Noname manuscript No.
(will be inserted by the editor)

Simultaneous multidiagonalization for the CS decomposition

Kingston Kang · William Lothian · Jessica Sears · Brian D. Sutton

the date of receipt and acceptance should be inserted later

Abstract When an orthogonal matrix is partitioned into a two-by-two block structure, its four blocks can be simultaneously bidiagonalized. This observation underlies numerically stable algorithms for the CS decomposition and the existence of CMV matrices for orthogonal polynomial recurrences. We discover a new matrix decomposition for simultaneous multidiagonalization, which reduces the blocks to any desired bandwidth. Its existence is proved, and a backward stable algorithm is developed. The resulting matrix with banded blocks is parameterized by a product of Givens rotations, guaranteeing orthogonality even on a finite-precision computer. The algorithm relies heavily on Level 3 BLAS routines and supports parallel computation.

Keywords CS decomposition · CMV matrix

Mathematics Subject Classification (2000) 15B10 · 15A23 · 65F30

1 Introduction

As a first step to the CS decomposition (CSD), we seek to simultaneously reduce the four blocks of a partitioned orthogonal matrix to low bandwidth. Specifically, the desired decomposition is

$$\underbrace{\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}}_{\text{orthogonal}} = \underbrace{\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}}_{\text{orthogonal}} \underbrace{\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}}_{\substack{\text{orthogonal,} \\ \text{low-bandwidth } B_{ij}}} \underbrace{\begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T}_{\text{orthogonal}}. \tag{1}$$

This material is based upon work supported by the National Science Foundation under Grant No. DMS-0914559.

Department of Mathematics, Randolph-Macon College, PO Box 5005, Ashland, VA 23005 USA

We ultimately obtain this decomposition with B_{11}, B_{21} upper b-diagonal and B_{12}, B_{22} lower b-diagonal. For example, if $b = 3$, then the reduced matrix has the form

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{ccccc|ccccc} \times & \times & \times & & & \times & & & & \\ & \times & \times & \times & & \times & \times & & & \\ & & \times & \times & \times & \times & \times & \times & & \\ & & & \times & \times & & \times & \times & \times & \\ & & & & \times & & & \times & \times & \times \\ \hline \times & \times & \times & & & \times & & & & \\ & \times & \times & \times & & \times & \times & & & \\ & & \times & \times & \times & \times & \times & \times & & \\ & & & \times & \times & & \times & \times & \times & \\ & & & & \times & & & \times & \times & \times \end{array}\right]. \quad (2)$$

By subsequently diagonalizing the blocks, the CSD could be obtained. An algorithm for the decomposition is presented. It generalizes earlier work for the bidiagonal, i.e., 2-diagonal, case [19, 20] and is inspired by recent communication-avoiding algorithms, especially band bidiagonal reduction for the singular value decomposition (SVD) [7, 13].

The resulting contributions are both theoretical and practical. On the theoretical side, a new matrix decomposition is discovered. It generalizes the CS decomposition pioneered by Davis-Kahan, Björck-Golub, and Stewart [3, 6, 17] and the bidiagonal-block form unveiled in stages and in different guises by Ammar-Gragg-Reichel, Bunse-Gerstner-Elner, Cantero-Moral-Velázquez, and Edelman-Sutton [1, 4, 5, 10]. Paige and Wei and Simon have written histories of these developments [15, 16]. The decomposition is especially interesting because it provides a parameterization of the Stiefel manifold. On the computational side, the new *simultaneous multidiagonalization* algorithm is backward stable and relies heavily on Level 3 BLAS routines (specifically the QR decomposition and matrix multiplication), opening new possibilities for cache utilization, parallel computation, and restrained communication. As far as we are aware, earlier work on the CSD does not address such computational concerns [2, 8, 9, 11, 14, 18, 19, 20, 22].

The algorithm does not actually produce the entries of (2). Rather, it returns a sequence of Givens rotations that generate this matrix, analogously to how latitude and longitude locate a point on the sphere. If desired, the entries can be recovered by applying the rotations to the identity matrix. However, as a product of Givens rotations, the matrix is exactly orthogonal, even if the angles are stored in finite precision. In addition, the angles may prove to have analytic or geometric significance, as the angles in the bidiagonal case reveal orthogonal polynomial recurrences. (Orthogonal matrices are not explicit in early work by Verblunsky and Szegő, but they appear in a 1993 paper by Watkins. More recently, the name *CMV matrix* has been attached to banded, orthogonal matrices in this context [16, 21, 23, 24].)

We assume that all four submatrices X_{ij} are n-by-n. Earlier work has dealt with nonuniform partitions [19].

The rest of the paper is organized as follows: Section 2 reviews the bidiagonal case from previous work. Section 3 defines and proves properties for a specific notion of a block Givens rotation, which is essential for numerical stability and for parameterizing the final result of the coming algorithm. Section 4 proves half of the existence theorem for the decomposition (1). Section 5 specifies the simultaneous multidiagonalization algorithm, completes the proof of existence, and proves numerical stability. Section 6 presents a numerical example.

2 The bidiagonal case

We review simultaneous bidiagonalization in preparation for simultaneous multidiagonalization.

Let

$$Y = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{3}$$

be an orthogonal matrix with n-by-n blocks, B_{11} and B_{21} upper-bidiagonal and B_{12} and B_{22} lower-bidiagonal. Such a matrix can be represented as a product

$$Y = (G_1 \cdots G_n) D (H_{n-1}^T \cdots H_1^T), \tag{4}$$

in which $G_1, \ldots, G_n$ and $H_1, \ldots, H_{n-1}$ are Givens rotations satisfying

$$G_i\,[i, n+i] = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}, \qquad H_i\,[i+1, n+i] = \begin{bmatrix} \cos\phi_i & -\sin\phi_i \\ \sin\phi_i & \cos\phi_i \end{bmatrix},$$

and $D = \operatorname{diag}(1, \ldots, 1, \pm 1)$ [19]. (Notationally, $A[i_1, \ldots, i_r]$ is the principal submatrix of A lying in rows and columns $i_1, \ldots, i_r$.) Vice versa, any product of the form (4) is an orthogonal matrix with bidiagonal blocks of the form (3).

An arbitrary orthogonal matrix X can be reduced to bidiagonal-block form by orthogonal transformations:

$$\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^T X \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^T \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

The reduction is illustrated by Fig. 1. First, a sequence of Householder reflectors and Givens rotations reduces X to a quasi-triangular matrix as in Fig. 1(a). The sequence is as follows: Householder from left (P_i), Givens from left (G_i), Householder from right (Q_i), Givens from right (H_i), repeat. This produces

$$(G_n^T P_n \cdots G_2^T P_2 G_1^T P_1) X (Q_1 H_1 Q_2 H_2 \cdots Q_{n-1} H_{n-1}) D = \begin{bmatrix} L_{11} & Z \\ 0 & L_{22} \end{bmatrix}$$

with L_{11} upper triangular and L_{22} lower triangular with nonnegative diagonals. (The diagonal signature matrix $D = \operatorname{diag}(1, \ldots, 1, \pm 1)$ accomplishes the final sign change in Fig. 1(a).) The form on the right-hand side is what we mean by *quasi-triangular*. If X is orthogonal, then the right-hand side must also be orthogonal and therefore the identity matrix I:

$$(G_n^T P_n \cdots G_2^T P_2 G_1^T P_1) X (Q_1 H_1 Q_2 H_2 \cdots Q_{n-1} H_{n-1}) D = I.$$

(See Fig. 1(b).) In addition, every Givens rotation G_i commutes with every direct sum of Householder reflectors P_j with $j > i$ and every H_i commutes with every Q_j with $j > i$ because they operate on different rows or columns. So, the Givens rotations and Householder reflectors can be separated:

$$(P_n \cdots P_2 P_1) X (Q_1 Q_2 \cdots Q_{n-1}) = (G_1 G_2 \cdots G_n) D (H_{n-1}^T \cdots H_2^T H_1^T).$$

(See Fig. 1(c).) By (4), this is equivalent to

$$\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^T X \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

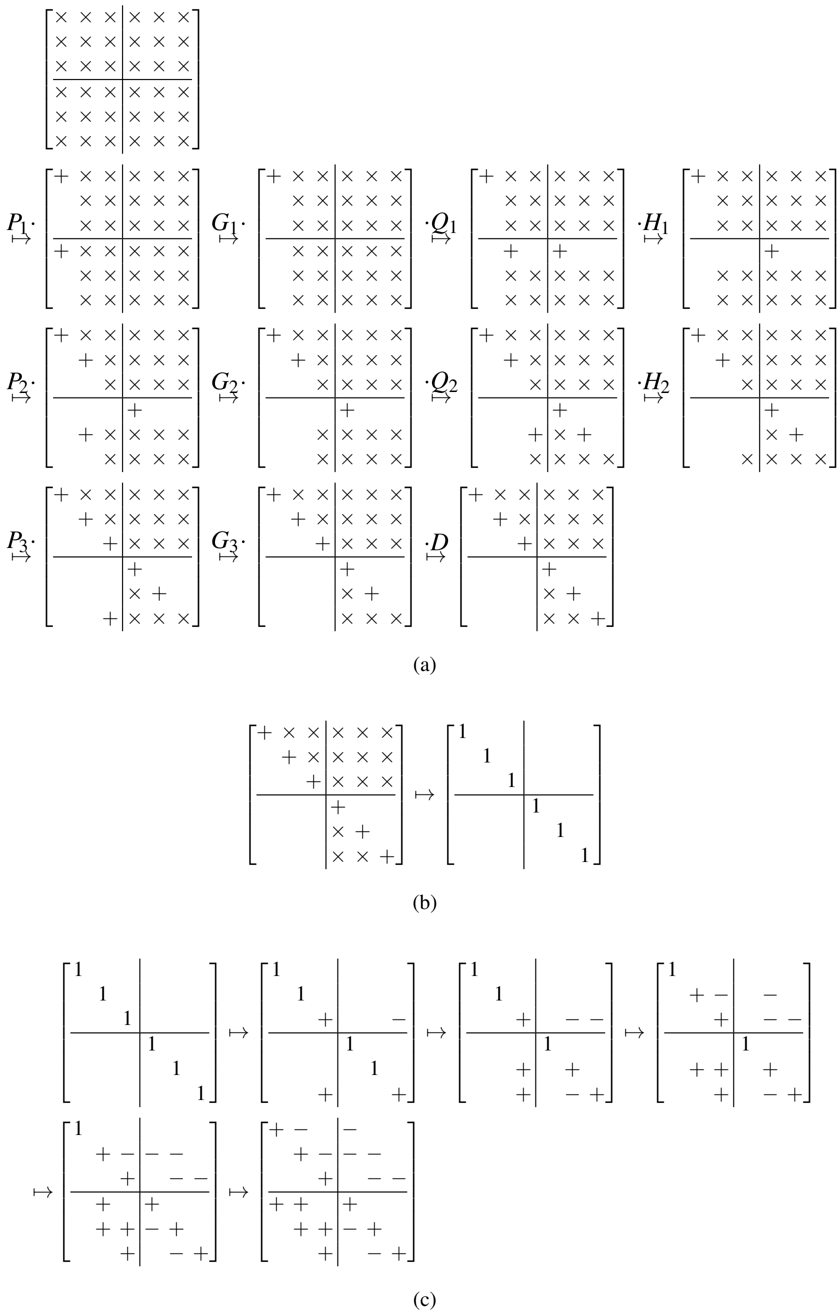

Fig. 1 Review of simultaneous bidiagonalization. (a) The input matrix is reduced to quasi-triangular form with Householder reflectors and Givens rotations. (b) A quasi-triangular matrix that is orthogonal and has positive diagonal is the identity matrix. (c) The Givens rotations are inverted.

In production code, there is no need to compute the off-diagonal entries in Fig. 1(a) that are truncated to zero in Fig. 1(b), or even the diagonal entries that are rounded to one. However, for the stability analysis, it is useful to imagine that these entries are explicitly computed.

In developing a simultaneous *multi*diagonalization algorithm, every Householder reflector will be replaced by the orthogonal factor in a QR decomposition and every Givens rotation by a block Givens rotation. Care needs to be taken, though. Although it is easy to replace scalar entries by $(b-1)$-by-$(b-1)$ matrices, achieving half-bandwidth b requires that those small matrices be triangular. It is not obvious how to simultaneously triangularize two submatrices—or if it is even possible. The next section develops a particular type of block Givens rotation with triangular structures that produces the desired form.

3 Block Givens rotations

In the familiar bidiagonal case, the input X is reduced to output Y with two properties: (1) Y is orthogonal and (2) the blocks of Y are bidiagonal. In the more general setting of simultaneous multidiagonalization, the desired output form is

$$\left[\begin{array}{ccccc|ccccc}
R & L & & & & L & & & & \\
& R & L & & & R & L & & & \\
& & \ddots & \ddots & & & \ddots & \ddots & & \\
& & & R & L & & & R & L & \\
& & & & R & & & & R & L \\
\hline
R & L & & & & L & & & & \\
& R & L & & & R & L & & & \\
& & \ddots & \ddots & & & \ddots & \ddots & & \\
& & & R & L & & & R & L & \\
& & & & R & & & & R & L
\end{array}\right],$$

in which R's represent $(b-1)$-by-$(b-1)$ upper-triangular matrices and L's represent $(b-1)$-by-$(b-1)$ lower-triangular ones. To achieve this, Givens rotations are replaced by *block Givens rotations*, defined below. *Three* properties are required: (1) the output is orthogonal, (2) the blank entries in the matrix above are zero, and (3) the R's are upper-triangular and the L's are lower-triangular.

The third property requires a special notion of a block Givens rotation that itself has triangular structures. This is the matrix G of the following lemma, or any permutation of such a matrix. The proof describes how to compute one.

Lemma 1 *Given a $2n$-by-n matrix with upper-triangular square blocks,*

$$\begin{bmatrix} R_1 \\ R_2 \end{bmatrix},$$

there exists an orthogonal matrix G satisfying

$$G = \begin{bmatrix} R_{11} & L_{12} \\ R_{21} & L_{22} \end{bmatrix} \quad \text{and} \quad G^T \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} R_3 \\ 0 \end{bmatrix},$$

with R_{11}, R_{21}, and R_3 upper-triangular and L_{12}, L_{22} lower-triangular.

Proof The proof is by induction on n.

When $n = 1$, let G equal a Givens rotation that eliminates R_2.

For $n > 1$, express the input matrix as

$$\left[\begin{array}{cc} A & c \\ & d \\ \hline B & e \\ & f \end{array}\right].$$

By induction, there exists a transformation satisfying

$$\left[\begin{array}{cc|cc} J & & M & \\ & 1 & & \\ \hline K & & N & \\ & & & 1 \end{array}\right]^T \left[\begin{array}{cc} A & c \\ & d \\ \hline B & e \\ & f \end{array}\right] = \left[\begin{array}{cc} P & J^T c + K^T e \\ 0 & d \\ \hline 0 & M^T c + N^T e \\ 0 & f \end{array}\right],$$

with J, K, and P upper-triangular and M and N lower-triangular. Construct a sequence of Givens rotations $Z_1, \ldots, Z_n$ so that

$$Z_n^T \cdots Z_1^T \left[\begin{array}{cc|cc} J & & M & \\ & 1 & & \\ \hline K & & N & \\ & & & 1 \end{array}\right]^T \left[\begin{array}{cc} A & c \\ & d \\ \hline B & e \\ & f \end{array}\right] = \left[\begin{array}{cc} P & J^T c + K^T e \\ 0 & q \\ \hline 0 & 0 \\ 0 & 0 \end{array}\right],$$

with Z_i acting on rows n and $2n+1-i$. The product of Givens rotations has the form

$$Z_1 \cdots Z_n = \left[\begin{array}{cc|cc} I & & & \\ & r & u & x \\ \hline & s & V & \\ & t & w & y \end{array}\right],$$

with V lower-triangular, so

$$\left[\begin{array}{cc|cc} J & & M & \\ & 1 & & \\ \hline K & & N & \\ & & & 1 \end{array}\right] Z_1 \cdots Z_n = \left[\begin{array}{cc|cc} J & Ms & MV & \\ & r & u & x \\ \hline K & Ns & NV & \\ & t & w & y \end{array}\right].$$

This is the desired transformation. □

4 Parameterization by block Givens rotations

Block Givens rotations provide a convenient, efficient, and robust representation for partitioned orthogonal matrices

$$X = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{5}$$

whose submatrices are b-diagonal. Let $k = \lceil n/(b-1) \rceil$. Let G_i, $i = 1, \ldots, k$, be a $2n$-by-$2n$ block Givens rotation

$$G_i = \left[\begin{array}{ccc|ccc} I_{b(i-1)} & & & & & \\ & R_{i,i} & & & L_{i,k+i} & \\ & & I_{n-bi} & & & \\ \hline & & & I_{b(i-1)} & & \\ & R_{k+i,i} & & & L_{k+i,k+i} & \\ & & & & & I_{n-bi} \end{array}\right], \tag{6}$$

and let H_i, $i = 1, \ldots, k-1$, be a $2n$-by-$2n$ block Givens rotation

$$H_i = \left[\begin{array}{ccc|ccc} I_{bi} & & & & & \\ & L_{i+1,i+1} & & & R_{i+1,k+i} & \\ & & I_{n-b(i+1)} & & & \\ \hline & & & I_{b(i-1)} & & \\ & L_{k+i,i+1} & & & R_{k+i,k+i} & \\ & & & & & I_{n-bi} \end{array}\right]. \tag{7}$$

($L_{\cdot,\cdot}$ and $R_{\cdot,\cdot}$ are lower triangular and upper triangular, respectively, as in Lemma 1.) Finally, let $D = \operatorname{diag}(D_1, \ldots, D_{2k})$ be a $2n$-by-$2n$ diagonal signature matrix.

Lemma 2 *Let G_i, H_i, and D be defined as above. Then*

$$X = (G_1 \cdots G_k) D (H_{k-1}^T \cdots H_1^T) \tag{8}$$

is an orthogonal matrix of the form

$$X = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{9}$$

with B_{11} and B_{21} upper b-diagonal and B_{12} and B_{22} lower b-diagonal.

Proof The block Givens rotations $G_1, \ldots, G_k$ act on distinct rows, and we find

$$G_1 \cdots G_k = \left[\begin{array}{cccc|cccc} R & & & & L & & & \\ & R & & & & L & & \\ & & \ddots & & & & \ddots & \\ & & & R & & & & L \\ \hline R & & & & L & & & \\ & R & & & & L & & \\ & & \ddots & & & & \ddots & \\ & & & R & & & & L \end{array}\right],$$

in which every R represents a (generally) different upper-triangular matrix and every L represents a (generally) different lower-triangular matrix. Every R and L is $(b-1)$-by-$(b-1)$ except possibly the four in the lower-right corners, which are $(n-(b-1)(k-1))$-by-$(n-(b-1)(k-1))$. Also, the block Givens rotations $H_1, \ldots, H_{k-1}$ act on distinct columns, and we find

$$H_1 \cdots H_{k-1} = \left[\begin{array}{cccc|cccc} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & L & & & R & & & 0 \\ 0 & & \ddots & & & \ddots & & 0 \\ 0 & & & L & & & R & 0 \\ \hline 0 & L & & & R & & & 0 \\ 0 & & \ddots & & & \ddots & & 0 \\ 0 & & & L & & & R & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \end{array}\right],$$

using the same notational conventions and block sizes. Hence, the product has the form

$$(G_1 \cdots G_k) D (H_{k-1}^T \cdots H_1^T) = \left[\begin{array}{cccc|cccc} RD & LDR^T & & & LDR^T & & & \\ & RDL^T & LDR^T & & RDL^T & LDR^T & & \\ & & \ddots & \ddots & & \ddots & \ddots & \\ & & & RDL^T & & & RDL^T & LD \\ \hline RD & LDR^T & & & LDR^T & & & \\ & RDL^T & LDR^T & & RDL^T & LDR^T & & \\ & & \ddots & \ddots & & \ddots & \ddots & \\ & & & RDL^T & & & RDL^T & LD \end{array}\right],$$

in which each D is a submatrix D_i of the $2n$-by-$2n$ diagonal signature matrix with the subscript suppressed. □

Theorem 1 will show that the reverse is true as well, specifically that any orthogonal X with b-diagonal blocks can be expressed as a product of block Givens rotations.

5 Simultaneous multidiagonalization

Now, we develop the new simultaneous multidiagonalization algorithm and prove that it is backward stable. The existence of the decomposition (1) follows.

The new algorithm follows the outline of Fig. 1. However, it deals with submatrices instead of individual entries, QR factorizations instead of single Householder reflectors, block Givens rotations instead of classical ones, and it relies on Lemmas 1 and 2 to produce banded forms.

Consider a $2n$-by-$2n$ matrix

$$A = \left[\begin{array}{ccc|ccc} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} & A_{36} \\ \hline A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} & A_{56} \\ A_{61} & A_{62} & A_{63} & A_{64} & A_{65} & A_{66} \end{array}\right] \tag{10}$$

in which A_{11}, A_{22}, A_{44}, and A_{55} are $(b-1)$-by-$(b-1)$ and A_{33} and A_{66} are $(n-2(b-1))$-by-$(n-2(b-1))$. For now, we do *not* assume that A is orthogonal. We specify an algorithm for eliminating certain submatrices of A. At the end, we shall see that when A is (close to) orthogonal, additional submatrices (of a small backward perturbation) are forced to be zero, and simultaneous multidiagonalization is achieved.

To begin, compute QR decompositions $A_{11} = P_1^{(a)} B_{11}$ and $A_{41} = P_1^{(b)} B_{41}$ so that

$$P_1^T A = \left[\begin{array}{ccc|ccc} B_{11} & B_{12} & B_{13} & B_{14} & B_{15} & B_{16} \\ 0 & B_{22} & B_{23} & B_{24} & B_{25} & B_{26} \\ 0 & B_{32} & B_{33} & B_{34} & B_{35} & B_{36} \\ \hline B_{41} & B_{42} & B_{43} & B_{44} & B_{45} & B_{46} \\ 0 & B_{52} & B_{53} & B_{54} & B_{55} & B_{56} \\ 0 & B_{62} & B_{63} & B_{64} & B_{65} & B_{66} \end{array}\right], \tag{11}$$

in which P_1 is the block-diagonal matrix $P_1^{(a)} \oplus P_1^{(b)}$. Then compute a block Givens rotation so that

$$G_1^T P_1^T A = \left[\begin{array}{ccc|ccc} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ 0 & B_{22} & B_{23} & B_{24} & B_{25} & B_{26} \\ 0 & B_{32} & B_{33} & B_{34} & B_{35} & B_{36} \\ \hline 0 & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ 0 & B_{52} & B_{53} & B_{54} & B_{55} & B_{56} \\ 0 & B_{62} & B_{63} & B_{64} & B_{65} & B_{66} \end{array}\right] \tag{12}$$

with C_{11} upper triangular. Next, attack from the right rather than the left. Compute LQ decompositions $C_{42} = D_{42}(Q_1^{(a)})^T$ and $C_{44} = D_{44}(Q_1^{(b)})^T$ so that

$$G_1^T P_1^T A Q_1 = \left[\begin{array}{ccc|ccc} C_{11} & D_{12} & D_{13} & D_{14} & D_{15} & D_{16} \\ 0 & D_{22} & D_{23} & D_{24} & D_{25} & D_{26} \\ 0 & D_{32} & D_{33} & D_{34} & D_{35} & D_{36} \\ \hline 0 & D_{42} & 0 & D_{44} & 0 & 0 \\ 0 & D_{52} & D_{53} & D_{54} & D_{55} & D_{56} \\ 0 & D_{62} & D_{63} & D_{64} & D_{65} & D_{66} \end{array}\right], \tag{13}$$

in which $Q_1 = 1 \oplus Q_1^{(a)} \oplus Q_1^{(b)}$. Then compute a block Givens rotation so that

$$G_1^T P_1^T A Q_1 H_1 = \left[\begin{array}{ccc|ccc} C_{11} & E_{12} & D_{13} & E_{14} & D_{15} & D_{16} \\ 0 & E_{22} & D_{23} & E_{24} & D_{25} & D_{26} \\ 0 & E_{32} & D_{33} & E_{34} & D_{35} & D_{36} \\ \hline 0 & 0 & 0 & E_{44} & 0 & 0 \\ 0 & E_{52} & D_{53} & E_{54} & D_{55} & D_{56} \\ 0 & E_{62} & D_{63} & E_{64} & D_{65} & D_{66} \end{array}\right] \tag{14}$$

with E_{44} lower triangular. (This can be done by exchanging D_{42} and D_{44}, transposing, and applying Lemma 1. After inverting the permutation and transposition, H_1 has the form of (7).) The road laid by Fig. 1(a) ends at

$$(G_k^T P_k^T \cdots G_1^T P_1^T) A (Q_1 H_1 \cdots Q_{k-1} H_{k-1}) = \left[\begin{array}{ccc|ccc} F_{11} & F_{12} & F_{13} & F_{14} & F_{15} & F_{16} \\ 0 & F_{22} & F_{23} & F_{24} & F_{25} & F_{26} \\ 0 & 0 & F_{33} & F_{34} & F_{35} & F_{36} \\ \hline 0 & 0 & 0 & F_{44} & 0 & 0 \\ 0 & 0 & 0 & F_{54} & F_{55} & 0 \\ 0 & 0 & 0 & F_{64} & F_{65} & F_{66} \end{array}\right]. \tag{15}$$

Many of the orthogonal transformations commute. Specifically, once a block Givens rotation manipulates a row or column, no subsequent QR or LQ factorization observes or modifies that row or column. Hence, the block Givens rotations can be moved outward and to the right-hand side of the equation, producing

$$\left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right]^T A \left[\begin{array}{c|c} V_1 & \\ \hline & V_2 \end{array}\right] = (G_1 \cdots G_k) F (H_{k-1}^T \cdots H_1^T),$$

in which F is the right-hand side of (15), $U_1 \oplus U_2 = P_1 \cdots P_k$, and $V_1 \oplus V_2 = Q_1 \cdots Q_{k-1}$. Notice that F is quasi-triangular.

In the following lemma, $\mathbf{u}$ is unit roundoff for a finite-precision arithmetic with the usual properties.

Lemma 3 *On a finite-precision computer, the above computation is backward stable. Given A, the procedure computes*

$$\left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right]^T (A + \Delta A) \left[\begin{array}{c|c} V_1 & \\ \hline & V_2 \end{array}\right] = (G_1 \cdots G_k) \hat{F} (H_{k-1}^T \cdots H_1^T)$$

for a small backward error ΔA satisfying $\|\Delta A\|_F \leq c_n \mathbf{u} \|A\|_F$, in which c_n is a constant that grows slowly with n. The symbols U_i, V_i, G_i, and H_i refer to the exact orthogonal transformations necessary to produce zeros at various intermediate stages, while $\hat{F}$ refers to the numerically computed quasi-triangular matrix.

Proof The reduction is a sequence of Householder reflectors and Givens rotations. See [12, chap. 19]. □

Theorem 1 *Given an input matrix*

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}, \qquad \left\|I - X^T X\right\|_2 = \varepsilon \leq \tfrac{1}{4},$$

the algorithm of (10–15) achieves simultaneous multidiagonalization in a backward stable way. Specifically, it finds

$$\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^T (X + \Delta X) \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix} = (G_1 \cdots G_k) D (H_{k-1}^T \cdots H_1^T) = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

with B_{11}, B_{21} upper b-diagonal, B_{12}, B_{22} lower b-diagonal, and $\|\Delta X\|_F \leq c_n \mathbf{u} + d_n \varepsilon$ for constants c_n, d_n that grow slowly with n.

Proof Replacing A by X and ΔA by E, the previous lemma gives

$$\left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right]^T (X + E) \left[\begin{array}{c|c} V_1 & \\ \hline & V_2 \end{array}\right] = (G_1 \cdots G_k) \hat{F} (H_{k-1}^T \cdots H_1^T)$$

with $\|E\|_F \leq c_n^{(1)} \mathbf{u} \|X\|_F$ for a constant $c_n^{(1)}$. The assumption $\left\|I - X^T X\right\|_2 \leq \frac{1}{4}$ keeps $\|X\|_F$ bounded, and therefore $\|E\|_F \leq c_n^{(2)} \mathbf{u}$ for a constant $c_n^{(2)}$ independent of $\|X\|_F$. The matrix $\hat{F}$ is quasi-triangular and satisfies

$$\left\|I - \hat{F}^T \hat{F}\right\|_F = \left\|I - (X + E)^T (X + E)\right\|_F.$$

Hence, $\hat{F}$ should be close to a diagonal signature matrix. This is made precise by Lemmas 7.1, 7.2, and 7.3 of [20]. They show that $\left\|I - \hat{F}^T \hat{F}\right\|_F \leq c_n^{(3)} \mathbf{u} + d_n^{(3)} \varepsilon$ for slowly growing constants $c_n^{(3)}$, $d_n^{(3)}$ and subsequently $\|D-\hat{F}\|_F \leq c_n^{(3)}\mathbf{u}+d_n^{(3)}\varepsilon$ for a diagonal signature matrix D. Hence,

$$\left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right]^T (X+\Delta X) \left[\begin{array}{c|c} V_1 & \\ \hline & V_2 \end{array}\right] = (G_1 \cdots G_k) D (H_{k-1}^T \cdots H_1^T)$$

with

$$\Delta X = E + \left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right] (G_1 \cdots G_k)(D-\hat{F})(H_{k-1}^T \cdots H_1^T) \left[\begin{array}{c|c} V_1 & \\ \hline & V_2 \end{array}\right]^T .$$

So, $\|\Delta X\|_F \leq \|E\|_F + \|D-\hat{F}\|_F \leq c_n\mathbf{u}+d_n\varepsilon$ with $c_n = c_n^{(2)} + c_n^{(3)}$ and $d_n = d_n^{(3)}$. $\square$

Of course, if arithmetic is exact and X is exactly orthogonal, then $\mathbf{u} = \varepsilon = 0$, the backward error is zero, and the existence of the decomposition

$$X = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T$$

with any desired half-bandwidth b is proved.

6 Example

Let

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

with

$$X_{11} = \begin{bmatrix} 0.0818768 & -0.2082266 & 0.1121874 & -0.0748685 & -0.2851551 & 0.1813253 \\ 0.2792668 & 0.0107689 & 0.2206289 & 0.3123625 & -0.2797934 & -0.0311379 \\ -0.3439806 & -0.1942959 & 0.2147062 & -0.3238816 & 0.2713634 & -0.2409807 \\ 0.1312931 & 0.0548822 & -0.0880027 & 0.6813623 & -0.1004676 & -0.4071046 \\ 0.0485421 & 0.0340342 & 0.0655397 & 0.0420500 & 0.4483995 & -0.3234565 \\ -0.1991368 & -0.4182146 & -0.1642219 & 0.2700508 & 0.2981295 & -0.2913474 \end{bmatrix},$$

$$X_{12} = \begin{bmatrix} 0.3438512 & -0.2641726 & 0.6076464 & -0.1595303 & -0.4809488 & -0.0594785 \\ -0.3129692 & 0.4786344 & -0.0179628 & -0.4929471 & -0.1835516 & 0.3038772 \\ 0.1303824 & 0.6133828 & 0.0447719 & 0.1747780 & -0.3656758 & -0.0413504 \\ 0.3241628 & 0.0733268 & 0.0089712 & 0.4395850 & -0.1579022 & -0.0567107 \\ 0.4914797 & -0.2118953 & -0.2168564 & -0.5800993 & -0.0142419 & 0.1205770 \\ -0.4021697 & -0.0656994 & 0.4851373 & -0.1846776 & 0.2230968 & -0.1627492 \end{bmatrix},$$

$$X_{21} = \begin{bmatrix} -0.0660281 & -0.3987007 & 0.2268668 & 0.1073527 & 0.0580933 & 0.2388745 \\ 0.0521754 & -0.4040447 & -0.2780677 & -0.1590016 & -0.4776298 & -0.1265961 \\ 0.5449237 & -0.2047616 & -0.3138040 & 0.0804998 & 0.4749267 & 0.4618872 \\ 0.4217341 & 0.3323236 & -0.2446359 & -0.3945466 & 0.0223984 & -0.3882945 \\ -0.2055628 & -0.1985483 & -0.6881404 & -0.0573065 & -0.0916941 & -0.0561258 \\ 0.4621628 & -0.4750778 & 0.3049300 & -0.2252498 & -0.0275146 & -0.3353926 \end{bmatrix},$$

$$X_{22} = \begin{bmatrix} 0.2545578 & 0.0352663 & 0.0968295 & 0.2118938 & 0.3854378 & 0.6666097 \\ 0.3352282 & 0.2765185 & -0.1509525 & -0.1945805 & 0.3787510 & -0.3071232 \\ 0.0738851 & 0.2574703 & -0.0106759 & 0.0523539 & -0.1095481 & -0.1751481 \\ 0.0256227 & 0.1293236 & 0.4349586 & 0.1011798 & 0.1990770 & 0.2974462 \\ -0.1082225 & -0.1134066 & -0.2121007 & -0.0184738 & -0.3996994 & 0.4475294 \\ -0.2518679 & -0.3109788 & -0.2819947 & 0.2026566 & -0.1515370 & -0.0064904 \end{bmatrix}.$$

The matrix is nearly orthogonal: $\left\|I - X^T X\right\|_2 < 1.4 \times 10^{-7}$. Using IEEE 754 single-precision arithmetic, the simultaneous multidiagonalization algorithm reduces X to $\hat{G}\hat{H}^T$, in which $\hat{G}$ and $\hat{H}$ are the following products of block Givens rotations:

$$\hat{G} = \left[\begin{array}{ccc|ccc} R_{11} & & & L_{14} & & \\ & R_{22} & & & L_{25} & \\ & & R_{33} & & & L_{36} \\ \hline R_{41} & & & L_{44} & & \\ & R_{52} & & & L_{55} & \\ & & R_{63} & & & L_{66} \end{array}\right],$$

$$R_{11} = \begin{bmatrix} -0.5121192 & -0.2830047 \\ 0 & 0.4245084 \end{bmatrix}, \qquad L_{14} = \begin{bmatrix} -0.8109514 & 0 \\ -0.1481443 & -0.8932222 \end{bmatrix},$$

$$R_{22} = \begin{bmatrix} -0.8939536 & 0.0392057 \\ 0 & -0.9718667 \end{bmatrix}, \qquad L_{25} = \begin{bmatrix} 0.4464413 & 0 \\ 0.0853476 & -0.2195242 \end{bmatrix},$$

$$R_{33} = \begin{bmatrix} -0.3075649 & 0.6039014 \\ 0 & 0.1386668 \end{bmatrix}, \qquad L_{36} = \begin{bmatrix} 0.7353278 & 0 \\ -0.1138826 & -0.9837694 \end{bmatrix},$$

$$R_{41} = \begin{bmatrix} 0.8589144 & -0.1687388 \\ 0 & 0.8433434 \end{bmatrix}, \qquad L_{44} = \begin{bmatrix} -0.4835218 & 0 \\ -0.2943089 & 0.4496155 \end{bmatrix},$$

$$R_{52} = \begin{bmatrix} -0.4481595 & -0.0782044 \\ 0 & 0.2186826 \end{bmatrix}, \qquad L_{55} = \begin{bmatrix} -0.8905264 & 0 \\ -0.0192043 & -0.9756070 \end{bmatrix},$$

$$R_{63} = \begin{bmatrix} -0.9515271 & -0.1952008 \\ 0 & 0.7602443 \end{bmatrix}, \qquad L_{66} = \begin{bmatrix} -0.2376822 & 0 \\ -0.6243644 & 0.1794372 \end{bmatrix},$$

and

$$\hat{H} = \left[\begin{array}{ccc|ccc} I & 0 & 0 & 0 & 0 & 0 \\ 0 & L_{22} & & R_{24} & & 0 \\ 0 & & L_{33} & & R_{35} & 0 \\ \hline 0 & L_{42} & & R_{44} & & 0 \\ 0 & & L_{53} & & R_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{array}\right],$$

$$L_{22} = \begin{bmatrix} 0.5455633 & 0 \\ -0.0429554 & -0.9613652 \end{bmatrix}, \qquad R_{24} = \begin{bmatrix} -0.8336260 & -0.0861880 \\ 0 & -0.2719042 \end{bmatrix},$$

$$L_{33} = \begin{bmatrix} 0.7302171 & 0 \\ -0.4228516 & -0.7254794 \end{bmatrix}, \qquad R_{35} = \begin{bmatrix} 0.3787578 & 0.5686172 \\ 0 & 0.5430251 \end{bmatrix},$$

$$L_{42} = \begin{bmatrix} -0.8234140 & 0 \\ 0.1500157 & -0.2752763 \end{bmatrix}, \qquad R_{44} = \begin{bmatrix} -0.5523294 & 0.1300827 \\ 0 & 0.9495884 \end{bmatrix},$$

$$L_{53} = \begin{bmatrix} -0.2988403 & 0 \\ 0.4457287 & -0.6882439 \end{bmatrix}, \qquad R_{55} = \begin{bmatrix} 0.9254958 & -0.2327058 \\ 0 & -0.5724039 \end{bmatrix}.$$

The product $\hat{G}\hat{H}^T$ has the correct structure:

$$\left[\begin{array}{cccccc|cccccc}
\times & \times & \times & & & & \times & & & & & \\
 & \times & \times & \times & & & \times & \times & & & & \\
 & & \times & \times & \times & & \times & \times & \times & & & \\
 & & & \times & \times & \times & & \times & \times & \times & & \\
 & & & & \times & \times & & & \times & \times & \times & \\
 & & & & & \times & & & & \times & \times & \times \\
\hline
\times & \times & \times & & & & \times & & & & & \\
 & \times & \times & \times & & & \times & \times & & & & \\
 & & \times & \times & \times & & \times & \times & \times & & & \\
 & & & \times & \times & \times & & \times & \times & \times & & \\
 & & & & \times & \times & & & \times & \times & \times & \\
 & & & & & \times & & & & \times & \times & \times
\end{array}\right].$$

The algorithm also computes $\hat{U}_1$, $\hat{U}_2$, $\hat{V}_1$, and $\hat{V}_2$ for which

$$\left\| X - \begin{bmatrix} \hat{U}_1 & \\ & \hat{U}_2 \end{bmatrix} (\hat{G}\hat{H}^T) \begin{bmatrix} \hat{V}_1 & \\ & \hat{V}_2 \end{bmatrix} \right\|_F < 8.7 \times 10^{-7},$$

$$\left\| I - \hat{U}_1^T \hat{U}_1 \right\|_2 < 3.1 \times 10^{-7}, \qquad \left\| I - \hat{V}_1^T \hat{V}_1 \right\|_2 < 2.1 \times 10^{-7},$$
$$\left\| I - \hat{U}_2^T \hat{U}_2 \right\|_2 < 6.8 \times 10^{-7}, \qquad \left\| I - \hat{V}_2^T \hat{V}_2 \right\|_2 < 6.3 \times 10^{-7}.$$

The algorithm is stable. The computed output $\hat{G}\hat{H}^T$ has the desired structure and is the correct reduction of a small perturbation of the input. The transformation matrices $\hat{U}_1$, $\hat{U}_2$, $\hat{V}_1$, and $\hat{V}_2$ are nearly orthogonal.

7 Conclusions

We have proved the existence of the matrix decomposition (1), described an algorithm for its computation, and proved the algorithm's numerical stability. The resulting matrix (2) is parameterized so that it is exactly orthogonal, even if the parameters are stored in finite precision. Because the algorithm spends most of its time on QR decompositions and matrix-matrix multiplications, it can take advantage of Level 3 BLAS and recent advances in communication-avoiding algorithms.

References

1. Ammar, G.S., Gragg, W.B., Reichel, L.: On the eigenproblem for orthogonal matrices. In: Proc. 25th Conference on Decision and Control, vol. 25, pp. 1963–1966. Athens (1986)
2. Bai, Z., Demmel, J.W.: Computing the generalized singular value decomposition. SIAM J. Sci. Comput. **14**(6), 1464–1486 (1993)
3. Björck, Å., Golub, G.H.: Numerical methods for computing angles between linear subspaces. Math. Comp. **27**(123), 579–594 (1973)
4. Bunse-Gerstner, A., Elsner, L.: Schur parameter pencils for the solution of the unitary eigenproblem. Linear Algebra and its Applications **154/156**, 741–778 (1991)
5. Cantero, M.J., Moral, L., Velázquez, L.: Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. Linear Algebra Appl. **362**, 29–56 (2003)
6. Davis, C., Kahan, W.M.: Some new bounds on perturbation of subspaces. Bull. Amer. Math. Soc. **75**, 863–868 (1969)

7. Demmel, J., Grigori, L., Hoemmen, M., Langou, J.: Communication-optimal parallel and sequential QR and LU factorizations. SIAM J. Sci. Comput. **34**(1), A206–A239 (2012)
8. Drmač, Z.: A tangent algorithm for computing the generalized singular value decomposition. SIAM J. Numer. Anal. **35**(5), 1804–1832 (1998)
9. Drmač, Z., Jessup, E.R.: On accurate quotient singular value computation in floating-point arithmetic. SIAM J. Matrix Anal. Appl. **22**(3), 853–873 (2000)
10. Edelman, A., Sutton, B.D.: The beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems. Found. Comput. Math. **8**(2), 259–285 (2008)
11. Hari, V.: Accelerating the SVD block-Jacobi method. Computing **75**(1), 27–53 (2005)
12. Higham, N.J.: Accuracy and stability of numerical algorithms, second edn. SIAM, Philadelphia, PA (2002)
13. Ltaief, H., Kurzak, J., Dongarra, J.: Parallel two-sided matrix reduction to band bidiagonal form on multicore architectures. IEEE Transactions on Parallel and Distributed Systems **21**(4), 417–423 (2010)
14. Paige, C.C.: Computing the generalized singular value decomposition. SIAM J. Sci. Statist. Comput. **7**(4), 1126–1146 (1986)
15. Paige, C.C., Wei, M.: History and generality of the CS decomposition. Linear Algebra Appl. **208/209**, 303–326 (1994)
16. Simon, B.: CMV matrices: five years after. J. Comput. Appl. Math. **208**(1), 120–154 (2007)
17. Stewart, G.W.: On the perturbation of pseudo-inverses, projections and linear least squares problems. SIAM Rev. **19**(4), 634–662 (1977)
18. Stewart, G.W.: Computing the CS decomposition of a partitioned orthonormal matrix. Numerische Mathematik **40**(3), 297–306 (1982)
19. Sutton, B.D.: Computing the complete CS decomposition. Numer. Algorithms **50**(1), 33–65 (2009)
20. Sutton, B.D.: Stable computation of the CS decomposition: simultaneous bidiagonalization. SIAM J. Matrix Anal. Appl. **33**(1), 1–21 (2012)
21. Szegő, G.: Orthogonal polynomials, fourth edn. American Mathematical Society, Providence, R.I. (1975)
22. Van Loan, C.: Computing the CS and the generalized singular value decompositions. Numerische Mathematik **46**(4), 479–491 (1985)
23. Verblunsky, S.: On positive harmonic functions: A contribution to the algebra of Fourier series. Proc. London Math. Soc. **S2-38**(1), 125–127 (1935)
24. Watkins, D.S.: Some perspectives on the eigenvalue problem. SIAM Review **35**(3), 430–471 (1993)