MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Simultaneous Singular Value Decomposition

Takanori MAEHARA and Kazuo MUROTA

METR 2009–14 April 2009

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

Simultaneous Singular Value Decomposition

Takanori Maehara* and Kazuo Murota*

April 2009

Abstract

We consider the following problem: Given a set of $m \times n$ real (or complex) matrices $A_1, \ldots, A_N$, find an $m \times m$ orthogonal (or unitary) matrix P and an $n \times n$ orthogonal (or unitary) matrix Q such that $P^* A_1 Q, \ldots, P^* A_N Q$ are in a common block-diagonal form with possibly rectangular diagonal blocks. We call this the *simultaneous singular value decomposition* (simultaneous SVD). The name is motivated by the fact that the special case with $N = 1$, where a single matrix is given, reduces to the ordinary SVD. With the aid of the theory of $*$-algebra and bimodule it is shown that a finest simultaneous SVD is uniquely determined. An algorithm is proposed for finding the finest simultaneous SVD on the basis of recent algorithms of Murota–Kanno–Kojima–Kojima and Maehara–Murota for simultaneous block-diagonalization of square matrices under orthogonal (or unitary) similarity.

Keywords: singular value decomposition, block-diagonalization, matrix $*$-algebra, bimodule, eigenvalue

*Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. maehara@misojiro.t.u-tokyo.ac.jp, murota@mist.i.u-tokyo.ac.jp

1 Introduction

Singular value decomposition (SVD) is one of the most fundamental tools in dealing with noisy data. It is useful, for instance, in least squares method, principal component analysis, and matrix approximations. Mathematically, the singular value decomposition of an $m \times n$ real matrix A is to transform A to a diagonal matrix, with nonnegative diagonal elements, through a transformation of the form $P^\top AQ$ with an $m \times m$ orthogonal matrix P and an $n \times n$ orthogonal matrix Q. Singular value decomposition can also be defined for a complex matrix A, where a unitary transformation $P^* AQ$ with unitary matrices P and Q is employed.

In this paper we consider such decompositions for a family of matrices, which we call the *simultaneous singular value decomposition*. We distinguish two cases, decompositions over $\mathbb{R}$ and over $\mathbb{C}$:

> Problem [$\mathbb{R}$]: Given a set of $m \times n$ real matrices $A_1, \ldots, A_N$, find an $m \times m$ orthogonal matrix P and an $n \times n$ orthogonal matrix Q such that $P^\top A_1 Q, \ldots, P^\top A_N Q$ are in a common block-diagonal form.

> Problem [$\mathbb{C}$]: Given a set of $m \times n$ complex matrices $A_1, \ldots, A_N$, find an $m \times m$ unitary matrix P and an $n \times n$ unitary matrix Q such that $P^* A_1 Q, \ldots, P^* A_N Q$ are in a common block-diagonal form.

Obviously, the special case with $N = 1$, where a single matrix is given, reduces to the ordinary singular value decomposition. In this special case we obtain a (genuine) diagonal matrix, which means that a family of orthogonal one-dimensional subspaces are identified as special directions of importance, and the singular vectors are the bases for these subspaces. For multiple matrices, we cannot hope for simultaneous diagonalization but we look for a common block-diagonal form, where the diagonal blocks are possibly rectangular matrices. This means that we are to identify a family of mutually orthogonal subspaces characteristic to the given family of matrices. It may be said that the diagonal blocks in our decomposition are higher dimensional extensions of singular values, which are scalars (or 1×1 matrices).

This paper shows, with the theory of $*$-algebra and bimodule, that a finest simultaneous singular value decomposition exists and is uniquely determined. Moreover, structure theorems will be established in both cases (see Theorems 2 and 7). As an immediate corollary of the structure theorems we obtain a necessary and sufficient condition for the simultaneous diagonalization under the transformation $P^\top A_i Q$ or $P^* A_i Q$ (see Corollaries 3 and 8).

Our construction of simultaneous SVD is a natural extension of the well-known fact that the SVD of a single (real) matrix A can be constructed from the eigenvalue decompositions of $AA^\top$ and $A^\top A$. In place of the eigenvalue decompositions of $AA^\top$ and $A^\top A$, we use the Wedderburn-type canonical decompositions of the $*$-algebra generated by $A_i A_j^\top$ $(i, j = 1, \ldots, N)$ and the $*$-algebra generated by $A_i^\top A_j$ $(i, j = 1, \ldots, N)$. Then using the theoretical framework of bimodule we can derive the desired simultaneous SVD. In the structure theorems for simultaneous SVD there is a substantial difference between $\mathbb{R}$ and $\mathbb{C}$, which stems from the difference in the structure theorems of matrix $*$-algebra over $\mathbb{R}$ and $\mathbb{C}$.

An algorithm is proposed for finding the simultaneous SVD. This is built upon recent algorithms of Murota–Kanno–Kojima–Kojima [8] and Maehara–Murota [7] for simultaneous block-diagonalization of square matrices, i.e., for finding, given a set of square matrices $B_1, \ldots, B_N$, an orthogonal (or unitary) matrix P such that $P^* B_1 P, \ldots, P^* B_N P$ are in a common block-diagonal form.

In the literature of semidefinite programming group representation theory and matrix $*$-algebra have been attracting research interest as effective tools for exploiting algebraic structures due to symmetry, sparsity, etc. [1, 2, 4, 5, 6, 8, 9]. Typically, we are given a family of symmetric (or Hermitian) matrices $B_1, \ldots, B_N$ such that each $B = B_i$ is endowed with invariance to a finite group G in the sense of $T(g)^* B T(g) = B$ $(g \in G)$ with respect to an orthogonal (or unitary) representation T. Then the problem is to find an orthogonal (or unitary) matrix P such that $P^* B_1 P, \ldots, P^* B_N P$ are in the same block-diagonal form. In contrast, the simultaneous SVD of the present paper corresponds to equivariance in the sense of $S(g)^* A T(g) = A$ $(g \in G)$ with respect to orthogonal (or unitary) representations S and T. A standard result in group representation theory affords a canonical decomposition for such matrices. Our contribution is to generalize this by means of bimodule, and also to give an algorithm for the decomposition.

The structure theorems of $*$-algebras form the foundation of the decomposition method for semidefinite programs. It is hoped that the structure theorems established in this paper trigger a new direction in some area of optimization or data science.

2 Structure theorem for simultaneous SVD over $\mathbb{C}$

Problem [$\mathbb{C}$] is considered in this section. As a preliminary the structure theorem of matrix $*$-algebras is described in §2.1 and the simultaneous SVD is constructed in §2.2.

2.1 Matrix $*$-algebra over $\mathbb{C}$

We denote by $\mathcal{M}_{m,n} = \mathcal{M}_{m,n}(\mathbb{C})$ the set of $m \times n$ complex matrices, and put $\mathcal{M}_n = \mathcal{M}_{n,n}$. A subset $\mathcal{T}$ of $\mathcal{M}_n$ is said to be a $*$-subalgebra (or a matrix $*$-algebra) over $\mathbb{C}$ if $I_n \in \mathcal{T}$ and $[A, B \in \mathcal{T}; \alpha, \beta \in \mathbb{C} \implies \alpha A + \beta B, AB, A^* \in \mathcal{T}]$. We say that a matrix $*$-algebra $\mathcal{T}$ is simple if $\mathcal{T}$ has no ideal other than $\{O\}$ and $\mathcal{T}$ itself, where an ideal of $\mathcal{T}$ means a submodule $\mathcal{I}$ of $\mathcal{T}$ such that $[A \in \mathcal{T}, B \in \mathcal{I} \implies AB, BA \in \mathcal{I}]$. A linear subspace W of $\mathbb{C}^n$ is said to be invariant with respect to $\mathcal{T}$, or $\mathcal{T}$-invariant, if $AW \subseteq W$ for every $A \in \mathcal{T}$. We say that $\mathcal{T}$ is irreducible if no $\mathcal{T}$-invariant subspace other than $\{\mathbf{0}\}$ and $\mathbb{C}^n$ exists.

The following is a standard result in $*$-algebra (e.g., [10, Chapter X]). Note that for a matrix $*$-algebra $\mathcal{T}$ and a unitary matrix P, the set $P^*\mathcal{T}P = \{P^*AP : A \in \mathcal{T}\}$ is another matrix $*$-algebra isomorphic to $\mathcal{T}$.

Theorem 1. Let $\mathcal{T}$ be a $*$-subalgebra of $\mathcal{M}_n(\mathbb{C})$.

(A) There exist a unitary matrix Q and simple $*$-subalgebras $\mathcal{T}_j$ of $\mathcal{M}_{\hat{n}_j}(\mathbb{C})$ for some $\hat{n}_j$ $(j = 1, 2, \ldots, \ell)$ such that $Q^*\mathcal{T}Q = \{\mathrm{diag}\,(S_1, S_2, \ldots, S_\ell) : S_j \in \mathcal{T}_j \ (j = 1, 2, \ldots, \ell)\}$.

(B) If $\mathcal{T}$ is simple, there exist a unitary matrix P and an irreducible $*$-subalgebra $\mathcal{T}'$ of $\mathcal{M}_{\bar{n}}(\mathbb{C})$ for some $\bar{n}$ such that $P^*\mathcal{T}P = \{\mathrm{diag}\,(B, B, \ldots, B) : B \in \mathcal{T}'\}$.

(C) If $\mathcal{T}$ is irreducible, $\mathcal{T} = \mathcal{M}_n(\mathbb{C})$.

2.2 Construction of simultaneous SVD over $\mathbb{C}$

For $*$-algebras $\mathcal{T}_L$ $(\subseteq \mathcal{M}_m(\mathbb{C}))$ and $\mathcal{T}_R$ $(\subseteq \mathcal{M}_n(\mathbb{C}))$ we call a submodule $\mathcal{A}$ of $\mathcal{M}_{m,n}(\mathbb{C})$ a matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule over $\mathbb{C}$ if $[A \in \mathcal{A}, L \in \mathcal{T}_L, R \in \mathcal{T}_R \implies LAR \in \mathcal{A}]$. Given a family of $m \times n$ complex matrices $A_1, \ldots, A_N$ we consider three algebraic structures:

(i) Matrix $*$-algebra $\mathcal{T}_L$ generated by $A_iA_j^*$ $(i, j = 1, \ldots, N)$.

(ii) Matrix $*$-algebra $\mathcal{T}_R$ generated by $A_i^*A_j$ $(i, j = 1, \ldots, N)$.

(iii) Matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule $\mathcal{A}$ generated by $A_1, \ldots, A_N$.

Note that $\mathcal{T}_L$ and $\mathcal{T}_R$ are determined by $\mathcal{A}$; that is, $\mathcal{T}_L$ and $\mathcal{T}_R$ are $*$-algebras generated, respectively, by $\mathcal{A}\mathcal{A}^*$ and $\mathcal{A}^*\mathcal{A}$. It is mentioned that if $A_i = O$ $(i = 1, \ldots, N)$, we have $\mathcal{A} = \{O\}$, and then both $\mathcal{T}_L$ and $\mathcal{T}_R$ are $*$-algebras generated by zero matrices, which means that $\mathcal{T}_L = \mathbb{C}I_m$ and $\mathcal{T}_R = \mathbb{C}I_n$, since a $*$-algebra (in our present definition) always contains the identity matrix. Such a degenerate case needs to be included as it may possibly occur as a result of our decomposition.

The fundamental fact underlying our approach is that decomposing the given matrices $A_1, \ldots, A_N$ by means of a transformation of the form P^*A_iQ

is equivalent to decomposing every element A of $\mathcal{A}$ by P^*AQ. Accordingly we assume that we are given a matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule $\mathcal{A}$ $(\subseteq \mathcal{M}_{m,n}(\mathbb{C}))$ such that $\mathcal{T}_L$ and $\mathcal{T}_R$ are $*$-algebras generated, respectively, by $\mathcal{A}\mathcal{A}^*$ and $\mathcal{A}^*\mathcal{A}$. Note that no reference is made to the generators $A_1, \ldots, A_N$ in this setting.

The following theorem shows that the simultaneous SVD, i.e., the finest decomposition under $P^*A_1Q, \ldots, P^*A_NQ$ can be constructed from the decompositions of $*$-algebras $\mathcal{A}\mathcal{A}^*$ and $\mathcal{A}^*\mathcal{A}$ in the sense of Theorem 1. Note that this construction generalizes the construction of the SVD of a single matrix A through the eigenvalue decompositions of AA^* and A^*A.

Theorem 2. Let $\mathcal{A} \subseteq \mathcal{M}_{m,n}(\mathbb{C})$, $\mathcal{A} \neq \{O\}$, be a matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule over $\mathbb{C}$ such that $\mathcal{T}_L$ and $\mathcal{T}_R$ are $*$-algebras generated, respectively, by $\mathcal{A}\mathcal{A}^*$ and $\mathcal{A}^*\mathcal{A}$.

(A) There exist unitary matrices P and Q and a natural number ℓ such that

$$P^*\mathcal{T}_L P = \mathcal{T}_{L1} \oplus \cdots \oplus \mathcal{T}_{L\ell}, \quad P^*\mathcal{A}Q = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_\ell, \quad Q^*\mathcal{T}_R Q = \mathcal{T}_{R1} \oplus \cdots \oplus \mathcal{T}_{R\ell}.$$

Here each $\mathcal{A}_j$ is a matrix $(\mathcal{T}_{Lj}, \mathcal{T}_{Rj})$-bimodule, and $\mathcal{T}_{Lj}$ and $\mathcal{T}_{Rj}$ are simple matrix $*$-algebras generated by $\mathcal{A}_j\mathcal{A}_j^*$ and $\mathcal{A}_j^*\mathcal{A}_j$, respectively.

(B) If $\mathcal{T}_L$ and $\mathcal{T}_R$ are simple, there exist unitary matrices P and Q and a natural number μ such that

$$P^*\mathcal{T}_L P = I_\mu \otimes \mathcal{T}'_L, \qquad P^*\mathcal{A}Q = I_\mu \otimes \mathcal{A}', \qquad Q^*\mathcal{T}_R Q = I_\mu \otimes \mathcal{T}'_R.$$

Here $\mathcal{A}'$ is a matrix $(\mathcal{T}'_L, \mathcal{T}'_R)$-bimodule, and $\mathcal{T}'_L$ and $\mathcal{T}'_R$ are irreducible matrix $*$-algebras generated by $\mathcal{A}'\mathcal{A}'^*$ and $\mathcal{A}'^*\mathcal{A}'$, respectively.

(C) If $\mathcal{T}_L$ and $\mathcal{T}_R$ are irreducible, there exist unitary matrices P and Q such that

$$P^*\mathcal{T}_L P = \mathcal{M}_m(\mathbb{C}), \qquad P^*\mathcal{A}Q = \mathcal{M}_{m,n}(\mathbb{C}), \qquad Q^*\mathcal{T}_R Q = \mathcal{M}_n(\mathbb{C}).$$

As an immediate corollary we obtain a necessary and sufficient condition for complex matrices $A_1, \ldots, A_N$ to have the same set of singular vectors in the conventional sense.

Corollary 3. For complex matrices $A_1, \ldots, A_N$, there exist unitary matrices P and Q such that P^*A_iQ $(i = 1, \ldots, N)$ are diagonal if and only if $A_iA_j^*$ $(i, j = 1, \ldots, N)$ are all normal and commute with each other, and $A_i^*A_j$ $(i, j = 1, \ldots, N)$ are all normal and commute with each other.

Example 4. For two 4×8 complex matrices

$$A_1 = \begin{bmatrix} -0.911 & 2.886 & 3.216 & 0.605 & 0.451 & 3.066 & 3.501 & 1.611 \\ -0.312 & -0.447 & 3.673 & -0.384 & 5.036 & -3.539 & 0.082 & 0.876 \\ -0.216 & 1.852 & 3.006 & 0.314 & 0.675 & 1.841 & 3.570 & 0.727 \\ -0.426 & 1.005 & -1.986 & -0.222 & -2.619 & 3.533 & 1.339 & -0.246 \end{bmatrix}$$

$$+\imath \begin{bmatrix} -0.996 & 2.329 & 2.650 & -0.319 & 1.869 & 1.383 & 4.423 & 0.508 \\ 0.275 & 1.580 & 4.492 & -3.558 & 3.038 & -1.790 & -0.327 & -0.271 \\ -1.755 & 3.714 & 1.724 & 0.421 & 0.955 & 0.678 & 4.692 & -0.333 \\ -0.888 & 1.692 & -1.371 & 0.958 & -1.540 & 0.285 & 1.185 & -0.440 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -1.256 & 0.400 & 0.921 & 0.451 & 1.980 & 1.332 & 4.859 & 1.503 \\ 2.415 & -0.090 & 4.228 & -1.484 & 2.073 & 1.328 & 0.005 & -0.171 \\ 0.388 & 0.789 & 2.320 & 0.661 & 2.843 & 2.068 & 2.573 & 1.357 \\ -0.242 & 1.220 & -2.288 & 2.998 & -0.249 & 0.194 & 1.445 & 2.126 \end{bmatrix}$$

$$+\imath \begin{bmatrix} -1.659 & 3.148 & 0.960 & -0.740 & 0.248 & 2.196 & 5.254 & 1.201 \\ -1.002 & 0.706 & 4.215 & -1.143 & 3.649 & -0.668 & 0.453 & 0.466 \\ -0.560 & 1.419 & 4.783 & -0.072 & 2.268 & 0.050 & 3.540 & -0.334 \\ -0.820 & 1.492 & -1.731 & 1.241 & -2.950 & 2.742 & 2.540 & 2.272 \end{bmatrix}$$

we have

$$P^*A_1Q = \left[\begin{array}{cccc|cccc} 0.243 & -0.037 & -0.845 & -0.929 & 0 & 0 & 0 & 0 \\ 0.105 & -0.555 & 0.801 & -2.216 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -2.127 & 0.016 & -3.388 & -0.245 \\ 0 & 0 & 0 & 0 & -0.470 & 0.793 & 0.545 & -0.857 \end{array}\right]$$

$$+\imath \left[\begin{array}{cccc|cccc} 12.720 & -0.608 & -0.572 & 0.413 & 0 & 0 & 0 & 0 \\ 0.706 & 0.456 & -0.554 & -0.575 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 9.610 & -0.003 & -1.702 & 0.103 \\ 0 & 0 & 0 & 0 & -2.709 & 0.415 & 1.246 & 0.166 \end{array}\right],$$

$$P^*A_2Q = \left[\begin{array}{cccc|cccc} 1.726 & 0.092 & 1.256 & 1.284 & 0 & 0 & 0 & 0 \\ -1.339 & -0.205 & -2.439 & 1.214 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1.072 & -0.471 & 4.684 & 0.053 \\ 0 & 0 & 0 & 0 & 1.419 & 1.463 & 0.749 & 0.082 \end{array}\right]$$

$$+\imath \left[\begin{array}{cccc|cccc} 12.150 & 0.636 & 0.416 & -0.461 & 0 & 0 & 0 & 0 \\ 0.743 & 0.397 & -0.864 & -0.881 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 8.616 & 0.296 & 1.072 & -0.073 \\ 0 & 0 & 0 & 0 & 2.307 & -1.488 & 0.142 & -0.321 \end{array}\right]$$

with suitable unitary matrices P and Q. We have $\ell = 2$, $\mu = 1$ in Theorem 2, and accordingly both P^*A_1Q and P^*A_2Q belong to $\mathcal{M}_{2,4}(\mathbb{C}) \oplus \mathcal{M}_{2,4}(\mathbb{C})$.

Example 5. Consider two 4×6 matrices

$$A_1 = \begin{bmatrix} -1.433 & 0.234 & -0.517 & -0.347 & 0.008 & 2.097 \\ 0.508 & 0.337 & -0.142 & -2.177 & -1.271 & -0.466 \\ -0.024 & 1.412 & 1.337 & 0.126 & 0.160 & -0.100 \\ -1.573 & -0.154 & 0.262 & 0.446 & 0.199 & -1.556 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -0.215 & 0.525 & 0.342 & 1.538 & 1.063 & 2.151 \\ 1.486 & 1.040 & 0.247 & -1.456 & -0.712 & 1.589 \\ 2.630 & -1.493 & -0.841 & 1.129 & 0.449 & -0.025 \\ 1.112 & 1.748 & 2.080 & 0.744 & 0.444 & -0.837 \end{bmatrix}.$$

Then $A_i A_j^*$ $(i, j = 1, 2)$ are all normal and commute each other and $A_i^* A_j$ $(i, j = 1, 2)$ are all normal and commute each other. Therefore, by Corollary 3, there exist unitary matrices P and Q such that $P^* A_i Q$ $(i = 1, 2)$ are diagonal matrices, which read as follows:

$$P^* A_1 Q = \begin{bmatrix} -1.666 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.896 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1.433 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2.493 & 0 & 0 \end{bmatrix} + \imath \begin{bmatrix} 0.959 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1.701 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.388 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.241 & 0 & 0 \end{bmatrix},$$

$$P^* A_2 Q = \begin{bmatrix} 2.972 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.377 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.230 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2.753 & 0 & 0 \end{bmatrix} + \imath \begin{bmatrix} 1.616 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.219 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.779 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.447 & 0 & 0 \end{bmatrix}.$$

3 Structure theorem for simultaneous SVD over $\mathbb{R}$

Problem [$\mathbb{R}$] is considered in this section. The structure theorem of $*$-algebras is modified for $\mathbb{R}$ in §3.1 and the simultaneous SVD over $\mathbb{R}$ is constructed in §3.2.

3.1 Matrix $*$-algebra over $\mathbb{R}$

Matrix $*$-algebra over $\mathbb{R}$ and the associated concepts such as irreducibility are defined similarly as in §2.1, where "unitary" is replaced by "orthogonal."

The structure theorem, however, needs a revision stated in Theorem 6 below (see, e.g., [6], [8]).

Let $\mathbb{H}$ denote the quaternion field, i.e., $\mathbb{H} = \{a+\imath b+\jmath c+kd : a,b,c,d \in \mathbb{R}\}$ with the multiplication defined as: $\imath = \jmath k = -k\jmath$, $\jmath = k\imath = -\imath k$, $k = \imath\jmath = -\jmath\imath$, $\imath^2 = \jmath^2 = k^2 = -1$. We regard $\mathbb{C}$ as a subset of $\mathbb{H}$ by identifying $\imath$ with the imaginary unit in $\mathbb{C}$.

We define three types of matrices: the set of $m \times n$ real matrices $\mathcal{M}_{m,n} = \mathcal{M}_{m,n}(\mathbb{R})$, the real representation of complex matrices $\mathcal{C}_{m,n} \subset \mathcal{M}_{2m,2n}(\mathbb{R})$ defined by

$$\mathcal{C}_{m,n} = \left\{ \begin{bmatrix} C(z_{11}) & \cdots & C(z_{1n}) \\ \vdots & \ddots & \vdots \\ C(z_{m1}) & \cdots & C(z_{mn}) \end{bmatrix} : z_{11}, z_{12}, \ldots, z_{mn} \in \mathbb{C} \right\}$$

with

$$C(a+\imath b) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

and the real representation of quaternion matrices $\mathcal{H}_{m,n} \subset \mathcal{M}_{4m,4n}(\mathbb{R})$ defined by

$$\mathcal{H}_{n,m} = \left\{ \begin{bmatrix} H(h_{11}) & \cdots & H(h_{1n}) \\ \vdots & \ddots & \vdots \\ H(h_{m1}) & \cdots & H(h_{mn}) \end{bmatrix} : h_{11}, h_{12}, \ldots, h_{mn} \in \mathbb{H} \right\}$$

with

$$H(a+\imath b+\jmath c+kd) = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}.$$

We put $\mathcal{M}_n = \mathcal{M}_{n,n}$, $\mathcal{C}_n = \mathcal{C}_{n,n}$, $\mathcal{H}_n = \mathcal{H}_{n,n}$ for notational simplicity.

Theorem 6. Let $\mathcal{T}$ be a $*$-subalgebra of $\mathcal{M}_n = \mathcal{M}_n(\mathbb{R})$.

(A) There exist an orthogonal matrix Q and simple $*$-subalgebras $\mathcal{T}_j$ of $\mathcal{M}_{\hat{n}_j}(\mathbb{R})$ for some $\hat{n}_j$ $(j = 1,2,\ldots,\ell)$ such that $Q^\top \mathcal{T} Q = \{\mathrm{diag}\,(S_1, S_2, \ldots, S_\ell) : S_j \in \mathcal{T}_j \ (j = 1,2,\ldots,\ell)\}$.

(B) If $\mathcal{T}$ is simple, there exist an orthogonal matrix P and an irreducible $*$-subalgebra $\mathcal{T}'$ of $\mathcal{M}_{\bar{n}}(\mathbb{R})$ for some $\bar{n}$ such that $P^\top \mathcal{T} P = \{\mathrm{diag}\,(B, B, \ldots, B) : B \in \mathcal{T}'\}$.

(C) If $\mathcal{T}$ is irreducible, there exists an orthogonal matrix P such that $P^\top \mathcal{T} P = \mathcal{M}_n, \mathcal{C}_{n/2}$ or $\mathcal{H}_{n/4}$.

3.2 Construction of simultaneous SVD over $\mathbb{R}$

The simultaneous SVD over $\mathbb{R}$ can be constructed in parallel with the case over $\mathbb{C}$. The result, however, has a significant difference due to the difference between the statements in (C) of Theorems 1 and 6.

For $*$-algebras $\mathcal{T}_L$ $(\subseteq \mathcal{M}_m(\mathbb{R}))$ and $\mathcal{T}_R$ $(\subseteq \mathcal{M}_n(\mathbb{R}))$ we call a submodule $\mathcal{A}$ of $\mathcal{M}_{m,n}(\mathbb{R})$ a matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule over $\mathbb{R}$ if $[A \in \mathcal{A}, L \in \mathcal{T}_L, R \in \mathcal{T}_R \implies LAR \in \mathcal{A}]$. Given a family of $m \times n$ real matrices $A_1, \ldots, A_N$ we consider three algebraic structures:

(i) Matrix $*$-algebra $\mathcal{T}_L$ generated by $A_i A_j^\top$ $(i, j = 1, \ldots, N)$.

(ii) Matrix $*$-algebra $\mathcal{T}_R$ generated by $A_i^\top A_j$ $(i, j = 1, \ldots, N)$.

(iii) Matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule $\mathcal{A}$ generated by $A_1, \ldots, A_N$.

Note that $\mathcal{T}_L$ and $\mathcal{T}_R$ are determined by $\mathcal{A}$; that is, $\mathcal{T}_L$ and $\mathcal{T}_R$ are $*$-algebras generated, respectively, by $\mathcal{A}\mathcal{A}^\top$ and $\mathcal{A}^\top\mathcal{A}$. It is mentioned that if $A_i = O$ $(i = 1, \ldots, N)$, we have $\mathcal{A} = \{O\}$, and then $\mathcal{T}_L = \mathbb{R}I_m$ and $\mathcal{T}_R = \mathbb{R}I_n$. Such a degenerate case needs to be included as it may possibly occur as a result of our decomposition.

The fundamental fact underlying our approach is, again, that decomposing the given matrices $A_1, \ldots, A_N$ by means of a transformation of the form $P^\top A_i Q$ is equivalent to decomposing every element A of $\mathcal{A}$ by $P^\top A Q$. Accordingly we assume that we are given a matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule $\mathcal{A}$ $(\subseteq \mathcal{M}_{m,n}(\mathbb{R}))$ such that $\mathcal{T}_L$ and $\mathcal{T}_R$ are $*$-algebras generated, respectively, by $\mathcal{A}\mathcal{A}^\top$ and $\mathcal{A}^\top\mathcal{A}$. Note that no reference is made to the generators $A_1, \ldots, A_N$ in this setting.

The following theorem shows that the simultaneous SVD, i.e., the finest decomposition under $P^\top A_1 Q, \ldots, P^\top A_N Q$ can be constructed from the decompositions of $*$-algebras $\mathcal{A}\mathcal{A}^\top$ and $\mathcal{A}^\top\mathcal{A}$ as given in Theorem 6. Note that this construction generalizes the construction of the SVD of a single matrix A through the eigenvalue decompositions of $AA^\top$ and $A^\top A$.

Theorem 7. Let $\mathcal{A} \subseteq \mathcal{M}_{m,n}(\mathbb{R})$, $\mathcal{A} \neq \{O\}$, be a matrix $(\mathcal{T}_L, \mathcal{T}_R)$-bimodule over $\mathbb{R}$ such that $\mathcal{T}_L$ and $\mathcal{T}_R$ are $*$-algebras generated, respectively, by $\mathcal{A}\mathcal{A}^\top$ and $\mathcal{A}^\top\mathcal{A}$.

(A) There exist orthogonal matrices P and Q and a natural number ℓ such that

$$P^\top \mathcal{T}_L P = \mathcal{T}_{L1} \oplus \cdots \oplus \mathcal{T}_{L\ell}, \quad P^\top \mathcal{A} Q = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_\ell, \quad Q^\top \mathcal{T}_R Q = \mathcal{T}_{R1} \oplus \cdots \oplus \mathcal{T}_{R\ell}.$$

Here each $\mathcal{A}_j$ is a matrix $(\mathcal{T}_{Lj}, \mathcal{T}_{Rj})$-bimodule, and $\mathcal{T}_{Lj}$ and $\mathcal{T}_{Rj}$ are simple matrix $*$-algebras generated by $\mathcal{A}_j \mathcal{A}_j^\top$ and $\mathcal{A}_j^\top \mathcal{A}_j$, respectively.

(B) If $\mathcal{T}_L$ and $\mathcal{T}_R$ are simple, there exist orthogonal matrices P and Q and a natural number μ such that

$$P^\top \mathcal{T}_L P = I_\mu \otimes \mathcal{T}_L', \qquad P^\top \mathcal{A} Q = I_\mu \otimes \mathcal{A}', \qquad Q^\top \mathcal{T}_R Q = I_\mu \otimes \mathcal{T}_R'.$$

Here $\mathcal{A}'$ is a matrix $(\mathcal{T}_L', \mathcal{T}_R')$-bimodule, and $\mathcal{T}_L'$ and $\mathcal{T}_R'$ are irreducible matrix $*$-algebras generated by $\mathcal{A}'\mathcal{A}'^\top$ and $\mathcal{A}'^\top\mathcal{A}'$, respectively.

(C) If $\mathcal{T}_L$ and $\mathcal{T}_R$ are irreducible, there exist orthogonal matrices P and Q such that

$$P^\top \mathcal{T}_L P = \mathcal{D}_{\hat{m}}, \qquad P^\top \mathcal{A} Q = \mathcal{D}_{\hat{m},\hat{n}}, \qquad Q^\top \mathcal{T}_R Q = \mathcal{D}_{\hat{n}}.$$

Here $\mathcal{D} = \mathcal{M}$, $\mathcal{C}$, or $\mathcal{H}$, and $(\hat{m}, \hat{n}) = (m, n)$ if $\mathcal{D} = \mathcal{M}$; $(\hat{m}, \hat{n}) = (m/2, n/2)$ if $\mathcal{D} = \mathcal{M}$; and $(\hat{m}, \hat{n}) = (m/4, n/4)$ if $\mathcal{D} = \mathcal{M}$.

As an immediate corollary we obtain a necessary and sufficient condition for real matrices $A_1, \ldots, A_N$ to have the same set of singular vectors in the conventional sense. Compare this with its $\mathbb{C}$-version given in Corollary 3.

Corollary 8. For real matrices $A_1, \ldots, A_N$, there exist orthogonal matrices P and Q such that $P^\top A_i Q$ $(i = 1, \ldots, N)$ are diagonal if and only if $A_i A_j^\top$, $A_i^\top A_j$ $(i, j = 1, \ldots, N)$ are symmetric matrices.

Example 9. For two 4×8 matrices

$$A_1 = \begin{bmatrix} 0.365 & 1.991 & -0.627 & 1.740 & -2.133 & 1.908 & 3.684 & 0.850 \\ 3.045 & -1.686 & 0.790 & -2.203 & -1.121 & -0.445 & 0.616 & 3.251 \\ 3.071 & -0.299 & 2.218 & 2.053 & -1.671 & -0.782 & -0.666 & -2.738 \\ -1.221 & 0.293 & 2.099 & -1.876 & -2.125 & 3.559 & -1.776 & -0.396 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -0.393 & 2.325 & 0.081 & 1.615 & 0.029 & 0.759 & 2.596 & 2.123 \\ 2.289 & -2.798 & 0.854 & 0.219 & -0.243 & -0.573 & 1.124 & 2.122 \\ 1.900 & 0.548 & 0.309 & 3.017 & -1.967 & 0.301 & -1.052 & -1.350 \\ -1.217 & -0.035 & 1.802 & -1.398 & -3.230 & 1.353 & -0.453 & 0.921 \end{bmatrix}$$

we have

$$P^\top A_1 Q = \left[\begin{array}{cccc|cccc} 5.475 & 0 & -0.131 & -0.051 & 0 & 0 & 0 & 0 \\ 0 & 5.475 & 0.051 & -0.131 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 5.475 & 0 & -0.131 & -0.051 \\ 0 & 0 & 0 & 0 & 0 & 5.475 & 0.051 & -0.131 \end{array}\right],$$

$$P^\top A_2 Q = \left[\begin{array}{cccc|cccc} 3.702 & 1.442 & 2.053 & 0 & 0 & 0 & 0 & 0 \\ -1.442 & 3.702 & 0 & 2.053 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 3.702 & 1.442 & 2.053 & 0 \\ 0 & 0 & 0 & 0 & -1.442 & 3.702 & 0 & 2.053 \end{array}\right]$$

with suitable orthogonal matrices P and Q. We have $\ell = 1$, $\mu = 2$, and $\mathcal{D} = \mathcal{C}$ in Theorem 7, and accordingly both $P^\top A_1 Q$ and $P^\top A_2 Q$ belong to $I_2 \otimes \mathcal{C}_{1,2}$. For instance, $P^\top A_1 Q$ consists of two copies of a 2×4 matrix $[C(5.475),\ C(-0.131 + \imath 0.051)] \in \mathcal{C}_{1,2}$.

Example 10. The decomposition in the previous example is the finest over

$\mathbb{R}$, but it can be decomposed further over $\mathbb{C}$. We have indeed

$$P^* A_1 Q = \left[\begin{array}{cccc|cccc} 5.475 & -0.131 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5.475 & -0.131 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 5.475 & -0.131 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5.475 & -0.131 \end{array}\right]$$
$$+\,\imath \left[\begin{array}{cccc|cccc} 0 & -0.051 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.051 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -0.051 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.051 \end{array}\right],$$

$$P^* A_2 Q = \left[\begin{array}{cccc|cccc} 3.702 & 2.053 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3.702 & 2.053 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 3.702 & 2.053 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3.702 & 2.053 \end{array}\right]$$
$$+\,\imath \left[\begin{array}{cccc|cccc} 1.442 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1.442 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1.442 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1.442 & 0 \end{array}\right]$$

with suitable unitary matrices P and Q, which are different from the orthogonal matrices P and Q in Example 9. This decomposition can be easily obtained from the decomposition in Example 9.

Example 11. Consider two 4×6 matrices

$$A_1 = \begin{bmatrix} -0.013 & 0.472 & -0.000 & 0.230 & 0.106 & -1.336 \\ -0.687 & -0.689 & -0.677 & 0.898 & 1.187 & 0.727 \\ 1.082 & -0.362 & -0.315 & -0.820 & 0.062 & 0.757 \\ -0.728 & 0.310 & 0.879 & 0.169 & -0.908 & 0.578 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0.462 & 1.459 & 1.052 & 1.034 & 0.193 & -0.382 \\ -0.283 & -0.495 & -0.683 & 1.349 & 1.815 & 1.069 \\ 1.720 & 0.359 & 0.194 & -0.128 & 0.567 & 1.244 \\ 0.424 & 1.081 & 0.884 & 0.561 & -0.094 & -0.146 \end{bmatrix}.$$

Then $A_i A_j^\top$ $(i, j = 1, 2)$ and $A_i^\top A_j$ $(i, j = 1, 2)$ are symmetric matrices. Therefore, by Corollary 8, there exist orthogonal matrices P and Q such that $P^\top A_i Q$ $(i = 1, 2)$ are diagonal matrices, which read as follows:

$$P^\top A_1 Q = \begin{bmatrix} -1.742 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.920 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.095 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.172 & 0 & 0 \end{bmatrix},$$

$$P^\top A_2 Q = \begin{bmatrix} -0.146 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.910 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2.793 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.785 & 0 & 0 \end{bmatrix}.$$

Example 12. Recall the matrices A_1 and A_2 in Example 5, which have been brought to diagonal matrices though a unitary transformation. They satisfy the conditions in Corollary 3 but not the conditions in Corollary 8. Indeed, $A_1 A_2^\top$ is not a symmetric matrix. Therefore they are not simultaneously diagonalizable over $\mathbb{R}$. The finest decomposition over $\mathbb{R}$ is

$$P^* A_1 Q = \begin{bmatrix} 2.756 & 0.397 & 0 & 0 & 0 & 0 \\ -0.397 & 2.756 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1.897 & 0.315 & 0 & 0 \\ 0 & 0 & -0.315 & -1.897 & 0 & 0 \end{bmatrix},$$

$$P^* A_2 Q = \begin{bmatrix} 2.473 & -1.288 & 0 & 0 & 0 & 0 \\ 1.288 & 2.473 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.211 & 2.556 & 0 & 0 \\ 0 & 0 & -2.556 & 2.217 & 0 & 0 \end{bmatrix}.$$

4 Proof of the structure theorems

In this section, we will prove the structure theorems (Theorem 2 and 7). We prove Theorem 7 only since the proof of Theorem 2 is similar and easier.

We first prove the following lemma, which shows the relation between the block diagonalization of $\mathcal{A}$ and the block diagonalizations of $\mathcal{T}_L$ and $\mathcal{T}_R$. This is an extension of the fact that the ordinary SVD of a matrix A can be constructed from the eigenvalue decompositions of $AA^\top$ and $A^\top A$.

Lemma 13. The following are equivalent:
(1) $\mathcal{A}$ does not have a nontrivial block diagonalization.
(2) Both $\mathcal{T}_L$ and $\mathcal{T}_R$ are irreducible.

Proof. If $\mathcal{A}$ has a nontrivial block diagonalization, at least one of $\mathcal{T}_L$ or $\mathcal{T}_R$ has also a nontrivial block diagonalization since $\mathcal{T}_L$ and $\mathcal{T}_R$ are generated by $A_i A_j^\top$ $(i, j = 1, \ldots, N)$ and $A_i^\top A_j$ $(i, j = 1, \ldots, N)$ respectively. This proves that (1) implies (2).

To prove the converse, we may assume that $\mathcal{T}_R$ is reducible; otherwise we transpose all matrices. In this case, $\mathcal{T}_R$ has a nontrivial invariant subspace $W \subset \mathbb{R}^n$. Let $U = \mathrm{span}(\mathcal{A}W) \subseteq \mathbb{R}^m$. We take an orthogonal basis for W, $W^\perp$ and U, $U^\perp$. Then we claim that for all $A \in \mathcal{A}$, we have

$$P^\top A Q = \begin{array}{c} \\ \updownarrow U \\ \updownarrow U^\perp \end{array} \begin{array}{c} \begin{array}{cc} \leftarrow W \rightarrow & \leftarrow W^\perp \rightarrow \end{array} \\ \begin{array}{|c|c|} \hline A_1 & O \\ \hline O & A_2 \\ \hline \end{array} \end{array}$$

where P is an orthogonal basis transformation for U and $U^\perp$, and Q is an orthogonal basis transformation for W and $W^\perp$. (Note that if $U = \{\mathbf{0}\}$

or $U^\perp = \{\mathbf{0}\}$, the corresponding part disappears but we still say that such decomposition is nontrivial.) Because of the definition of U, the lower-left part is clearly zero. To prove that the upper-right part is zero, it is sufficient to check $u^\top A v = 0$ for all $v \in W^\perp$ and $u \in U$. By the definition of U, we have $u = \sum_j B_j w_j$ for some $B_j \in \mathcal{A}$ and $w_j \in W$. Therefore $u^\top A v = \sum_j w_j^\top B_j^\top A v$. Since $A^\top B_j \in \mathcal{T}_R$, we have $A^\top B_j w_j \in W$. Therefore $u^\top A v = 0$. □

Structure theorem (A). There exist orthogonal matrices P and Q and a natural number ℓ such that

$$P^\top \mathcal{T}_L P = \mathcal{T}_{L1} \oplus \cdots \oplus \mathcal{T}_{L\ell}, \quad P^\top \mathcal{A} Q = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_\ell, \quad Q^\top \mathcal{T}_R Q = \mathcal{T}_{R1} \oplus \cdots \oplus \mathcal{T}_{R\ell}.$$

Here each $\mathcal{A}_j$ is a matrix $(\mathcal{T}_{Lj}, \mathcal{T}_{Rj})$-bimodule, and $\mathcal{T}_{Lj}$ and $\mathcal{T}_{Rj}$ are simple matrix $*$-algebras generated by $\mathcal{A}_j \mathcal{A}_j^\top$ and $\mathcal{A}_j^\top \mathcal{A}_j$, respectively.

Proof. Take any minimal block diagonalization of $\mathcal{A}$, by which we mean a decomposition with diagonal blocks that cannot be decomposed further. Then $\mathcal{T}_L$ and $\mathcal{T}_R$ are decomposed accordingly into irreducible components. Then by collecting equivalent irreducible components, we obtain the decomposition in the above form. □

Structure theorem (C) . If $\mathcal{T}_L$ and $\mathcal{T}_R$ are irreducible, there exist orthogonal matrices P and Q such that

$$P^\top \mathcal{T}_L P = \mathcal{D}_{\hat{m}}, \qquad P^\top \mathcal{A} Q = \mathcal{D}_{\hat{m},\hat{n}}, \qquad Q^\top \mathcal{T}_R Q = \mathcal{D}_{\hat{n}}.$$

Here $\mathcal{D} = \mathcal{M}$, $\mathcal{C}$, or $\mathcal{H}$, and $(\hat{m}, \hat{n}) = (m, n)$ if $\mathcal{D} = \mathcal{M}$; $(\hat{m}, \hat{n}) = (m/2, n/2)$ if $\mathcal{D} = \mathcal{C}$; and $(\hat{m}, \hat{n}) = (m/4, n/4)$ if $\mathcal{D} = \mathcal{H}$.

Proof. By of the structure theorem for matrix $*$-algebras (Theorem 6), there exist orthogonal matrices P and Q such that $P^\top \mathcal{T}_L P = \mathcal{D}_{\hat{m}}$ and $Q^\top \mathcal{T}_R Q = \mathcal{D}'_{\hat{n}}$. Therefore we can assume, without loss of generality, $\mathcal{T}_L = \mathcal{D}_{\hat{m}}$ and $\mathcal{T}_R = \mathcal{D}'_{\hat{n}}$.

Let $d = 1$, 2 or 4 for $\mathcal{D} = \mathcal{M}$, $\mathcal{C}$ or $\mathcal{H}$ respectively, and let $d' = 1$, 2 or 4 for $\mathcal{D}' = \mathcal{M}$, $\mathcal{C}$, $\mathcal{H}$ respectively. Put $\hat{m} = m/d$ and $\hat{n} = n/d'$. We divide $A \in \mathcal{A}$ into $\hat{m} \times \hat{n}$ blocks of size $d \times d'$, whose (i, j) block is denoted $A_{[i,j]}$. Similarly, we divide $L \in \mathcal{T}_L$ into $\hat{m} \times \hat{m}$ blocks of size $d \times d$ and $R \in \mathcal{T}_R$ into $\hat{n} \times \hat{n}$ blocks of size $d' \times d'$.

Since $\mathcal{T}_L = \mathcal{D}_{\hat{m}}$, it contains the matrix, say E_{Li}, of which the i-th diagonal block is I_d and the other blocks are O_d. Similarly, $\mathcal{T}_R$ has the matrix, say E_{Rj}, of which the j-th diagonal block is $I_{d'}$ and the other blocks are $O_{d'}$. Therefore, for all $A \in \mathcal{A}$, $\mathcal{A}$ has the matrix $E_{Li} A E_{Rj}$ of which the (i, j) block is $A_{[i,j]}$ and the other blocks are $O_{d,d'}$. Noting that $\mathcal{T}_L$ and $\mathcal{T}_R$ contain block-wise permutation matrices, we see that for all $A, A' \in \mathcal{A}$, $A_{[i,j]} A'^\top_{[k,l]} \in \mathcal{D}$ and $A^\top_{[i,j]} A'_{[k,l]} \in \mathcal{D}'$.

Pick a nonzero matrix $A \in \mathcal{A}$, and let $A_{[i,j]}$ be one of the nonzero blocks of A. Since $A_{[i,j]}A_{[i,j]}^\top \in \mathcal{D}$ and a symmetric matrix in $\mathcal{D}$ is necessarily a scalar matrix, we have $A_{[i,j]}A_{[i,j]}^\top = \alpha I_d$ for some $\alpha > 0$. Similarly, we also have $A_{[i,j]}^\top A_{[i,j]} = \alpha' I_{d'}$ for some $\alpha' > 0$. These imply that $A_{[i,j]}$ has full row rank and full column rank. Therefore we have $d = d'$, and $\mathcal{D} = \mathcal{D}'$ in particular. Note also $\alpha = \alpha'$.

Next, we construct an orthogonal transformation from the nonzero matrix $A \in \mathcal{A}$ (chosen above). Let $P' = \mathrm{diag}(A_{[i,j]}, \ldots, A_{[i,j]})/\sqrt{\alpha}$, which is an orthogonal matrix. We claim the following equalities:

$$P'^\top \mathcal{T}_L P' = \mathcal{D}_{\hat{m}}, \qquad P'^\top \mathcal{A} = \mathcal{D}_{\hat{m},\hat{n}}.$$

The first equality is clear since $\mathcal{T}_L = \mathcal{D}_{\hat{m}}$ and $P' \in \mathcal{T}_L$. The second equality can be shown as follows: For all $A' \in \mathcal{A}$, the (k,l) block of $P'^\top A'$ is $A_{[i,j]}^\top A'_{[k,l]}/\sqrt{\alpha}$, which is an element of $\mathcal{D}$. Therefore $P'^\top A' \in \mathcal{D}_{\hat{m},\hat{n}}$, and hence $P'^\top \mathcal{A} = \mathcal{D}_{\hat{m},\hat{n}}$. □

Structure theorem (B) If $\mathcal{T}_L$ and $\mathcal{T}_R$ are simple, there exist orthogonal matrices P and Q and a natural number μ such that

$$P^\top \mathcal{T}_L P = \mathcal{T}'_L \otimes I_\mu, \qquad P^\top \mathcal{A} Q = \mathcal{A}' \otimes I_\mu, \qquad Q^\top \mathcal{T}_R Q = \mathcal{T}'_R \otimes I_\mu.$$

Here $\mathcal{A}'$ is a matrix $(\mathcal{T}'_L, \mathcal{T}'_R)$-bimodule, and $\mathcal{T}'_L$ and $\mathcal{T}'_R$ are irreducible matrix $*$-algebras generated by $\mathcal{A}'\mathcal{A}'^\top$ and $\mathcal{A}'^\top\mathcal{A}'$, respectively.

Proof. It turns out to be convenient to prove the above claim by showing

$$P^\top \mathcal{T}_L P = \mathcal{T}'_L \otimes I_\mu, \qquad P^\top \mathcal{A} Q = \mathcal{A}' \otimes I_\mu, \qquad Q^\top \mathcal{T}_R Q = \mathcal{T}'_R \otimes I_\mu.$$

Note that $\mathcal{T}'_L \otimes I_\mu$ and $I_\mu \otimes \mathcal{T}'_L$, for example, are connected by permutations of row and columns. The proof goes in a similar way as the proof of structure theorem (C).

By of the structure theorem for matrix $*$-algebras (Theorem 6), there exist orthogonal matrices P and Q such that $P^\top \mathcal{T}_L P = \mathcal{D}_{\hat{m}} \otimes I_\mu$ and $Q^\top \mathcal{T}_R Q = \mathcal{D}_{\hat{n}} \otimes I_{\mu'}$. Therefore we can assume, without loss of generality, $\mathcal{T}_L = \mathcal{D}_{\hat{m}} \otimes I_\mu$ and $\mathcal{T}_R = \mathcal{D}_{\hat{n}} \otimes I_{\mu'}$. Note that $\mathcal{D}$ is common in these equalities by structure theorem (C).

Let $d = 1$, 2 or 4 for $\mathcal{D} = \mathcal{M}$, $\mathcal{C}$ or $\mathcal{H}$ respectively. Put $\hat{m} = m/d\mu$ and $\hat{n} = n/d\mu'$. We divide $A \in \mathcal{A}$ into $\hat{m} \times \hat{n}$ blocks of size $d\mu \times d\mu'$, whose (i,j) block is denoted $A_{[i,j]}$. Similarly, we divide $L \in \mathcal{T}_L$ into $\hat{m} \times \hat{m}$ blocks of size $d\mu \times d\mu$ and $R \in \mathcal{T}_R$ into $\hat{n} \times \hat{n}$ blocks of size $d\mu' \times d\mu'$.

Since $\mathcal{T}_L = \mathcal{D}_{\hat{m}} \otimes I_\mu$, it contains the matrix, say E_{Li}, of which the i-th diagonal block is $I_{d\mu}$ and the other blocks are $O_{d\mu}$. Similarly, the $\mathcal{T}_R$ has the matrix, say E_{Rj}, of which j-th diagonal block is $I_{d\mu'}$ and the other blocks are $O_{d\mu'}$. Therefore, for all $A \in \mathcal{A}$, $\mathcal{A}$ has the matrix $E_{Li}AE_{Rj}$, of which the (i,j) block is $A_{[i,j]}$ and the other blocks are $O_{d\mu,d\mu'}$. Noting that $\mathcal{T}_L$ and $\mathcal{T}_R$ contain block-wise permutation matrices, we see that for all $A, A' \in \mathcal{A}$, $A_{[i,j]}A'^{\top}_{[k,l]} \in \mathcal{D} \otimes I_\mu$ and $A^{\top}_{[i,j]}A'_{[k,l]} \in \mathcal{D} \otimes I_{\mu'}$.

Pick a nonzero matrix $A \in \mathcal{A}$, and let $A_{[i,j]}$ be one of the nonzero blocks of A. Since $A_{[i,j]}A^{\top}_{[i,j]} \in \mathcal{D} \otimes I_\mu$ and a symmetric matrix in $\mathcal{D}$ is necessarily a scalar matrix, we have $A_{[i,j]}A^{\top}_{[i,j]} = \alpha I_{d\mu}$ for some $\alpha > 0$. Similarly, we also have $A^{\top}_{[i,j]}A_{[i,j]} = \alpha' I_{d\mu'}$ for some $\alpha' > 0$. These imply that $A_{[i,j]}$ has full row rank and full column rank. Therefore we have $\mu = \mu'$. Note also $\alpha = \alpha'$.

Next, we construct an orthogonal transformation from the nonzero matrix $A \in \mathcal{A}$ (chosen above). Let $P' = \mathrm{diag}(A_{[i,j]}, \ldots, A_{[i,j]})/\sqrt{\alpha}$, which is an orthogonal matrix. We claim the following equalities:

$$P'^{\top}\mathcal{T}_L P' = \mathcal{D}_{\hat{m}} \otimes I_\mu, \qquad P'^{\top}\mathcal{A} = \mathcal{D}_{\hat{m},\hat{n}} \otimes I_\mu.$$

The first equality is clear since $\mathcal{T}_L = \mathcal{D}_{\hat{m}} \otimes I_\mu$ and $P' \in \mathcal{T}_L$. The second equality can be shown as follows: For all $A' \in \mathcal{A}$, the (k,l) block of $P'^{\top}A'$ is $A^{\top}_{[i,j]}A'_{[k,l]}/\sqrt{\alpha}$, which is an element of $\mathcal{D} \otimes I_\mu$. Therefore $P'^{\top}A' \in \mathcal{D}_{\hat{m},\hat{n}} \otimes I_\mu$, and hence $P'^{\top}\mathcal{A} = \mathcal{D}_{\hat{m},\hat{n}} \otimes I_\mu$. □

5 Algorithms

The proofs of the structure theorems (Theorems 2 and 7) for simultaneous SVD are constructive, so that they can readily be turned into algorithms.

In this section, we describe an algorithm for Problem [$\mathbb{R}$] only, whereas an algorithm for Problem [$\mathbb{C}$] is similar and simpler, and hence omitted. The algorithm assumes subroutines for the decomposition of $*$-algebras into simple and irreducible components. Such algorithms for $*$-algebras are indeed available; see Murota–Kanno–Kojima–Kojima [8] and Maehara–Murota [7] as well as Eberly–Giesbrecht [3].

The decomposition in Part (A) of Theorem 7 can be carried out by the following algorithm. Recall that $\mathcal{T}_L$ is the $*$-algebra generated by $A_iA_j^{\top}$ $(i,j = 1,\ldots,N)$ and $\mathcal{T}_R$ is generated by $A_i^{\top}A_j$ $(i,j = 1,\ldots,N)$.

Algorithm 1.

Step 1: Find an orthogonal matrix P that decomposes the $*$-algebra $\mathcal{T}_L$ into simple components as in Theorem 6 (A). Also find an orthogonal matrix Q that decomposes the $*$-algebra $\mathcal{T}_R$ into simple components.

Step 2: Find permutations Π_L and Π_R such that $\Pi_L(P^{\top}A_iQ)\Pi_R$ for $i = 1,\ldots,N$ are in the same block-diagonal form, say $\bar{A}_{i1} \oplus \cdots \oplus \bar{A}_{i\ell}$.

For each $k = 1, \ldots, \ell$, $\mathcal{A}_k$, $\mathcal{T}_{Lk}$ and $\mathcal{T}_{Rk}$ are generated by $\bar{A}_{ik}$ $(i = 1, \ldots, N)$, $\bar{A}_{ik}\bar{A}_{jk}^\top$ $(i, j = 1, \ldots, N)$, and $\bar{A}_{ik}^\top \bar{A}_{jk}$ $(i, j = 1, \ldots, N)$, respectively. The validity of this algorithm is guaranteed by the fact that the orthogonal matrix denoted as "Q" in Theorem 6 (A) for $*$-algebras is unique up to a permutation of simple components and transformations within simple components.

The decompositions in Parts (B) and (C) of Theorem 7 can be carried out by the following algorithm, which should be applied to each $\mathcal{A}_k$ obtained in Part (A). To simply notation we omit the subscript k and assume that $\mathcal{A}$ satisfies the premise in (B) that $\mathcal{T}_L$ and $\mathcal{T}_R$ are simple $*$-algebras with multiplicity μ of irreducible components. We define $d = 1, 2, 4$ according to whether $\mathcal{D} = \mathcal{M}$, $\mathcal{C}$, or $\mathcal{H}$ in (C).

Algorithm 2.

Step 1: Find an orthogonal matrix P that decomposes the $*$-algebra $\mathcal{T}_L$ into irreducible components as in Theorem 6 (B). Also find an orthogonal matrix Q that decomposes the $*$-algebra $\mathcal{T}_R$ into irreducible components.

Step 2: Pick a nonzero matrix A_i from among the input matrices, and regard it as a $d\mu \times d\mu$ block-matrix. Let B be one of the nonzero blocks of A_i, where B is $m/(d\mu) \times n/(d\mu)$ if A_i is $m \times n$.

Step 3: Set $P' = \mathrm{diag}(B, B, \ldots, B)/c$, where c is a constant such that $c^2 I = B^\top B$.

Step 4: Find permutations Π_L and Π_R such that $\Pi_L(P'^\top A_i)\Pi_R$ for $i = 1, \ldots, N$ are in the same block-diagonal form.

The performance of this algorithm depends strongly on the performance of the subroutines. Currently, all algorithms for the decomposition of $*$-algebras into simple and irreducible components are sensitive to numerical errors, and as a consequence the proposed algorithm is also sensitive to numerical errors and accordingly it can only solve not too large instances, e.g., with n, m and N less than a few hundreds. To solve larger instances, an improvement of the subroutines is needed.

Acknowledgments

This work is supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan and by the Global COE "The Research and Training Center for New Development in Mathematics."

References

[1] E. de Klerk, D.V. Pasechnik and A. Schrijver: Reduction of symmetric semidefinite programs using the regular $*$-representation, Mathematical Programming, Series B, Vol. 109 (2007), pp. 613–624.

[2] E. de Klerk and R. Sotirov: Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment, CentER Discussion paper 2007-44, Center for Economic Research, Tilburg University, June 2007.

[3] W. Eberly and M. Giesbrecht: Efficient decomposition of separable algebras, Journal of Symbolic Computation, Vol. 37 (2004), pp. 35–81.

[4] K. Gatermann and P.A. Parrilo: Symmetry groups, semidefinite programs, and sums of squares, Journal of Pure and Applied Algebra, Vol. 192 (2004), pp. 95–128.

[5] Y. Kanno, M. Ohsaki, K. Murota and N. Katoh: Group symmetry in interior-point methods for semidefinite program, Optimization and Engineering, Vol. 2 (2001), pp. 293–320.

[6] M. Kojima, S. Kojima and S. Hara: Linear algebra for semidefinite programming, Research Report B-290, Tokyo Institute of Technology, October 1994; also in RIMS Kokyuroku 1004, Kyoto University, pp. 1–23, 1997.

[7] T. Maehara and K. Murota: A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with general irreducible components, METR 2008-26, Department of Mathematical Informatics, University of Tokyo, May 2008.

[8] K. Murota, Y. Kanno, M. Kojima and S. Kojima: A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras, METR 2007-52, Department of Mathematical Informatics, University of Tokyo, September 2007. (Revised version available at http://www.misojiro.t.u-tokyo.ac.jp/~murota/paper/MKKKrev.pdf)

[9] F. Vallentin: Symmetry in semidefinite programs: Linear Algebra and Its Applications, Vol. 430 (2009), pp. 360–369.

[10] J.H.M. Wedderburn: Lectures on Matrices, American Mathematical Society, New York, 1934; Dover, Mineola, N.Y., 2005.