89-215 IJ 90

A Tree of Generalizations of the Ordinary Singular Value Decomposition

Bart De Moor*
Department of Electrical Engineering
Katholieke Universiteit Leuven
Kardinaal Mercierlaan 94
B-3030 Leuven, Belgium

and

Hongyuan Zha†
Department of Computer Science
Stanford University
Stanford, California 94305

Submitted by Roger A. Horn

ABSTRACT

It is shown how to generalize the ordinary singular value decomposition of a matrix into a combined factorization of any number of matrices. We propose to call these factorizations *generalized singular value decompositions*. For two matrices, this reduces to the product and quotient singular value decompositions. One of the

*Telephone: 32/16/220931; fax: 32/16/221855; E-mail: demoor@esat.kuleuven.ac.be or na.demoor@na-net.stanford.edu.

Dr. De Moor is a research associate of the Belgian National Fund for Scientific Research (NFWO). Part of this work was done while he was a visiting research associate at the departments of Computer Science and Electrical Engineering of Stanford University. Part of this research was supported by an Advanced Research Fellowship in Science and Technology of the NATO Science Fellowships Programme, the US Army under contract DAAL03-87-K-0095, the Belgian Program on Interuniversity Attraction Poles initiated by the Belgian State Science Policy Programming (Prime Minister's Office) and the European Community Research Program ESPRIT, Basic Research Action 3280. The scientific responsibility is assumed by its authors.

†Telephone: (415) 723-4269; fax: (415) 725-7411; E-mail: na.zha@na-net.stanford.edu.

Part of the work of Zha was done while he was a visitor at the Konrad-Zuse-Zentrum fur Informationstechnik Berlin, West Berlin, and the department of Electrical Engineering of the Katholieke Universiteit Leuven, Belgium.

LINEAR ALGEBRA AND ITS APPLICATIONS 147:469–500 (1991) 469

655 Avenue of the Americas, New York, NY 10010 0024-3795/91/$3.50

factorizations for three matrices is the restricted singular value decomposition. These generalizations form a tree of factorizations, where at level k, for k matrices, there are 2^k factorizations, not all of which are independent. The different levels are related to each other in a recursive fashion. Any generalized singular value decomposition for k matrices can be constructed from a decomposition for $k-1$ matrices. This results in an inductive proof which uses only the ordinary singular value decomposition. Several examples are analysed in detail.

1. INTRODUCTION

The ordinary singular value decomposition (OSVD) has become an important tool in the analysis and numerical solution of numerous problems (see e.g. [6, 8] for properties and applications). Not only does it allow for an elegant problem formulation, but at the same time it provides geometrical and algebraic insight together with an immediate numerically robust implementation [8]. In [11, p. 78], credit for the first proofs of the OSVD is given to Beltrami [2], Jordan [10], Sylvester [13], and Autonne [1].

Recently, several generalizations of the OSVD have been proposed and their properties analysed. The best known example is the generalized SVD for two matrices, as introduced in [14] and refined in [12], which we propose to rename as the *quotient SVD* (QSVD) [4]. One reason for this name is the relation of this matrix factorization to the SVD of the "quotient" of two matrices; the main motivation is of course the fact that there are several other similar generalizations. For instance, a *product induced SVD*, also for two matrices, was proposed in [7], where it was called the ΠSVD. It was a refinement of ideas in [9]. We shall refer to it as the PSVD (see [4]). In [15], another generalization, this time for three matrices, was proposed. In [3] we have called it the *restricted SVD* (RSVD) and analysed its properties in detail.

One of the main results of this paper is that all of these decompositions can be organized into a tree of generalizations of the OSVD. The OSVD resides at the top of this tree. The PSVD and QSVD are two different decompositions for two matrices at the next level of generality. The RSVD is one of the four possible factorizations for three matrices. It will be shown that only three of these are theoretically essential. Similar generalizations for $4,5,\ldots$ matrices are obtained in a completely structured manner.

The main idea can be easily understood from the special case of real nonsingular $n\times n$ matrices $A,B,C,\ldots$. At the top of the tree, we have the OSVD of A. At the next level we have the OSVD of AB and AB^{-t}. Note that we use B^{-t} rather than B^{-1}, so that in the general case of rectangular matrices, all dimensions will be compatible. The OSVD of AB leads to the PSVD of the matrix pair (A,B) as follows. Consider the OSVD of the product

AB, and insert a nonsingular matrix X and its inverse in order to obtain a combined factorization of the two matrices:

$$AB = USV^t = UI_nX^{-1}XSV^t \quad\Rightarrow\quad \begin{cases} A = UI_nX^{-1}, \\ B = XSV^t. \end{cases}$$

Our main theorem states that such a common factor in the decompositions of A and B can always be found.

The QSVD of the matrix pair follows from the OSVD of AB^{-t} in a similar fashion as

$$AB^{-t} = USV^t = UI_nX^{-1}XSV^t \quad\Rightarrow\quad \begin{cases} A = UI_nX^{-1}, \\ B = X^{-t}S^{-t}V^t, \end{cases}$$

where again, X is a nonsingular matrix. The factor X in the PSVD differs completely from the one in the QSVD, and the algebraic, structural, and geometric properties of the two decompositions are likewise different.

At the next level, we have the OSVDs of products and quotients for three matrices: ABC, ABC^{-t}, $AB^{-t}C^{-t}$, and $AB^{-t}C$. For instance, from the OSVD of $AB^{-t}C^{-t}$ we find, by inserting nonsingular matrices X and Y and their inverses, a decomposition of the matrix triplet (A,B,C) as

$$AB^{-t}C^{-t} = USV^t = UX^{-1}XY^tY^{-t}SV^t \quad\Rightarrow\quad \begin{cases} A = UI_nX^{-1}, \\ B = X^{-t}I_nY^{-1}, \\ C = YS^{-t}V^t. \end{cases} \qquad (1)$$

We shall define a terminology such that a combination between two matrices F and G is of *type P* if they occur as FG or $F^{-t}G^{-t}$, and of *type Q* if they occur as FG^{-t} or $F^{-t}G$. Here P stands for "product" [note that $F^{-t}G^{-t}=(FG)^{-t}$] and Q stands for "quotient." The decomposition of (1) will therefore be called a QP-SVD. The remaining factorizations for three matrices result then in a PP-SVD for ABC, a PQ-SVD for ABC^{-t}, and a QQ-SVD for $AB^{-t}C$.

Note that the QP-SVD of (A,B,C) can be found from the PQ-SVD of (C^t,B^t,A^t). Hence, for three matrices, there are only three theoretically distinct generalized SVDs.

It is clear that we could find similar factorizations for $4,5,\ldots$ real nonsingular matrices.

One contribution of this paper is to show how to obtain (at least theoretically) explicit factorizations of the *individual* matrices $A,B,C,\ldots$ corresponding to an OSVD of products and quotients of these matrices. Much of the complexity of this paper is introduced in order to handle fully general matrices of compatible dimensions which can be rectangular and possibly rank deficient. In this last case, the inverse in a quotient is to be replaced with some generalized inverse. If the reader feels uncertain about our general derivation, it might be instructive to return to the real nonsingular case.

This paper is organized as follows:

(1) The main theorem is stated in Section 2, and some nomenclature is proposed.

(2) A block factorization lemma is proved in Section 3. It is one of the key results that will be needed.

(3) A constructive proof of the main theorem is given in Section 4.

NOTATION AND ABBREVIATIONS. Throughout the paper, matrices are denoted by capitals, and vectors by lowercase letters other than $i, j, k, l, m, n, p, q, r, s$, which are nonnegative integers. Scalars (complex) are denoted by Greek letters. The notation $A\ (m \times n)$ refers to a complex matrix with m rows and n columns. The conjugate of A is $\overline{A}$, while the complex conjugate transpose is $A^* = \text{conj}(A^+)$. A^{-*} is the inverse of A^* (nonsingular). I_k is the $k \times k$ identity matrix. a_i is the ith column of the matrix A. We shall denote different matrices with subscripts, as $A_1, A_2, A_3, \ldots$. Quantities related to the jth matrix will have a subscript j. For instance, the rank of A_j will be denoted by r_j. We adopt the following convention for block matrices: Any (possibly rectangular) block of zeros is denoted by 0, the precise dimensions being obvious from the context. The symbol I means an identity matrix of appropriate dimensions. Whenever a dimension indicating integer in a block matrix is negative or zero, the corresponding block row or block column should be omitted, and all expressions and equations in which a block matrix of that block row or block column appears can be discarded. An equivalent formulation would be that we allow $0 \times n$ or $n \times 0$ $(n \neq 0)$ to appear in matrices. This allows an elegant general treatment of several cases at once. The following abbreviations will be used throughout: GSVD: generalized singular value decomposition; OSVD: ordinary singular value decomposition; PSVD: product singular value decomposition; QSVD: quotient singular value decomposition; RSVD: restricted singular value decomposition.

2. THE MAIN THEOREM

THEOREM 1 (Generalized singular value decompositions for k matrices). *Consider a set of k matrices with compatible dimensions:* $A_1\ (n_0 \times n_1)$,

$A_2(n_1 \times n_2), \ldots, A_{k-1}(n_{k-2} \times n_{k-1}), A_k(n_{k-1} \times n_k)$. *Then there exist*

(a) *Unitary matrices* $U_1\ (n_0 \times n_0)$ *and* $V_k\ (n_k \times n_k)$,
(b) *matrices* D_j, $j = 1, 2, \ldots, k-1$, *of the form*

$$\underset{n_{j-1} \times n_j}{D_j} = \begin{array}{c} r_j^1 \\ r_{j-1}^1 - r_j^1 \\ r_j^2 \\ r_{j-1}^2 - r_j^2 \\ r_j^3 \\ \vdots \\ r_j^j \\ n_{j-1} - r_{j-1} - r_j^j \end{array} \overset{\begin{array}{cccccc} r_j^1 & r_j^2 & r_j^3 & \cdots & r_j^j & n_j - r_j \end{array}}{\begin{pmatrix} I & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & I & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}, \tag{2}$$

where

$$r_0 = 0, \qquad r_j = \sum_{i=1}^{j} r_j^i = \text{rank}(A_j), \tag{3}$$

(c) *a matrix* S_k *of the form*

$$\underset{n_{k-1} \times n_k}{S_k} = \begin{array}{c} r_k^1 \\ r_{k-1}^1 - r_k^1 \\ r_k^2 \\ r_{k-1}^2 - r_k^2 \\ r_k^3 \\ \vdots \\ r_k^k \\ n_{k-1} - r_{k-1} - r_k^k \end{array} \overset{\begin{array}{cccccc} r_k^1 & r_k^2 & r_k^3 & \cdots & r_k^k & n_k - r_k \end{array}}{\begin{pmatrix} S_k^1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & S_k^2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & S_k^3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & S_k^k & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}, \tag{4}$$

where

$$r_k = \sum_{i=1}^{k} r_k^i = \text{rank}(A_k) \tag{5}$$

and the $r_k^i \times r_k^i$ *matrices* S_k^i *are diagonal with positive diagonal elements (the integers* r_j^i *are ranks of certain matrices as given in the constructive proof in Section* 4),

(d) *nonsingular matrices* X_j $(n_j \times n_j)$ *and* Z_j, $j = 1, 2, \ldots, k-1$, *where* Z_j *is either* $Z_j = X_j^{-*}$ *or* $Z_j = X_j$ *(i.e. both choices are always possible),*

such that the given matrices can be factorized as

$$A_1 = U_1 D_1 X_1^{-1},$$

$$A_2 = Z_1 D_2 X_2^{-1},$$

$$A_3 = Z_2 D_3 X_3^{-1},$$

$$\vdots$$

$$A_i = Z_{i-1} D_i X_i^{-1},$$

$$\vdots$$

$$A_k = Z_{k-1} S_k V_k^*.$$

Observe that the matrices D_j in (2) and S_k in (4) are in general not diagonal. Their only nonzero blocks however are diagonal block matrices. We propose to call them *quasidiagonal* matrices. The matrices D_j, $j = 1, \ldots, k-1$, are quasidiagonal, their only nonzero blocks being identity matrices. The matrix S_k is quasidiagonal, and its nonzero blocks are diagonal matrices with positive diagonal elements. Observe that we always take the last factor in every factorization as the inverse of a nonsingular matrix, which is only a matter of convention (another convention would result in a modified definition of the matrices Z_i). As to the name of a certain GSVD, we propose to adopt the following convention:

Definition 1 (The nomenclature for GSVDs). If $k = 1$ in Theorem 1, then the corresponding factorization of the matrix A_1 will be called the ordinary singular value decomposition.

If for a matrix pair A_i, A_{i+1}, $1 \leqslant i \leqslant k-1$, in Theorem 1, we have that

$$Z_i = X_i,$$

then the factorization of the pair will be said to be of *P-type*. If, on the other

hand, for a matrix pair A_i, A_{i+1}, $1 \leqslant i \leqslant k-1$, in Theorem 1, we have that

$$Z_i = X_i^{-*},$$

the factorization of the pair will be said to be of *Q-type*.

The name of a GSVD of the matrices A_i, $i = 1, 2, \ldots, k > 1$, as in Theorem 1, is then obtained by simply enumerating the different factorization types.

Let us give some examples.

Example 1. Consider two matrices A_1 $(n_0 \times n_1)$ and A_2 $(n_1 \times n_2)$. Then we have two possible GSVDs:

	P-type	Q-type
A_1	$U_1 D_1 X_1^{-1}$	$U_1 D_1 X_1^{-1}$
A_2	$X_1 S_2 V_2^*$	$X_1^{-*} S_2 V_2^*$

The P-type factorization corresponds to the PSVD as in [7] (called ΠSVD there) and [5], while the Q-type factorization is nothing else than the QSVD in [8, 12, 14] (called generalized SVD there). This justifies the choice of names for the factorization of pairs: A P-type factorization is precisely the kind of transformation that occurs in the PSVD, while a Q-type factorization occurs in the QSVD.[1]

Example 2. The RSVD for three matrices (A_1, A_2, A_3) as introduced and analysed in [3, 15] has the form

$$A_1 = U_1 S_1 X_1^{-1},$$

$$A_2 = X_1^{-*} S_2 X_2^{-1},$$

$$A_3 = X_2^{-*} S_3 V_3^*,$$

where S_1, S_2, S_3 are certain quasidiagonal matrices. It can be verified that this RSVD can be rearranged into a QQ-SVD that has the structure described in Theorem 1.

[1]The interested reader may wish to consult [4], where we have proposed a standardized and mnemonic nomenclature for the several possible generalizations of the singular value decomposition.

EXAMPLE 3. Let us write down the PQQP-SVD for five matrices, together with the structure of the matrices D_i, $i = 1,2,3,4$, and S_5 (we have omitted the block dimensions):

$$A_1 = U_1\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}X_1^{-1}, \qquad A_2 = X_1\begin{pmatrix} I & 0 & \\ 0 & & \\ 0 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{pmatrix}X_2^{-1},$$

$$A_3 = X_2^{-*}\begin{pmatrix} I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}X_3^{-1}, \qquad A_4 = X_3^{-*}\begin{pmatrix} I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}X_4^{-1},$$

$$A_5 = X_4\begin{pmatrix} S_5^1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & S_5^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & S_5^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & S_5^4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & S_5^5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}V_5^*.$$

We also introduce the following notation, using powers, which symbolize a certain repetition of a letter or of a sequence of letters:

$$P^3Q^2\text{-SVD} = PPPQQ\text{-SVD},$$

$$(PQ)^2Q^3(PPQ)^2\text{-SVD} = PQPQQQQPPQPPQ\text{-SVD}.$$

Despite the fact that there are 2^{k-1} different sequences of letters P and Q at level $k > 1$, not all of these sequences correspond to different GSVDs. The reason for this is that for instance the QP-SVD of (A^1, A^2, A^3) can be

obtained from the PQ-SVD of $((A^3)^*, (A^2)^*, (A^1)^*)$. Similarly, the $P^2(QP)^3$-SVD of $(A^1, \dots, A^9)$ is essentially the same as the $(PQ)^3P^2$-SVD of $((A^9)^*, \dots, (A^1)^*)$. The following table gives the number of *different* factorizations (GSVDs) for k matrices:

k even	k odd
$\frac{1}{2}(2^{k-1}+2^{k/2})$	$\frac{1}{2}(2^{k-1}+2^{(k-1)/2})$

A possible way to visualize Theorem 1 is to build a tree with all different factorizations for 1, 2, 3, etc... matrices as follows:

```
1                    O
2               P         Q
3          P²        PQ        Q²
4  P³  P²Q     PQP        PQ²       QPQ  Q³
⋮                    ⋮
```

3. A BLOCK FACTORIZATION LEMMA

In this section, we shall construct a factorization of a block matrix into three factors that have some special properties: The first factor is lower block triangular, the second one is quasidiagonal, and the third one is unitary. This result will be of key importance in our constructive proof of the GSVD, which will be developed in Section 4.

LEMMA 1 (On the factorization of a block matrix). *Consider a $p \times q$ matrix B with k block rows:*

$$B = \begin{matrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{matrix}\overset{q}{\begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_k \end{pmatrix}},$$

where $\sum_{i=1}^k p_i = p$. The matrix B can be factorized as

$$B = TSV^*,$$

where:

(a) *T is a* $p\times p$ *block lower triangular matrix*

$$\begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_k \end{array}\begin{array}{c} \begin{array}{ccccccccc} r_1 & p_1-r_1 & r_2 & p_2-r_2 & r_3 & p_3-r_3 & \cdots & r_k & p_k-r_k \end{array} \\ \begin{pmatrix} T_{11} & T_{12} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ T_{21} & 0 & T_{23} & T_{24} & 0 & 0 & \cdots & 0 & 0 \\ T_{31} & 0 & T_{33} & 0 & T_{35} & T_{36} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ T_{k1} & 0 & T_{k3} & 0 & T_{k5} & 0 & \cdots & T_{k\,2k-1} & T_{k\,2k} \end{pmatrix}\end{array},$$

and the $p_i\times p_i$ *matrices* $(T_{i(2i-1)}\ T_{i(2i)})$ *are unitary.*

(b) S *is a* $p\times q$ *quasidiagonal matrix with the following block structure*:

$$S=\begin{array}{c} \\ r_1 \\ p_1-r_1 \\ r_2 \\ p_2-r_2 \\ r_3 \\ p_3-r_3 \\ \vdots \\ r_k \\ p_k-r_k \end{array}\begin{array}{c} \begin{array}{cccccc} r_1 & r_2 & r_3 & \cdots & r_k & q-\sum_{i=1}^k r_i \end{array} \\ \begin{pmatrix} S^1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & S^2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & S^3 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & S^k & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}\end{array},$$

and the $r_i\times r_i$ *matrices* S^i *are diagonal with positive diagonal elements.*

(c) V *is a* $q\times q$ *unitary matrix.*

(d) *The integers* r_i *are defined recursively as*

$$r_1=\operatorname{rank}(B_1),$$

$$r_i=\operatorname{rank}\begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_i \end{pmatrix}-\operatorname{rank}\begin{pmatrix} B_1 \\ \vdots \\ B_{i-1} \end{pmatrix}, \qquad i=2,3,\dots,k,$$

and satisfy

$$\sum_{i=1}^k r_i=\operatorname{rank}(B).$$

Proof. Let the OSVD of B_1 be

$$B_1=U_1S_1V_1^*=\begin{pmatrix} U_1^1 & U_1^2 \end{pmatrix}\begin{pmatrix} S^1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} (V_1^1)^* \\ (V_1^2)^* \end{pmatrix}$$

with S^1 an $r_1\times r_1$ diagonal matrix with positive diagonal elements that are the singular values of B_1, and $r_1=\operatorname{rank}(B_1)$. The matrix $(U_1^1\ U_1^2)$ is $p_1\times p_1$ unitary, while $(V_1^1\ V_1^2)$ is $q\times q$ unitary. It is straightforward to obtain the following factorization of the matrix B:

$$B=\begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ \vdots \\ B_k \end{pmatrix}=\begin{pmatrix} U_1^1 & U_1^2 & 0 & 0 & \cdots & 0 \\ B_2V_1^1(S^1)^{-1} & 0 & I_{p_2} & 0 & \cdots & 0 \\ B_3V_1^1(S^1)^{-1} & 0 & 0 & I_{p_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ B_kV_1^1(S^1)^{-1} & 0 & 0 & 0 & \cdots & I_{p_k} \end{pmatrix}$$

$$\times\begin{pmatrix} S^1 & 0 \\ 0 & 0 \\ 0 & B_2V_1^2 \\ 0 & B_3V_1^2 \\ \vdots & \vdots \\ 0 & B_kV_1^2 \end{pmatrix}\begin{pmatrix} (V_1^1)^* \\ (V_1^2)^* \end{pmatrix}. \tag{6}$$

In case $r_1=q$, there is no matrix V_1^2. In that case, simply omit the corresponding blocks in the second and third factors. The obtained factorization is then the one of the lemma, and we may stop here. (Observe that it is also possible that there is no matrix U_1^2, namely if $r_1=p_1$. In that case, we simply follow our convention regarding matrix blocks with possibly zero dimensions. In this case, the construction does not stop, however.)

Assume that $r_1<q$. The proof now proceeds by considering the block matrix

$$\begin{pmatrix} B_2V_1^2 \\ B_3V_1^2 \\ \vdots \\ B_kV_1^2 \end{pmatrix} \tag{7}$$

and applying a similar decomposition to that in (6), using the OSVD of the matrix $B_2V_1^2$:

$$B_2V_1^2 = \begin{pmatrix} U_2^1 & U_2^2 \end{pmatrix} \begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (V_2^1)^* \\ (V_2^2)^* \end{pmatrix} \tag{8}$$

with S^2 an $r_2 \times r_2$ diagonal matrix with positive diagonal elements and

$$r_2 = \operatorname{rank}(B_2V_1^2) = \operatorname{rank}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} - \operatorname{rank}(B_1). \tag{9}$$

We then find that

$$\begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ \vdots \\ B_k \end{pmatrix} = \begin{pmatrix} U_1^1 & U_1^2 & 0 & 0 & 0 & 0 & \cdots & 0 \\ B_2V_1^1(S^1)^{-1} & 0 & U_2^1 & U_2^2 & 0 & 0 & \cdots & 0 \\ B_3V_1^1(S^1)^{-1} & 0 & B_3V_1^2V_2^1(S^2)^{-1} & 0 & I_{p_3} & 0 & \cdots & 0 \\ B_4V_1^1(S^1)^{-1} & 0 & B_4V_1^2V_2^1(S^2)^{-1} & 0 & 0 & I_{p_4} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ B_kV_1^1(S^1)^{-1} & 0 & B_kV_1^2V_2^1(S^2)^{-1} & 0 & 0 & 0 & \cdots & I_{p_k} \end{pmatrix}$$

$$\times \begin{pmatrix} S^1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & S^2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & B_3V_1^2V_2^2 \\ 0 & 0 & B_4V_1^2V_2^2 \\ \vdots & \vdots & \vdots \\ 0 & 0 & B_kV_1^2V_2^2 \end{pmatrix} \begin{pmatrix} (V_1^1)^* \\ (V_2^1)^*(V_1^2)^* \\ (V_2^2)^*(V_1^2)^* \end{pmatrix}.$$

In case $r_1 + r_2 = q$, there is no matrix V_2^2. The block rows and columns in the factorization that contain V_2^2 should then be omitted. In that case, we have obtained the decomposition as in the lemma and we may stop here.

Assume that $r_1 + r_2 < q$. Next consider a similar factorization of

$$\begin{pmatrix} B_3V_1^2V_2^2 \\ B_4V_1^2V_2^2 \\ \vdots \\ B_kV_1^2V_2^2 \end{pmatrix}$$

using the OSVD of $B_3V_1^2V_2^2$, and so on. In general, the recursive construction will stop after $j \leqslant k$ steps, when we have

$$\sum_{i=1}^{j} r_i = q,$$

which completes the proof. ■

Observe that in the extreme case where $j = k$, we obtain a block factorization of the matrix B as $B = TSV^*$ where

(a) the matrix T is given by

$$T = \begin{pmatrix} U_1^1 & U_1^2 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ B_2V_1^1(S^1)^{-1} & 0 & U_2^1 & U_2^2 & 0 & 0 & \cdots & 0 & 0 \\ B_3V_1^1(S^1)^{-1} & 0 & B_3V_1^2V_2^1(S^2)^{-1} & 0 & U_3^1 & U_3^2 & \cdots & 0 & 0 \\ B_4V_1^1(S^1)^{-1} & 0 & B_4V_1^2V_2^1(S^2)^{-1} & 0 & B_4V_1^2V_2^2V_3^1(S^3)^{-1} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ B_kV_1^1(S^1)^{-1} & 0 & B_kV_1^2V_2^1(S^2)^{-1} & 0 & B_kV_1^2V_2^2V_3^1(S^3)^{-1} & 0 & \cdots & U_k^1 & U_k^2 \end{pmatrix}, \tag{10}$$

in which we have the OSVDs

$$B_iV_1^2V_2^2 \cdots V_{i-1}^2 = \begin{pmatrix} U_i^1 & U_i^2 \end{pmatrix} \begin{pmatrix} S^i & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (V_i^1)^* \\ (V_i^2)^* \end{pmatrix}, \qquad i = 2, \ldots, k, \tag{11}$$

and S^i is $r_i \times r_i$ diagonal with positive diagonal elements;

(b) the matrix S has the structure as stated in the lemma;

(c) the matrix V is then given by

$$V = \begin{pmatrix} V_1^1 & V_1^2 V_2^1 & V_1^2 V_2^2 V_3^1 & \cdots & V_1^2 V_2^2 \cdots V_{k-1}^2 V_k^1 & V_1^2 V_2^2 \cdots V_{k-1}^2 V_k^2 \end{pmatrix} \tag{12}$$

and is unitary.

4. A CONSTRUCTIVE PROOF OF THE MAIN THEOREM

In this section, we derive a constructive proof of the main theorem. The proof is based upon two ideas:

(1) An inductive argument: Any GSVD at level k for k matrices $A_1, A_2, \ldots, A_k$ can be constructed from a corresponding GSVD of the first $k-1$ matrices $A_1, A_2, \ldots, A_{k-1}$. We basically use one of the factors of the matrix A_{k-1} and the block factorization lemma of the previous section, to find a factorization of the matrix A_k.

(2) Once such a factorization of the matrix A_k has been obtained, the combined factorizations of the matrices $A_1, \ldots, A_{k-1}$ have to be modified, a procedure which will be described in terms of the so-called *ripple-through phenomenon*.

The inductive argument will be described in Section 4.1, and the ripple-through phenomenon in Section 4.2. The details of both basic ideas are developed in Section 4.3. A summary of the algorithmic proof is provided in Section 4.4.

4.1. *The Induction Step*

In order to construct a GSVD of the set of matrices $(A_1, A_2, \ldots, A_k)$, we assume that we have obtained the corresponding GSVD for the matrices $(A_1, A_2, \ldots, A_{k-1})$:

$$\begin{aligned} A_1 &= U_1 D_1 X_1^{-1}, \\ A_2 &= Z_1 D_2 X_2^{-1}, \\ &\vdots \\ A_{k-1} &= Z_{k-2} S_{k-1} V_{k-1}^*. \end{aligned} \tag{13}$$

Recall that the matrices D_j, $j = 1, \ldots, k-2$, are quasidiagonal with nonzero blocks equal to identity matrices. The matrix S_{k-1} is quasidiagonal; its nonzero blocks are diagonal matrices with positive diagonal elements. Also recall that for a P-type step, $Z_{k-1} = X_{k-1}$, while for a Q-type step, $Z_{k-1} = X_{k-1}^{-*}$.

Our first step is to rewrite the last equation as follows: Partition the matrix V_{k-1} according to the block columns of S_{k-1} (4) as

$$V_{k-1} = \begin{pmatrix} \overset{r_{k-1}^1}{V_{k-1}^1} & \overset{r_{k-1}^2}{V_{k-1}^2} & \overset{\cdots}{\cdots} & \overset{r_{k-1}^{k-1}}{V_{k-1}^{k-1}} & \overset{n_{k-1}-r_{k-1}}{V_{k-1}^k} \end{pmatrix},$$

where

$$r_{k-1} = \sum_{i=1}^{k-1} r_{k-1}^i.$$

Then the matrix Y_{k-1} in

$$S_{k-1} V_{k-1}^* = D_{k-1} Y_{k-1} \tag{14}$$

can be chosen to be nonsingular. This matrix is obtained from a scaling of the rows of V_{k-1}^* with the diagonal blocks in S_{k-1}, and replacing these diagonal blocks with identity matrices in order to obtain the quasidiagonal matrix D_{k-1} in (14). A matrix X_{k-1} is defined as

$$X_{k-1} = Y_{k-1}^{-1}, \tag{15}$$

so that we can rewrite the factorization for A_{k-1} as

$$A_{k-1} = Z_{k-2} S_{k-1} V_{k-1}^* = Z_{k-2} D_{k-1} X_{k-1}^{-1}.$$

Consider now the matrix $Z_{k-1}^{-1} A_k$. It can be partitioned into k block rows B_i, $i = 1, \ldots, k$, having the same dimensions as the block columns of S_{k-1}:

$$Z_{k-1}^{-1} A_k = \begin{matrix} r_{k-1}^1 \\ r_{k-1}^2 \\ \vdots \\ r_{k-1}^{k-1} \\ n_{k-1}-r_{k-1} \end{matrix} \overset{n_k}{\begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_{k-1} \\ B_k \end{pmatrix}}. \tag{16}$$

Suppose now, for the time being, that we had a way of factorizing the matrix $Z_{k-1}^{-1}A_k$ as

$$Z_{k-1}^{-1}A_k = P_k S_k V_k^*, \qquad (17)$$

where P_k is nonsingular, S_k is quasidiagonal with the structure in (4), and V_k is an $n_k \times n_k$ unitary matrix. Then we would have a factorization of the matrix A_k as

$$A_k = (Z_{k-1}P_k)S_k V_k^*. \qquad (18)$$

However, the set of factorizations of the matrices $A_1, \dots, A_{k-1}$ (13), completed with this factorization for the matrix A_k (18), is not a GSVD as in the main theorem, unless the matrix P_k has some special structure, as will now be explained via a so-called *ripple-through phenomenon*.

The result of this discussion will be that P_k can be obtained by applying the block factorization lemma of Section 3 to a certain permutation of the block rows of $Z_{k-1}^{-1}A_k$ that is determined by the sequence of letters P and Q in the GSVD name.

4.2. *The Ripple-Through Phenomenon*

Assume that we have obtained a factorization of the matrix A_k as in (18). The combined factorizations of the matrices A_{k-1}, A_k are now

$$A_{k-1} = Z_{k-2}D_{k-1}X_{k-1}^{-1}, \qquad (19)$$

$$A_k = Z_{k-1}P_k S_k V_k^*, \qquad (20)$$

where $Z_{k-1} = X_{k-1}$ for a P-type and $Z_{k-1} = X_{k-1}^{-*}$ for a Q-type factorization. Because of the presence of the factor P_k in (20) and its absence in (19), the combined factorization of A_{k-1}, A_k does not have the structure required by the main theorem. The idea is now to preserve the factorization of A_k and modify the factorization of A_{k-1} as follows:

(a) If the combined factorization is of P-type, the factorization of A_{k-1} is modified by introducing a nonsingular matrix P_{k-1} as

$$A_{k-1} = Z_{k-2}P_{k-1}D_{k-1}P_k^{-1}X_{k-1}^{-1},$$

where

$$P_{k-1}D_{k-1}P_k^{-1} = D_{k-1}.$$

(b) If the combined factorization is of Q-type, we modify the factorization of A_{k-1} by introducing a nonsingular matrix P_{k-1} as

$$A_{k-1} = Z_{k-2}P_{k-1}D_{k-1}P_k^* X_{k-1}^{-1},$$

where

$$P_{k-1}D_{k-1}P_k^* = D_{k-1}.$$

Assume that such a modification is always possible (and it is, as we shall show). Then, in either case, we have changed the first factor of A_{k-1} from Z_{k-2} to $Z_{k-2}P_{k-1}$. But, in order to conform with the structure of the GSVD as in the main theorem, we need to change the factorization of A_{k-2} by introducing a nonsingular matrix P_{k-2} that satisfies:

(a) for a combined factorization of A_{k-2}, A_{k-1} of P-type,

$$P_{k-2}D_{k-2}P_{k-1}^{-1} = D_{k-2};$$

(b) for a combined factorization of A_{k-2}, A_{k-1} of Q-type,

$$P_{k-2}D_{k-2}P_{k-1}^* = D_{k-2}.$$

However, we also have to change the factorization of $A_{k-3}, \dots$. This goes on, backwards through the sequence of factorizations of $A_{k-1}, A_{k-2}, A_{k-3}, \dots$ till we arrive at the factorization of the matrix A_1, where it will be necessary to change the matrix U_1 into U_1P_1, where P_1 must be unitary (because U_1P_1 must be unitary from Theorem 1) and

(a) for a combined factorization of A_1, A_2 of P-type,

$$P_1D_1P_2^{-1} = D_2, \qquad P_1 \text{ unitary};$$

(b) for a combined factorization of A_1, A_2 of Q-type,

$$P_1D_1P_2^* = D_1, \qquad P_1 \text{ unitary}.$$

The conclusion is that the inductive argument of Section 4.1, where we assumed that we had a factorization of the matrix A_k as in (18), leads to a necessary modification of the factorizations of the matrices A_{j-1}, $j = 2, \dots, k$, by inserting nonsingular matrices P_{j-1}, $j = 2, \dots, k$, such that

(a) for a combined factorization of A_{j-1}, A_j of P-type,

$$P_{j-1} D_{j-1} P_j^{-1} = D_{j-1}; \tag{21}$$

(b) for a combined factorization of A_{j-1}, A_j of Q-type,

$$P_{j-1} D_{j-1} P_j^{*} = D_{j-1}; \tag{22}$$

(c) P_1 is a unitary matrix.

This backward modification will be called the *ripple-through phenomenon*.

4.3. *A Detailed Analysis of the Inductive Argument and the Ripple-Through Phenomenon*

Having stated the two main ideas of the proof, we shall now show how the GSVD of $A_1, \dots, A_k$ can be obtained from the GSVD of $A_1, \dots, A_{k-1}$, essentially by using the block factorization lemma of Section 3. We'll show that:

1. The matrix P_k in (18) has a certain zero structure which is determined by the specific sequence of P's and Q's in the GSVD name. This will be investigated in Section 4.3.1.
2. The zero structure of P_k corresponds to a block row, block column permutation of a lower block triangular matrix T, i.e., there exists a block permutation matrix N such that

$$P_k = N^t T N. \tag{23}$$

The block permutation N is determined by the specific sequence of P's and Q's in the GSVD name. This is demonstrated in Section 4.3.2.

3. In Section 4.3.3 we show that the block diagonal blocks of T must be unitary.
4. From (18) and (23) we find

$$N Z_{k-1}^{-1} A_k = T N S_k V_k^{*}. \tag{24}$$

While the matrix NS_k is not a quasidiagonal matrix, it will be shown that there always exists a block permutation matrix M such that $NS_k M$ is quasidiagonal. Then (24) becomes

$$N Z_{k-1}^{-1} A_k = T (N S_k M)(M^t V_k^{*}). \tag{25}$$

The factorization in (25) can be obtained from a straightforward application of the block factorization lemma of Section 3. The algorithm to find the block permutation matrix M is given in Section 4.3.4.

We shall now demonstrate that P_k is a block row, block column permutation of a lower block triangular matrix as in (23), and we shall find the precise structure of the block permutation matrices N and M.

4.3.1. *Fact 1: The Matrix P_k Has a Certain Block Zero Structure.* First, we will show that the equations (21) and (22) impose a certain block zero structure on the matrix P_k.

Consider the combined factorization of the matrices A_{j-1}, A_j. There are two possible cases: The combined factorization is of Q-type or it is of P-type.

First assume that it is of P-type. Then, from (21) it follows that

$$P_{j-1} D_{j-1} = D_{j-1} P_j. \tag{26}$$

The idea is to find a nonsingular solution P_{j-1} when P_j and D_{j-1} are known, and to discuss the conditions for consistency of the matrix equation (26). Recall the structure of the matrix D_{j-1}:

$$\underset{n_{j-2} \times n_{j-1}}{D_{j-1}} = \begin{array}{r} \\ r_{j-1}^{1} \\ r_{j-2}^{1} - r_{j-1}^{1} \\ r_{j-1}^{2} \\ r_{j-2}^{2} - r_{j-1}^{2} \\ \vdots \\ r_{j-1}^{j-1} \\ n_{j-2} - r_{j-2} - r_{j-1}^{j-1} \end{array} \begin{array}{c} \begin{array}{ccccc} r_{j-1}^{1} & r_{j-1}^{2} & \cdots & r_{j-1}^{j-1} & n_{j-1} - r_{j-1} \end{array} \\ \left(\begin{array}{ccccc} I & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{array} \right) \end{array}. \tag{27}$$

We partition the matrix P_j according to the block columns of D_{j-1} as

$$\underset{n_{j-1}\times n_{j-1}}{P_j} = \begin{array}{c} r^1_{j-1} \\ r^2_{j-1} \\ \vdots \\ n_{j-1}-r_{j-1} \end{array} \overset{\begin{array}{ccccc} r^1_{j-1} & r^2_{j-1} & \cdots & r^{j-1}_{j-1} & n_{j-1}-r_{j-1} \end{array}}{\begin{pmatrix} P_j^{11} & P_j^{12} & \cdots & P_j^{1(j-1)} & P_j^{1j} \\ P_j^{21} & P_j^{22} & \cdots & P_j^{2(j-1)} & P_j^{2j} \\ \vdots & \vdots & & \vdots & \vdots \\ P_j^{j1} & P_j^{j2} & \cdots & P_j^{j(j-1)} & P_j^{jj} \end{pmatrix}}. \qquad (28)$$

Hence we find that

$$D_{j-1}P_j = \begin{array}{c} r^1_{j-1} \\ r^1_{j-2}-r^1_{j-1} \\ r^2_{j-1} \\ r^2_{j-2}-r^2_{j-1} \\ \vdots \\ r^{j-1}_{j-1} \\ n_{j-2}-r_{j-2}-r^{j-1}_{j-1} \end{array} \overset{\begin{array}{ccccc} r^1_{j-1} & r^2_{j-1} & \cdots & r^{j-1}_{j-1} & n_{j-1}-r_{j-1} \end{array}}{\begin{pmatrix} P_j^{11} & P_j^{12} & \cdots & P_j^{1(j-1)} & P_j^{1j} \\ 0 & 0 & \cdots & 0 & 0 \\ P_j^{21} & P_j^{22} & \cdots & P_j^{2(j-1)} & P_j^{2j} \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ P_j^{(j-1)1} & P_j^{(j-1)2} & \cdots & P_j^{(j-1)(j-1)} & P_j^{(j-1)j} \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}} \qquad (29)$$

The following two observations are crucial:

(a) It follows from (26) and the fact that the last block column of D_{j-1} (27) is zero that the last block column in (29) must be zero. This implies that the matrix equation (26) is consistent only if

$$P_j^{1j} = 0, \qquad P_j^{2j} = 0, \ldots, \qquad P_j^{(j-1)j} = 0. \qquad (30)$$

(b) It can be seen from (26), (28), and (29) that the last block row of P_j plays no role in the determination of P_{j-1}.

If the consistency condition (30) is satisfied, a nonsingular solution P_{j-1} of the matrix equation (26) is given by Table 1. The nonsingularity of P_{j-1} follows from the factorization lemma as shown below. Also note that P_{j-1} is partitioned into $j-1$ block columns and block rows.

TABLE 1

$$\underset{n_{j-2}\times n_{j-2}}{P_{j-1}} = \begin{array}{c} r^1_{j-2} \\ r^2_{j-2} \\ \vdots \\ r^{j-2}_{j-2} \\ n_{j-2}-r_{j-2} \end{array} \overset{\begin{array}{ccccc} r^1_{j-2} & r^2_{j-2} & \cdots & r^{j-2}_{j-2} & n_{j-2}-r_{j-2} \end{array}}{\begin{pmatrix} \begin{pmatrix} P_j^{11} & 0 \\ 0 & I \end{pmatrix} & \begin{pmatrix} P_j^{12} & 0 \\ 0 & 0 \end{pmatrix} & \cdots & \begin{pmatrix} P_j^{1(j-2)} & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} P_j^{1(j-1)} & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} P_j^{21} & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} P_j^{22} & 0 \\ 0 & I \end{pmatrix} & \cdots & \begin{pmatrix} P_j^{2(j-2)} & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} P_j^{2(j-1)} & 0 \\ 0 & 0 \end{pmatrix} \\ \vdots & \vdots & & \vdots & \vdots \\ \begin{pmatrix} P_j^{(j-2)1} & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} P_j^{(j-2)2} & 0 \\ 0 & 0 \end{pmatrix} & \cdots & \begin{pmatrix} P_j^{(j-2)(j-2)} & 0 \\ 0 & I \end{pmatrix} & \begin{pmatrix} P_j^{(j-2)(j-1)} & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} P_j^{(j-1)1} & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} P_j^{(j-1)2} & 0 \\ 0 & 0 \end{pmatrix} & \cdots & \begin{pmatrix} P_j^{(j-1)(j-2)} & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} P_j^{(j-1)(j-1)} & 0 \\ 0 & I \end{pmatrix} \end{pmatrix}}$$

Next, assume that the combined factorization of A_{j-1}, A_j is of Q-type. Then from (22) we find

$$D_{j-1} P_j^* = P_{j-1}^{-1} D_{j-1}. \tag{31}$$

Exploiting the structure of D_{j-1} (27), we arrive at the following observations as above:

(a) The matrix equation (31) is consistent only if

$$P_j^{j1} = 0, \qquad P_j^{j2} = 0, \ldots, \qquad P_j^{j(j-1)} = 0. \tag{32}$$

(b) The last block column of P_j plays no role in the determination of P_{j-1}.

From (31) we can find a solution for P_{j-1}^{j-1} as shown in Table 2. Observe that for this Q-type factorization, we determine the inverse P_{j-1}^{-1} from D_{j-1} and P_j, instead of the matrix P_{j-1} itself as with the P-type factorization. It is not necessary to determine P_{j-1} explicitly, because we can easily determine P_{j-2} or its inverse from either P_{j-1} or P_{j-1}^{-1}, as can be seen as follows: Again, consider the two cases of a P-type and a Q-type factorization of A_{j-2}, A_{j-1}.

(a) Assume that the combined factorization of A_{j-2}, A_{j-1} is of P-type. Then $D_{j-2} P_{j-1}^{-1} = P_{j-2}^{-1} D_{j-2}$, which is consistent only if the first $j-2$ blocks of the last block column of P_{j-1}^{-1} are zero. But from Table 2 it then follows that $P_j^{(j-1)1} = 0$, $P_j^{(j-1)2} = 0, \ldots, P_j^{(j-1)(j-2)} = 0$. The last block row of P_{j-1}^{-1} plays no role in the determination of P_{j-2}^{-1}.

(b) Assume that the combined factorization of A_{j-2}, A_{j-1} is of Q-type. Then $D_{j-2} P_{j-1}^* = P_{j-2}^{-1} D_{j-2}$, which is consistent only if the first $j-2$ blocks of the last block row of P_{j-1}^{-1} are zero: $P_j^{1(j-1)} = 0$, $P_j^{2(j-1)} = 0, \ldots, P_j^{(j-2)(j-1)} = 0$. The last block column of P_{j-1}^{-1} plays no role in the determination of P_{j-2}^{-1}.

The conclusion is that there is no need for explicit inversion of the matrix P_{j-1} but that we can reach similar conclusions from the matrix P_{j-1}^{-1}.

Depending on the sequence in which the letters P and Q appear, we now apply, for $j = k, k-1, \ldots, 2$, either equation (21) or (22) to determine the matrices $P_{k-1}, P_{k-2}, \ldots, P_1$ or their inverses. The consistency of these equations implies that certain blocks of P_k must be zero, in a structured manner that we shall elucidate. Let us first give an example.

TABLE 2

$$\underset{n_{j-2} \times n_{j-2}}{P_{j-1}^{-1}} = \begin{array}{c|ccccc} & r_{j-2}^{1} & r_{j-2}^{2} & \cdots & r_{j-2}^{j-2} & n_{j-2}-r_{j-2} \\ \hline r_{j-2}^{1} & \begin{pmatrix} (P_j^{11})^* & 0 \\ 0 & I \end{pmatrix} & \begin{pmatrix} (P_j^{21})^* & 0 \\ 0 & 0 \end{pmatrix} & \cdots & \begin{pmatrix} (P_j^{(j-2)1})^* & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} (P_j^{(j-1)1})^* & 0 \\ 0 & 0 \end{pmatrix} \\ r_{j-2}^{2} & \begin{pmatrix} (P_j^{12})^* & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} (P_j^{22})^* & 0 \\ 0 & I \end{pmatrix} & \cdots & \begin{pmatrix} (P_j^{(j-2)2})^* & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} (P_j^{(j-1)2})^* & 0 \\ 0 & 0 \end{pmatrix} \\ \cdots & \vdots & \vdots & & \vdots & \vdots \\ r_{j-2}^{j-2} & \begin{pmatrix} (P_j^{1(j-2)})^* & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} (P_j^{2(j-2)})^* & 0 \\ 0 & 0 \end{pmatrix} & \cdots & \begin{pmatrix} (P_j^{(j-2)(j-2)})^* & 0 \\ 0 & I \end{pmatrix} & \begin{pmatrix} (P_j^{(j-1)(j-2)})^* & 0 \\ 0 & 0 \end{pmatrix} \\ n_{j-2}-r_{j-2} & \begin{pmatrix} (P_j^{1(j-1)})^* & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} (P_j^{2(j-1)})^* & 0 \\ 0 & 0 \end{pmatrix} & \cdots & \begin{pmatrix} (P_j^{(j-2)(j-1)})^* & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} (P_j^{(j-1)(j-1)})^* & 0 \\ 0 & I \end{pmatrix} \end{array}$$

EXAMPLE 4. In order to illustrate the ripple-through phenomenon, we derive the QPQ-SVD of the matrices (A_1, A_2, A_3, A_4) from a QP-SVD of the matrices (A_1, A_2, A_3):

$$A_1 = U_1 D_1 X_1^{-1},$$

$$A_2 = X_1^{-*} D_2 X_2^{-1},$$

$$A_3 = X_2 S_3 V_3^*.$$

First, the factorization for A_3 is rewritten as $A_3 = X_2 D_3 X_3^{-1}$, where D_3 is a quasidiagonal matrix the nonzero blocks of which are identity matrices. Next, assume that we have a factorization of $X_3^* A_1$ as $X_3^* A_4 = P_4 S_4 V_4^*$, where P_4 is nonsingular, S_4 is quasidiagonal, and V_4 is unitary. The ripple-through phenomenon now consists of a modification of the factorizations of A_1, A_2, A_3 by introducing nonsingular matrices P_3, P_2, P_1 such that

$$A_4 = X_3^{-*} P_4 S_4 V_4^*,$$

$$A_3 = X_2 P_3 D_3 P_4^* X_3^{-1} \quad \rightarrow \quad P_3 D_3 P_4^* = D_3,$$

$$A_2 = X_1^{-*} P_2 D_2 P_3^{-1} X_2^{-1} \quad \rightarrow \quad P_2 D_2 P_3^{-1} = D_2,$$

$$A_1 = U_1 P_1 D_1 P_2^* X_1^{-1} \quad \rightarrow \quad P_1 D_1 P_2^* = D_1.$$

The equations on the right are consistent only if the matrix P_4 has a certain structure. Partition P_4 according to the block dimensions of S_4 as

$$P_4 = \begin{pmatrix} P_4^{11} & P_4^{12} & P_4^{13} & P_4^{14} \\ P_4^{21} & P_4^{22} & P_4^{23} & P_4^{24} \\ P_4^{31} & P_4^{32} & P_4^{33} & P_4^{34} \\ P_4^{41} & P_4^{42} & P_4^{43} & P_4^{44} \end{pmatrix}.$$

Then, using the definitions of the matrices P_3, P_2, P_1 as above, we find that:

(a) From $P_3 D_3 P_4^* = D_3$, it follows that $P_4^{41} = 0$, $P_4^{42} = 0$, $P_4^{43} = 0$.
(b) From $P_2 D_2 P_3^{-1} = D_2$, it follows that $P_4^{31} = 0$, $P_4^{32} = 0$.
(c) From $P_1 D_1 P_2^* = D_1$, it follows that $P_4^{12} = 0$.

Hence

$$P_4 = \begin{pmatrix} P_4^{11} & 0 & P_4^{13} & P_4^{14} \\ P_4^{21} & P_4^{22} & P_4^{23} & P_4^{24} \\ 0 & 0 & P_4^{33} & P_4^{34} \\ 0 & 0 & 0 & P_4^{44} \end{pmatrix}.$$

4.3.2. *Fact 2: The Matrix P_k is a Block Row, Block Column Permutation of a Lower Block Triangular Matrix.* The following observation follows immediately from the recursive definition of the matrices P_j, $j = 1, \dots, k-1$, in Section 4.3.1:

The net effect of the consistency condition (21) *or* (22) *is that the first $j-1$ blocks of the jth block row or column of P_k are zero, with $j = 1, \dots, k$.*

The sequence in which block rows or columns are zeroed is determined by the sequence in which the letters P and Q appear in the GSVD name. Consider as an illustration all possible GSVDs for four matrices:

EXAMPLE 5. The reader may wish to verify that the zero structure of the matrix P_4 in every GSVD for four matrices is given by Table 3 (the zero

TABLE 3

QQQ	QQP	QPQ	QPP
$\begin{bmatrix} x & x & 0 & x \\ 0 & x & 0 & x \\ x & x & x & x \\ 0 & 0 & 0 & x \end{bmatrix}$	$\begin{bmatrix} x & 0 & x & 0 \\ x & x & x & 0 \\ 0 & 0 & x & 0 \\ x & x & x & x \end{bmatrix}$	$\begin{bmatrix} x & 0 & x & x \\ x & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{bmatrix}$	$\begin{bmatrix} x & x & 0 & 0 \\ x & x & 0 & 0 \\ 0 & x & x & 0 \\ x & x & x & x \end{bmatrix}$
PQQ	PQP	PPQ	PPP
$\begin{bmatrix} x & 0 & 0 & x \\ x & x & 0 & x \\ x & x & x & x \\ 0 & 0 & 0 & x \end{bmatrix}$	$\begin{bmatrix} x & x & x & 0 \\ 0 & x & x & 0 \\ 0 & 0 & x & 0 \\ x & x & x & x \end{bmatrix}$	$\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{bmatrix}$	$\begin{bmatrix} x & 0 & 0 & 0 \\ x & x & 0 & 0 \\ x & x & x & 0 \\ x & x & x & x \end{bmatrix}$

elements are represented by zeros). Observe that in each of the cases, there are $j-1$ zero blocks in either block row j *or* block column j, $j=1,2,3,4$.

An important consequence of this observation also holds for the general case:

The matrix P_k is always a block row, block column permutation of a lower block triangular matrix T, i.e., there exists an $n_{k-1}\times n_{k-1}$ block permutation matrix N such that

$$P_k = N^t T N. \tag{33}$$

A careful analysis reveals that in general the block permutation matrix N can be obtained from the following algorithm:

ALGORITHM (For the block permutation matrix N).

1. Denote by $N_{(i)}$ the block permutation matrix obtained in step i of the recursion that follows.
2. Define the *block row reverse* of a partitioned identity matrix

$$I_{p_1+\cdots+p_q} = \begin{pmatrix} I_{p_1} & 0 & 0 & \cdots & 0 \\ 0 & I_{p_2} & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & I_{p_q} \end{pmatrix}$$

as

$$\begin{pmatrix} 0 & \cdots & 0 & I_{p_q} \\ 0 & \cdots & I_{p_{q-1}} & 0 \\ \vdots & & \vdots & \vdots \\ I_{p_1} & \cdots & 0 & 0 \end{pmatrix}.$$

Obviously, the block row reverse of the second matrix is again the identity matrix.

3. *Initialization*. Let $N_{(k)}$ be equal to the partitioned identity matrix

$$N_{(k)} = \begin{pmatrix} I_{r^1_{k-1}} & 0 & 0 & \cdots & 0 & 0 \\ 0 & I_{r^2_{k-1}} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_{r^{k-1}_{k-1}} & 0 \\ 0 & 0 & 0 & \cdots & 0 & I_{n_{k-1}-r_{k-1}} \end{pmatrix}, \tag{34}$$

in which the block dimensions correspond to the block dimensions of the partitioned matrix P_k.

4. For $i=k-1,k-2,\dots,1$: If the factorization of A_i, A_{i+1} is of
P-type, then $N_{(i)} = N_{(i+1)}$;
Q-type, then, $N_{(i)}$ is obtained from $N_{(i+1)}$ as follows:
Take the submatrix of $N_{(i+1)}$ formed by its first $i+1$ block columns.
Reverse the block rows that have a nonzero block.
5. The matrix $N_{(1)}$ is the desired block permutation matrix: $N = N_{(1)}$.

EXAMPLE 6. The block permutation matrix for a QPQ-SVD is obtained as

$$N_{(4)} = \begin{pmatrix} I_{r^1_3} & 0 & 0 & 0 \\ 0 & I_{r^2_3} & 0 & 0 \\ 0 & 0 & I_{r^3_3} & 0 \\ 0 & 0 & 0 & I_{n_3-r_3} \end{pmatrix} \to N_{(3)} = \begin{pmatrix} 0 & 0 & 0 & I_{n_3-r_3} \\ 0 & 0 & I_{r^3_3} & 0 \\ 0 & I_{r^2_3} & 0 & 0 \\ I_{r^1_3} & 0 & 0 & 0 \end{pmatrix}$$

$$\to N_{(2)} = \begin{pmatrix} 0 & 0 & 0 & I_{n_3-r_3} \\ 0 & 0 & I_{r^3_3} & 0 \\ 0 & I_{r^2_3} & 0 & 0 \\ I_{r^1_3} & 0 & 0 & 0 \end{pmatrix} \to N_{(1)} = \begin{pmatrix} 0 & 0 & 0 & I_{n_3-r_3} \\ 0 & 0 & I_{r^3_3} & 0 \\ I_{r^1_3} & 0 & 0 & 0 \\ 0 & I_{r^2_3} & 0 & 0 \end{pmatrix}.$$

4.3.3. *Fact 3: The Block Diagonal Blocks of T Must Be Unitary.* An important consequence of the fact that P_k is a block row, block column permutation of a lower block triangular matrix T is that the block diagonal blocks of P_k and T must be the same (but possibly ordered differently). The matrix P_1 or its inverse at the end of the ripple-through phenomenon in Section 4.3.1 is of the form

$$\begin{pmatrix} P_k^{11} & 0 \\ 0 & I \end{pmatrix},$$

where P_k^{11} is the first block diagonal block of P_k, which can be any block diagonal block of T, depending on the specific sequence of P's and Q's in the GSVD name. But because U_1P_1 must be unitary, it follows that P_k^{11} must be unitary; hence, all block diagonal blocks of T must be unitary.

4.3.4 Fact 4: A Block Permutation Such That NSM Is Quasidiagonal. Let us first summarize what we have obtained so far.

Consider the GSVD of $A_1,\ldots,A_{k-1}$ as in (13). If a nonsingular matrix P_k can be obtained such that $A_k = Z_{k-1}P_kS_kV_k^*$, then P_k must be a block row, block column permutation of a lower block triangular matrix T, i.e., there exists a block permutation matrix N such that $P_k = N^tTN$. An algorithm to obtain N was given in Section 4.3.2. We then find that

$$NZ_{k-1}^{-1}A_k = TNS_kV_k^*.$$

However, if S_k has the quasidiagonal structure in (4), then NS_k is a block row permutation of that quasidiagonal structure. It is, however, always possible to find an $n_k \times n_k$ block permutation matrix M such that NS_kM is quasidiagonal. We then have

$$\begin{aligned} NZ_{k-1}^{-1}A_k &= TNS_kMM^tV_k^* \\ &= T(NS_kM)(V_kM)^* \\ &= T\tilde{S}\tilde{V}^*, \quad \text{say}, \end{aligned}$$

where $\tilde{S}$ is quasidiagonal and $\tilde{V}$ is unitary. Obviously, the matrices T, $\tilde{S}$, and $\tilde{V}$ follow immediately from the block factorization lemma of Section 3 applied to the matrix $NZ_{k-1}^{-1}A_k$, in which the scalars p_i (which are the block row dimensions) follow from the dimensions of the block rows of N.

The block permutation matrix M can be obtained as follows:

Algorithm (For block permutation matrix M).

1. The upper $r_k \times r_k$ block of M^t (the transpose of M) has the same zero-identity block structure as the matrix N, the identity matrix in block column j of M^t being $I_{r_k^j}$.
2. In block position $(k+1, k+1)$ of M^t (lower right hand corner), there is an $(n_k - r_k)\times(n_k - r_k)$ identity matrix $I_{n_k - r_k}$.

Example 7. Let us proceed with the QPQ-SVD of Example 6. From the block factorization lemma, one obtains a decomposition as

$$N_{(1)}Z_3^{-1}A_4 = T\tilde{S}\tilde{V}^*$$

$$= T \begin{array}{r} \\ \tilde r_1 \\ n_3-r_3-\tilde r_1 \\ \tilde r_2 \\ r_3^3-\tilde r_2 \\ \tilde r_3 \\ r_3^1-\tilde r_3 \\ \tilde r_4 \\ r_3^2-\tilde r_4 \end{array} \begin{pmatrix} \tilde r_1 & \tilde r_2 & \tilde r_3 & \tilde r_4 & n_4-r_4 \\ S^1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & S^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & S^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & S^4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tilde{V}^*,$$

where the integers $\tilde r_i$, $i = 1,2,3,4$, are the recursively defined integers of the block factorization lemma.

It is easily found that

$$N^t\tilde{S} = \begin{array}{r} \\ \tilde r_3 \\ n_3-r_3 \\ \tilde r_4 \\ r_3^3-\tilde r_4 \\ \tilde r_2 \\ r_3^2-\tilde r_2 \\ \tilde r_1 \\ r_3^1-\tilde r_1 \end{array} \begin{pmatrix} \tilde r_1 & \tilde r_2 & \tilde r_3 & \tilde r_4 & n_4-r_4 \\ 0 & 0 & S^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & S^4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & S^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ S^1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The dimensions r_4^j, $j = 1,2,3,4$, are now defined as

$$r_4^1 = \tilde r_3, \qquad r_4^2 = \tilde r_4, \qquad r_4^3 = \tilde r_2, \qquad r_4^4 = \tilde r_1.$$

The matrix M is constructed via

$$M^t = \begin{pmatrix} 0 & 0 & 0 & I_{r_4^4} & 0 \\ 0 & 0 & I_{r_4^3} & 0 & 0 \\ I_{r_4^1} & 0 & 0 & 0 & 0 \\ 0 & I_{r_4^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{n_4 - r_4} \end{pmatrix}.$$

4.4. *Summary of the Constructive Proof*

Putting together all elements of Section 4.3, we find the following algorithm, which at the same time is a constructive proof, to derive the GSVD of the matrices $A_1, \dots, A_k$ from a corresponding one of $A_1, \dots, A_{k-1}$ as in (13):

1. Determine the block permutation matrices N and M from the algorithms in Sections 4.3.2 and 4.3.4.
2. Apply the block factorization lemma of Section 3 to the partitioned matrix

$$NZ_{k-1}^{-1} A_k = T\tilde{S}_k \tilde{V}_k^*,$$

where T is lower block triangular with unitary block diagonal matrices, $\tilde{S}_k$ is quasidiagonal, and $\tilde{V}_k$ is unitary.

3. Determine the matrices $P_k = N^t TN$, $S_k = N^t \tilde{S}_k M^t$, and $V_k = \tilde{V}_k M^t$. The factorization of the matrix A_k becomes $A_k = (Z_{k-1} P_k) S_k V_k^*$.
4. Determine the matrices $P_{k-1}, P_{k-2}, \dots, P_1$ from the ripple-through recursion of Section 4.3.1, and update the factorizations of the matrices $A_{k-1}, \dots, A_1$ accordingly.
5. The obtained factorizations of $A_1, \dots, A_k$ constitute a GSVD, the structure and properties of which conform with the statement of the main theorem.

5. CONCLUSIONS

In this paper, we have proposed a tree of generalizations of the singular value decomposition, with at the top the OSVD for one matrix and the PSVD and the QSVD for two matrices. Our proof is constructive, the main idea being an induction step that permits us to find a GSVD at level k from a corresponding one at level $k-1$. This necessitates a detailed analysis of the backward modification of the chain of decompositions at level $k-1$, which we have called the ripple-through phenomenon. A key role is played by a certain block factorization lemma.

We are convinced that this tree of generalized singular value decompositions will open new and exciting fields of research with respect to numerical algorithms, geometrical interpretations, and applications. For instance, the uniqueness issues, the relation to generalized eigenvalue problems, genericity and sensitivity analysis, and numerical and implementational aspects definitely deserve more attention. We have already found and collected several interesting applications of these generalizations, which will be reported in subsequent publications. We have also derived expressions for the scalars r_j^i, which are the dimensions of the blocks in the quasidiagonal matrices D_j and S_k in the main theorem, in terms of ranks of the matrices A_j, $k = 1, \dots, k$, and products and concatenations thereof.

Bart De Moor would like to thank Professor Gene Golub of the department of Computer Science and Professor Thomas Kailath of the department of Electrical Engineering for the opportunity they gave him to spend a wonderful year at such a marvelous place as Stanford University. The detailed comments and suggestions of two anonymous referees have greatly improved the presentation of this paper.

1 L. Autonne, Sur les groupes linéaires, réels et orthogonaux, *Bull. Soc. Math. France* 30:121–133 (1902).

2 E. Beltrami, Sulle funzione bilineari, *Giornale Math.* 11:98–106 (1873).

3 B. De Moor and G. H. Golub, The Restricted Singular Value Decomposition: Properties and Applications, ESAT-SISTA Report 1989-09, Dept. of Electrical Engineering, Katholieke Univ. Leuven, Belgium, Apr. 1989, 55 pp.; Numerical Analysis Project Manuscript NA-89-03, Dept. of Computer Science, Stanford Univ.; *SIAM J. Matrix Anal. Appl.*, to appear.

4 B. De Moor and G. Golub, Generalized Singular Value Decompositions: A Proposal for a Standardized Nomenclature, ESAT-SISTA Report 1989-10, Dept. of Electrical Engineering, Katholieke Univ. Leuven, Belgium, Apr. 1989, 15 pp.; Numerical Analysis Project Manuscript NA-89-04, Dept. of Computer Science, Stanford Univ.

5 B. De Moor, On the Structure and Geometry of the Product Singular Value Decomposition, ESAT-SISTA Report 1989-12, Dept. of Electrical Engineering, Katholieke Univ. Leuven, Belgium, May 1989, 60 pp.; Numerical Analysis Project Manuscript NA-89-05, Dept. of Computer Science, Stanford Univ.; submitted for publication.

6 E. Deprettere (Ed.); *SVD and Signal Processing: Algorithms, Applications and Architectures*, North Holland, 1988.

7 K. V. Fernando and S. J. Hammarling, A Product Induced Singular Value Decomposition for Two Matrices and Balanced Realisation, NAG Technical Report TR8/87.

8 G. H. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins U.P., Baltimore, 1983.

9 M. T. Heath, A. J. Laub, C. C. Paige, and R. C. Ward, Computing the singular value decomposition of a product of two matrices, *SIAM J. Sci. Statist. Comput.* 7:1147–1159 (1986).

10 C. J. Jordan, Mémoire sur les formes bilinéaires, *J. Math. Pures Appl. II* 19:35–54 (1874).

11 C. C. MacDuffee, *The Theory of Matrices*, Springer-Verlag, Berlin, 1933; corrected reprint, Chelsea, New York.

12 C. C. Paige and M. A. Saunders, Towards a generalized singular value decomposition, *SIAM J. Numer. Anal.* 18:398–405 (1981).

13 J. J. Sylvester, Sur la réduction biorthogonale d'une forme linéo-linéaire à sa forme canonique, *C.R. Acad. Sci. Paris* CVIII:651–653 (1889).

14 C. F. Van Loan, Generalizing the singular value decomposition, *SIAM J. Numer. Anal.* 13:76–83 (1976).

15 H. Zha, Restricted SVD for Matrix Triplets and Rank Determination of Matrices, Scientific Report 89-2, ZIB, Berlin; submitted for publication.

Received 15 November 1989; final manuscript accepted 16 August 1990