THE GENERALIZED NEWTON ITERATION FOR THE MATRIX SIGN FUNCTION*

XIAOBAI SUN† AND ENRIQUE S. QUINTANA-ORTÍ‡

Abstract. In this paper we present modified algorithms for computing deflating subspaces of matrix pencils by means of the matrix sign function. Specifically, our new algorithms reduce the number of iterations to half, cut the cost of each Newton iteration by more than 50%, and improve the accuracy of the computed deflating subspaces. The matrix sign function is thus revealed as an effective technique for applications where a part of the spectrum has to be identified or only the deflating subspaces are required. When the complete spectrum is desired, the matrix sign function can be used as an initial divide-and-conquer technique. The high performance of the basic kernels involved in this iteration are also specially appropriate for current parallel architectures.

Key words. matrix Sign function, spectral decomposition, nonsymmetric eigenvalue problem, invariant subspaces.

AMS subject classifications. 65F15, 47A75.

1. Introduction

In the last years the matrix sign function has received a considerable attention as a divide and conquer method for obtaining complete or partial information about the eigenstructure of a matrix pencil (see, among others, [4,5,6,15]). Given a matrix pair (A, B), A, $B \in \mathbb{C}^{n\times n}$, the matrix sign function can be used to obtain two unitary matrices, Q, $Z \in \mathbb{C}^{n\times n}$, such that

$$
\begin{aligned}
Q^{\mathrm{H}} A Z &= \begin{pmatrix} Q_1^{\mathrm{H}} \\ Q_2^{\mathrm{H}} \end{pmatrix} A\,(Z_1, Z_2) = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \\
Q^{\mathrm{H}} B Z &= \begin{pmatrix} Q_1^{\mathrm{H}} \\ Q_2^{\mathrm{H}} \end{pmatrix} B\,(Z_1, Z_2) = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}.
\end{aligned} \tag{1}
$$

Let $\lambda(A, B)$ be the set of generalized eigenvalues of the matrix pencil $A - \lambda B$ (hereafter, when it is clear from the context, we will shorten generalized eigenvalue/spectrum as eigenvalue/spectrum). Then, $\lambda(A_{22}, B_{22}) = \lambda(A, B) - \lambda(A_{11}, B_{11})$ and the spectrum of (A, B) is decoupled into the disjoint spectrums of two smaller subproblems, (A_{11}, B_{11}) and (A_{22}, B_{22}). A subspace $\mathcal{X}$ is a right-deflating subspace of (A, B) if $\dim(A\mathcal{X} + B\mathcal{X}) \leq \dim(\mathcal{X})$, where the equality holds if (A, B) is regular [18]. A left-deflating subspace of (A, B) is a right-deflating subspace of $(A^{\mathrm{H}}, B^{\mathrm{H}})$. In (1), Q_2^{H} and Z_1 are orthonormal basis for the left and right-deflating subspaces, respectively, associated with the eigenvalues of (A_{11}, B_{11}).

By recursively applying the spectral decoupling in (1), we can identify a specific part of the spectrum of (A, B). In case the complete spectrum is desired, the matrix sign function can be used as an initial spectral divide-and-conquer technique. On the other hand, several applications in linear control systems design only require an orthonormal base for a certain deflating subspace [2,7,10,14, 17,24].

The matrix sign function was first defined by Roberts [24] as

$$
\operatorname{sign}(A) \stackrel{\text{def}}{=} Y \begin{pmatrix} -I_r & 0 \\ 0 & I_{n-r} \end{pmatrix} Y^{-1}, \tag{2}
$$

*X. Sun was supported in part by the Advanced Research Projects Agency, under contract P-95006. E. S. Quintana received support from the CICYT project TIC 96-1062-C03-03.

†Department of Computer Science, Duke University, Durham, NC 27708-0129, xiaobai@cs.duke.edu.

‡Departamento de Informática, Universidad Jaime I, 12080 Castellón, Spain, quintana@inf.uji.es.

where A has exactly r eigenvalues in the left complex plane, I_r is the identity matrix of order r, and Y is the matrix of eigenvectors of A. If A has an eigenvalue on the imaginary axis then $\mathrm{sign}(A)$ is undefined.

A practical procedure for computing $\mathrm{sign}(A)$, assuming A is nonsingular, is to use the Newton iteration for the equation $X^2 - I_n = 0$, with A as the initial iterate. The following classical Newton iteration for the matrix sign function is thus obtained

$$A_{k+1} = (A_k + A_k^{-1})/2, \quad A_0 = A. \tag{3}$$

The iteration converges if A has no eigenvalue on the imaginary axis, and the asymptotic convergence rate is quadratic [24].

The definition of the matrix sign function and practical procedures for its computation can also be extended for matrix pairs. Gardiner and Laub proposed in [17] the generalized Newton iteration

$$A_{k+1} = (A_k + BA_k^{-1}B)/2, \quad A_0 = A. \tag{4}$$

Under certain conditions, this iteration converges to $B\,\mathrm{sign}(B^{-1}A) = \mathrm{sign}(AB^{-1})B$. There are inverse-free iteration schemes for the matrix sign function based on polynomial approximations to the inverse [20]. Such schemes, however, have bounded convergence regions. Other inverse-free iterative schemes overcome these problems, though at the expense of a much higher computational cost [6,8].

The generalized Newton iteration is attractive in that unlike QZ algorithm [22], the computation of spectral subspaces does not depend on the computation of the eigenvalues and the reordering of the computed eigenvalues, and the algorithmic building blocks are mainly triangular factorizations and matrix multiplications. High performance software for such matrix computations is available on a broad class of high performance computers and, especially, on parallel computers [1,13,23,29].

In this paper we describe several modifications to the generalized Newton iteration and the subspace extraction stage that allow a reduction to 25% in the computational cost and a remarkable improvement in the numerical accuracy.

First, in order to obtain the basis for the left and right-deflating subspaces, in [3] it is proposed to perform two independent generalized Newton iterations, with initial matrix pairs (A, B) and $(A^{\mathrm{H}}, B^{\mathrm{H}})$, followed by a subspace extraction procedure which involves the converged matrices. Following our results in [28], we show that a single iteration provides basis for both the left and right-deflating subspaces. Thus, we reduce the number of iterations of the approach in [3] by half.

Secondly, we propose an iteration on an equivalent matrix pair, where one of the matrices is reduced to a very simple form (bidiagonal form). This reduction is carried out before the iteration is started and cuts the computational cost of each iteration by half.

Finally, we describe a new subspace extraction scheme that provides more accurate deflating subspaces. Our experimental results confirm the improvements over the traditional generalized Newton iteration, both from the point of view of numerical accuracy and performance. These contributions report the matrix sign function as a wide-appeal numerical tool for the spectral division problem.

We use the following notation and conventions in addition to what is introduced so far. The matrix norm $\|\cdot\|$ denotes the 2-norm. We denote by $\mathcal{R}(X)$ and $\mathcal{N}(X)$ the range and the null space of matrix X, respectively. In some cases it is necessary to write the generalized eigenvalue problem in the symmetrical cross-product form, as suggested in [26]. If $(\alpha, \beta) \neq 0$ and $\det(\beta\, A - \alpha\, B) = 0$, then (α, β) represents an eigenvalue of (A, B) in the sense that, for any $\delta \neq 0$, we have $\det(\delta\beta\, A - \delta\alpha\, B) = 0$. The corresponding eigenvalue-eigenvector problem is

$$\beta\, A\, x = \alpha\, B\, x, \quad (\alpha, \beta) \neq 0,\, x \neq 0.$$

A matrix pencil $\beta A - \alpha B$, or the matrix pair (A, B), is singular if $\det(\beta\, A - \alpha\, B) \equiv 0$, for all (α, β); it is regular otherwise. In other cases we use for convenience the traditional, asymmetrical forms

$Ax = \lambda Bx$, or, $\mu Ax = Bx$. When $\beta \neq 0$, we simply set $\beta = 1$ and hence $\alpha = \lambda$. When $\beta = 0$, we set $\alpha = 1$.

2. The Generalized Newton Iteration

In this section we provide an analysis on the convergence of the generalized Newton iteration, and describe the algebraic structure of the converged matrix, including the case in which matrix B is singular. Note that the generalized Newton iteration treats A and B asymmetrically.

We first introduce the following Proposition, the proof of which is straightforward and therefore skipped.

PROPOSITION 2.1. *Let $(F, G) = (YAX^{-1}, YBX^{-1})$ be an equivalent matrix pair of (A, B), where X and Y are nonsingular matrices. Then, the generalized Newton iteration with (the initial matrix pair) (F, G) converges if and only if the iteration with (A, B) converges. Assume that the iteration with (A, B) converges. Let $A_\infty \stackrel{\mathrm{def}}{=} \lim_{k\to\infty} A_k$ with (A, B) and $F_\infty \stackrel{\mathrm{def}}{=} \lim_{k\to\infty} F_k$ with (F, G). Then, $A_\infty = Y^{-1}F_\infty X$.* If (F, G) is equivalent to (A, B), we say, for simplicity, the iteration with (F, G) is equivalent to the iteration with (A, B). We say the iteration with (A, B) converges if any of its equivalent iterations converges.

2.1. Case: both A and B are nonsingular

If A and B are nonsingular, then (A, B) is equivalent to (AB^{-1}, I_n), and $(B^{-1}A, I_n)$. The generalized Newton iteration (4) with (A, B) is therefore equivalent to the Newton iteration (3) with the single matrix $A_0 = B^{-1}A$ or $A_0 = AB^{-1}$. It is known that iteration (3) converges if and only if there is no eigenvalue of A_0 on the imaginary axis. Moreover, the matrix at convergence is $B\,\mathrm{sign}(A_0)$ and the asymptotic convergence rate is quadratic. We assume for the rest of this subsection that there are no eigenvalues of (A, B) on the imaginary axis. By Proposition 2.1, the generalized Newton iteration in (4) with (A, B) converges. Denote by A_∞ the matrix at converge. Then,

$$A_\infty = B\,\mathrm{sign}(B^{-1}A) = \mathrm{sign}(AB^{-1})B,$$

and

$$B \pm A_\infty = B(I \pm \mathrm{sign}(B^{-1}A) = (I \pm \mathrm{sign}(AB^{-1}))B.$$

The matrices $I + B^{-1}A_\infty$ and $I - B^{-1}A_\infty$ are the projectors for the right-deflating subspaces of (A, B) corresponding, respectively, to the eigenvalues in the right complex plan, $\mathbb{C}^+$, and left complex plane, $\mathbb{C}^-$. The matrices $I + A_\infty B^{-1}$ and $I - A_\infty B^{-1}$ are the projectors for the left-deflating subspaces corresponding to the eigenvalues in $\mathbb{C}^+$ and $\mathbb{C}^-$, respectively. Given a projector P for a subspace, since $\mathcal{R}(P)$ is the subspace, one can find a basis for $\mathcal{R}(P)$, using, for example, rank-revealing techniques [12, 16,19].

To avoid the inverse of B and the subsequent matrix multiplication with A_∞ and B^{-1} in forming the above projectors, one can extract the deflating subspaces using the following relationships between the matrix pair at convergence (A_∞, B) and the projectors.

$$\begin{aligned} &\mathcal{N}(B \pm A_\infty) = \mathcal{N}(I \pm \mathrm{sign}(B^{-1}A)) = \mathcal{R}(I \mp \mathrm{sign}(B^{-1}A)), \\ &\mathcal{N}(B^{\mathrm{H}} \pm A_\infty^{\mathrm{H}}) = \mathcal{N}(I \pm \mathrm{sign}(AB^{-1})^{\mathrm{H}}) = \mathcal{R}(I \mp \mathrm{sign}(AB^{-1})^{\mathrm{H}}). \end{aligned} \tag{5}$$

The right-deflating subspace of (A, B) corresponding to the eigenvalues in $\mathbb{C}^-$, for example, can be obtained by finding the null space of $(B + A_\infty)$. If no further information about the null space is known, rank-revealing algorithms are the methods of choice. For example, let $X^{\mathrm{H}}\Pi =$

$(Q_1, Q_2)\begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix}$ be a QR factorization of X^{H} with column pivoting, where R_{11} is nonsingular; then Q_2 is an orthonormal basis for $\mathcal{N}(X)$.

Equation (5) also shows that the basis for both the left and right-deflating subspaces can be obtained from a single iteration starting with (A, B). To avoid inconsistences in subspace extraction, we adapt a technique proposed in [28] for a related matrix function (see Fig. 1). In the following sections we elaborate and improve the technique.

Input : The converged matrix pair (A_∞, B) from (4).
A tolerance threshold for the numerical rank τ_{rank}.
Output: Unitary basis Q and Z for the left and right-deflating subspaces.
1. Compute $((B + A_\infty)^{\mathrm{H}})\Pi_Z = ZR_Z$ (rank-revealing QR factorization)
2. Let $r = \mathrm{rank}(R_Z, \tau_{rank})$,
and $Z = (Z_1, Z_2)$, $Z_1 \in \mathbb{C}^{n\times r}$ and $Z_2 \in \mathbb{C}^{n\times(n-r)}$
3. Compute $(AZ_2, BZ_2)\Pi_Q = QR_Q$ (rank-revealing QR factorization)
4. Set $Z = (Z_2, Z_1)$

Fig. 1: Subspace extraction of the left and right-deflating subspaces from a single generalized Newton iteration

To the contrary of a common belief, it is shown in [28] that the condition of B numerically affects the iteration as well as the subspace extraction using the null-space approach in (5). Note, however, that the inverse of B is not required at either of the two stages. In the following we review and elaborate the analysis. First, there are two unitary matrices Q and Z so that

$$A = Q\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} Z^{\mathrm{H}}, \quad \text{and} \quad B = Q\begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix} Z^{\mathrm{H}}, \tag{6}$$

where $\lambda(A_{11}, B_{11}) \subset \mathbb{C}^-$, $\lambda(A_{22}, B_{22}) \subset \mathbb{C}^+$, and $\lambda(A_{11}, B_{11}) \cup \lambda(A_{22}, B_{22}) = \lambda(A, B)$. Furthermore, there are two nonsingular matrices X and Y so that

$$A = Y\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} X^{-1}, \quad \text{and} \quad B = Y\begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix} X^{-1}, \tag{7}$$

where

$$X = Z\begin{pmatrix} I & F \\ 0 & I \end{pmatrix}, \quad Y = Q\begin{pmatrix} I & G \\ 0 & I \end{pmatrix}, \tag{8}$$

and F and G are the solution to the following system of matrix equations

$$GA_{22} - A_{11}F = A_{12}, \quad GB_{22} - B_{11}F = B_{12}. \tag{9}$$

For references, see [27,28].

By (7) and Proposition 2.1, the generalized Newton iteration with (A, B) is equivalent to two simultaneous iterations with (A_{11}, B_{11}) and (A_{22}, B_{22}). In detail, we have

$$\begin{aligned} A_{k+1} &= Y\begin{pmatrix} A_{11}^{(k+1)} & 0 \\ 0 & A_{22}^{(k+1)} \end{pmatrix} X^{-1}, \quad \text{and} \\ A_{k+1} &= Q\begin{pmatrix} A_{11}^{(k+1)} & GA_{22}^{(k+1)} - A_{11}^{(k+1)}F \\ 0 & A_{22}^{(k+1)} \end{pmatrix} Z^{\mathrm{H}}, \end{aligned} \tag{10}$$

where

$$A_{ii}^{(k+1)} = (A_{ii}^{(k)} + B_{ii}(A_{ii}^{(k)})^{-1}B_{ii}), \quad i = 1, 2.$$

As $k \to \infty$, $A_{11}^{(k)}$ and $A_{22}^{(k)}$ converge to $-B_{11}$ and B_{11}, respectively. In other words, the matrix at convergence by the generalized Newton iteration with (A, B) can be represented in the equivalent-to-block diagonal form,

$$A_\infty = Y \begin{pmatrix} -B_{11} & 0 \\ 0 & B_{22} \end{pmatrix} X^{-1}, \tag{11}$$

or, in the unitary-block triangular form,

$$A_\infty = Q \begin{pmatrix} -B_{11} & GB_{22} + B_{11}F \\ 0 & B_{22} \end{pmatrix} Z^{\mathrm{H}}. \tag{12}$$

The algebraic structure of A_∞ revealed in (11) relates the matrix pair (A_∞, B) to (A, B). The matrix on the right hand side of (11) is defined in [28] as the *generalized matrix sign function of A with respect to B* and denoted by $\mathrm{sign}(A, B)$. The matrix B is called the *pivot* matrix. In particular, $\mathrm{sign}(A, B)$ extends to $\mathrm{sign}(A) = \mathrm{sign}(A, I_n)$, when the pivot matrix is the identity matrix. We have proved the following theorem.

THEOREM 2.2. *If A and B are nonsingular, and (A, B) has no eigenvalue(s) on the imaginary axis, the generalized Newton iteration with (A, B) converges, and the matrix at convergence is the generalized matrix sign function of A with respect to B,* $\mathrm{sign}(A, B)$.

If the condition of B_{11} or B_{22} is poor, the effect on the computation of the deflating subspace is twofold. First, by (10), the condition of A_k is poor as $k \to \infty$, even if A is well-conditioned. The deflating subspaces can therefore be perturbed significantly during the iteration. Secondly, even if the matrix sign function is accurately computed, the extraction of the deflating subspaces using rank-revealing techniques based on (5) can be inaccurate or fail. We have from (6) and (12),

$$B + A_\infty = 2Q \begin{pmatrix} 0 & GB_{22} \\ 0 & B_{22} \end{pmatrix} Z^{\mathrm{H}}, \quad B - A_\infty = 2Q \begin{pmatrix} B_{11} & -B_{11}F \\ 0 & 0 \end{pmatrix} Z^{\mathrm{H}}. \tag{13}$$

The nonzero singular values of $B + A_\infty$ are those of $B_{22}^{\mathrm{H}}(G^{\mathrm{H}}, I)$. If B_{22} has small singular values, the numerical gap between the zero singular values and the nonzero singular values may be small and rank-revealing may become a difficult task. Similarly, the rank-revealing of $B - A_\infty$ may be affected by the poor condition of B_{11}.

2.2. Case: B is singular

In this subsection we are concerned with the case in which B is singular and A is nonsingular. The understanding of the case in which B is singular sheds light on the case in which B is near-singular. When B is singular, (A, B) has zero eigenvalue(s) in the μ-from.

We introduce the following lemma first.

LEMMA 2.3. *Let J_{nil} be a Jordan block of dimension n associated with the eigenvalue 0. The generalized Newton iteration with (I, J_{nil}) converges if and only if $n \leq 2$. Moreover, if $n \leq 2$, then A_k converges to 0 at a linear rate.* *Proof*: Let $A_0 = I_n$. In case $n \leq 2$, $J_{\mathrm{nil}}^2 = 0$, and $A_{k+1} = (A_k + J_{\mathrm{nil}} A_k^{-1} J_{\mathrm{nil}})/2 = A_k/2 = I/2^{k+1}$, $k = 1, 2, \ldots$. Thus, A_0 is nonsingular, but A_k converge to 0 as $k \to \infty$. If $n > 2$, we use the case $n = 3$ to illustrate that A_k do not converge. In this case, $J^2 \neq 0$ and $J_{\mathrm{nil}}^l = 0$ for any $l \geq 3$. It can be checked that $A_k = (I + \alpha_k J_{\mathrm{nil}}^2)/2^k$ with $\alpha_k = \sum_{i=0}^{k-1} 4^i$. ∎

THEOREM 2.4. *If (A, B) is a matrix pair with no eigenvalue(s) on the imaginary axis in the λ-form, the generalized Newton iteration converges if and only if B has no zero eigenvalue(s) of multiplicity greater than 2. Moreover,*

(i) if B has no zero eigenvalue(s), the convergence is quadratic;

(ii) if B has zero eigenvalue(s) of multiplicity less than or equal to 2, the convergence is linear, and the matrix at convergence is singular.

Proof: The condition that (A, B) has no eigenvalue(s) on the imaginary axis in the λ-form implies that A is nonsingular. Similarly to (7), we have the following canonical decomposition of (A, B) with respect to the spectral division between the zero eigenvalues and the nonzero eigenvalues in the μ-form

$$A = Y \begin{pmatrix} A_{\beta 0} & 0 \\ 0 & A_{\omega} \end{pmatrix} X^{-1}, \quad B = Y \begin{pmatrix} B_{\beta 0} & 0 \\ 0 & B_{\omega} \end{pmatrix} X^{-1}, \tag{14}$$

where X and Y are nonsingular, $\lambda(A_{\beta 0}, B_{\beta 0}) = \{(0, 1)\}$, and $\lambda(A_{\omega}, B_{\omega}) = \{(\alpha, \beta), \alpha\beta \neq 0\}$. By Proposition 2.1, the generalized Newton iteration with (A, B) converges if and only if the iteration with $(A_{\beta 0}, B_{\beta 0})$ and the iteration with (A_{ω}, B_{ω}) converge. The iteration with (A_{ω}, B_{ω}) converges as this matrix pair has no eigenvalue(s) on the imaginary axis. We shall give the convergence conditions on the iteration with $(A_{\beta 0}, B_{\beta 0})$. Since $A_{\beta 0}$ is nonsingular, $(A_{\beta 0}, B_{\beta 0})$ is equivalent to $(I, \mathrm{diag}(J_i))$, where J_i are the Jordan blocks associated with the eigenvalue 0. The iteration with $(A_{\beta 0}, B_{\beta 0})$ converges if and only if the iterations with (I, J_i) for all the Jordan blocks J_i converge. The proof is completed with the use of Lemma 2.3. ■

COROLLARY 2.5. *Let (A, B) be a matrix pair with no eigenvalue(s) on the imaginary axis. Assume that the generalized Newton iteration with (A, B) converges to A_{∞}. Then,*

- *If the zero eigenvalues in the μ-form are all simple, $\mathcal{N}(B + A_{\infty})$ is the right-deflating subspace corresponding to the zero eigenvalues and the eigenvalues in $\mathbb{C}^{-}$, and $\mathcal{N}(B^{\mathrm{H}} - A_{\infty}^{\mathrm{H}})$ is the left-deflating subspace corresponding to the zero eigenvalues and the eigenvalues in $\mathbb{C}^{+}$.*
- *$\mathcal{N}(A_{\infty})$ is the right-deflating subspace corresponding to the zero eigenvalues in the λ-form.*

3. Modified Algorithms

The analysis presented in Section 2 reveals the cause of certain computational difficulties. In this section we present variants for the generalized Newton iteration and modified algorithms for the extraction of the deflating subspaces from the computed generalized matrix sign function and its pivot matrix.

3.1. Reductions of the pivot matrix

The arithmetic cost per iteration of the generalized Newton iteration (4) is significantly higher than that of the Newton iteration (3), and the gap is bigger if the matrices are complex. The major difference lies in forming $BA^{-1}B$. In Table 1, we give the comparison in each substep per iteration, measuring the arithmetic cost in terms of floating point operations (flops). We assume that all matrices involved are real (the extension to complex matrices is straightforward) and the use of the LU factorization [18] for computing A^{-1} and solving $BA^{-1}B$.

We note that the pivot matrix remains unchanged in the iteration. If B is full, we propose to reduce B to a compact form with orthogonal transformations at the initial step. The compact form can be, for example, bidiagonal.

Given a matrix pair (A, B), the initial reduction of the pivot matrix B to a compact form requires an initial equivalence transformation of the matrix pair and a backward transformation at the subspace extraction stage. Specifically, let $B = YCX^{-1}$ be a decomposition of B, where X and Y are nonsingular matrices, and C presents the desired compact form. Set $(A_0, B) = (Y^{-1}AX, C)$

FORM OF B	BA^{-1}	$(BA^{-1})B$	TOTAL
$B = I$	$2n^3$	0	$2n^3$
FULL	$8n^3/3$	$2n^3$	$14n^3/3$
TRIANGULAR	$3n^3$	n^3	$4n^3$
NARROW BANDED	$2n^3 + 2\,b\,n^2$	$2\,b\,n^2$	$2n^3 + 4\,b\,n^2$
BIDIAGONAL	$2n^3 + 4n^2$	$4n^2$	$2n^3 + 8n^2$

Table 1: Arithmetic cost (in flops) for computing $BA^{-1}B$ (b is the bandwith for narrow band B).

for the generalized Newton iteration, and denote by A_∞ the matrix at convergence. Let $Z = (Z_1, Z_2)$ be a nonsingular matrix so that $(A_\infty + B)Z_1 = 0$. Then XZ_1 is the right-deflating subspace of (A, B) corresponding to the eigenvalues in $\mathbb{C}^-$. If X and Y are orthogonal matrices, the initial transformation of A and the backward transformation of Z are computationally stable.

Among many choices for the initial reduction, a QR/QL factorization of B renders a triangular pivot matrix and does not require a backward transformation of Z_1. The bidiagonal reduction of B gives the most sparse and structured compact form for a general pivot matrix B. The overhead cost (see Table 2) at the initial transformation and the backward transformation is well justified by the significant reduction in the arithmetic cost per iteration (see Table 1). Computationally, a bidiagonal

Factorization of B	$2n(2n - b_u - b_l)^2(b_u + b_l)/3$
Transformation of A	$4(n^2(n - b_u) - n(n - b_u)^2 + (n - b_u)^3/3)$
Backward transformation of Z	$2n^3 - 6n^2(n - b_u) + 2n(n - b_u)^2$

Table 2: Overhead arithmetic cost involved in the reduction.

reduction of B has little scope for Level 3 BLAS operations [11]. A relaxation on the bandwidth allows the use of Level 3 BLAS operations and introduces a tradeoff between the arithmetic cost and the memory access cost. Denote by bu and bl the upper bandwidth and the lower bandwidth of a band matrix, respectively. Let $b = bu + bl + 1$. To maintain the arithmetic cost of the multiplication of the pivot matrix with another matrix at $O(n^2)$, b must be kept small; that is, B must be narrow banded (independent of the matrix size). If B is bidiagonal, $b = 2$. Notice that the cost per iteration of the generalized Newton iteration with a bidiagonal pivot matrix is about the same as that of the Newton iteration (see Table 1).

3.2. Preconditioning for subspace extraction

We are concerned with general techniques for subspace extraction, assuming no special structure or a prior knowledge of the deflating space of interest. Rank-revealing algorithms are the method of choice. We use the following example to illustrate that a rank-revealing algorithm for the subspace extraction based on the null-space-finding approach, see (5), may fail if the pivot matrix B is ill-conditioned.

EXAMPLE 3.1. *Let*

$$A = \begin{pmatrix} \cos(\pi/6) & 0 & 0 & \sin(\pi/6) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sin(\pi/6) & 0 & 0 & \cos(\pi/6) \end{pmatrix},$$

and

$$B = A\begin{pmatrix} (5/3)\,e{-}06 & 0 & 0 & 0 \\ 0 & (10/3)\,e{-}11 & 0 & 0 \\ -4.3664\,e{-}01 & -7.5171\,e{-}01 & -(10/3)\,e{+}02 & 0 \\ -5.8242\,e{-}01 & -9.9153\,e{-}01 & 0 & -4.0000\,e{+}02 \end{pmatrix}.$$

The matrix pair (A, B) has two eigenvalues in $\mathbb{C}^-$ and two eigenvalues in $\mathbb{C}^+$. Let A_∞ be the computed matrix at the end of 35 iteration steps by the generalized Newton iteration with (A, B), using MATLAB 4.2C *with machine precision $\epsilon \approx 2,2\,e{-}16$. Then, $\kappa(A_\infty) \approx \kappa(B) \approx 1.2001\,e{+}12$ and the singular values of $B + A_\infty$ are approximately* $\{3.3333\,e{-}06, 7.0702\,e{-}10,\ 1.4246\,e{-}14, 6.3582\,e{-}34\}$. *The numerical rank of $B + A_\infty$ is not clear.*

We note two factors, among others, affecting the effectiveness of subspace extraction: the gap between the near-zero singular values and the nonzero singular values of $B + A_\infty$, and the threshold for numerical rank revealing. We know by the analysis presented in Section 2, see (10) and (13), that in case the pivot matrix B is ill-conditioned, A_k also becomes ill-conditioned and perturbed, and the perturbation blurs the boundary between the zero singular values and the non-zero singular values of $B + A_\infty$. If the perturbation is relatively smaller than the non-zero singular values of $B + A_\infty$, a rank-revealing algorithm can still succeed with a rank-revealing threshold adjusted accordingly. To reduce the perturbation in the computed generalized matrix sign function and then extract subspace accurately, we propose the following combined strategies.

(1) Stop the generalized Newton iteration halfway, namely, the following stopping criterion is employed

$$\|A_{k+1} - A_k\| < \sqrt{\epsilon}\|A_k\|, \text{ or } niter > max\_niter, \tag{15}$$

where *max_iter* is a prespecified upper bound on the number of iterations, *niter*. We denote by $A_{\infty/2}$ the matrix at halfway.

(2) Use one Malyshev iteration step on the Cayley transformed matrix pair $(A_{\infty/2}, B_{\infty/2})$. In detail,

$$\begin{array}{l} \text{Set } \ A_{\infty/2} = B + A_{\infty/2}, \quad B_{\infty/2} = B - A_{\infty/2} \ \ \text{(Cayley transform)} \\ \text{Compute } \ (Q_1, Q_2) \ \text{ so that } \ (Q_1, Q_2)\begin{pmatrix} A_{\infty/2} \\ B_{\infty/2} \end{pmatrix} = 0, \quad Q_1 Q_1^{\mathrm{T}} + Q_2 Q_2^T = I_n. \\ \text{Set } \ A_{\infty/2} = Q_2 A_{\infty/2}, \quad B_{\infty/2} = Q_1 B_{\infty/2} \ \ \text{(cross-update)} \end{array}$$

(3) Employ the subspace extraction method with preconditioning, cf. [28]. Specifically, compute

$$\begin{pmatrix} A^{\mathrm{T}}_{\infty/2} \\ B^{\mathrm{T}}_{\infty/2} \end{pmatrix} = \begin{pmatrix} Z_- \\ Z_+ \end{pmatrix} R \quad \text{(QR factorization)},$$

with $Z_-,\ Z_+ \in \mathbb{C}^{n\times n}$. Then, $\mathcal{N}(Z_-)$ and $\mathcal{N}(Z_+)$ are the right-deflating subspaces corresponding to the eigenvalues of (A, B) in $\mathbb{C}^-$ and $\mathbb{C}^+$, respectively.
Next, compute

$$Z_- \Pi_Z = Z R_Z \quad \text{(rank-revealing QR factorization)},$$

and let $r = \mathrm{rank}(R_Z, \tau_{rank})$, $Z = (Z_1, Z_2)$, $Z_1 \in \mathbb{C}^{n\times r}$, and $Z_2 \in \mathbb{C}^{n\times(n-r)}$.
Finally, compute

$$(AZ_2, BZ_2)\Pi_Q = QR_Q \quad \text{(rank-revealing QR factorization)},$$

and set $Z = (Z_2, Z_1)$.

Stopping the generalized Newton iteration within half of the machine precision attempts to avoid the catastrophic effect of the inverse of an ill-conditioned matrix. The Malyshev iteration step on the Cayley transformed matrix pair is equivalent to one step of the generalized Newton iteration, but does not require the inverse (see [28] for details). That is, if the convergence of the generalized Newton iteration is asymptotically quadratic, $(B+A_\infty, B-A_\infty)$ by the generalized Newton iteration is equivalent to $(A_{\infty/2}, B_{\infty/2})$ at the end of the Malyshev step within the machine precision. If there are eigenvalues close to the imaginary axis or if B is singular or near singular, then the generalized Newton iteration converges at a linear or nearly-linear rate and the prespecified upper bound *max_niter* on the number of iterations will terminate the iteration.

4. Numerical Experiments

In this section we compare the numerical results and the performance of the following spectral division algorithms:

- The QZ algorithm [22].
- GNEWTON: The generalized Newton iteration, with both A and B full-dense matrices.
- NEW-GNEWTON: The new generalized Newton iteration, with the pivot matrix reduced to bidiagonal form, and the subspace extraction with preconditioning (in case A or B are ill-conditioned).

In some cases we use a determinantal scaling [14] to accelerate the convergence of the generalized Newton iteration:

$$\begin{aligned} \gamma &= (|\det(A_k)|/|\det(B_k)|)^{-1/n}, \\ A_{k+1} &= (A_k + \gamma B\, A_k^{-1}\, B)/(2\gamma^2). \end{aligned}$$

4.1. Numerical accuracy

The computed basis for the deflating subspaces, Q and Z, satisfy

$$Q^{\mathrm{H}} A Z = \begin{pmatrix} A_{11} & A_{12} \\ E_{21} & A_{22} \end{pmatrix},$$

$$Q^{\mathrm{H}} B Z = \begin{pmatrix} B_{11} & B_{12} \\ F_{21} & B_{22} \end{pmatrix},$$

with $\|E_{21}\|$ and $\|F_{21}\|$ small-norm matrices. We use these norms to measure and compare the backward accuracy of the algorithms in our experiments; thus, we evaluate the *relative decoupling residual*:

$$\mathrm{rdr} = \|(E_{21}, F_{21})\|_F / \|(A, B)\|_F. \tag{16}$$

We relate this result to the condition number of the matrices A and B ($\kappa(A)$ and $\kappa(B)$, respectively) and the gap between the (real part of the) eigenvalues of the matrix pencil and the imaginary axis:

$$\Delta(A) = \min_{\lambda_j \in \lambda(A)} |Re(\lambda_j)|.$$

To estimate the sensitivity of the spectral subspaces to perturbations in a matrix pair, we compute a lower bound of the quantity *dif* introduced in [25]:

$$\mathrm{dif}[(A_{11}, B_{11}), (A_{22}, B_{22})] \stackrel{\mathrm{def}}{=} \inf_{\max\{\|P\|_{\mathrm{F}}, \|Q\|_{\mathrm{F}}\}=1} \max\{\|QA_{11} + A_{22}P\|_{\mathrm{F}}, \|QB_{11} + B_{22}P\|_{\mathrm{F}}\}.$$

All the numerical experiments are performed in MATLAB 4.2 on a Sun Sparc-20 workstation using IEEE arithmetic and $\epsilon \approx 2.2 \times 10^{-16}$.

We conducted two sets of experiments. In both cases, we applied the spectral division algorithms for the division along the imaginary axis.

The first set of experiments consists of matrix pairs (A, I_n) with parameter-dependent eigenstructures which control the difficulty of the spectral division [2,6]. We use the matrix pair (R, Q^{H}), where $A = QR$ is a QR factorization (no noticiable differences were found in the numerical accuracy when the equivalent pair (A, I_n) was used).

The tables in this section list the gap between the eigenvalues and the imaginary axis $\Delta(A)$, the condition number of A, $\kappa(A)$ ($\kappa(B) = 1$), the relative decoupling quantity dif, the relative decoupling residual rdr of the algorithms, and the number of iterations required by the generalized Newton iterations without/with determinantal scaling (#iter./#itsc., respectively).

Example 1. Consider the following matrix arising in a linear-quadratic optimal control problem [2,21]:

$$A = Q^{\mathrm{T}} \begin{pmatrix} F & S \\ G & -F^{\mathrm{T}} \end{pmatrix} Q,$$

where

$$F = \begin{pmatrix} -\eta & 1 & 0 & 0 \\ -1 & -\eta & 0 & 0 \\ 0 & 0 & \eta & 1 \\ 0 & 0 & -1 & \eta \end{pmatrix}, \quad S = G = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix},$$

and Q is a random orthogonal matrix (obtained from the QR factorization of a random matrix). The eigenvalues of A are symmetric with respect to the imaginary axis as well as the real axis, and are in a neighborhood of $\pm\eta^2 \pm i$. As $\eta \to 0$ the eigenvalues at about $\pm\eta^2 \pm i$ approach two points $\pm i$ on the imaginary axis and the division along this axis becomes more difficult. Table 3 reports that, as $\eta \to 0$,

$\Delta(A) \approx$ $\eta^2/2$	$\kappa(A)$	dif	QZ rdr	GNEWTON rdr (#iter.)	NEW-GNEWTON rdr (#iter./#itsc.)
5.0e-1	9.8	4.2e-1	3.2e-16	7.0e-16 (7)	3.3e-16 (7/7)
5.0e-3	1.7e+1	5.0e-3	4.5e-16	3.3e-15 (14)	2.6e-16 (14/9)
5.0e-5	1.7e+1	5.0e-5	3.2e-16	1.4e-13 (21)	3.3e-16 (21/8)
5.0e-7	1.7e+1	5.0e-7	2.5e-16	1.7e-11 (27)	3.0e-16 (27/8)

Table 3: Experimental results with Example 1.

the generalized Newton iteration (GNEWTON) obtains less accurate decoupling residuals. Our new algorithm (NEW-GNEWTON) obtains similar rdr to those of the numerically stable QZ algorithm in all cases. The determinantal scaling technique obtains a remarkable acceleration of the convergence of NEW-GNEWTON, while maintaining the numerical behavior (similar acceleration was achieved in GNEWTON).

Example 2. The matrix in this example presents a Hamiltonian-like structure:

$$A = Q^{\mathrm{T}} \begin{pmatrix} A_{11} & A_{12} \\ 0 & -A_{11}^T \end{pmatrix} Q,$$

with

$$A_{11} = \begin{pmatrix} 1-\alpha & & & \alpha \\ \alpha & 1-\alpha & & \\ & \ddots & \ddots & \\ & & \alpha & 1-\alpha \end{pmatrix},$$

A_{12} a $k \times k$ random matrix (not necessarily symmetric), Q a random orthogonal matrix, and A_{11} a circulant matrix with parameter α. The eigenvalues of A_{11} are evenly distributed on a circle with center at $1-\alpha$ and radius α. The eigenvalues of A are therefore distributed on two circles symmetric to the imaginary axis for any $\alpha \in (0, 0.5)$.

$\Delta(A)$	$\kappa(A)$	dif	QZ rdr	GNEWTON rdr (#iter.)	NEW-GNEWTON rdr (#iter./#itsc.)
10^{-1}	8.1e+2	1.4e-1	8.2e-16	7.0e-16 (9)	1.6e-16 (9/9)
10^{-3}	6.6e+6	1.4e-3	7.6e-16	3.6e-12 (16)	1.3e-16 (15/12)
10^{-5}	6.6e+10	1.4e-5	8.1e-16	3.6e-8 (22)	1.3e-16 (21/15)
10^{-7}	6.6e+14	1.4e-7	7.1e-16	–	1.3e-16 (22/16)
10^{-3}	4.1e+4	1.4e-1	7.5e-16	2.6e-14 (15)	1.2e-16 (15/13)
10^{-5}	4.1e+6	1.4e-1	7.8e-16	1.4e-12 (22)	1.6e-16 (21/17)
10^{-7}	4.1e+8	1.4e-1	7.8e-16	1.9e-10 (29)	1.7e-16 (25/20)

Table 4: Experimental results with Example 2 ($k = 10$).

Table 4 lists two groups of experimental results. In the first group, we let $\alpha \to 0.5$ so that the two circles get closer to each other toward the imaginary axis. Both $\Delta(A)$ and dif decrease at the same time. In the second group, we fix α at 0.45 and shift A so that $\Delta(A)$ gets smaller and dif remains the same. Our algorithm NEW-GNEWTON obtains close decoupling residuals to those of the QZ algorithm and much better than GNEWTON. In some cases NEW-GNEWTON required a smaller number of iterations to converge than GNEWTON and, in one case (dif$= 1.4 \times 10^{-7}$), NEW-GNEWTON solves the lack of convergence of GNEWTON. The use of the determinantal scaling technique accelerates the convergence of this example.

Example 3. The matrix in this example is constructed as

$$A = Q^{\mathrm{T}} \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} Q,$$

where Q is a random orthogonal matrix, A_{12} is a 5×5 random matrix, and A_{11} and A_{22} are random upper triangular matrices except that $\mathrm{diag}(A_{11}) = -\mathrm{diag}(A_{22}) = \beta D$, with D a random positive diagonal matrix, and β a positive parameter. As β gets smaller, dif decreases quickly but $\Delta(A)$ remains of the same order.

Table 5 reports that, for this example, the decoupling residual of our algorithm is slightly affected by the progresive ill-conditioning of the problem but NEW-GNEWTON still performs much better than GNEWTON.

β	$\Delta(A)$	$\kappa(A)$	dif	QZ rdr	GNEWTON rdr (#iter.)	NEW-GNEWTON rdr (#iter./#itsc.)
1.0	2.4e-1	2.8e+3	2.5e-2	4.4e-16	2.0e-15 (8)	2.1e-16 (7/7)
0.5	1.2e-1	7.3e+4	2.1e-3	3.8e-16	1.4e-14 (9)	2.7e-16 (9/8)
0.3	7.2e-2	8.8e+5	2.3e-4	5.1e-16	3.7e-13 (9)	9.5e-16 (9/8)
0.2	4.8e-2	6.7e+6	2.4e-5	4.5e-16	1.4e-11 (10)	2.1e-15 (10/8)
0.1	2.4e-2	7.1e+8	1.2e-7	4.6e-16	3.5e-9 (11)	5.9e-14 (11/8)

Table 5: Experimental results with Example 3.

Finally, we also conducted a second set of experiments with 100×100 random matrices and entries uniformly distributed within $[0, 1]$. All our executions reported roughly similar results for rdr, both from the QZ algorithm and the generalized Newton iterations.

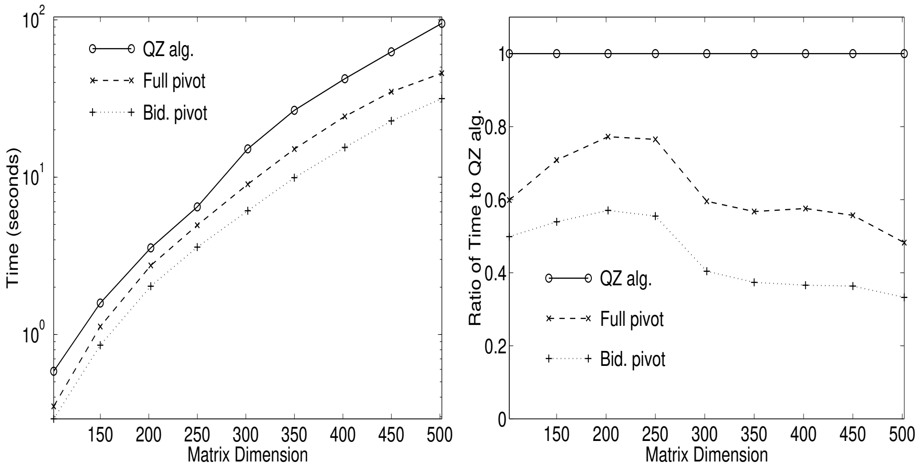

Fig. 2: Spectral division of random problem. Global execution time and ratio between the execution time of the QZ algorithm and the generalized Newton iterations on an SGI R10000

4.2. Serial perfomance

In this section we analyze the performance of the spectral division methods. We compare the QZ algorithm and two implementations of our generalized Newton iteration NEW-GNEWTON. One of these implementations works on the original matrix pair (A, B) while the other initially reduces the pivot matrix to bidiagonal form. The purpose here is to measure the reduction of the execution time obtained by the initial transformation to condensed form.

We implemented all our algorithms on an SGI R10000 platform, using Fortran-77, the vendor-supplied BLAS, and the LAPACK [1] library.

All the figures in this section contain two different plots. The left-hand plot reports the overall execution time (in seconds) of the QZ algorithm and the generalized Newton iterations. A semilogarithmic scale is used for axis y. In the right-hand plot, we define the execution time of the QZ algorithm as the unit of time and report the reduction of the cost of the generalized Newton iterations with respect to the QZ algorithm.

We use the spectral division methods in two different applications. In the first application, the goal is to divide the spetrum of a matrix pencil along the imaginary axis. We report on matrix pairs of size $100, 150, \ldots, 500$, with entries uniformly distributed within $[0, 1]$. The experimental results are averaged for 5 executions on different random matrices.

Figure 2 reports that the execution time of the algorithm which employs a full pivot matrix is around 50%-60% of that of the QZ algorithm for large matrices. By transforming the pivot matrix to bidiagonal form, this cost is further reduced to 35%-40%.

In our second application we compute the solution matrix, X, of a generalized Riccati matrix equation [21]

$$A^T X E + E^T X A - E^T X S X E + Q = 0,$$

where the matrices involved are all square, say $m \times m$. This quadratic matrix equation arises in control problems [14], and its solution can be obtained by computing the right-deflating "stable"

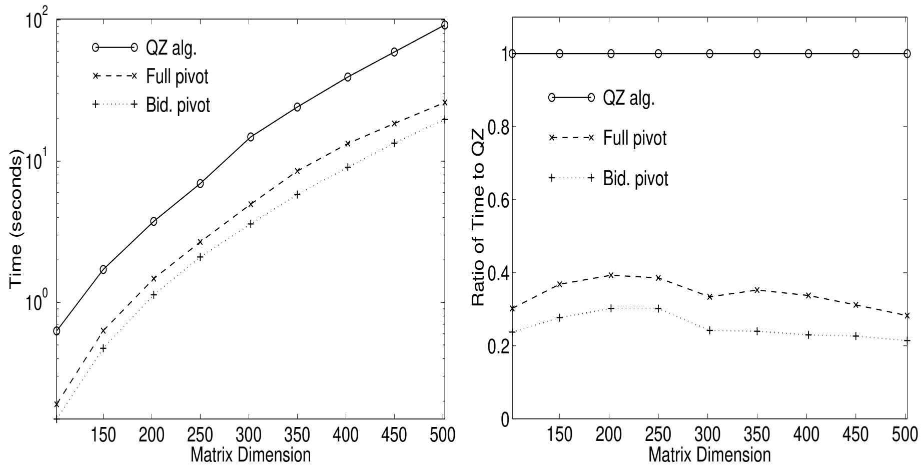

Fig. 3: Solution of GARE for the CHSV problem. Global execution time and ratio between the execution time of the QZ algorithm and the generalized Newton iteration on an SGI R10000

subspace of the $2m \times 2m$ matrix pair defined by

$$H = \begin{pmatrix} A & S \\ Q & -A^{\mathrm{T}} \end{pmatrix} \quad \text{and} \quad K = \begin{pmatrix} E & 0 \\ 0 & E^{\mathrm{T}} \end{pmatrix}. \tag{17}$$

To measure the performance of the algorithms we employ two classical, well-conditioned control problems. The first one arises in the control of a string of high speed vehicles (hereafter, CHSV); the second problem is constructed with an a priori known *circulant* solution matrix (see, among many, [9, 14]). In both examples, $E = I_n$ and an standard eigenvalue problem (H_{std}, I_n) is involved. For our performance evaluation purposes, we work on the equivalent matrix pair (H_{gen}, K_{gen}), where $A_{gen} := LA$, $S_{gen} := LAL^T$, $Q_{gen} := Q$, and $E_{gen} = L$ a random, well-conditioned lower unit triangular matrix.

Figures 3 and 4 report an important advantage of the generalized Newton iterations over the QZ algorithm, which is increased with the problem dimension. The generalized Newton iterations require only 20%-40% of the execution time of QZ algorithm for the CHSV problem. When applied to the circulant problem, the difference is progresively increased for larger problem sizes. The advantage is partially due to the fast convergence of the generalized Newton iteration for both problems. Also, for the same reason, there is only a small difference between both generalized Newton iterations.

Two improvements on the generalized Newton iterations can produce implementations which further outperform the QZ algorithm. First, in the described control problems, we can exploit the special structure of matrix K when computing its bidiagonal form. Second, in general spectral division problems, there are efficient algorithms that reduce a matrix to narrow banded form, and allow the introduction of high-performance BLAS level 3 kernels [11].

5. Concluding Remarks

We have presented several modified algorithms for computing deflating subspaces that reduce by 75% the cost of the traditional approach based on the matrix sign function. First, the left and right-deflating subspaces are extracted from a single iteration, reducing the number of iterations to half. Secondly, the cost of each Newton iteration is cut by more than 50% by transforming the pivot

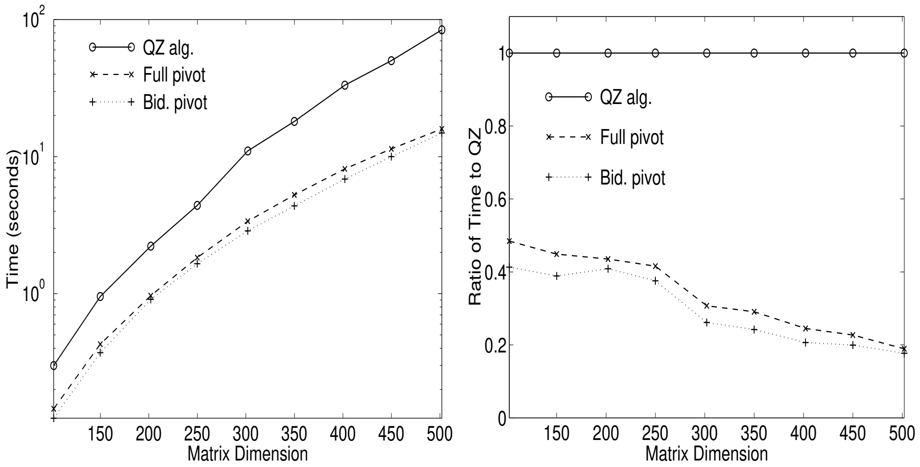

Fig. 4: Solution of GARE for the problem with circulant solution matrix. Global execution time and ratio between the execution time of the QZ algorithm and the generalized Newton iteration on an SGI R10000

matrix to bidiagonal form. Our new subspace extraction techniques also improve the accuracy of the computed deflating subspaces.

Numerical experiments report the matrix sign fucntion as a fast, parallel, and reliable approach for computing partial information of the eigenspectrum of matrix pencils.

REFERENCES

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen, *LAPACK User's Guide*, SIAM, 1995.

[2] W. F. Arnold and A. J. Laub, *Generalized eigenproblem algorithms and software for algebraic Riccati equations*, Proc. IEEE, 72 (1984), pp. 1746–1754.

[3] Z. Bai and J. W. Demmel, *Design of a parallel nonsymmetric eigenroutine toolbox, part II*, tech. report, Department of Mathematics, University of Kentucky, 1996.

[4] ———, *Using the matrix sign function to compute invariant subspaces*, SIAM J. Mat. Anal. Appl., 19 (1998), pp. 205–225.

[5] Z. Bai, J. W. Demmel, J. Dongarra, A. Petitet, and H. Robinson, *The spectral decomposition of nonsymmetric matrices on distributed memory multiprocessors*, SIAM J. Sci. Comp., 18 (1997), pp. 1446–1461.

[6] Z. Bai, J. W. Demmel, and M. Gu, *Inverse free parallel spectral divide and conquer algorithms for nonsymmetric eigenproblems*, Numerische Mathematik, 76 (1997), pp. 279–308.

[7] P. Benner, *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems*, dissertation, Fak. f. Mathematik, TU Chemnitz–Zwickau, Chemnitz, FRG, 1997.

[8] P. Benner and R. Byers, *An arithmetic for matrix pencils*, in Proceedings MTNS'98, Padova, Italy, 1998.

[9] 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*, Tech. Report SPC 95-22, Faculty of Mathematics, Technische Universität Chemnitz-Zwickau, 1995.

[10] P. Benner and E. S. Quintana-Ortí, *Solving stable generalized Lyapunov equations with the matrix sign function*, Tech. Report SFB393/97-23, Faculty of Mathematics, Technische Universität Chemnitz-Zwickau, 1997. To appear in *Numer. Alg.*

[11] C. H. Bischof, B. Lang, and X. Sun, *The SBR toolbox - software for succesive band reduction*, tech. report, PRISM Working Note #24, 1996.

[12] C. H. Bischof and G. Quintana-Ortí, *Computing rank-revealing QR factorizations of dense matrices*, Tech. Report MCS-P559-0196, Mathematics and Computer Science Division, Argonne National Laboratory, 1996. To appear in ACM Trans. Math. Soft.

[13] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley, *ScaLAPACK user's guide*, SIAM,

1997.
[14] R. Byers, *Solving the algebraic Riccati equation with the matrix sign function*, Linear Algebra and Its Appl., 85 (1987), pp. 267–279.
[15] R. Byers, C. He, and V. Mehrmann, *The matrix sign function and the computation of invariant subspaces*, SIAM J. Mat. Anal. Appl., 18 (1997), pp. 615–632.
[16] T. Chan, *Rank-revealing QR factorizations*, Lin. Alg. Appl., 88/89 (1987), pp. 67–82.
[17] J. Gardiner and A. J. Laub, *A generalization of the matrix sign function solution for algebraic Riccati equations*, Int. J. Control, 44 (1986), pp. 823–832.
[18] G. Golub and C. V. Loan, *Matrix Computations*, The Johns Hopkins University Press, 1989.
[19] P. Hong and C. T. Pan, *The rank-revealing QR and SVD*, Math. Comp., 58 (1992), pp. 213–232.
[20] C. Kenney and A. Laub, *Rational iterative methods for the matrix sign function*, SIAM J. Mat. Anal. Appl., 12 (1991), pp. 273–291.
[21] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.
[22] C. Moler and G. W. Stewart, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), pp. 241–256.
[23] E. S. Quintana-Ortí, G. Quintana-Ortí, X. Sun, and R. van de Geijn, *Efficient matrix inversion via Gauss-Jordan elimination and its parallelization*, Tech. Report TR 98-, Department of computer Sciences, The University of Texas at Austin, 1998.
[24] J. Roberts, *Linear model reduction and solution of algebraic Riccati equations by the use of the sign function*, Int. J. Control, 32 (1980), pp. 677–687.
[25] G. W. Stewart, *On the sensitivity of the eigenvalue problem*, SIAM J. Numer. Anal., 9 (1972), pp. 669–686.
[26] ———, *Perturbation theory for the generalized eigenvalue problem*, in Recent Advances in Numerical Analysis, G. G. C. de Boor, ed., Academic Press, 1978, pp. 193–206.
[27] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
[28] X. Sun and E. S. Quintana-Ortí, *Spectral division methods for block generalized schur decompositions*, tech. report, PRISM Working Note #32, 1996. Available from `http://www-c.mcs.anl.gov/Projects/PRISM`.
[29] R. van de Geijn, *Using PLAPACK*, The MIT Press, 1997.