (M - \sigma N)^{-1}(M + \hat{\lambda}_\pm N)V_m s \frac{1}{\hat{\lambda}_\pm + \sigma}.\end{aligned}$$

If σ is extremely close to $\hat{\lambda}_+$, then $(M - \sigma N)^{-1}(M + \hat{\lambda}_- N) \approx I$ and $\hat{x}_-(\sigma) \propto V_m s$. Therefore, in general, the eigenvector approximation $\hat{x}_-(\sigma)$ cannot be accurate, since the columns of V_m alone do not span a close to deflating subspace of the pencil (M, N). As a result, if σ is very close to an approximate eigenvalue $\hat{\lambda}_+$, then the matched pair $(\hat{\lambda}_-, \hat{x}_-(\sigma))$ does not yield a small residual, in spite of a possibly good approximate eigenvalue. To derive an accurate approximate eigenvector for $\hat{\lambda}_-$, we exploit the fact that the columns of $[V_m, W_m(-\sigma)]$ also span an approximate invariant subspace of (M, N) (cf. Proposition 3.1 and subsequent discussion). Since in such a case the role of $\hat{\lambda}_-$ and $\hat{\lambda}_+$ is reversed, the approximation of $\hat{x}_-(-\sigma) = [V_m, W_m(-\sigma)]\hat{z}_m(-\sigma)$ will now be more accurate. In summary, the following two matched pairs are computed

$$\left(\hat{\lambda}_+, \frac{\hat{x}_+(\sigma)}{\|\hat{x}_+(\sigma)\|}\right), \quad \left(\hat{\lambda}_-, \frac{\hat{x}_-(-\sigma)}{\|\hat{x}_-(-\sigma)\|}\right), \tag{4.1}$$

(cf. also (3.5)-(3.6)) which will both tend to eigenpairs of (M, N) also for σ extremely close to one of the eigenvalues.

The resulting Algorithm is summarized as follows.

Algorithm Even-IRA.

Require: M symmetric, N skew-symmetric, shift $\sigma \in \mathbb{R}$ or $\sigma \in \imath\mathbb{R}$, unit start vector vector v_1, maximum search space dimension m_{max}, restart size m_{res} (m_{res} greater than or equal to the number of requested matched pairs)

1: $V \leftarrow [v_1]$

2: $m \leftarrow 0$

3: **while** cycle 1,2,3,... **do**
4: % Generation of the approximation space
5: **while** $m < m_{max}$ **do**
6: $m \leftarrow m + 1$
7: $v \leftarrow K v_m$
8: Orthogonalize v against V giving $H_{1:m,m}$
9: Orthogonalize v against NV
10: $h_{m+1,m} \leftarrow \|v\|$, $v_{m+1} \leftarrow v/h_{m+1,m}$
11: $V \leftarrow [V, v_{m+1}]$
12: **end while**
13: % Contraction of approximation space and matrix
14: $H_{1:m,1:m} \to QTQ^\top$ (real Schur form)
15: Partition $T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$, $Q = [Q_1, Q_2]$, with $T_{11} \in \mathbb{R}^{m_{res} \times m_{res}}$, $Q_1 \in \mathbb{R}^{n \times m_{res}}$
16: $V \leftarrow [V_m Q_1, v_{m+1}]$, $H \leftarrow \begin{bmatrix} T_{11} \\ h_{m+1,m} e_m^\top Q_1 \end{bmatrix}$, $m \leftarrow m_{res}$
17: % Eigenpair extraction
18: Compute eigenpairs (μ, s) of $H_{1:m,1:m}$
19: Compute approximate eigenpairs $(+\hat{\lambda}, \frac{\hat{x}_+(\sigma)}{\|\hat{x}_+(\sigma)\|})$ $(-\hat{\lambda}, \frac{\hat{x}_-(-\sigma)}{\|\hat{x}_-(-\sigma)\|})$ (cf. (4.1))
20: Check for convergence
21: **end while**

The space generation phase in the algorithm differs from the standard restarted Arnoldi method only in line 9, where N-neutrality of the basis is also enforced. A double sweep of modified Gram-Schmidt is employed to ensure orthogonality in lines 8 and 9. Moreover, for $\sigma \in \mathbb{R}$ the matrix-vector product Kv_m only requires a single real LU factorization $(M - \sigma N) = LU$, because $(M + \sigma N) = (M - \sigma N)^\top = U^\top L^\top$. For $\sigma \in \imath\mathbb{R}$, the matrix $(M - \sigma N)$ is Hermitian, and an LDL^* decomposition may be performed, thus requiring the same amount of memory as a real LU decomposition. The matrix $(M + \sigma N)$ can then be written as $(M + \sigma N) = (M - \sigma N)^\top = \overline{L} D^\top L^\top$.

REMARK 4.1. Refined projection methods ([3, Section 3.2]) could be used to compute eigenvector approximations in place of the procedure leading to (3.6). However serious problems may arise when seeking vectors associated with close eigenvalues, since multiple copies or close vectors may be obtained.

REMARK 4.2. As an alternative approximation strategy to our N-neutral space, one could require that the Krylov subspace basis is N-orthogonal. Upon permutation, the resulting representation matrix would have a convenient structure that would allow one to use structure preserving dense eigenvalue solvers [15]. However, we found the overall procedure to be unreliable, since such dense solvers require that the structure is preserved to full machine accuracy, and this cannot be ensured by the Arnoldi recurrence.

5. A related approach: The rational Krylov method. Eigenvector approximations are sought in the space generated by the reduced basis $[V_{m_{res}}, W_{m_{res}}(\sigma)]$. The question naturally arises whether one could directly build the search space generated by $[V_m, (M - \sigma N)^{-1} N V_m]$, bypassing the computation of the N-neutral search space. The construction of this subspace is particularly convenient, since the next vector v_{m+1} in the "V" basis may be obtained from the vector $w_m = (M - \sigma N)^{-1} N v_m$ by means of a single system solve, namely, $v_{m+1} = (M + \sigma N)^{-1} N w_m$. The search space is thus generated by alternate multiplications by $(M - \sigma N)^{-1} N$ or by $(M + \sigma N)^{-1} N$, yielding the basis

$$\begin{aligned}&\{v_1, (M-\sigma N)^{-1}Nv_1, (M+\sigma N)^{-1}N(M-\sigma N)^{-1}Nv_1, \dots,\\ &\quad (M-\sigma N)^{-1}N\big((M+\sigma N)^{-1}N(M-\sigma N)^{-1}N\big)^{m-1}v_1, \dots\}\end{aligned} \tag{5.1}$$

A closer look reveals that this is nothing but a particular form of the rational Krylov method [41], where the shifts σ and $-\sigma$ are used alternatingly, so that the rational function has only the two poles $\sigma, -\sigma$ with high multiplicity.

The following result specializes the general Arnoldi-type relation in [3, Section 8.5] to our multiple pole case. Its simple proof is omitted.

PROPOSITION 5.1. *Given a normalized vector v_1, let $[v_1, w_1, v_2, w_2, \dots, v_m, w_m]$ be the matrix obtained by the Gram-Schmidt process applied to the first $2m$ columns of the matrix in (5.1). Then the matrices $V_m = [v_1, \dots, v_m]$, $W_m = [w_1, \dots, w_m]$ satisfy the relations*

$$(M+\sigma N)^{-1}NW_m = V_{m+1}\underline{H} + W_mR, \quad \textit{and} \quad (M-\sigma N)^{-1}NV_m = W_mT + V_mS,$$

where the upper triangular matrices $R, S, T \in \mathbb{R}^{m\times m}$ and the Hessenberg matrix $\underline{H} \in \mathbb{R}^{(m+1)\times m}$ contain the orthogonalization coefficients from the Gram-Schmidt process. Moreover,

$$M[V_m, W_m]\begin{bmatrix} S & H \\ T & R\end{bmatrix} = N[V_m, W_m]\left(I - \sigma\begin{bmatrix} -S & H \\ -T & R\end{bmatrix}\right) \tag{5.2}$$

$$+[O_m, -(M+\sigma N)v_{m+1}h_{m+1,m}e_m^\top]. \tag{5.3}$$

Relation (5.2) suggests using the eigensolutions of the problem

$$\begin{bmatrix} S & H \\ T & R\end{bmatrix} z = \mu\left(I - \sigma\begin{bmatrix} -S & H \\ -T & R\end{bmatrix}\right) z. \tag{5.4}$$

to generate approximate eigenpairs of (M, N). Proposition 5.1 emphasizes the similarities between the rational Krylov procedure and our new algorithm. However, a few major differences make our new method particularly appealing for the considered eigenvalue problem: 1) The reduced eigenvalue problem in (3.4) generates paired eigenvalue approximations $\hat{\lambda}_+, \hat{\lambda}_-$, which approximate paired eigenvalues of (M, N), whereas the eigenvalues of (5.4) are in general not paired; 2) In the new method, one can chose between two ways to compute approximate eigenvectors (cf. (4.1)). In the rational Krylov method there is only one way to compute eigenvectors; 3) For $\sigma \in \imath\mathbb{R}$ the search space V_m in (3.1) is real, whereas the search space of the rational Krylov method is complex in general. The relevance of these properties will become apparent in the numerical experiments.

6. Applications. There are several major classes of applications in control theory that lead to eigenvalue problems of the form (1.1).

6.1. Linear quadratic optimal control. The classical linear quadratic optimal control problem for descriptor systems consists of minimizing the cost functional

$$\int_{t_0}^{t_1} x^\top Qx + 2u^\top Sx + u^\top Ru \; dt$$

(with $Q = Q^\top$ and $R = R^\top$) subject to the descriptor system

$$\begin{aligned} E\dot{x} &= Ax + Bu, \;\; x(0) = x^0 \\ y &= Cx, \end{aligned} \tag{6.1}$$

where $E, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, and where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the control input vector, and $y(t) \in \mathbb{R}^p$ contains measured outputs. The solution to this problem is important in the design of a feedback controller so that the closed loop system is *(asymptotically) stable*.

Under some further stabilizability and detectability conditions [34], a necessary condition for the existence of a stabilizing feedback controller is that there exists a costate function $\mu(t)$ with values in $\mathbb{C}^n$, such that $x(t), \mu(t), u(t)$ satisfy the boundary value problem:

$$\mathcal{L}(x,\mu,u) = \begin{bmatrix} 0 & E & 0 \\ -E^\top & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\mu}(t) \\ \dot{x}(t) \\ \dot{u}(t) \end{bmatrix} - \begin{bmatrix} 0 & A & B \\ A^\top & C^\top QC & C^\top S \\ B^\top & S^\top C & R \end{bmatrix} \begin{bmatrix} \mu(t) \\ x(t) \\ u(t) \end{bmatrix} = 0,$$

with boundary conditions $x(t_0) = x^0$, $E^\top \mu(t_1) = 0$.

This boundary value problem can be turned into two initial value problems by decoupling the forward and backward integration. In turn this may be performed by the computation of the deflating subspace associated with the eigenvalues in the left half plane of the matrix pencil

$$L(\lambda) = \lambda N - M = \lambda \begin{bmatrix} 0 & E & 0 \\ -E^\top & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & A & B \\ A^\top & C^\top QC & C^\top S \\ B^\top & S^\top C & R \end{bmatrix}. \tag{6.2}$$

If the pencil $L(\lambda) = \lambda N - M$ is regular and of index at most 1, and if there exists an $r = \operatorname{rank}(E)$-dimensional N-neutral deflating subspace associated with the open left half plane eigenvalues, then the optimal control can be directly obtained from this deflating subspace [34].

To check that the index of $L(\lambda)$ is at most 1 we need to compute a matrix T whose columns span the right null-space of N and then to check whether $T^\top MT$ is nonsingular. This can be done via (Cholesky-like or spectral) factorizations, see [6, 16].

To check the existence of the stabilizing controller, it is sufficient to determine the eigenvalues close to the imaginary axis. To compute the optimal controller, one would need to compute the full deflating subspace associated with the eigenvalues in the left half plane. However, in many applications it is sufficient to only approximate the subspace associated with the eigenvalues closest to the imaginary axis and after projection into this subspace to solve the smaller boundary value problem to compute a feedback control [46].

6.2. Optimal H_∞ control. In the optimal H_∞ control problem for descriptor systems one considers

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + B_1 w(t) \;\; + B_2 u(t), \qquad x(t_0) = x^0, \\ z(t) &= C_1 x(t) + D_{11} w(t) + D_{12} u(t), \\ y(t) &= C_2 x(t) + D_{21} w(t) + D_{22} u(t), \end{aligned} \tag{6.3}$$

with $E, A \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{n\times m_i}$, $C_i \in \mathbb{R}^{p_i\times n}$, and $D_{ij} \in \mathbb{R}^{p_i,m_j}$ for $i,j = 1,2$, where, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^{m_2}$ is the control input vector, $w(t) \in \mathbb{R}^{m_1}$ is an exogenous input, $y(t) \in \mathbb{R}^{p_2}$ contains measured outputs, and $z(t) \in \mathbb{R}^{p_1}$ is a regulated output or an estimation error. In such a system one tries to stabilize the system by a controller (dynamic compensator)

$$\begin{aligned} \hat{E}\dot{\hat{x}}(t) &= \hat{A}\hat{x}(t) + \hat{B}y(t), \\ u(t) &= \hat{C}\hat{x}(t) + \hat{D}y(t) \end{aligned}$$

with transfer function $K(s) = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B} + \hat{D}$ such that the closed-loop system that is given by

$$\begin{aligned} E\dot{x}(t) &= (A + B_2\hat{D}Z_1C_2)x(t) + (B_2Z_2\hat{C})\hat{x}(t) + (B_1 + B_2\hat{D}Z_1D_{21})w(t), \\ \hat{E}\dot{\hat{x}}(t) &= \hat{B}Z_1C_2x(t) + (\hat{A} + \hat{B}Z_1D_{22}\hat{C})\hat{x}(t) + \hat{B}Z_1D_{21}w(t), \\ z(t) &= (C_1 + D_{12}Z_2\hat{D}C_2)x(t) + D_{12}Z_2\hat{C}\hat{x}(t) + (D_{11} + D_{12}\hat{D}Z_1D_{21})w(t), \end{aligned}$$

with $Z_1 = (I_{p_2} - D_{22}\hat{D})^{-1}$ and $Z_2 = (I_{m_2} - \hat{D}D_{22})^{-1}$, is *internally stable*, and the closed-loop transfer function $T_{zw}(s)$ from w to z satisfies $T_{zw} \in \mathcal{H}_\infty^{p_1,m_1}$ and is minimized in the $\mathcal{H}_\infty$-norm. Here, the space $\mathcal{H}_\infty^{p,m}$ consists of all $\mathbb{C}^{p,m}$-valued functions that are analytic and bounded in the open complex right half plane. For $F \in \mathcal{H}_\infty^{p,m}$ the $\mathcal{H}_\infty$-norm is given by

$$\|F\|_\infty = \sup_{s\in\mathbb{C}^+} \sigma_{\max}(F(s)),$$

where $\sigma_{\max}(F(s))$ denotes the maximal singular value of the matrix $F(s)$.

This is a difficult non-convex optimization problem, hence one alternatively solves the *modified optimal $\mathcal{H}_\infty$ control problem* of determining $\gamma_{mo} = \inf \Gamma$, where Γ is the set of positive real numbers γ for which there exists an internally stabilizing dynamic controller of the form (6.2) so that the transfer function $T_{zw}(s)$ of the closed loop system satisfies $T_{zw} \in \mathcal{H}_\infty^{p_1,m_1}$ with $\|T_{zw}\|_\infty < \gamma$. Note that it is possible that there is no internally stabilizing dynamic controller, i. e. that $\Gamma = \emptyset$ and $\gamma_{mo} = \infty$. It is shown in [31] that for the solution of the modified optimal $\mathcal{H}_\infty$ control problem one has to form the two skew-symmetric/symmetric pencils

$$\lambda N_H + M_H(\gamma) = \left[\begin{array}{cc|ccc} 0 & -\lambda E^\top - A^\top & 0 & 0 & -C_1^\top \\ \lambda E - A & 0 & -B_1 & -B_2 & 0 \\ \hline 0 & -B_1^\top & -\gamma^2 I_{m_1} & 0 & -D_{11}^\top \\ 0 & -B_2^\top & 0 & 0 & -D_{12}^\top \\ -C_1 & 0 & -D_{11} & -D_{12} & -I_{p_1} \end{array}\right]$$

and

$$\lambda N_J + M_J(\gamma) = \left[\begin{array}{cc|ccc} 0 & -\lambda E - A & 0 & 0 & -B_1 \\ \lambda E^\top - A^\top & 0 & -C_1^\top & -C_2^\top & 0 \\ \hline 0 & -C_1 & -\gamma^2 I_{p_1} & 0 & -D_{11} \\ 0 & -C_2 & 0 & 0 & -D_{21} \\ -B_1^\top & 0 & -D_{11}^\top & -D_{21}^\top & -I_{m_1} \end{array}\right]$$

and to check the existence of deflating subspaces

$$X_H(\gamma) = \begin{bmatrix} X_{H,1}(\gamma) \\ X_{H,2}(\gamma) \\ X_{H,3}(\gamma) \\ X_{H,4}(\gamma) \\ X_{H,5}(\gamma) \end{bmatrix}, \quad X_J(\gamma) = \begin{bmatrix} X_{J,1}(\gamma) \\ X_{J,2}(\gamma) \\ X_{J,3}(\gamma) \\ X_{J,4}(\gamma) \\ X_{J,5}(\gamma) \end{bmatrix},$$

with

$$\begin{aligned} & X_{H,1}(\gamma),\ X_{H,2}(\gamma), X_{J,1}(\gamma),\ X_{J,2}(\gamma) \in \mathbb{R}^{n\times r}, X_{H,4}(\gamma) \in \mathbb{R}^{m_2\times r}, \\ & X_{J,4}(\gamma) \in \mathbb{R}^{p_2\times r}, \quad X_{H,3}(\gamma), X_{J,5}(\gamma) \in \mathbb{R}^{m_1\times r}, X_{H,5}(\gamma), X_{J,3}(\gamma) \in \mathbb{R}^{p_1\times r}, \end{aligned}$$

where $r = \operatorname{rank}(E)$ associated with the eigenvalues in the left half plane. If no such deflating subspace exists then the corresponding γ-value is smaller than the optimal value. This is the case if the index of one of the pencils $\lambda N_H + M_H(\gamma)$ or $\lambda N_J + M_J(\gamma)$ is larger than 1 or if there are purely imaginary eigenvalues. Again checking the index can be performed as in the previous section, while the existence of purely imaginary eigenvalues can be obtained via our new algorithm.

It should be noted that for the computation of the optimal γ further conditions need to be checked (see, e.g., [31]), and then explicit formulas for the optimal controllers can be determined [8].

6.3. Passivity checking and passivation. A third application of our new method arises in passivity checking and passivation. Consider a control system $E\dot{x} = Ax + Bu,\ x(0) = 0,\ y = Cx + Du$, and suppose that the homogeneous system is asymptotically stable and that D is square and nonsingular. Defining a real scalar valued *supply function* $s(u, y)$, the system is called *dissipative* if there exists a nonnegative scalar valued function Θ such that the *dissipation inequality*

$$\Theta(x(t_1)) - \Theta(x(t_0)) \le \int_{t_0}^{t_1} s(u(t), y(t))dt$$

holds for all $t_1 \ge t_0$, i. e. the system absorbs supply energy.

A dissipative system with the supply function $s(u, y) = \|u\|_2 - \|y\|_2$ is called *contractive* and with the supply function $s(u, y) = u^\top y + y^\top u$ it is called *passive*. It is well-known for systems with $E = I_n$ to be passive it is necessary that the skew-symmetric/symmetric pencil

$$\lambda N - M := \lambda \begin{bmatrix} 0 & E & 0 \\ -E^\top & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & A & B \\ A^\top & 0 & C \\ B^\top & C^\top & D + D^\top \end{bmatrix}$$

is regular, of index at most 1 and has no purely imaginary eigenvalues, see [1]. The proof of the corresponding results for general E is currently under investigation. The same problem arises as before and after the index check it is sufficient to determine the eigenvalues close to or on the imaginary axis.

If the system is not passive, then typically the system is made passive by determining a small perturbation $(\Delta E, \Delta A, \Delta B, \Delta C, \Delta D)$ that makes the system passive, see [17, 20, 24, 42, 44]. Again the main task is to determine the deflating subspace associated with the eigenvalues on or close to the imaginary axis and to carry out a sequence of perturbations so that all eigenvalues are moved off the imaginary axis and the index is at most one.

7. Numerical experiments. In this section we report on some of our numerical experience with the new method. We consider various examples with medium size matrices, most of which stem from benchmark application problems. We compare the new algorithm with its closest relatives, namely the Rational Krylov method outlined in section 5, and the matlab function `eigs` [33], which implements the general purpose ARPACK method for the approximation of eigenpairs of a given matrix or pencil [30]. Neither of these two algorithms is structure preserving, therefore matching pairs are approximated separately. We mention that `eigs` was called with the shift-and-invert matrix $(M-\sigma N)^{-1}N$, so that the largest eigenvalues in modulo of this matrix were in fact approximated. Therefore, whenever `eigs` was used, the approximate eigenvalues of the original problem were derived a posteriori, and then the corresponding residual checked.

To make our experiments fair, we also compared with our own implementation of the IRA method in [30], which computes the true residual. Unfortunately, results were not always consistent.

As an unstructured method, `eigs` either approximates both matching eigenvalues $\lambda, -\lambda$, if the mapped eigenvalues are both among the largest eigenvalues of $(M - \sigma N)^{-1}N$, or by means of two different runs of `eigs`, whenever the shift σ is significantly closer to either of λ and $-\lambda$. Such a behavior is a clear disadvantage of `eigs` in our context. A particularly poor performance is observed in the case of purely imaginary nonzero shifts, where both copies of matching eigenvalues are unnecessarily captured in all runs. A natural way out would be to set $\sigma = 0$, however in such a case clusters may not be easily identified, and in addition eigenvalues close to the imaginary axis, but not close to the origin will typically not be obtained. Therefore, our strategy for `eigs` is to chose the shift close to the expected target eigenvalues.

To make comparisons fair in terms of memory requirements, in case of real shifts we always use an approximation space of dimension m and $m/2$ for `eigs` and Even-IRA, respectively. (This is only fair if one stores NV_m in Even-IRA, otherwise Even-IRA should have the same search space dimension as `eigs` when a real shift is used.) For this reason, the two methods have the same number of solves per cycle, so that their main computational costs are comparable.

For a given problem, before starting the approximation, balancing of the matrices was carried out. For both the Rational Krylov method and Even_IRA, a preprocessing of one step of inverse iteration with the matrix $(M+\sigma N)^{-1}N$ is carried out, to mitigate the possible influence of the null space of N; see also [4].

Our stopping criterion is based on the true absolute residual of the original problem. A more robust criterion would require the use of a relative residual, in which the denominator contains the absolute value of the eigenvalue condition number, $|y^*x|$, where x, y are right and left eigenvector of the given eigenvalue. Because of the special structure of the problem, the left eigenvector coincides with the right eigenvector of the matching eigenvalue, and thus it is available in our case. This shows that the new algorithm provides additional spectral information for free, as opposed to standard procedures. Since the other algorithms do not produce approximations to the matching eigenvalue, we have decided to use only the absolute residual as stopping criterion for our performance comparison.

We also recall that the stopping test is performed only on the remaining eigenpairs after truncation of the Schur decomposition of H_m. In particular, this implies that a first decision on which eigenpairs should be retained is performed on H_m, i.e. on the transformed problem K. As already mentioned, good eigenpairs of K might not correspond to accurate eigenpairs of (M, N), therefore the computation of the original residuals is recommended. Using (3.1) and (3.7), the following explicit relation between the original and transformed residuals holds for each pair $(\hat{\lambda}, \hat{x}_+(\sigma)/\|\hat{x}_+(\sigma)\|)$:

$$r_{m,+} = \frac{1}{\|\hat{x}_+(\sigma)\|}(M + \sigma N)\mathbf{r}_m,$$

where $\mathbf{r}_m = KV_m s - V_m s\lambda_+$. A similar relation holds for $(-\hat{\lambda}, \hat{x}_-(-\sigma)/\|\hat{x}_-(-\sigma)\|)$: $r_{m,-} = \frac{1}{\|\hat{x}_-(-\sigma)\|}(M - \sigma N)\mathbf{r}_m$ with $\mathbf{r}_m = KV_m s - V_m s\lambda_-$ (cf. Proposition 3.1 with $\alpha = -\sigma$).

EXAMPLE 7.1. We consider the matrix $A \in \mathbb{R}^{6400\times 6400}$ stemming from the finite difference discretization of the operator $\mathcal{L}(u) = -\Delta u + 10(u_x + u_y)$ on the unit square $[0, 1]^2$ with Dirichlet homogeneous boundary conditions. We define $M = (A + A^\top)/2$ and $N = (A - A^\top)/2$, the symmetric and skew-symmetric parts of A, respectively. We are thus interested in analyzing whether the symmetric part "dominates" the skew-symmetric part of the matrix, by checking whether all eigenvalues of the pencil (M, N) are greater than one in modulo. This is of interest in the analysis of linear system solvers and of certain structured preconditioners; see, e.g., [45, 13]. We thus consider $\sigma = 1$ (a completely analogous scenario results from taking $\sigma = 10^{-3}$) and approximate the first matching 50 eigenpairs, that is, 25 eigenvalues in the right half complex plane, with a tolerance of tol=10^{-10} and $m = 80$ (we recall here that Even-IRA then uses a maximum subspace dimension equal to $m/2 = 40$). All sought after eigenvalues have zero real part and most of them are double. The left plot of Figure 7.1 shows that the new method ('$\times$' symbols) is able to capture all wanted matched eigenpairs; the algorithm required 6 restarts. On the other hand, although requiring only 3 restarts, `eigs` fails to compute all 25 eigenvalues because the method captures some of the "matching" imaginary eigenvalues. The right plot shows the approximate eigenvalues for $\sigma = \imath$. In this case, `eigs` captures most, but not all, eigenvalues with positive imaginary part.

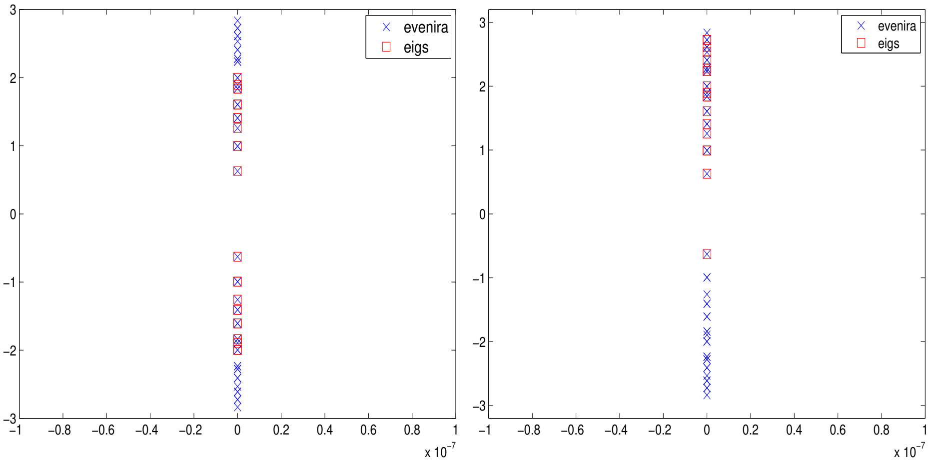

FIG. 7.1. *Example 7.1. Eigenvalues of (M, N) closest to $\sigma = 1$ (left) and to $\sigma = \imath$ (right) approximated with* `eigs` *and with Even-IRA*

.

EXAMPLE 7.2. We consider the benchmark problem `rail_1357` from the Oberwolfach Model Reduction Benchmark Collection [18], describing a semi-discretized

Table 7.1

Example 7.2. Sought after eigenvalues and Even-IRA residual norms, for purely imaginary and real shift.

eigenvalue	residual norm $\sigma = i10^{-5}$	residual norm $\sigma = 10^{-5}$
2.7062e-05	1.1674e-17	3.0662e-17
-2.7062e-05	7.7496e-18	6.7802e-18
8.8841e-05	6.0929e-17	6.2008e-17
-8.8841e-05	6.9922e-17	4.6029e-17
2.2710e-04	5.9494e-16	1.4799e-14
-2.2710e-04	3.2735e-15	6.9790e-14

heat transfer problem for the optimal cooling of steel profiles. The descriptor system is as in (6.1), and $R = 0 = Q$, $S = I$. The matrices A and Q have both size 1357 while R has size 7, yielding matrices M and N of size 2721 each. We look for the four matching eigenvalues closest to the shift $\sigma = i10^{-5}$. The eigenvalues of the pencil (M, N) closest to σ are $\pm$ 2.7062e-05 , $\pm$ 8.8841e-05 , $\pm$ 2.2710e-04 , $\pm$(-1.3161e-04 + 2.1279e-04i), $\pm$(1.3161e-04 + 2.1279e-04i). The estimated condition number (Matlab `condest`) of $M - \sigma N$ is $5.5\,10^{12}$.

For $m = 20$, after 3 restarts, Even-IRA with purely imaginary shift finds the requested (real) eigenvalues as shown in the second column of Table 7.1. With imaginary shift, both the Rational Krylov method and `eigs` are unable to find the third closest matched pair within 100 cycles, and a stopping tolerance of 10^{-12}. After 100 cycles, our implementation of IRA gives a residual norm for $\lambda = \pm 2.2710e-04$ of the order of 10^{-8} and 10^{-5}, thus highly above the requested tolerance. Similar figures are found with the Rational Krylov method. We stress that for purely imaginary shifts, all methods generate subspaces of maximum size 10 because all methods except Even-IRA need to use complex arithmetic.

For real shift $\sigma = 10^{-5}$, all methods stop before the maximum number of cycles is reached. Even-IRA takes 3 restarts to obtain accurate eigenpairs (cf. Table 7.1), whereas the Rational Krylov method takes 2 cycles (a subspace of dimension 20 is generated). On the other hand, although `eigs` stops after 1 cycle, corresponding to a subspace of size 20, the obtained eigenvalues are 2.7062e-05 (residual norm 8.9218e-13), -2.7063e-05 (residual norm 2.2999e-12) and 8.8842e-05 (residual norm 1.6850e-11). These digits shows that `eigs` detects the paired eigenvalue $\pm$ 2.7062e-05 instead of only the one with positive sign. On the other hand, it does not deliver the third positive eigenvalue closest to the shift. Note also that the residual norms are in two cases larger than the requested tolerance. Our implementation of IRA computes all requested eigenvalues at high accuracy, using 3 cycles.

Example 7.3. our next example stems from the model reduction analysis of a descriptor system associated with Navier-Stokes equations [40, 43] (cf. section 6.3). Here A has size 1159 and D is the zero matrix of dimension 5.

We consider detecting the eigenvalues closest to $\sigma = 10i$, namely the real matching eigenvalues $\pm$10.544. We use a convergence tolerance equal to 10^{-12} and $m = 20$, so that all methods generate a subspace of maximum size 10. Even-IRA converges to the wanted eigenvalue in 2 cycles, with residuals 3.6595e-13, 3.6885e-14 for the positive and negative eigenvalues. After 5 cycles, the call to `eigs` yields the following two eigenvalues and associated residual norms: $\hat{\lambda}_1$= 1.0544e+01 + 2.2318e-10i (1.0013e-

13), $\hat{\lambda}_2$ = -1.0544e+01 + 2.5349e-12i (5.7810e-14). Although the imaginary parts of both eigenvalues are small, these are above the residual norm, therefore one would be lead to think that these are two distinct eigenvalues, and that the method did not capture the eigenvalues on both sides of the complex plane. Completely analogous results are obtained with our implementation of IRA, and with the Rational Krylov method.

EXAMPLE 7.4. Finally, we perform a pure passivity test. We consider the problem described in [44, section 5], called `coax1`. The problem structure resembles that in section 6.3, and the pencil is given by

$$M - \lambda N = \begin{bmatrix} 0 & A & B & 0 \\ A^\top & 0 & 0 & C^\top \\ B^\top & 0 & -I & D^\top \\ 0 & C & D & -I \end{bmatrix} - \lambda \begin{bmatrix} 0 & I & 0 & 0 \\ -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

In this context, we are merely interested in detecting whether there are purely imaginary eigenvalues, therefore we will be content with very inaccurate eigenvalues. Even-IRA is particularly appropriate for this task, as it can readily detect imaginary eigenvalues. We seek imaginary eigenvalues around $\sigma = 6i$. After a rather cheap cycle with $m = 12$, Even-IRA finds the eigenvalues $\pm 6.0377i$, $\pm 6.0681i$ with a residual norm below 10^{-5} (both eigenvalue signs are delivered). After a single cycle with $m = 20$, both imaginary pairs are captured, with residual norms below 10^{-8}. On the other hand, our implementation of IRA determines the following spectral information

m (tol)	eigenvalue	residual norm	n. cycles
12 (10^{-5})	9.5567e-09 + 6.0377 i	2.3684e-08	2
	-1.5054e-06 + 6.0681 i	3.9961e-07	2
20 (10^{-8})	3.2863e-13 + 6.0377 i	3.8536e-16	2
	-4.8055e-12 + 6.0681 i	3.9425e-13	2

Similar results were obtained with `eigs`, although residuals were even less reliable, due to their a-posteriori computation. Note that the real part of the second eigenvalue is larger than the residual norm, therefore one is unable to identify such as eigenvalue as purely imaginary. Even larger real parts were delivered by `eigs`. The situation does not improve by allowing the approximation of more eigenvalues. Similar disappointing figures were obtained with the Rational Krylov method.

EXAMPLE 7.5. We conclude with an example of an artificially constructed system, where our new method does not perform so well. We consider the Benchmark problem Carex-16 [11, Example 16], of total size 2400. The problem is extended with two ill-conditioned extra matching eigenvalues ±20 i.

We then set $\sigma = 22$ i and look for one of these extra eigenvalues. All other relevant eigenvalues are real. In addition to the sought after eigenvalues, soon after the first few iterations Even-IRA spots a rough approximation to a matched pair at about 19.9998i. After that stagnation occurs, showing no recovery from the approximation of an apparently *ghost* matched pair at ±19.9998 i. This ghost eigenvalue does not emerge during the approximation process of the other methods. Such an unwelcome event may be related to the sensitivity of the problem, which is exacerbated by the squaring K^2 performed in our method. This behavior deserves further analysis, since a sensitivity analysis of similar methods using K^2 cannot be found in the relevant literature.

8. Conclusions. We proposed a new iterative method for the approximation of matched eigenvalues and corresponding eigenvectors of symmetric skew-symmetric pencils (M, N), where N can be highly singular. The method is an improvement and generalization of a method proposed in [36], and can efficiently handle large problems whose size allows for sparse system solves. Typical examples from benchmark problems are proposed to show the usefulness of the new approach for certain crucial tasks in control applications.

Aknowledgments. We thank Tatjana Stykel for providing us with the data of Example 7.3. This work was mostly performed while the third author was visiting the TU Berlin.

REFERENCES

[1] T. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM Publications, Philadelphia, PA, USA, 2005.

[2] T. Apel, V. Mehrmann, and D. Watkins. Structured eigenvalue methods for the computation of corner singularities in 3D anisotropic elastic structures. *Comp. Meth. Appl. Mech. Eng.*, 191:4459–4473, 2002.

[3] A. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the Solution of Algebraic Eigenvalue Problems*. SIAM, Philadelphia, 2000.

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

[5] C. A. Beattie, M. Embree, and D. C. Sorensen. Convergence of polynomial restart Krylov methods for eigenvalue computation. *SIAM Review*, 47:492–515, 2005.

[6] P. Benner, R. Byers, H. Fassbender, V. Mehrmann, and D. Watkins. Cholesky-like factorizations of skew-symmetric matrices. *ETNA*, 11:85–93, 2000.

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

[8] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Robust method for robust control. Preprint 2004-06, Institut für Mathematik, TU Berlin, D-10623 Berlin, FRG, 2004. `url: http://www.math.tu-berlin.de/preprints/`.

[9] P. Benner, R. Byers, V. Mehrmann, and H. Xu. A robust numerical method for the γ-iteration in H_∞-control. *Linear Algebra Appl.*, 425:548–570, 2007.

[10] P. Benner and D. Kressner. Algorithm 854: Fortran 77 subroutines for computing the eigenvalues of Hamiltonian matrices II. *ACM Trans. Math. Software*, 32:352–373, 2006.

[11] P. Benner, A.J. Laub, and V. Mehrmann. A collection of benchmark examples for the numerical solution of algebraic Riccati equations I: Continuous-time case. Technical Report SPC 95_22, Fakultät für Mathematik, TU Chemnitz–Zwickau, 09107 Chemnitz, FRG, 1995. Available from `http://www.tu-chemnitz.de/sfb393/spc95pr.html`.

[12] P. Benner, V. Mehrmann, and H. Xu. A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numer. Math.*, 78(3):329–358, 1998.

[13] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[14] T. Brüll and V. Mehrmann. STCSSP: A FORTRAN 77 routine to compute a structured staircase form for a (skew-)symmetric/(skew-)symmetric pencil. Preprint 2008-31, Institut für Mathematik, TU Berlin, D-10623 Berlin, FRG, 2008. `url: http://www.math.tu-berlin.de/preprints/`.

[15] A. Bunse-Gerstner, R. Byers, and V. Mehrmann. A chart of numerical methods for structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 13:419–453, 1992.

[16] R. Byers, V. Mehrmann, and H. Xu. A structured staircase algorithm for skew-symmetric/symmetric pencils. *Electr. Trans. Num. Anal.*, 26:1–33, 2007.

[17] C.P. Coelho, J.R. Phillips, and L.M. Silveira. Robust rational function approximation algorithm for model generation. In *Proceedings of the 36th DAC*, pages 207–212, New Orleans, Louisiana, USA, 1999.

[18] Benchmark Collection. Oberwolfach model reduction benchmark collection, 2003. `http://www.imtek.de/simulation/benchmark`.

[19] R.W. Freund. On Pade-type model order reduction for j-hermitian linear dynamical systems. *Linear Algebra Appl.*, 429:2451–2464, 2008.
[20] R.W. Freund and F. Jarre. An extension of the positive real lemma to descriptor systems. *Optimization Methods and Software*, 19:69–87, 2004.
[21] R.W. Freund, F. Jarre, and C. Vogelbusch. Nonlinear semidefinite programming: Sensitivity, convergence and an application in passive reduced-order modeling. *Math. Programming*, 109:581–611, 2007.
[22] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.
[23] E.J. Grimme, D.C. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numer. Algorithms*, 12:1–31, 1996.
[24] S. Grivet-Talocia. Passivity enforcement via perturbation of Hamiltonian matrices. *IEEE Trans. Circuits and Systems*, 51:1755–1769, 2004.
[25] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM Publications, Philadelphia, PA, 1996.
[26] T.-M. Hwang, W.-W. Lin, and V. Mehrmann. Numerical solution of quadratic eigenvalue problems with structure-preserving methods. *SIAM J. Sci. Comput.*, 24:1283–1302, 2003.
[27] D. Kressner, C. Schröder, and D.S. Watkins. Implicit QR algorithms for palindromic and even eigenvalue problems. Technical Report 432, DFG Research Center Matheon, *Mathematics for key technologies* in Berlin, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2008. `url: http://www.matheon.de/` , to appear in Numer. Alg.
[28] P. Lancaster and L. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Review*, 47:407–443, 2005.
[29] P. Lancaster and L. Rodman. Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence. *Linear Algebra Appl.*, 406:1–76, 2005.
[30] R. B. Lehoucq, D. C. Sorensen, and C. Yang. *ARPACK Users Guide: Solution of large scale Eigenvalue Problems by Implicitly Restarted Arnoldi Methods*. SIAM, Philadelphia, 1998.
[31] P. Losse, V. Mehrmann, L.K. Poppe, and T. Reis. The modified optimal $\mathcal{H}_\infty$ control problem for descriptor systems. *SIAM J. Cont.*, 2795–2811:47, 2008.
[32] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Structured polynomial eigenvalue problems: Good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28:1029–1051, 2006.
[33] The MathWorks, Inc. *MATLAB 7.5*, September 2007.
[34] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*. Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.
[35] V. Mehrmann and T. Stykel. Balanced truncation model reduction for large-scale systems in descriptor form. In P. Benner, V. Mehrmann, and D. Sorensen, editors, *Dimension Reduction of Large-Scale Systems*, volume 45 of *Lecture Notes in Computational Science and Engineering*, pages 83–115. Springer-Verlag, Berlin, Heidelberg, 2005.
[36] V. Mehrmann and D. Watkins. Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltoninan/Hamiltonian pencils. *SIAM J. Sci. Comput.*, 22:1905–1925, 2001.
[37] V. Mehrmann and D. Watkins. Polynomial eigenvalue problems with Hamiltonian structure. *Electr. Trans. Num. Anal.*, 13:106–113, 2002.
[38] V. Mehrmann and H. Xu. Numerical methods in control. *J. Comput. Appl. Math.*, 123:371–394, 2000.
[39] T. Reis and T. Stykel. Passivity preserving balanced truncation for electric circuits. Preprint 32, Inst. f. Mathematik, TU Berlin, TU Berlin, Straße des 17. Juni 136, 10623 Berlin, FRG, 2008. `http://www.math.tu-berlin.de/`.
[40] T. Reis and T. Stykel. Positive real and bounded real balancing for model reduction of descriptor systems. Preprint 2008/25, Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, 2008. To appear in Int. J. Control.
[41] A. Ruhe. Rational Krylov sequence methods for eigenvalue computation. *Lin. Alg. Appl.*, 58:391–405, 1984.
[42] D. Saraswat, R. Achar, and M. Nakhia. On passivity check and compensation of macromodels from tabulated data. In *7th IEEE Workshop on Signal Propagation on Interconnects*, pages 25–28, Siena, Italy, 2003.
[43] M. Schmidt. *Systematic Discretization of Input/Output Maps and other Contributions to the Control of Distributed Parameter Systems*. Ph.D. thesis, Institut für Mathematik, Technische Universität Berlin, Berlin, Germany, 2007.
[44] C. Schröder and T. Stykel. Passivation of LTI systems. Technical Report 368, DFG Research

Center Matheon, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, 2007.

[45] V. Simoncini and D. B. Szyld. New conditions for non-stagnation of minimal residual methods. *Numer. Math.*, 109(3):477–487, 2008.

[46] D. Sorensen. Passivity preserving model reduction via interpolation of spectral zeros. *Systems and Control Letters*, 54:347–360, 2005.

[47] G.W. Stewart. A Krylov-Schur algorithm for large eigenproblems. *SIAM J. Matrix Anal. Appl.*, 23:601–614, 2001.

[48] G.W. Stewart. Addendum to 'a Krylov-Schur algorithm for large eigenproblems'. *SIAM J. Matrix Anal. Appl.*, 24:599–601, 2002.

[49] R. C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.

[50] P. Van Dooren. A generalized eigenvalue approach for solving Riccati equations. *SIAM J. Sci. Statist. Comput.*, 2:121–135, 1981.

[51] D.S. Watkins. On Hamiltonian and symplectic Lanczos processes. *Linear Algebra Appl.*, 385:23–45, 2004.

[52] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.