Linear Algebra and its Applications 738 (2026) 25–44

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

On commuting matrices and pencils

Peter Benner [a], Paul Van Dooren [b,*]

[a] *Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany*
[b] *Department of Mathematical Engineering, Université catholique de Louvain, Belgium*

A R T I C L E   I N F O

Article history:
Received 17 January 2025
Received in revised form 5 February 2026
Accepted 23 February 2026
Available online 25 February 2026
Submitted by V. Mehrmann

MSC:
15A18
15A22
47AXX

Keywords:
Commuting descriptor systems
Simultaneous diagonalization
Simultaneous triangularization
Normal pencils

A B S T R A C T

In this paper we extend the classical notion of commuting matrices to that of commuting pencils. This definition requires the pencils $zB_j - A_j$ to be regular but they may have singular B_j matrices. If the B_j matrices happen to be invertible, our extension is equivalent to the commutativity of the (left) quotients $B_j^{-1}A_j$. This choice is linked to the application of commuting sets of descriptor systems [1]. We show that the invertibility condition of the B_j matrices can be circumvented by using Möbius transforms of the pencils $zB_j - A_j$. We also introduce minimization problems that can serve as necessary and sufficient condition for the commutativity of certain classes of diagonalizable pencils. The more challenging non-diagonalizable case is not considered here.

© 2026 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

1. Introduction

Engineering applications are sometimes making use of sets of commuting matrices, such as the set of circulant matrices [7] or the set of functions of a given base matrix

* Corresponding author.
E-mail addresses: benner@mpi-magdeburg.mpg.de (P. Benner), paul.vandooren@uclouvain.be (P. Van Dooren).

https://doi.org/10.1016/j.laa.2026.02.032
0024-3795/© 2026 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

[12]. One is then, for instance, interested in sets of nearly commuting matrices such as in the analysis of particular tensor decompositions [8], [14]. In this paper, we first revisit this matrix problem and show then how it can be extended to sets of pencils. In order to define commutativity of pencils correctly, one has to first describe the setting in which this property makes sense. Such a setting was already described in an earlier publication [1], where we considered the set of consecutive discrete-time systems of equations in descriptor form

$$F_i y_{i+1} = E_i y_i + u_i,\ i = 1, \ldots, m, \quad (E_i, F_i) \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}. \tag{1}$$

In the context of dynamical systems, we could write these equations as

$$(zF_i - E_i) y_i = u_i, \quad i = 1, \ldots, m,$$

with z the delay operator. Such equations evolving in time, can also be written as a system of equations linking the compound vectors

$$u_{[1:m]}^\top := \left[u_1^\top, \ldots, u_m^\top\right], \quad y_{[1:m+1]}^\top := \left[y_1^\top, \ldots, y_{m+1}^\top\right]$$

to each other, using the compound system of equations [15]

$$\left[\begin{array}{c|cccc|c} -E_1 & F_1 & & & & \\ & -E_2 & \ddots & & & \\ & & \ddots & F_{m-1} & & \\ & & & -E_m & & F_m \end{array}\right] \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{m+1} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}. \tag{2}$$

The solution $y_{[1:m+1]}$ of this system of equations can be made unique by appending to it the two-point boundary conditions at time steps 1 and $m+1$:

$$W_1 y_1 + W_{m+1} y_{m+1} = w, \quad \text{where} \quad (W_1, W_{m+1}) \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}. \tag{3}$$

Commutativity of a set of pencils $\mathcal{P} := \{(zB_j - A_j), j = 1, \ldots, m\}$, can then be defined in terms of the zero input solutions of the above two-point boundary value problem built on an arbitrary sequence $(zF_i - E_i)$ of m pencils, drawn from the set $\mathcal{P}$. The pencils in the set $\mathcal{P}$ are said to commute if the two-point boundary value problem yields the same end point solutions y_1 and y_{m+1}, irrespective of the order of the m chosen pencils $(zF_i - E_i)$. It is easy to see that when all the matrices $F_i = I_n$, then all matrices E_i have to commute for this property to hold. The problem we are considering can therefore be interpreted as the commutativity of pencils of matrices.

Notation. For simplicity of the arguments, we will assume that all matrices and pencils are square and complex, so that their spectrum belongs to the same field as the matrix elements. We will make use of the so-called "off"-norm and "low"-norm of a set of matrices $\mathcal{M} := \{M_1, \ldots, M_k\}$ and define it as the Frobenius norm of the compound matrices

$$\| [M_1, \ldots, M_m] \|_{of} := \| [(M_1)_{of}, \ldots, (M_m)_{of}] \|_F,$$

and

$$\| [M_1, \ldots, M_m] \|_{lo} := \| [(M_1)_{lo}, \ldots, (M_m)_{lo}] \|_F,$$

respectively, where $(M)_{of}$ and $(M)_{lo}$ means the off-diagonal part and the strictly lower-triangular part of a matrix M. In this paper, we will consider $n \times n$ regular matrix pencils. An $n \times n$ pencil $zB - A$ is said to be regular if $\det(zB - A)$ is not identically zero, which implies that its set of eigenvalues, denoted as $\Lambda(A, B) := \{\lambda_1, \ldots, \lambda_n\}$, is a self-conjugate set when the pencil is real. The finite eigenvalues are the roots of the polynomial $\det(zB - A)$ and the nonzero vectors x_i in the kernel of $(\lambda_i B - A)$ are the corresponding eigenvectors. When B is singular this pencil also has infinite eigenvalues. For this reason, it is preferred to define the eigenvalues as the ratios $\lambda_i := a_i/b_i$ where the pairs (a_i, b_i) satisfy the homogeneous equation $\det(aB - bA) = 0$. A subspace $\mathcal{T}$ of $\mathbb{C}^n$ is called an invariant subspace of an $n \times n$ matrix A when $A\mathcal{T} \subset \mathcal{T}$, where the notation $A\mathcal{T}$ means the subspace of all vectors $Ax, x \in \mathcal{T}$. One says then that $\mathcal{T}$ is A-invariant. A deflating subspace of a regular pencil $zB - A$ is a subspace that satisfies the equation

$$\dim (B\mathcal{T} + A\mathcal{T}) = \dim \mathcal{T}.$$

The spaces $\mathcal{S} := (B\mathcal{T} + A\mathcal{T})$ and $\mathcal{T}$ are called a pair of deflating subspaces. When B is invertible, one can rewrite this equation as

$$\dim B(\mathcal{T} + B^{-1}A\mathcal{T}) = \dim (\mathcal{T} + B^{-1}A\mathcal{T}) = \dim \mathcal{T},$$

which implies that the subspace $(B^{-1}A)\mathcal{T}$ must be a subspace of $\mathcal{T}$, and hence that $\mathcal{T}$ is $(B^{-1}A)$-invariant. The deflating subspaces of a monic pencil $zI_n - A$ are therefore also the invariant subspaces of A. Subspaces spanned by a set of eigenvectors x_i of the pencil $zB - A$ are examples of deflating subspaces. We refer to [18, Section 6.6] for a more elaborate discussion.

The paper is organized as follows. In Section 2, we recall several properties of commuting matrices and formulate optimization problems that verify if a set of matrices commutes. If not, the optimization also gives an estimate of how close the matrices are to a commuting set of matrices. The details of the numerical procedures are also discussed. In Section 3, we show how to extend this to a set of commuting pencils. This includes a discussion of the properties of commuting pencils and the corresponding optimization problems to verify commutativity of a set of pencils or of nearness to such a set. In Section 4, we give a number of numerical experiments to illustrate the theoretical results, and in Section 5, we conclude with a few final remarks.

2. Commuting matrices

2.1. Decompositions

We briefly recall here some classical results regarding decompositions for commuting matrices in $\mathbb{C}^{n\times n}$. They are based on the existence of invariant subspaces with disjoint spectrum. All the results of this section are known and can be found in [10,13] but we give concise proofs when useful for our discussion.

Theorem 2.1 *(see [10], Section VIII.2). Let the matrix $M \in \mathbb{C}^{n\times n}$ have two complementary invariant subspaces $\mathcal{T}_1$ and $\mathcal{T}_2$ of respective dimensions n_1 and $n_2 = n - n_1$, and with disjoint spectrum, then there exists an invertible transformation T such that*

$$T^{-1}MT = M_T = \begin{bmatrix} M_{1,1} & 0 \\ 0 & M_{2,2} \end{bmatrix}, \quad M_{1,1} \in \mathbb{C}^{n_1\times n_1}, M_{2,2} \in \mathbb{C}^{n_2\times n_2},$$

where the spectra of $M_{1,1}$ and $M_{2,2}$ are disjoint.

Proof. The matrix $T = \begin{bmatrix} T_1 & T_2 \end{bmatrix}$ whose block columns span the two complementary invariant subspaces, yields the result. $\square$

Definition 2.2. A set of matrices $\mathcal{M} := \{A_j, j = 1, \ldots, m\}$ is said to be commuting if every pair of $\mathcal{M}$ commutes, i.e. $A_iA_j = A_jA_i$ for all $i, j \in \{1, \ldots, m\}$.

A very useful property of commuting matrices is provided in the following result.

Theorem 2.3 *(see [10], Theorem 3, page 223). If a matrix $M \in \mathcal{M}$ from a commuting set has a block-diagonal decomposition*

$$T^{-1}MT = \begin{bmatrix} M_{1,1} & 0 \\ 0 & M_{2,2} \end{bmatrix}, \quad M_{1,1} \in \mathbb{C}^{n_1\times n_1}, M_{2,2} \in \mathbb{C}^{n_2\times n_2},$$

where the spectra of $M_{1,1}$ and $M_{2,2}$ are disjoint, then every other matrix in that set has a similar block-diagonal decomposition, using the same similarity transformation T. When partitioned as $T = \begin{bmatrix} T_1 & T_2 \end{bmatrix}$, the column spaces of T_1 and T_2 span the spaces $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively. These spaces are also invariant for every matrix in the set $\mathcal{M}$.

Corollary 2.4. *If a matrix $M \in \mathcal{M}$ from a commuting set has a block-diagonal decomposition*

$$T^{-1}MT = \mathrm{diag}(D_{1,1}, \ldots, D_{k,k}), \quad D_{i,i} \in \mathbb{C}^{n_i\times n_i},$$

where the spectra of the diagonal blocks $D_{i,i}$ are disjoint, then every matrix in $\mathcal{M}$ has such a decomposition, with the same similarity matrix T, and the same block sizes n_i.

Theorem 2.5 *(see [13] Th 1.3.19). A set of diagonalizable matrices is a commuting set if and only if they are simultaneously diagonalizable with the same similarity matrix T.*

Notice that any linear combination or product of matrices from the set $\mathcal{M}$ also commutes with all matrices of $\mathcal{M}$. If we can find such a matrix that has distinct eigenvalues, then its eigenvectors determine T up to diagonal scaling (see also [11]).

Corollary 2.6. *If the matrices of a commuting set $\mathcal{M}$ are all Hermitian, then every matrix in $\mathcal{M}$ can be diagonalized with the same unitary similarity transformation U.*

Proof. This follows from Theorem 2.5 since Hermitian matrices are diagonalizable. □

And finally, the following property follows from the fact that commuting matrices must have a common eigenvector (see [13], Lemma 1.3.19, p.63).

Theorem 2.7 *(see [13], Theorem 2.3.3, p. 81). If $\mathcal{M}$ is a set of commuting matrices, then there exists a unitary matrix U that simultaneously triangularizes all matrices in $\mathcal{M}$.*

Proof. All matrices M_i of a commuting set must have a common eigenvector u_1, which we assume normalized. Then there exists a unitary matrix U that has this normalized vector as first column: $U\mathbf{e}_1 = u_1$. Therefore the first column of each matrix $\hat{M}_i := U^H M_i U$ is proportional to $\mathbf{e}_1$, and the matrices $\hat{M}_i$ are thus block-triangular with a leading 1×1 block, and a trailing $(n-1) \times (n-1)$ block. Moreover, these diagonal blocks are commuting as well, and we can therefore repeat this argument in a recursive fashion on the trailing submatrices, like in the proof of the Schur form of a single matrix. The similarity matrices that are constructed at each step, are unitary, and so is their product since unitary matrices form a multiplicative group. □

2.2. Optimization

Two of the decompositions of the previous subsection lead to a necessary and sufficient condition for commutativity of a particular set $\mathcal{M}$.

The first one is a set of *commuting Hermitian matrices.* Such matrices can simultaneously be diagonalized over the group $\mathcal{U}$ of unitary transformations. Therefore, the following minimization problem

$$\min_{\mathcal{U}} \| \left[U^{\mathsf{H}} A_1 U, \ldots, U^{\mathsf{H}} A_m U \right] \|_{of} \tag{4}$$

has a zero minimum if and only if the given set of Hermitian matrices commutes.

The second one is a set of *commuting diagonalizable matrices.* Such a set of matrices can simultaneously be diagonalized over the group of general linear similarity transformations $\mathcal{G}$. Therefore the minimization problem

$$\min_{\mathcal{G}} \| \left[T^{-1} A_1 T, \ldots, T^{-1} A_m T \right] \|_{of} \tag{5}$$

has a zero minimum if and only if the given set of diagonalizable matrices commutes.

Notice that Theorem 2.7 only yields a necessary condition for an arbitrary set of matrices to commute. The minimization problem

$$\min_{\mathcal{U}} \| \left[U^{\mathsf{H}} A_1 U, \ldots, U^{\mathsf{H}} A_m U \right] \|_{lo} \tag{6}$$

will yield a zero minimum if the matrix set commutes, but the reverse claim is not true, otherwise any set of triangular matrices would always commute. The triangularization can also be obtained in another way, but only for any set of commuting diagonalizable matrices. Let T be the transformation obtained from the optimization problem (5), and let $T = QR$ be its QR factorization. Then the matrices $Q^{\mathsf{H}} A_i Q$ are upper triangular since the matrices $T^{-1} A_i T = R^{-1} Q^{\mathsf{H}} A_i Q R$ are diagonal, and the matrices R^{-1} and R are upper-triangular. The Q factor of the diagonalizing transformation is thus the unitary transformation that triangularizes the set. We point out that there are constructive methods to compute such a common unitary similarity that simultaneously triangularizes commuting matrices, but they require clustering techniques to detect potential groups of multiple eigenvalues [6].

These optimization problems can also be used to find a nearby set of commuting matrices to a given set of matrices. As soon as one encounters a (possibly local) minimizer of the above minimization problems, one can view the corresponding off-diagonal or lower triangular "residuals" as the corrections needed to make the matrices commute. For the optimization problem (4), let $R_i := (U^{\mathsf{H}} A_i U)_{of}$, then the matrices $A_i - U R_i U^{\mathsf{H}}$ are Hermitian and commute, and for optimization (5), let $R_i := (T^{-1} A_i T)_{of}$, then the matrices $A_i - T R_i T^{-1}$ are diagonalizable and commute. For Hermitian matrices, this point was already made in [3], [4], [11]. These papers also propose algorithms for solving the minimization problem (4) and recommend (for efficiency) to first apply a unitary diagonalization of a linear combination of the matrices, before switching to the minimization. As a result, the convergence of the optimization procedure will significantly improve [4].

For these two problems, we will consider Jacobi-like algorithms such as those developed for the symmetric and non-symmetric eigenvalue problem [19], [9]. At each step of these algorithms a 2×2 problem is solved with the same structure as the general problem, and propagated through the matrix or pencil. The matrix representation of such a Jacobi-like transformation T affecting rows and columns (i, j) is thus

$$T = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & t_{i,i} & 0 & t_{i,j} & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & t_{j,i} & 0 & t_{j,j} & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}.$$

It was proven in [19] that the Jacobi algorithm converges quadratically if one implements the 2×2 transformations in a cyclic fashion between every possible pair of rows and columns (i, j). There are $n(n-1)/2$ such pairs but they must be ordered in a consistent manner in what is called a "forward sweep" or "backward sweep" to guarantee that the "off"-norm gets squared after each sweep. The complexity for such a sweep is thus of the order of n^3 floating point operations since one pair of left and right 2×2 rotations requires of the order of n floating point operations. In [2] it was then shown that the cyclic sweep scheme could be reorganized, using well chosen permutations, such that all 2×2 rotations were applied to adjacent rows and columns $(2i-1, 2i)$ or $(2i, 2i+1)$ only, starting with "odd" and "even" index.[1] This reordered Jacobi-like algorithm now has the following properties:

1. only adjacent rows and columns are selected for each 2×2 subproblem
2. only orthogonal or invertible 2×2 transformations are used
3. the choice of 2×2 transformations is such that they are as close as possible to permutation matrices (see below).

2.3. Hermitian matrices

In this section, we re-derive a Jacobi like method for the minimization problem (4), which is heavily inspired by [3], [4], but we chose to simplify the derivation, under the assumption that a preliminary common transformation U was already applied to the Hermitian matrices A_i, so that the off-diagonal part is already small. The updating of the Jacobi rotations is then much easier to derive since it suffices to use first order approximations of the 2×2 diagonalizing transformations in order to observe quadratic convergence. This choice was made, because it also carries over to the transformation group $\mathcal{G}$ of invertible similarity transformations that will be done in the next section.

We recall the technique of applying 2×2 diagonalizing Jacobi rotations to a general Hermitian matrix. Let us consider a 2×2 Hermitian submatrix

$$A := \begin{bmatrix} a & e \\ \bar{e} & b \end{bmatrix}, \quad \text{where} \quad |e| \ll |b-a|.$$

The Jacobi rotation algorithm would diagonalize the matrix A using a unitary similarity transformation U, but since we want to apply this within an iterative process, it is sufficient to apply an approximate diagonalization that makes sure that the off-diagonal element e converges at least quadratically to 0. This can be obtained as follows. If $|e|$ is small, b and a are good (quadratic) approximations of the eigenvalues of A. This follows

[1] This new scheme was in fact developed for multi-processor arrays but is now the preferred scheme on most computer architectures because the data communication between matrix elements is simpler. A pair of forward and backward sweeps is then replaced by n odd-even steps, but the complexity of the floating point operations remains exactly the same. We refer to [9], [2], [19] for more details.

from the general theory of perturbations of Hermitian matrices (see [17, Theorem 4.15]). Moreover, the columns of the matrix

$$U = c\begin{bmatrix} \bar{t} & 1 \\ 1 & -t \end{bmatrix}, \quad \text{where} \quad t := \frac{e}{(b-a)}, \quad c := \frac{1}{\sqrt{1+|t|^2}},$$

are good approximations for the corresponding eigenvectors of these eigenvalues (in that order). Using $s := tc$, it then follows that

$$\begin{aligned} U^{\mathsf{H}}AU &= c^2\begin{bmatrix} t & 1 \\ 1 & -\bar{t} \end{bmatrix}\begin{bmatrix} a & e \\ \bar{e} & b \end{bmatrix}\begin{bmatrix} \bar{t} & 1 \\ 1 & -t \end{bmatrix} \\ &= \begin{bmatrix} c^2b + |s|^2a + 2c\Re(se) & -s^2e \\ -s^2e & c^2a + |s|^2b - 2c\Re(se) \end{bmatrix} \\ &= \begin{bmatrix} b & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} \mathcal{O}(e^2) & \mathcal{O}(e^3) \\ \mathcal{O}(e^3) & \mathcal{O}(e^2) \end{bmatrix}. \end{aligned}$$

This shows that we are nearly diagonalizing A and are moreover swapping the order of eigenvalues (which is needed for an efficient implementation of the Jacobi algorithm). Notice also that two such matrices commute

$$\begin{bmatrix} a_1 & e_1 \\ \bar{e}_1 & b_1 \end{bmatrix}\begin{bmatrix} a_2 & e_2 \\ \bar{e}_2 & b_2 \end{bmatrix} = \begin{bmatrix} a_2 & e_2 \\ \bar{e}_2 & b_2 \end{bmatrix}\begin{bmatrix} a_1 & e_1 \\ \bar{e}_1 & b_1 \end{bmatrix},$$

if and only if the corresponding tangents t are equal:

$$a_1e_2 + e_1b_2 = a_2e_1 + e_2b_1, \quad \Leftrightarrow \quad \frac{e_1}{(b_1-a_1)} = \frac{e_2}{(b_2-a_2)}.$$

In order to compute the tangent t defining the rotation described above, we can as well solve the equation $ec - (b-a)s = 0$ to define s and c and eventually $t = s/c$. When we have several of these tangents, we can then find the vector $[c; -s]$ satisfying

$$\begin{bmatrix} e_1 & (b_1 - a_1) \\ \vdots & \vdots \\ e_m & (b_m - a_m) \end{bmatrix}\begin{bmatrix} c \\ -s \end{bmatrix} \approx 0$$

in a total least squares sense, i.e. we take the right singular vector corresponding to the smallest singular value of this stacked matrix, and scale it so that c is real. Since the rows of largest norm also yield a larger contribution to the singular value decomposition of the data matrix, they will also have a bigger influence on the solution. Notice that only the quantities e_i and s are complex. In the symmetric case, they also will be real.

2.4. Diagonalizable matrices

Here we give the analogous non-orthogonal Jacobi-like algorithm for diagonalizing an arbitrary complex 2×2 matrix

$$B := \begin{bmatrix} a & e \\ f & b \end{bmatrix}, \quad \text{where} \quad |e|, |f| \ll |b-a|,$$

using a similarity transformation $T^{-1}BT$. The Jacobi-like algorithm would diagonalize the matrix B using an invertible transformation T. If $|e|$ and $|f|$ are small, b and a are good approximations of the eigenvalues of B, and the (scaled) columns of

$$T = \begin{bmatrix} y & 1 \\ 1 & -x \end{bmatrix}, \quad \text{where} \quad x := \frac{f}{(b-a)}, \; y := \frac{e}{(b-a)}$$

are good approximations of the eigenvectors corresponding to these eigenvalues (again, in that order). It then follows that

$$\begin{aligned} T^{-1}BT &= \frac{1}{1+xy} \begin{bmatrix} x & 1 \\ 1 & -y \end{bmatrix} \begin{bmatrix} a & e \\ f & b \end{bmatrix} \begin{bmatrix} y & 1 \\ 1 & -x \end{bmatrix} \\ &= \frac{1}{1+xy} \begin{bmatrix} b + xya + xe + yf & -x^2 e \\ -y^2 f & a + xyb - yf - xe \end{bmatrix} \\ &= \begin{bmatrix} b & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} \mathcal{O}(e^2) & \mathcal{O}(e^3) \\ \mathcal{O}(e^3) & \mathcal{O}(e^2) \end{bmatrix}. \end{aligned}$$

This shows again that we are nearly diagonalizing B and are moreover swapping the eigenvalues. Notice also that two such matrices commute

$$\begin{bmatrix} a_1 & e_1 \\ f_1 & b_1 \end{bmatrix} \begin{bmatrix} a_2 & e_2 \\ f_2 & b_2 \end{bmatrix} = \begin{bmatrix} a_2 & e_2 \\ f_2 & b_2 \end{bmatrix} \begin{bmatrix} a_1 & e_1 \\ f_1 & b_1 \end{bmatrix}$$

if and only if the corresponding values of x and y are equal:

$$a_1 e_2 + e_1 b_2 = a_2 e_1 + e_2 b_1, \quad \Leftrightarrow \quad \frac{e_1}{(b_1 - a_1)} = \frac{e_2}{(b_2 - a_2)}.$$

$$f_1 a_2 + b_1 f_2 = f_2 a_1 + b_2 f_1, \quad \Leftrightarrow \quad \frac{f_1}{(b_1 - a_1)} = \frac{f_2}{(b_2 - a_2)}.$$

It is obvious how to extend the total least squares technique of the previous subsection to the calculation of x and y of the non-Hermitian problem.

3. Commuting pencils

We recall here first a necessary and sufficient test that was presented in [1] for the commutativity of two general $n \times n$ pencils $zB_1 - A_1$ and $zB_2 - A_2$. It was shown there that in order for this problem to make sense, one first has to assume that the pencil $\begin{bmatrix} -A_1 & zB_1 \\ zB_2 & -A_2 \end{bmatrix}$ must be regular. Under this condition, the pencils $zB_1 - A_1$ and $zB_2 - A_2$ commute if and only if

$$\operatorname{rank}(\begin{bmatrix} -A_{+2}A_1 & B_{+1}B_2 \\ -A_{+1}A_2 & B_{+2}B_1 \end{bmatrix}) = n, \tag{7}$$

where A_{+1}, A_{+2}, B_{+1} and B_{+2} are defined via

$$\ker \begin{bmatrix} A_{+2} & B_{+1} \end{bmatrix} = \operatorname{Im}(\begin{bmatrix} B_1 \\ -A_2 \end{bmatrix}), \quad \ker \begin{bmatrix} A_{+1} & B_{+2} \end{bmatrix} = \operatorname{Im}(\begin{bmatrix} B_2 \\ -A_1 \end{bmatrix}).$$

Moreover, the pencils $zB_1 - A_1$ and $zB_2 - A_2$ will then also be regular. Such a test is very useful since it also applies to the case where B_1 and B_2 are singular. But if we are considering a set $\mathcal{P}$ with m pencils $zB_j - A_j$, then such a rank test has to be performed on all $m(m-1)/2$ distinct pairs of pencils. Notice that if the matrices B_1 and B_2 are invertible, then the rank condition (7) is equivalent to the commutativity of the equivalent monic pencils $zI_n - B_1^{-1}A_1$ and $zI_n - B_2^{-1}A_2$, meaning

$$B_2^{-1}A_2B_1^{-1}A_1 = B_1^{-1}A_1B_2^{-1}A.$$

This thus shows that it is a natural extension of the commutativity of matrices. We will assume in Subsection 3.1 that the B_j matrices of the set $\mathcal{P}$ are all invertible. The pencils $zB_j - A_j$ of a commuting set are then just characterized by the fact that the matrices $B_j^{-1}A_j$ are commuting matrices. We can then straightforwardly apply the properties of commuting matrices to derive similar properties for commuting pencils. The case of singular B_j matrices will be treated afterwards in Subsection 3.2 using Möbius transformations.

3.1. Decompositions

Theorem 3.1. *Let the $n \times n$ regular pencil $zB - A$ have two complementary deflating subspaces $\mathcal{T}_1$ and $\mathcal{T}_2$ of respective dimensions n_1 and $n_2 = n - n_1$, and with disjoint spectrum, then there exist invertible transformations S and T such that*

$$S(zB - A)T = \begin{bmatrix} zB_{1,1} - A_{1,1} & 0 \\ 0 & zB_{2,2} - A_{2,2} \end{bmatrix}, \tag{8}$$

where the spectra of the $n_i \times n_i$ subpencils $zB_{i,i} - A_{i,i}$, $i = 1, 2$, are disjoint. The block columns of $T := \begin{bmatrix} T_1 & T_2 \end{bmatrix}$ are bases for the deflating subspaces $\mathcal{T}_1$ and $\mathcal{T}_2$, and the block columns of $S^{-1} := \begin{bmatrix} S_1 & S_2 \end{bmatrix}$ are bases for the corresponding spaces $\mathcal{S}_1 := A\mathcal{T}_1 + B\mathcal{T}_1$ and $\mathcal{S}_2 := A\mathcal{T}_2 + B\mathcal{T}_2$.

Proof. The matrix $T = \begin{bmatrix} T_1 & T_2 \end{bmatrix}$ whose block columns span the two complementary deflating subspaces, yields

$$T^{-1}B^{-1}AT = \begin{bmatrix} M_{1,1} & 0 \\ 0 & M_{2,2} \end{bmatrix},$$

since these are also invariant subspaces for $B^{-1}A$. Assuming B invertible, we also have

$$B \begin{bmatrix} T_1 & T_2 \end{bmatrix} \begin{bmatrix} M_{1,1} & 0 \\ 0 & M_{2,2} \end{bmatrix} = A \begin{bmatrix} T_1 & T_2 \end{bmatrix}$$

and we can construct arbitrary bases for the spaces $\mathcal{S}_1$ and $\mathcal{S}_2$ using

$$\begin{bmatrix} S_1 & S_2 \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} := B \begin{bmatrix} T_1 & T_2 \end{bmatrix}, \quad \det X_1 \neq 0, \quad \det X_2 \neq 0.$$

If we then define $S^{-1} := \begin{bmatrix} S_1 & S_2 \end{bmatrix}$, we obtain

$$SBT = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}, \quad SAT = \begin{bmatrix} X_1 M_{1,1} & 0 \\ 0 & X_2 M_{2,2} \end{bmatrix}.$$

We point out that $X_1 = I$ and $X_2 = I$ are obvious solutions. □

With this we can extend the properties of commuting matrices to that of commuting pencils.

Definition 3.2. A set of pencils $\mathcal{P} := \{zB_j - A_j, j = 1, \ldots, m\}$ with B_j invertible is said to be commuting if every pair of left quotients $(B_i^{-1}A_i, B_j^{-1}A_j)$ commutes.

Theorem 3.3. *If a pencil $zB - A$ from a commuting set $\mathcal{P} := \{zB_j - A_j, j = 1, \ldots, m\}$ has a block-diagonal decomposition*

$$S(zB - A)T = \begin{bmatrix} zB_{1,1} - A_{1,1} & 0 \\ 0 & zB_{2,2} - A_{2,2} \end{bmatrix}, \tag{9}$$

where the spectra of the $n_i \times n_i$ subpencils $zB_{i,i} - A_{i,i}$, $i = 1, 2$, are disjoint, then all pencils $zB_j - A_j$ in that set have such a block diagonal decomposition, using equivalence transformations (S_j, T) with a common right transformation T.

Proof. The block diagonal decomposition (9) with disjoint spectra implies that the deflating subspaces $\mathcal{T}_1$, $\mathcal{T}_2$, are unique and hence, following Theorem 3.1, we can choose a common right transformation $T = \begin{bmatrix} T_1 & T_2 \end{bmatrix}$ for all quotients

$$T^{-1}B_j^{-1}A_jT = \begin{bmatrix} M_{j_{1,1}} & 0 \\ 0 & M_{j_{2,2}} \end{bmatrix},$$

and hence also for the right transformation of all pencils. For the left transformations S_j it suffices to choose $S_jB_jT = I_n$ to obtain the resulting block diagonal decomposition with $B_{j_{i,i}} = I_{n_i}$, and $A_{j_{i,i}} = M_{j_{i,i}}$, $i = 1, 2$. □

Corollary 3.4. *If a pencil $(zB - A) \in \mathcal{P}$ from a commuting set has a block-diagonal decomposition*

$$S(zB - A)T = \operatorname{diag}(zB_{1,1} - A_{1,1}, \ldots, zB_{k,k} - A_{k,k}), \quad B_{i,i}, A_{i,i} \in \mathbb{C}^{n_i \times n_i},$$

where the spectra of the diagonal pencils $zB_{i,i} - A_{i,i}$ are disjoint, then every pencil in $\mathcal{P}$ has such a decomposition, using equivalence transformations (S_j, T) with the same right transformation T.

Corollary 3.5. *A set of diagonalizable pencils $\mathcal{P}$ is a commuting set if and only if the pencils are simultaneously diagonalizable using equivalence transformations (S_j, T) with the same right transformation T.*

Proof. The "if" direction is obvious since diagonal pencils commute and this property is preserved under equivalence transformation with the same right transformation T. For the "only if" part we rely on Theorem 2.5 which says that since the diagonalizable matrices $B_j^{-1}A_j$ commute, there is a common transformation T that diagonalizes them all, i.e. $T^{-1}B_j^{-1}A_jT = D_j$. But then setting $S_j := (B_jT)^{-1}$ yields $S_jA_jT = D_j$, which shows that

$$S_j(zB_j - A_j)T = zI - D_j. \quad \square$$

And finally, the following property follows from the fact that commuting pencils must have a common right eigenvector.

Theorem 3.6. *If $\mathcal{P}$ is a set of commuting pencils, then there exist unitary equivalence transformations (V_j^{H}, U) with a common right transformation U, that triangularize each pencil in $\mathcal{P}$ via $V_j^{\mathsf{H}}(zB_j - A_j)U$.*

Proof. All pencils of a commuting set must have a common (normalized) right eigenvector u_1 because the kernel of any $B_j^{-1}A_j - \lambda I_n$ at an eigenvalue λ of $B_j^{-1}A_j$, is also $B_i^{-1}A_i$-invariant when the pencils commute. Let U have this vector u_1 as first column.

Moreover, the vectors $(\lambda B_j - A_j)u_1$ are zero and hence $B_j u_1$ and $A_j u_1$ are parallel to a normalized vector v_{j_1} that we choose as first column of the unitary transformations V_j. It then follows that the first column of each pencil $z\hat{B}_j - \hat{A}_j := V_j^{\mathsf{H}}(zB_j - A_j)U$ is parallel to $\mathbf{e}_1$. Therefore these pencils are block-triangular with a leading 1×1 block, and a trailing $(n-1) \times (n-1)$ block. Moreover, these diagonal blocks are commuting as well, and we can therefore repeat this argument in a recursive fashion on the trailing subpencils, like in the proof of Theorem 2.7. The similarity matrices that are constructed at each step, are unitary, and so is their product. □

Definition 3.7 *([5]).* A pencil $zB - A$ is normal if it has a decomposition $V^{\mathsf{H}}(zB - A)U = zD_b - D_a$, where U and V are unitary, and D_b and D_a are diagonal. If B is invertible, then it follows that $B^{-1}A$ and AB^{-1} are both normal, whence the name "normal pencil".

Corollary 3.8. *If the pencils of a commuting set $\mathcal{P}$ are all normal, then every pencil in $\mathcal{P}$ can be diagonalized using unitary equivalence transformation (V_j, U) with the same right transformation U.*

Proof. This follows from the triangular decomposition of Theorem 3.6, since an upper triangular normal pencil is also diagonal. □

3.2. Möbius transformation

We show that the condition of invertibility of the B_j matrices can be removed in all statements of the previous subsection, by making use of a special Möbius transform.

Lemma 3.9 *(see [16]). Let the pencil $zB - A$ be regular, then there always exists an "orthogonal" Möbius transformation*

$$\begin{bmatrix} B^{(\theta)} & A^{(\theta)} \end{bmatrix} := \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} cI_n & sI_n \\ -sI_n & cI_n \end{bmatrix}, \quad \text{where} \quad \begin{matrix} c := \cos(\theta) \\ s := \sin(\theta) \end{matrix}, \tag{10}$$

such that $B^{(\theta)}$ is now invertible. Moreover, the eigenvalues $z_i^{(\theta)}$ of the transformed pencil also satisfy $z_i^{(\theta)} = (cz_i - s)/(sz_i + s)$, while the corresponding eigenvectors or deflating subspaces do not change.

Invertible transformations S and T applied to $zB - A$ and $zB^{(\theta)} - A^{(\theta)}$ can be chosen to be the same, since the block transformation $\operatorname{diag}(T, T)$ commutes with the block transformation $\begin{bmatrix} cI_n & sI_n \\ -sI_n & cI_n \end{bmatrix}$. Therefore, (10) implies that

$$\begin{bmatrix} SB^{(\theta)}T & SA^{(\theta)}T \end{bmatrix} := \begin{bmatrix} SBT & SAT \end{bmatrix} \begin{bmatrix} cI_n & sI_n \\ -sI_n & cI_n \end{bmatrix}, \tag{11}$$

and the inverse transformation yields

$$\begin{bmatrix} SB^{(\theta)}T & SA^{(\theta)}T \end{bmatrix} \begin{bmatrix} cI_n & -sI_n \\ sI_n & cI_n \end{bmatrix} = \begin{bmatrix} SBT & SAT \end{bmatrix}.$$

Therefore, all decompositions derived in Section 3.1 still hold, even when the matrices B_j are singular, because the invertibility of the B_j matrices can always be ensured by first performing a Möbius transform, without altering the form of their decomposition.

3.3. Optimization

We again select two decompositions that can lead to optimization problems with ultimate quadratic convergence. The first one is the set of *commuting diagonalizable pencils.* They can be diagonalized simultaneously over the group $\mathcal{G}_{m+1} := \{S_1, \ldots, S_m, T\}$ of $m+1$ invertible linear transformations. Therefore the minimization problem

$$\min_{\mathcal{G}_{m+1}} \| [S_1A_1T, S_1B_1T, \ldots, S_mA_mT, S_mB_mT] \|_{of} \tag{12}$$

has a zero minimum if and only if this set commutes. The triangularization of diagonalizable pencils will again not be considered because this only gives a necessary condition for commutativity.

The second one is the set of *commuting normal pencils*, which can simultaneously be diagonalized over the group $\mathcal{U}_{m+1} := \{V_1, \ldots, V_m, U\}$ of $m+1$ unitary transformations. Therefore the following minimization problem

$$\min_{\mathcal{U}_{m+1}} \| \left[V_1^{\mathsf{H}}A_1U, V_1^{\mathsf{H}}B_1U, \ldots, V_m^{\mathsf{H}}A_mU, V_m^{\mathsf{H}}B_mU\right] \|_{of} \tag{13}$$

has a zero minimum if and only if the given set of normal pencils commutes.

3.4. Diagonalizable pencils

We use the idea that the B_i matrices are invertible, which is a reasonable assumption since we can always perform a Möbius transform to make sure that the coefficients B_i are invertible. Next we make sure that the B_i matrix of the pencils is upper triangular, which can be obtained by left-multiplying the pencil $zB_i - A_i$ with Q_i^{H}, as obtained from the QR factorization of B_i. We then implement the above ideas for matrices to the products $B_i^{-1}A_i$, or rather to their 2×2 diagonal sub-matrices. If we can assume that the lower-triangular part of the A_i matrices is small with respect to the gaps $|r_jb_j - q_ja_j|$ of all diagonal subblocks, then the 2×2 diagonal subblocks of $B_i^{-1}A_i$ are reasonably well approximated by the quotients of the corresponding blocks of B_i and A_i. Let us denote them as follows

$$\begin{bmatrix} r & s \\ 0 & q \end{bmatrix}^{-1} \begin{bmatrix} a & e \\ f & b \end{bmatrix} = \frac{1}{rq} \begin{bmatrix} q & -s \\ 0 & r \end{bmatrix} \begin{bmatrix} a & e \\ f & b \end{bmatrix} = \frac{1}{rq} \begin{bmatrix} qa - sf & qe - sb \\ rf & rb \end{bmatrix},$$

from which we obtain the elements x and y to form the updating transformation T, using the total least squares solution of the following overdetermined systems

$$\begin{bmatrix} r_1 f_1 & (r_1 b_1 - q_1 a_1 + s_1 f_1) \\ \vdots & \vdots \\ r_m f_m & (r_m b_m - q_m a_m + s_m f_m) \end{bmatrix} \begin{bmatrix} 1 \\ -x \end{bmatrix} = 0$$

and

$$\begin{bmatrix} (q_1 a_1 - r_1 b_1 - s_1 f_1) & (q_1 e_1 - s_1 b_1) \\ \vdots & \vdots \\ (q_m a_m - r_m b_m - s_m f_m) & (q_m e_m - s_m b_m) \end{bmatrix} \begin{bmatrix} y \\ 1 \end{bmatrix} = 0.$$

The transformation T (scaled to have column norms 1) is then applied to the right of the pencils $\lambda B_i - A_i$, and little 2×2 rotations Z_i are then applied to the left of each pencil $\lambda B_i - A_i$ to restore the triangular form of B_i.

3.5. Normal pencils

This is a special case since the transformation group should be $\mathcal{G}_{m+1}$ while the little 2×2 diagonal pencils are not necessarily normal. We then propose to try to put the matrix in upper Schur form instead, which amounts to selecting x as in the previous section and take $y = x$. With the triangularization of the coefficient of z, this then will put that pencil near to its Schur form, and in the limit, it will tend to a diagonal pencil.

4. Numerical examples

In this section we give numerical evidence of the ultimate quadratic convergence of the Jacobi-like methods for the four optimization problems developed in this paper. They each provide a necessary and sufficient test for the commutativity of certain classes of matrices or pencils: (4) for Hermitian matrices, (5) for diagonalizable matrices, (12) for diagonalizable pencils, and (13) for normal pencils. All computations were performed using MATLAB version 2023a running on a Dell notebook with Intel Core i5 vPro 8th Gen CPU with Windows 10. The machine precision was $\epsilon \approx 2.210^{-16}$.

In the Jacobi-like methods, for sets of matrices $A_j,\ j = 1, \ldots, m$, we propagate $(n-1)$ 2×2 transformations on selected rows and columns of each matrix A_j which therefore requires $12mn(n-1)$ flops. And n such odd-even iterations amount to two sweeps of the Jacobi-like algorithm (see [2]). Each sweep therefore requires roughly $6mn^3$ flops in the matrix case. For pencils of matrices $zB_j - A_j, j = 1, \ldots, m$, one has to add another

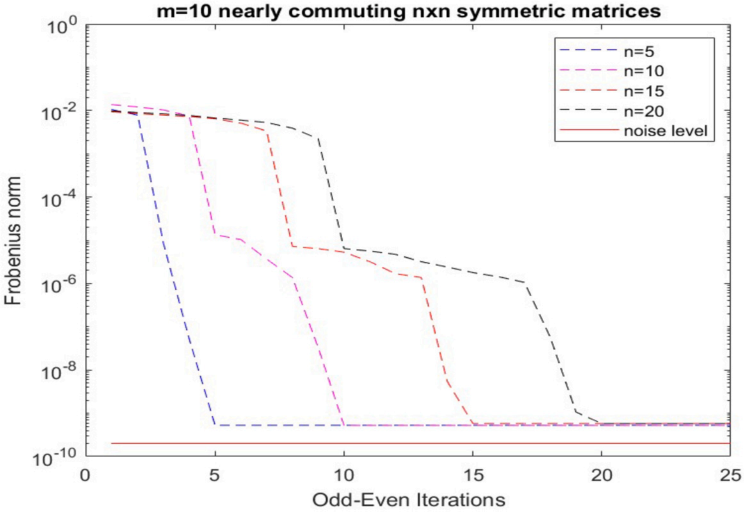

Fig. 1. Evolution of the Off-norm in the Jacobi algorithm for 10 symmetric matrices of dimensions $n = 5, 10, 15, 20$.

$3mn^3$ flops for updating the B_j matrices since they are upper triangular. If one wants to update also the common right transformation matrices, one has to add again $3mn^3$ flops per sweep. But all together, the complexity of the method is $\mathcal{O}(mn^3)$ flops per sweep. As we will see below, the number of sweeps needed for convergence is very reasonable, once we have a relative off-norm of the order of 10^{-2}. So the method is definitely to be recommended for refining solutions.

In the first set of experiments we ran the Jacobi algorithm on a set $\mathcal{M}$ of m real symmetric matrices, generated as $A_j = Q^T(diag(d_j) + \delta_1 N_j)Q$ where

- d_j, $j = 1, \ldots, m$, are random vectors of Euclidian norm 1 and N_j, $j = 1, \ldots, m$, are random (noise) matrices of Frobenius norm 1; the level δ_1 indicates that the matrices in the set $\mathcal{M}$ are nearly commuting, with residual of the order of δ_1 (chosen as $2 \cdot 10^{-2}$).
- Q is an orthogonal matrix δ_2-close to the identity ($\delta_2 = 10^{-2}$); this is to emulate that a first estimate of the common diagonalizing transformation was obtained with a different method yielding a common diagonalizing transformation for the first two digits.

In Fig. 1, we indicate the number of odd-even iterations, which have to be divided by $n/2$ to know the number of Jacobi "sweeps". One then observes that convergence occurs in essentially two "sweeps", irrespective of the matrix dimension n.

In the second experiment we ran the Jacobi-like algorithm on a set $\mathcal{M}$ of m real diagonalizable matrices, generated as $A_j = T^{-1}(diag(d_j) + \delta_1 N_j)T$, where

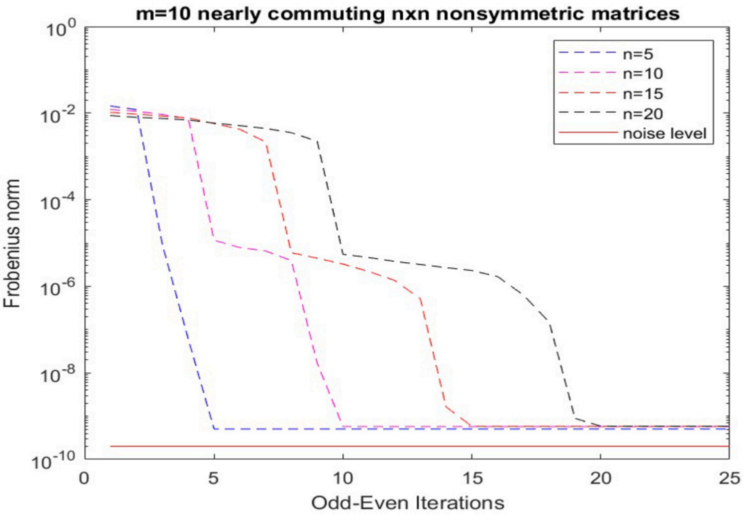

Fig. 2. Evolution of the Off-norm in the Jacobi-like algorithm for 10 unsymmetric matrices of dimensions $n = 5, 10, 15, 20$.

- the d_j vectors, the N_j matrices and δ_1 are chosen as in Experiment 1,
- T is an orthogonal matrix δ_2-close to the identity ($\delta_2 = 10^{-2}$) in order to emulate that a first estimate of the diagonalizing transformation was obtained with 2 digits of accuracy.

Again one observes that convergence occurs roughly in two sweeps for different choices of the matrix dimension n. It is interesting to point out that once converged, the "residual" obtained in both experiments is a fair estimate of the noise level injected in the matrix set (Fig. 2).

The next two experiments are for pencils of matrices. For Experiment 3, we consider a set $\mathcal{P}$ of m real diagonalizable pencils generated using $A_j = S_j(diag(d_{a_j}) + \delta_1 N_{a_j})T$, $B_j = S_j(diag(d_{b_j}) + \delta_1 N_{b_j})T$, where

- the d_{a_j}, d_{b_j} vectors, the N_{a_j}, N_{b_j} matrices and δ_1 are chosen as before,
- T is an invertible matrix δ_2-close to the identity, and the S_j matrices are orthogonal matrices that triangularize the B_j matrices (Fig. 3).

For Experiment 4, we consider a set $\mathcal{P}$ of m real normal pencils generated using $A_j = S_j(diag(d_{a_j}) + \delta_1.N_{a_j})Q$, $B_j = S_j(diag(d_{b_j}) + \delta_1.N_{b_j})Q$, where

- the d_{a_j}, d_{b_j} vectors, the N_{a_j}, N_{b_j} matrices and δ_1 are chosen as before,
- Q is an orthogonal matrix δ_2-close to the identity, and the S_j matrices are orthogonal matrices that triangularize the B_j matrices (Fig. 4).

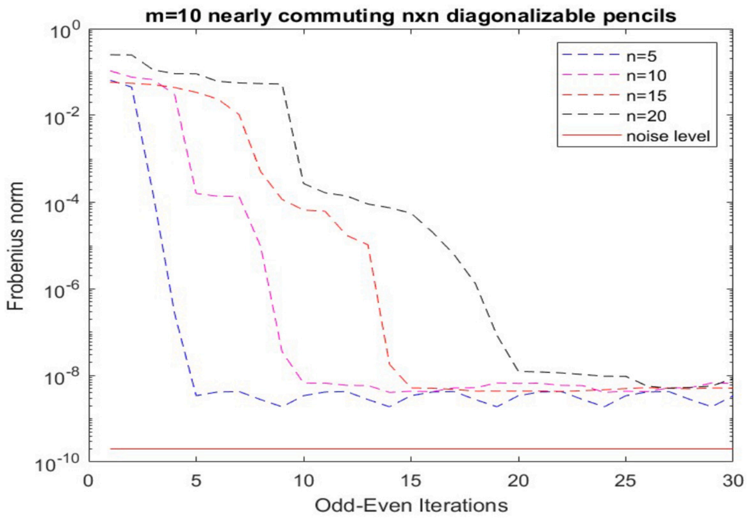

Fig. 3. Evolution of the Off-norm in the Jacobi-like algorithm for 10 diagonalizable pencils of dimensions $n = 5, 10, 15, 20$.

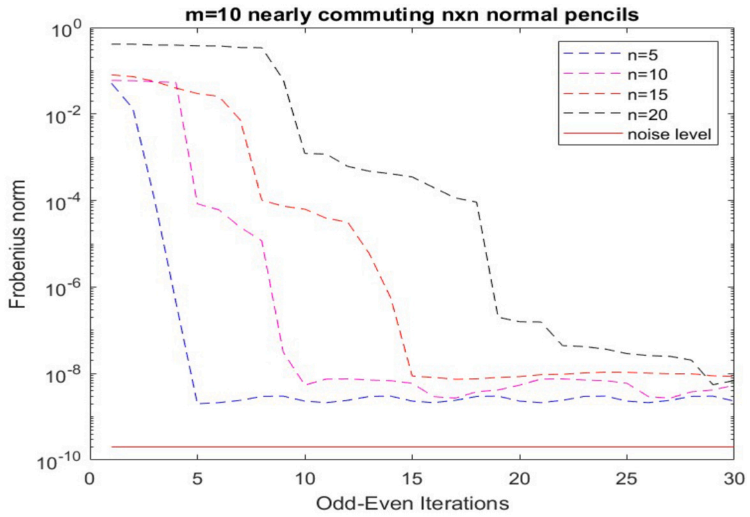

Fig. 4. Evolution of the Off-norm in the Jacobi-like algorithm for 10 normal pencils of dimensions $n = 5, 10, 15, 20$.

In both pencil experiments, we can again observe quadratic convergence, but the off-norm of the residual clearly oscillates, because the inverses of the diagonal blocks of the B_j matrices affect the norm of these blocks.

5. Concluding remarks

In this paper, we extended the classical notion of commuting matrices to that of commuting pencils. This extension was originally introduced in the context of invariance properties of the solution of two-point boundary value problems described by descriptor systems. For pencils $zB_j - A_j$ with invertible matrices B_j, the commutativity boils down to the commutativity of the (left) quotients $B_j^{-1}A_j$, which makes it a natural extension of the matrix problem. It was shown that the use of Möbius transforms allows us to circumvent the invertibility condition of the B_j matrices. We also introduced minimization problems that can serve as necessary and sufficient condition for the commutativity of certain classes of pencils.

Declaration of competing interest

None declared.

Acknowledgement

The research was performed during a visit of the second author to the Max Planck Institute in Magdeburg. We also want to thank Daniel Kressner for useful discussions that led to the research reported here.

Data availability

No data was used for the research described in the article.

References

[1] P. Benner, P. Van Dooren, Periodic two-dimensional descriptor systems, Electron. J. Linear Algebra 39 (2023) 472–490.
[2] R. Brent, F. Luk, The solution of singular value and symmetric eigenvalue problems on multiprocessor arrays, SIAM J. Sci. Stat. Comput. 6 (1) (1986) 69–84.
[3] A. Bunse-Gerstner, R. Byers, V. Mehrmann, Numerical methods for simultaneous diagonalization, SIAM J. Matrix Anal. Appl. 14 (4) (1993) 927–949.
[4] J.F. Cardoso, A. Souloumiac, Jacobi angles for simultaneous diagonalization, SIAM J. Matrix Anal. Appl. 17 (1) (1996) 161–164.
[5] J.P. Charlier, P. Van Dooren, A Jacobi-like algorithm for computing the generalized Schur form of a regular pencil, J. Comput. Appl. Math. 27 (1989) 17–36.
[6] Robert M. Corless, Gianni M. Patrizia, Barry M. Trager, A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots, in: Proc. 1997 International Symposium on Symbolic and Algebraic Computation (ISSAC), 1997, pp. 133–140.
[7] Philip J. Davis, Circulant Matrices, Wiley-Interscience, 1979.
[8] L. De Lathauwer, A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization, SIAM J. Matrix Anal. Appl. 28 (3) (2006) 642–666.
[9] P. Eberlein, J. Boothroyd, Solution to the eigenproblem by a norm reducing Jacobi type method, Numer. Math. 11 (1968) 1–12.
[10] F.R. Gantmacher, The Theory of Matrices, vol. I, Chelsea Publ., New York, 1959.

[11] H. He, D. Kressner, Randomized joint diagonalization of symmetric matrices, SIAM J. Matrix Anal. Appl. 45 (1) (2024) 661–684.
[12] N. Higham, Functions of Matrices: Theory and Computation, SIAM Publications, Philadelphia, 2008.
[13] R. Horn, C. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[14] M. Kilmer, C. Martin, Factorization strategies for third order tensors, Linear Algebra Appl. 435 (2011) 641–658.
[15] D.G. Luenberger, Dynamic equations in descriptor form, IEEE Trans. Autom. Control 22 (6) (1977) 312–321.
[16] D.S. Mackey, N. Mackey, C. Mehl, V. Mehrmann, Möbius transformations of matrix polynomials Lin, Algebra Appl. 470 (2015) 120–184.
[17] G.W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, SIAM Rev. 15 (4) (1973) 727–764.
[18] David S. Watkins, The Matrix Eigenvalue Problem, SIAM Publications, Philadelphia, 2007.
[19] J.H. Wilkinson, Note on the quadratic convergence of the cyclic Jacobi process, Numer. Math. 4 (1962) 296–300.