2019

Schur's Lemma For Coupled Reducibility And Coupled Normality

D. Lahat

C. Jutten

Helene Shapiro
Swarthmore College, hshapir1@swarthmore.edu

Recommended Citation

D. Lahat, C. Jutten, and Helene Shapiro. (2019). "Schur's Lemma For Coupled Reducibility And Coupled Normality". *SIAM Journal On Matrix Analysis And Applications.* Volume 40, Issue 3. DOI: 10.1137/18M1232462
https://works.swarthmore.edu/fac-math-stat/257

Schur's Lemma for Coupled Reducibility and Coupled Normality

*Dana Lahat[†] Christian Jutten[‡] Helene Shapiro[§]

December 3, 2018

Abstract

Let $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$, where $\mathcal{I}$ is an index set, be a doubly indexed family of matrices, where A_{ij} is $n_i \times n_j$. For each $i \in \mathcal{I}$, let $\mathcal{V}_i$ be an n_i-dimensional vector space. We say $\mathcal{A}$ is *reducible in the coupled sense* if there exist subspaces, $\mathcal{U}_i \subseteq \mathcal{V}_i$, with $\mathcal{U}_i \neq \{0\}$ for at least one $i \in \mathcal{I}$, and $\mathcal{U}_i \neq \mathcal{V}_i$ for at least one i, such that $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ for all i, j. Let $\mathcal{B} = \{B_{ij}\}_{i,j\in\mathcal{I}}$ also be a doubly indexed family of matrices, where B_{ij} is $m_i \times m_j$. For each $i \in \mathcal{I}$, let X_i be a matrix of size $n_i \times m_i$. Suppose $A_{ij}X_j = X_iB_{ij}$ for all i, j. We prove versions of Schur's Lemma for $\mathcal{A}$, $\mathcal{B}$ satisfying coupled irreducibility conditions. We also consider a refinement of Schur's Lemma for sets of normal matrices and prove corresponding versions for $\mathcal{A}$, $\mathcal{B}$ satisfying coupled normality and coupled irreducibility conditions.

*Funding: The work of D. Lahat and C. Jutten was supported by the project CHESS, 2012- ERC-AdG-320684. GIPSA-lab is a partner of the LabEx PERSYVAL-Lab (ANR11-LABX-0025).

[†]IRIT, Université de Toulouse, CNRS, Toulouse, France (Dana.Lahat@irit.fr). This work was carried out when D. Lahat was with GIPSA-lab, Grenoble, France.

[‡]Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France. C. Jutten is a senior member of Institut Univ. de France (christian.jutten@gipsa-lab.grenoble-inp.fr).

[§]Department of Mathematics and Statistics, Swarthmore College, Swarthmore, PA 19081 (hshapir1@swarthmore.edu).

1 Introduction

Let K be a positive integer and let $n_1, \ldots, n_K$ and $m_1, \ldots, m_K$ be positive integers. Consider two doubly indexed families of matrices, $\mathcal{A} = \{A_{ij}\}_{i,j=1}^K$ and $\mathcal{B} = \{B_{ij}\}_{i,j=1}^K$, where A_{ij} is $n_i \times n_j$ and B_{ij} is $m_i \times m_j$. Put $N = \sum_{i=1}^K n_i$ and $M = \sum_{i=1}^K m_i$. Arrange the A_{ij}'s into an $N \times N$ matrix, A, with A_{ij} in block i, j of A.

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1K} \\ A_{21} & A_{22} & \cdots & A_{2K} \\ \vdots & \vdots & & \vdots \\ A_{K1} & A_{K2} & \cdots & A_{KK} \end{pmatrix}.$$

Similarly, form an $M \times M$ matrix B with B_{ij} in block i, j. Let X_i be an $n_i \times m_i$ matrix and form an $N \times M$ matrix X with blocks $X_1, \ldots, X_K$ down the diagonal and zero blocks elsewhere. Thus,

$$X = \begin{pmatrix} X_1 & 0 & 0 & \cdots & 0 \\ 0 & X_2 & 0 & \cdots & 0 \\ 0 & 0 & X_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & X_K \end{pmatrix},$$

where the 0 in position i, j represents an $n_i \times m_i$ block of zeroes. Then $AX = XB$ if and only if

$$A_{ij}X_j = X_i B_{ij}, \tag{1}$$

for all $i, j = 1, \ldots, K$. We may rewrite $AX = XB$ as $AX - XB = 0$, a homogeneous Sylvester equation [13].

We define several versions of coupled reducibility and prove corresponding versions of Schur's Lemma [12] for pairs $\mathcal{A}$, $\mathcal{B}$. Imposing coupled irreducibility constraints on $\mathcal{A}$ and $\mathcal{B}$ restricts the possible solutions, $X_1, \ldots, X_K$, to the equations (1). We also discuss a refinement of Schur's Lemma for normal matrices, and prove corresponding versions for $\mathcal{A}$, $\mathcal{B}$ satisfying coupled normality conditions.

The system of coupled matrix equations (1) arises in a recent model for multiset data analysis [5, 8], called *joint independent subspace analysis*, or JISA. This model consists of K datasets, where each dataset is an unknown mixture of several latent stochastic multivariate signals. The blocks of A and B represent statistical links among these datasets. More specifically, A_{ij} and B_{ij} represent statistical correlations among latent signals in the

ith and jth datasets. The multiset joint analysis framework, as opposed to the analysis of K distinct unrelated datasets, arises when a sufficient number of cross-correlations among datasets, i.e., A_{ij} and $B_{ij}, i \neq j$, are not zero. It turns out [6, 9, 10] that the identifiability and uniqueness of this model, in its simplest form, boil down to characterizing the set of solutions to the system of matrix equations (1), when the cross-correlations among the latent signals are subject to coupled (ir)reducibility. In other words, the coupled reducibility conditions that we introduce in this paper can be attributed with a physical meaning, and can be applied to real-world problems. We refer the reader to [6, 10] for further details on the JISA model, and on the derivation of (1). In this paper, we consider scenarios more general than those required to address the signal processing problem in [6, 10]. The analysis in Section 6, which focuses on coupled normal matrices, is inspired by the JISA model in [6, 10], in which the matrices A and B are Hermitian, and thus a special case of coupled normal. A limited version of some of the results in this manuscript was presented orally in, e.g., [7, 4, 1] and in a technical report [6]; however, they were never published.

While motivated by the matrix equation, $AX = XB$, the main definitions, theorems, and proofs do not require K to be finite. Thus, for a general index set, $\mathcal{I}$, we consider doubly indexed families, $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in\mathcal{I}}$. For the situation described above, $\mathcal{I} = \{1, 2, \ldots, K\}$.

We use $\mathbb{F}$ to denote the field of scalars. For some results, $\mathbb{F}$ can be any field. For results involving unitary and normal matrices, $\mathbb{F} = \mathbb{C}$, the field of complex numbers. For each $i \in \mathcal{I}$, let n_i and m_i be positive integers, let $\mathcal{V}_i$ be an n_i-dimensional vector space over $\mathbb{F}$, and let $\mathcal{W}_i$ be an m_i-dimensional vector space over $\mathbb{F}$. (Essentially, $\mathcal{V}_i = \mathbb{F}^{n_i}$ and $\mathcal{W}_i = \mathbb{F}^{m_i}$.) For all $i, j \in \mathcal{I}$, let A_{ij} be an $n_i \times n_j$ matrix and let B_{ij} be $m_i \times m_j$. View A_{ij} as a linear transformation from $\mathcal{V}_j$ to $\mathcal{V}_i$, and B_{ij} as a linear transformation from $\mathcal{W}_j$ to $\mathcal{W}_i$. For each $i \in \mathcal{I}$, let X_i be an $n_i \times m_i$ matrix; view X_i as a linear transformation from $\mathcal{W}_i$ to $\mathcal{V}_i$. We are interested in families $\mathcal{A}$, $\mathcal{B}$ satisfying the equations (1) for all $i, j \in \mathcal{I}$.

For some results, all of the n_i's will be equal, and all of the m_i's will be equal. In this case, we use n for the common value of the n_i's and m for the common value of the m_i's. We then set $\mathcal{V} = \mathbb{F}^n$ and $\mathcal{W} = \mathbb{F}^m$. Each X_i is then $n \times m$. For $\mathcal{I} = \{1, \ldots, K\}$, we have $N = Kn$, and $M = Km$. All of the blocks A_{ij} in A are $n \times n$, while all of the blocks B_{ij} in B are $m \times m$.

Section 2 reviews the usual matrix version of Schur's Lemma and its proof. Section 3 defines coupled reducibility and two restricted versions, called *proper* and *strong* reducibility. Section 4 states and proves versions of Schur's Lemma for coupled reducibility and proper reducibility, Theorem 4.2. In section 5, we define some graphs associated with the pair $\mathcal{A}$, $\mathcal{B}$ and use them for versions of Schur's Lemma corresponding to strongly coupled reducibility, Theorems 5.1 and 5.2. Section 6 deals with a refinement of Schur's Lemma for sets of normal matrices and corresponding versions for pairs $\mathcal{A}$, $\mathcal{B}$ which are coupled normal, Theorem 6.2. The Appendix presents examples to support some claims made in Section 3.

2 Reducibility and Schur's Lemma

We begin by reviewing the ordinary notion of reducibility for a set of linear transformations and Schur's Lemma.

Definition 2.1. A set, $\mathcal{T}$, of linear transformations, on an n-dimensional vector space, $\mathcal{V}$, is *reducible* if there is a proper, non-zero subspace $\mathcal{U}$ of $\mathcal{V}$ such that $T(\mathcal{U}) \subseteq \mathcal{U}$ for all $T \in \mathcal{T}$. If $\mathcal{T}$ is not reducible, we say it is *irreducible.*

The subspace $\mathcal{U}$ is an invariant subspace for each transformation T in $\mathcal{T}$. We say $\mathcal{T}$ is *fully reducible* if it is possible to decompose $\mathcal{V}$ as a direct sum $\mathcal{V} = \mathcal{U} \oplus \hat{\mathcal{U}}$, where $\mathcal{U}$ and $\hat{\mathcal{U}}$ are both nonzero, proper invariant subspaces of $\mathcal{T}$.

Alternatively, one can state this in matrix terms. The linear transformations in $\mathcal{T}$ may be represented as $n \times n$ matrices, relative to a choice of basis for $\mathcal{V}$. Let d be the dimension of $\mathcal{U}$; choose a basis for $\mathcal{V}$ in which the first d basis vectors are a basis for $\mathcal{U}$. Since $\mathcal{U}$ is an invariant subspace of each T in $\mathcal{T}$, the matrices representing $\mathcal{T}$ relative to this basis are block upper triangular with square diagonal blocks of sizes $d \times d$ and $(n-d) \times (n-d)$. The first diagonal block, of size $d \times d$, represents the action of the transformation on the invariant subspace $\mathcal{U}$. The $(n-d) \times d$ block in the lower left hand corner consists of zeroes. Since changing basis is equivalent to applying a matrix similarity, we have the following matrix version of Definition 2.1.

Definition 2.2. (Matrix version of reducibility.) A set $\mathcal{M}$ of $n \times n$ matrices is *reducible* if, for some d, with $0 < d < n$, there is a nonsingular matrix S such that, for each A in $\mathcal{M}$, the matrix $S^{-1}AS$ is block upper triangular with square diagonal blocks of sizes $d \times d$ and $(n-d) \times (n-d)$.

When $\mathcal{T}$ is fully reducible, we can use a basis for $\mathcal{V}$ in which the first d basis vectors are a basis for $\mathcal{U}$, and the remaining $n-d$ basis vectors are a basis for $\hat{\mathcal{U}}$. The corresponding matrices for $\mathcal{T}$ are then block diagonal, with diagonal blocks of sizes $d \times d$ and $(n-d) \times (n-d)$.

If $\mathbb{F} = \mathbb{C}$ and $\mathcal{M}$ is reducible, the S in Definition 2.2 can be chosen to be a unitary matrix, by using an orthonormal basis for $\mathbb{C}^n$ in which the first d basis vectors are an orthonormal basis for $\mathcal{U}$. If $\mathcal{M}$ is fully reducible, and the subspaces $\mathcal{U}$ and $\hat{\mathcal{U}}$ are orthogonal, use an orthonormal basis of $\mathcal{U}$ for the first d columns of S and an orthonormal basis of $\hat{\mathcal{U}}$ for the remaining $n-d$ columns. Then S is unitary, and $S^{-1}MS$ is block diagonal, with diagonal blocks of sizes $d \times d$ and $(n-d) \times (n-d)$.

The following fact is well known; we include the proof because the idea is used later. For this fact, we assume we are working over $\mathbb{C}$, or at least over a field that contains the eigenvalues of the transformations.

Proposition 2.1. Let $\mathcal{M}$ be an irreducible set of $n \times n$ complex matrices. Suppose the $n \times n$ matrix C commutes with every element of $\mathcal{M}$. Then C is a scalar matrix.

Proof. Let λ be an eigenvalue of C, and let $\mathcal{U}_\lambda$ denote the corresponding eigenspace. Let $A \in \mathcal{M}$. For any $\mathbf{v} \in \mathcal{U}_\lambda$, we have $C(A\mathbf{v}) = A(C\mathbf{v}) = \lambda(A\mathbf{v})$. Hence $A\mathbf{v}$ is in $\mathcal{U}_\lambda$, and so $\mathcal{U}_\lambda$ is invariant under each element of $\mathcal{M}$. Since an eigenspace is nonzero, and $\mathcal{M}$ is irreducible, we must have $\mathcal{U}_\lambda = \mathbb{C}^n$. Hence $C = \lambda I_n$. □

Schur's Lemma plays a key role in group representation theory. It is used to establish uniqueness of the decomposition of a representation of a finite group into a sum of irreducible representations. However, one need not have a matrix group; the result holds for irreducible sets of matrices. We include the usual proof [2], because the same idea is used to prove our versions for coupled reducibility.

Theorem 2.1 (Schur's Lemma)**.** Let $\{A_i\}_{i \in \mathcal{I}}$ be an irreducible family of $n \times n$ matrices, and let $\{B_i\}_{i \in \mathcal{I}}$ be an irreducible family of $m \times m$ matrices. Suppose P is an $n \times m$ matrix such that $A_i P = P B_i$ for all $i \in \mathcal{I}$. Then, either $P = 0$, or P is nonsingular; in the latter case we must have $m = n$. For matrices of complex numbers, if $A_i = B_i$ for all $i \in \mathcal{I}$, then P is a scalar matrix.

Proof. View the A_i's as linear transformations on an n-dimensional vector space $\mathcal{V}$, and the B_i's as linear transformations on an m-dimensional vector space $\mathcal{W}$. The $n \times m$ matrix P represents a linear transformation from $\mathcal{W}$ to $\mathcal{V}$. So $ker(P)$ is a subspace of $\mathcal{W}$ and $range(P)$ is a subspace of $\mathcal{V}$.

Let $\mathbf{w} \in ker(P)$. Then $P(B_i\mathbf{w}) = A_iP\mathbf{w} = 0$. Hence, $ker(P)$ is invariant under $\{B_i\}_{i\in\mathcal{I}}$. Since $\{B_i\}_{i\in\mathcal{I}}$ is irreducible, $ker(P)$ is either $\{0\}$ or $\mathcal{W}$. In the latter case, $P = 0$. If $P \neq 0$, then $ker(P) = \{0\}$. Now consider the range space of P. For any $\mathbf{w} \in \mathcal{W}$, we have $A_i(P\mathbf{w}) = P(B_i\mathbf{w}) \in range(P)$, so the range space of P is invariant under A_i for each i. Since $\{A_i\}_{i\in\mathcal{I}}$ is irreducible, $range(P)$ is either $\{0\}$ or $\mathcal{V}$. But we are assuming $P \neq 0$ so $range(P) = \mathcal{V}$. Since we also have $ker(P) = \{0\}$, the matrix P must be nonsingular and $m = n$.

If $A_i = B_i$ for all $i \in \mathcal{I}$, then P commutes with each A_i. If each A_i is a complex matrix, then, since $\{A_i\}_{i\in\mathcal{I}}$ is irreducible, P must be a scalar matrix. □

For nonsingular P, we have $P^{-1}A_iP = B_i$ for all $i \in \mathcal{I}$, so $\{A_i\}_{i\in\mathcal{I}}$ and $\{B_i\}_{i\in\mathcal{I}}$ are simultaneously similar.

3 Coupled Reducibility

For simultaneous similarity of $\{A_i\}_{i\in\mathcal{I}}$ and $\{B_i\}_{i\in\mathcal{I}}$, there is a nonsingular matrix, P, such that $P^{-1}A_iP = B_i$ for all i. We now define a "coupled" version of similarity for two doubly indexed families, $\mathcal{A}$ and $\mathcal{B}$, with $n_i = m_i$ for all $i \in \mathcal{I}$. In this case, A_{ij} and B_{ij} are matrices of the same size, and in the equations (1), each X_i is a square matrix.

Definition 3.1. Let $\mathcal{A} = \{A_{ij}\}_{i,j\in I}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in I}$, where $n_i = m_i$ for all $i \in \mathcal{I}$. We say $\mathcal{A}$ and $\mathcal{B}$ are *similar in the coupled sense* if there exist nonsingular matrices $\{T_i\}_{i\in I}$, where T_i is $n_i \times n_i$, such that $T_i^{-1}A_{ij}T_j = B_{ij}$ for all $i, j \in \mathcal{I}$.

For a finite index set $\mathcal{I} = \{1, \ldots, K\}$, this can be stated in terms of the matrices, A and B. Let T be the block diagonal matrix $T_1 \oplus T_2 \oplus \cdots \oplus T_K$. Then $AT = TB$ if and only if $A_{ij}T_j = T_iB_{ij}$ for all i, j. Hence, $\mathcal{A}$ and $\mathcal{B}$ are similar in the coupled sense if and only if T is nonsingular and $T^{-1}AT = B$.

We define several versions of "reducible in the coupled sense" for a doubly indexed family $\mathcal{A}$. The basic idea is that there are subspaces, $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$, where

$\mathcal{U}_i \subseteq \mathcal{V}_i$, such that $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ for all $i, j \in \mathcal{I}$. This holds trivially when $\mathcal{U}_i$ is zero for all i, and when $\mathcal{U}_i = \mathcal{V}_i$ for all i, so we shall insist that at least one subspace is nonzero, and at least one is not $\mathcal{V}_i$. We are also interested in two more restrictive versions: the case where at least one $\mathcal{U}_i$ is a nonzero, proper subspace, and the case where every $\mathcal{U}_i$ is a nonzero, proper subspace of $\mathcal{V}_i$.

Definition 3.2. Let $\mathcal{A} = \{A_{ij}\}_{i,j \in I}$ where A_{ij} is $n_i \times n_j$. We say $\mathcal{A}$ is *reducible in the coupled sense* if there exist subspaces $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$, where $\mathcal{U}_i \subseteq \mathcal{V}_i$, such that the following hold.

1. For at least one i, we have $\mathcal{U}_i \neq \{0\}$.
2. For at least one i, we have $\mathcal{U}_i \neq \mathcal{V}_i$.
3. For all $i, j \in \mathcal{I}$ we have $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$.

We say $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$ is a reducing set of subspaces for $\mathcal{A}$, or that $\mathcal{A}$ is reduced by $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$. If $\mathcal{A}$ is not reducible in the coupled sense, we say it is *irreducible* in the coupled sense. We say $\mathcal{A}$ is *properly reducible in the coupled sense* if at least one $\mathcal{U}_i$ is a nonzero, proper subspace of $\mathcal{V}_i$. We say $\mathcal{A}$ is *strongly reducible in the coupled sense* if every $\mathcal{U}_i$ is a nonzero, proper subspace of $\mathcal{V}_i$.

Remark 3.1. If $n_i = 1$ for all i, the one-dimensional spaces $\mathcal{V}_i$ have no nonzero proper invariant subspaces, so $\mathcal{A}$ cannot be properly or strongly irreducible. If $K = 1$, then $\mathcal{A}$ consists of a single $n \times n$ matrix.

Remark 3.2. If $n_i = n$ for all i, and the subspaces $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$ are all the same nonzero proper subspace, i.e, for all i, we have $\mathcal{U}_i = \mathcal{U}$ where $\mathcal{U}$ is a nonzero, proper subspace of $\mathcal{V}$, then $\mathcal{A}$ is reducible in the ordinary sense, given in Definition 2.1.

Note the following facts.

1. If $\mathcal{U}_j = \{0\}$, then $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ holds for any A_{ij} and any $\mathcal{U}_i$.
2. If $\mathcal{U}_i = \mathcal{V}_i$, then $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ holds for any A_{ij} and any $\mathcal{U}_j$.
3. If $\mathcal{U}_j = \mathcal{V}_j$ and $\mathcal{U}_i = \{0\}$, then $A_{ij} = 0$.
4. For $i = j$, we have $A_{ii}(\mathcal{U}_i) \subseteq \mathcal{U}_i$, so $\mathcal{U}_i$ is an invariant subspace of A_{ii}.

An equivalent matrix version of Definition 3.2 is obtained by choosing an appropriate basis for each $\mathcal{V}_j$. Let d_j be the dimension of the subspace $\mathcal{U}_j$. We have $0 \leq d_j \leq n_j$. If d_j is positive, let $\mathbf{v}_{j,1}, \ldots, \mathbf{v}_{j,d_j}$ be a basis for $\mathcal{U}_j$ and let T_j be a nonsingular $n_j \times n_j$ matrix which has $\mathbf{v}_{j,1}, \ldots, \mathbf{v}_{j,d_j}$ in the first d_j columns. If $d_j = 0$, we may use any nonsingular $n_j \times n_j$ matrix for T_j. Set $B_{ij} = T_i^{-1} A_{ij} T_j$; equivalently,

$$A_{ij} T_j = T_i B_{ij}.$$

The first d_j columns of T_j are a basis for $\mathcal{U}_j$, so $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ tells us the first d_j columns of $A_{ij}T_j$ are in $\mathcal{U}_i$. Hence, the first d_j columns of $T_i B_{ij}$ are in $\mathcal{U}_i$, so by the definition of T_i, each of the first d_j columns of $T_i B_{ij}$ is a linear combination of the first d_i columns of T_i. Therefore, each of the first d_j columns of B_{ij} will have zeroes in all entries below row d_i, and the lower left hand corner of B_{ij} is a block of zeroes of size $(n_i - d_i) \times d_j$. When $0 < d_i < n_i$ and $0 < d_j < n_j$, the matrix B_{ij} has the form

$$B_{ij} = \begin{pmatrix} C & D \\ 0_{(n-d_i)\times d_j} & E \end{pmatrix}, \tag{2}$$

where C is size $d_i \times d_j$ and represents the action of A_{ij} on the subspace $\mathcal{U}_j$. The zero block in the lower left hand block has size $(n - d_i) \times d_j$, while D is $d_i \times (n_j - d_j)$ and E is $(n_i - d_i) \times (n_j - d_j)$.

Remark 3.3. The block matrix (2) has a block of zeroes in the lower left hand corner of size $(n - d_i) \times d_j$. We also use this terminology for the cases $d_i = 0,\ d_i = n_i,\ d_j = 0,\ d_j = n_j$, in which case we mean the following. If $d_i = 0$, the first d_j columns of B_{ij} are zero. If $d_i = n_i$ there is no restriction on the form of B_{ij}. If $d_j = 0$ there is no restriction on the form of B_{ij}. If $d_j = n_j$ the last $n - d_i$ rows of B_{ij} are zero.

Conversely, if each $B_{ij} = T_i^{-1} A_{ij} T_j$ has the block form in (2), define $\mathcal{U}_i$ to be the subspace spanned by the first d_i columns of T_i. The subspaces $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$ then satisfy $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$. Hence, we have the following equivalent matrix form of Definition (3.2).

Definition 3.3 (Matrix version of coupled reducibility)**.** Let $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$. We say $\mathcal{A}$ is *reducible in the coupled sense*, or, *reducible by coupled similarity*, if there exist integers $\{d_i\}_{i\in\mathcal{I}}$, with $0 \leq d_i \leq n_i$, and nonsingular $n_i \times n_i$ matrices T_i, such that the following hold.

1. At least one d_i is positive.
2. At least one d_i is less than n_i.
3. Each matrix $B_{ij} = T_i^{-1} A_{ij} T_j$ has a block of zeroes in the lower left hand corner of size $(n_i - d_i) \times d_j$.

We say $\mathcal{A}$ is *properly reducible in the coupled sense* if $0 < d_i < n_i$ for at least one value of i. We say $\mathcal{A}$ is *strongly reducible in the coupled sense* if $0 < d_i < n_i$ for every i.

Full reducibility by coupled similarity occurs when, for each i, there is also a subspace $\hat{\mathcal{U}}_i$ such that $\mathcal{V}_i = \mathcal{U}_i \oplus \hat{\mathcal{U}}_i$, and $A_{ij}(\hat{\mathcal{U}}_j) \subseteq \hat{\mathcal{U}}_i$ for all $i, j \in \mathcal{I}$. For the corresponding matrix version, use a basis for $\mathcal{U}_j$ in the first d_j columns of T_j and a basis for $\hat{\mathcal{U}}_j$ in the remaining $n_j - d_j$ columns. For $0 < d_i < n_i$ and $0 < d_j < n_j$, the matrix $B_{ij} = T_i^{-1} A_{ij} T_j$ has the block form

$$B_{ij} = \begin{pmatrix} C & 0_{d_i \times (n-d_j)} \\ 0_{(n-d_i) \times d_j} & E \end{pmatrix}. \tag{3}$$

The $d_i \times d_j$ matrix C represents the action of A_{ij} on $\mathcal{U}_j$ and the $(n_i - d_i) \times (n_j - d_j)$ matrix E represents the action of A_{ij} on $\hat{\mathcal{U}}_j$.

For the field of complex numbers we have unitary versions.

Definition 3.4 (Unitary version of reducible in the coupled sense)**.** Let $\mathcal{A}$ be a family of complex matrices. We say $\mathcal{A}$ is *unitarily reducible in the coupled sense* if the conditions of Definition 3.3 are satisfied with unitary matrices T_i.

For complex $\mathcal{A}$, reducibility by coupled similarity implies reducibility by coupled unitary similarity. Simply use an orthonormal basis for each $\mathcal{U}_i$, and extend it to an orthonormal basis for $\mathcal{V}_i$ to obtain a unitary matrix for T_i. If $\mathcal{A}$ is fully reducible, and $\mathcal{U}_i$ and $\hat{\mathcal{U}}_i$ are orthogonal subspaces, then, for each $\mathcal{V}_i = \mathcal{U}_i \oplus \hat{\mathcal{U}}_i$, we can form a unitary matrix T_i using an orthonormal basis for $\mathcal{U}_i$ for the first d_i columns and an orthonormal basis for $\hat{\mathcal{U}}_i$ for the remaining $n_i - d_i$ columns. Each B_{ij} then has the block form (3).

Unitary reducibility matters in the JISA model, because A and B are correlation matrices, and the appropriate linear change of variable leads to a congruence, rather than a similarity. When T_i is unitary, $T_i^{-1} = T_i^*$. For $T = T_1 \oplus T_2 \oplus \cdots \oplus T_k$ we then have $T^{-1}AT = T^*AT$.

From the definition, it is clear that if $\mathcal{A}$ is strongly reducible, then it is also properly reducible, and if it is properly reducible, it is reducible. We introduce some notation. Fix an index set, $\mathcal{I}$. Use $|\mathcal{I}|$ to denote the size of $\mathcal{I}$; when $\mathcal{I}$ is a finite set with K elements, we assume $\mathcal{I} = \{1, 2, \ldots, K\}$. Fix a family $\{n_i\}_{i\in\mathcal{I}}$ of positive integers, and a field $\mathbb{F}$. Consider the set of all $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$, where A_{ij} is an $n_i \times n_j$ matrix with entries from $\mathbb{F}$. We use $\mathit{Red}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$ to denote the set of all such families $\mathcal{A}$ which are reducible in the coupled sense. We use $\mathit{PropRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$ for the set of all such $\mathcal{A}$ which are properly reducible in the coupled sense, and $\mathit{StrRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$ for the set of all such $\mathcal{A}$ that are strongly reducible in the coupled sense. When all n_i's have the same value, n, and $|\mathcal{I}| = K$, we use the notations $\mathit{Red}(\mathbb{F}, n, K)$, $\mathit{PropRed}(\mathbb{F}, n, K)$, and $\mathit{StrRed}(\mathbb{F}, n, K)$.

When $n_i = 1$, the space $\mathcal{V}_i = \mathbb{F}$ is one dimensional and has no nonzero proper subspaces. Hence, $\mathit{StrRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$ is the empty set if $n_i = 1$ for some i, and $\mathit{PropRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$ is the empty set when $n_i = 1$ for all i.

From Definition 3.2 it is obvious that

$$\mathit{StrRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subseteq \mathit{PropRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subseteq \mathit{Red}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}). \tag{4}$$

Using the superscript "C" to indicate the complement of a set, we then have

$$\mathit{Red}^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subseteq \mathit{PropRed}^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subseteq \mathit{StrRed}^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}). \tag{5}$$

The symbol "$\subseteq$" means "subset of or equal to." We use "$\subset$" to indicate "proper subset of." One might expect that "$\subseteq$" can generally be replaced by "$\subset$" in (4) and (5). This is correct when $\mathcal{I}$ has at least four elements, and $n_i \geq 2$ for at least one value of i. Furthermore, for $|\mathcal{I}| \geq 2$, we have $\mathit{StrRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset \mathit{PropRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$, provided $n_i \geq 2$ for at least one i. However, for $|\mathcal{I}| = 2$ and $|\mathcal{I}| = 3$, whether $\mathit{PropRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$ is equal to, or is a proper subset of, $\mathit{Red}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$ depends on the field $\mathbb{F}$, and on the n_i's. The appendix treats this in more detail. Here is a summary of what is shown there.

1. For any field $\mathbb{F}$, if $|\mathcal{I}| \geq 4$ and $n_i \geq 2$ for at least one i,

$$\mathit{StrRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset \mathit{PropRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset \mathit{Red}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}).$$

   Consequently,

$$\mathit{Red}^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset \mathit{PropRed}^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset \mathit{StrRed}^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}).$$

2. For any field $\mathbb{F}$, if $|\mathcal{I}| \geq 2$ and $n_i \geq 2$ for at least one i,

$$\mathit{StrRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset \mathit{PropRed}(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}).$$

3. If $\mathbb{F}$ is algebraically closed and $n \geq 2$, then

   $$PropRed(\mathbb{F}, n, 2) = Red(\mathbb{F}, n, 2) \text{ and } PropRed(\mathbb{F}, n, 3) = Red(\mathbb{F}, n, 3).$$

4. For the field, $\mathbb{R}$, of real numbers, when $n = 2$, we have

   $$PropRed(\mathbb{R}, 2, 2) \subset Red(\mathbb{R}, 2, 2) \text{ and } PropRed(\mathbb{R}, 2, 3) \subset Red(\mathbb{R}, 2, 3).$$

   For $n \geq 3$,

   $$PropRed(\mathbb{R}, n, 2) = Red(\mathbb{R}, n, 2) \text{ and } PropRed(\mathbb{R}, n, 3) = Red(\mathbb{R}, n, 3).$$

4 A coupled version of Schur's Lemma

The main result of this section is Theorem 4.2, a coupled version of Schur's Lemma for reducibility and proper reducibility. Section 5 deals with the more complicated version for strong reducibility.

Consider families $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in I}$, where A_{ij} is $n_i \times n_j$ and B_{ij} is $m_i \times m_j$, linked by equations $A_{ij}X_j = X_iB_{ij}$, where X_i is $n_i \times m_i$. Recall that A_{ij} is a linear transformation from $\mathcal{V}_j$ to $\mathcal{V}_i$, and B_{ij} is a linear transformation from $\mathcal{W}_j$ to $\mathcal{W}_i$ The matrix X_i is a linear transformation from $\mathcal{W}_i$ to $\mathcal{V}_i$. Note that $ker(X_i)$ is a subspace of $\mathcal{W}_i$ and $range(X_i)$ is a subspace of $\mathcal{V}_i$.

Reviewing the proof of Schur's Lemma (Theorem 2.1), the key facts are that $ker(P)$ is an invariant subspace of $\{B_i\}_{i\in\mathcal{I}}$, and $range(P)$ is an invariant subspace of $\{A_i\}_{i\in\mathcal{I}}$. For the case of complex matrices with $A_i = B_i$ for all i, any eigenspace of P is an invariant subspace of $\{A_i\}_{i\in\mathcal{I}}$. The following "coupled" versions of these facts are used to prove coupled versions of Schur's lemma for $\mathcal{A}$, $\mathcal{B}$. In the coupled versions, the X_i's play the role of the P.

Lemma 4.1. Let $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in I}$, where A_{ij} is $n_i \times n_j$ and B_{ij} is $m_i \times m_j$. Let X_i be $n_i \times m_i$ and suppose for all $i, j \in \mathcal{I}$, we have $A_{ij}X_j = X_iB_{ij}$. If $m_i = n_i$ for some i, then, for any scalar α, define $\mathcal{U}_i(\alpha) = \{\mathbf{v} \mid X_i\mathbf{v} = \alpha\mathbf{v}\}$. The following hold for all $i, j \in \mathcal{I}$.

1. $B_{ij}(ker(X_j)) \subseteq ker(X_i)$.

2. $A_{ij}(range(X_j)) \subseteq range(X_i)$.

3. If $\mathcal{A} = \mathcal{B}$, then $A_{ij}(\mathcal{U}_j(\alpha)) \subseteq \mathcal{U}_i(\alpha)$.

Proof. For any $\mathbf{w} \in ker(X_j)$, we have $X_i(B_{ij}\mathbf{w}) = A_{ij}X_j\mathbf{w} = 0$. Hence, $B_{ij}\mathbf{w} \in ker(X_i)$. This proves 1.

For $\mathbf{w} \in \mathcal{W}$, we have $A_{ij}(X_j\mathbf{w}) = X_i(B_{ij}\mathbf{w}) \in range(X_i)$, proving 2.

Finally, suppose $\mathcal{A} = \mathcal{B}$. Then $A_{ij}X_j = X_iA_{ij}$ for all i, j. Let $\mathbf{v} \in \mathcal{U}_j(\alpha)$. Then $X_i(A_{ij}\mathbf{v}) = A_{ij}X_j\mathbf{v} = \alpha(A_{ij}\mathbf{v})$, showing $A_{ij}\mathbf{v} \in \mathcal{U}_i(\alpha)$. $\square$

If $m_i = n_i$ and α is an eigenvalue of X_i, with $\alpha \in \mathbb{F}$, then $\mathcal{U}_i(\alpha)$ is the corresponding eigenspace. If α is not an eigenvalue of X_i, then $\mathcal{U}_i(\alpha)$ is the zero subspace.

We now state a version of Schur's Lemma for families that are irreducible in the coupled sense. The proofs simply extend the argument used to prove the usual Schur Lemma.

Theorem 4.2. Let $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in I}$, where A_{ij} is $n_i \times n_j$ and B_{ij} is $m_i \times m_j$. Let X_i be $n_i \times m_i$ and suppose for all $i, j \in \mathcal{I}$, we have $A_{ij}X_j = X_iB_{ij}$.

1. Suppose both $\mathcal{A}$ and $\mathcal{B}$ are irreducible in the coupled sense. Then either $X_i = 0$ for all i, or X_i is nonsingular for all i. In the latter case, $m_i = n_i$ for all i. If $\mathcal{A} = \mathcal{B}$, and $\mathcal{A}$ is a family of complex matrices, then there is a scalar α such that $X_i = \alpha I_{n_i}$ for all i.

2. Suppose neither $\mathcal{A}$ nor $\mathcal{B}$ is properly reducible in the coupled sense. Then for each i, either $X_i = 0$ or X_i is nonsingular. If X_i is nonzero we must have $m_i = n_i$. If $\mathcal{A} = \mathcal{B}$ and consists of complex matrices, then any nonzero X_i is a scalar multiple of I_{n_i}.

Proof. For part 1, assume $\mathcal{A}$ and $\mathcal{B}$ are both coupled irreducible. Consider the subspaces $ker(X_i), i \in \mathcal{I}$. Since $\mathcal{B}$ is irreducible in the coupled sense, statement 1 of Lemma 4.1 tells us there are only two possibilities: either $ker(X_i) = \{0\}$ for all i, or $ker(X_i) = \mathcal{W}_i$ for all i. In the latter case, $X_i = 0$ for all i and so we are done.

Suppose now that $ker(X_i) = \{0\}$ for all i. We now use the subspaces $range(X_i), i \in \mathcal{I}$. Since $\mathcal{A}$ is irreducible in the coupled sense, part 2 of Lemma 4.1 tells us the only possibilities are $range(X_i) = \{0\}$ for all i or $range(X_i) = \mathcal{V}_i$ for all i. If $range(X_i) = \{0\}$ for all i, then $X_i = 0$ for all i. Otherwise, we have both $kerX_i = \{0\}$ and $range(X_i) = \mathcal{V}_i$ for all i. Hence each X_i is nonsingular and $m_i = n_i$.

Now suppose $\mathcal{A} = \mathcal{B}$ and $\mathbb{F} = \mathbb{C}$. Let λ be an eigenvalue of X_p for some fixed $p \in \mathcal{I}$. Part 3 of Lemma 4.1 tells us $A_{ij}(\mathcal{U}_j(\lambda)) \subseteq \mathcal{U}_i(\lambda)$ for all i, j. Since $\mathcal{A}$ is irreducible in the coupled sense, there are then only two possibilities for the subspaces $\mathcal{U}_i(\lambda)$: either they are all zero, or $\mathcal{U}_i = \mathcal{V}_i$ for all i. Since λ was chosen to be an eigenvalue of X_p, we know $\mathcal{U}_p(\lambda)$ is not zero. Therefore, $\mathcal{U}_i(\lambda) = \mathcal{V}_i$ for all i and hence $X_i = \lambda I_{n_i}$ for all i.

For part 2, assume neither $\mathcal{A}$ nor $\mathcal{B}$ is properly reducible in the coupled sense. Consider $ker(X_i)$. Since $\mathcal{B}$ is not properly reducible in the coupled sense, Lemma 4.1 tells us $ker(X_i)$ cannot be a nonzero, proper subspace of $\mathcal{W}_i$. Hence, for each particular i, either $ker(X_i) = \{0\}$ or $ker(X_i) = \mathcal{W}_i$. In the latter case, $X_i = 0$.

Suppose $ker(X_i) = \{0\}$ for some i. Since $\mathcal{A}$ is not properly reducible in the coupled sense, Lemma 4.1 tells us $range(X_i)$ is either $\{0\}$ or $\mathcal{V}_i$. If $range(X_i) = \{0\}$ then $X_i = 0$. Otherwise, we have both $ker X_i = \{0\}$ and $range(X_i) = \mathcal{V}_i$, so X_i is nonsingular and $m_i = n_i$.

Now suppose $\mathcal{A} = \mathcal{B}$ is a family of complex matrices. Suppose $X_p \neq 0$ for some p. Let λ_p be an eigenvalue of X_p. Note $\lambda_p \neq 0$ because X_p is nonsingular. By Lemma 4.1, we have $A_{ij}(\mathcal{U}_j(\lambda_p)) \subseteq \mathcal{U}_i(\lambda_p)$ for all i, j. Since $\mathcal{A}$ is not properly reducible in the coupled sense, each $\mathcal{U}_i(\lambda_p)$ is either zero or the full vector space $\mathcal{V}_i$. Since λ_p is an eigenvalue of X_p, the space $\mathcal{U}_p(\lambda_p)$ is not zero. Therefore, $\mathcal{U}_p(\lambda_j) = \mathcal{V}_p$ and $X_p = \lambda_p I_{n_p}$. □

Remark 4.1. The ordinary version of Schur's Lemma, Theorem 2.1, applies to the case where both $\mathcal{A}$ and $\mathcal{B}$ are irreducible in the sense of Definition 2.1, and $X_i = P$ for all i.

Note the different conclusions for the two parts of Theorem 4.2. For part 1, either *all* X_i's are zero, or *all* are nonsingular. When $\mathcal{A} = \mathcal{B}$ and $\mathbb{F} = \mathbb{C}$, *all* the X_i's are the same scalar multiple of the identity matrix. In part 2, there are more options for the X_i's. Each X_i is either zero or nonsingular, but some can be zero and others nonsingular. For $\mathcal{A} = \mathcal{B}$ and $\mathbb{F} = \mathbb{C}$, the proof for part 2 gives $X_p = \lambda_p I_{n_p}$ for a particular value of p; it does not show every nonzero X_i equals the same scalar multiple of the identity matrix.

The broader range of options for the X_i's in part 2 makes sense when we consider that, at least for $|\mathcal{I}| \geq 4$, we have $PropRed(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset Red(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$, and hence $Red^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset PropRed^C(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$. Part 2 applies to a broader set of pairs $\mathcal{A}$, $\mathcal{B}$ than part 1.

Consider the situation in part 2 of Theorem 4.2. Suppose $X_i = 0$ and X_j is nonsingular. The equation $A_{ij}X_j = X_iB_{ij}$ then tells us $A_{ij} = 0$, while

$A_{ji}X_i = X_jB_{ji}$ gives $B_{ji} = 0$. Set

$$\mathcal{I}_0 = \{i \in \mathcal{I} \mid X_i = 0\} \quad \text{and} \quad \mathcal{I}_{non} = \{i \in \mathcal{I} \mid X_i \text{ is nonsingular}\}.$$

We have $A_{ij} = 0$ and $B_{ji} = 0$ whenever $i \in \mathcal{I}_0$ and $j \in \mathcal{I}_{non}$. For example, suppose $\mathcal{I} = \{1, 2, \ldots, K\}$, and, for some $0 < s < K$, we have $\mathcal{I}_0 = \{1, 2, \ldots, s\}$ and $\mathcal{I}_{non} = \{s + 1, s + 2, \ldots, K\}$. The $N \times N$ matrix A then has only zero blocks in the upper right hand corner formed from the first s rows and last $K - s$ columns. The $M \times M$ matrix B has zero blocks in the lower left hand corner formed by the last $K - s$ rows and first s columns.

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1s} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{s1} & \cdots & A_{ss} & 0 & \cdots & 0 \\ A_{(s+1)1} & \cdots & A_{(s+1)s} & A_{(s+1)(s+1)} & \cdots & A_{(s+1)K} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{K1} & \cdots & A_{Ks} & A_{K(s+1)} & \cdots & A_{KK} \end{pmatrix}.$$

Returning to the case of general $\mathcal{I}$, one can check that $\mathcal{A}$ is coupled reducible via the subspaces $\mathcal{U}_i = \{0\}$ for $i \in \mathcal{I}_0$, and $\mathcal{U}_i = \mathcal{V}_i$ for $i \in \mathcal{I}_{non}$. The family $\mathcal{B}$ is coupled reducible via the subspaces $\mathcal{U}_i = \mathcal{W}_i$ when $i \in \mathcal{I}_0$, and $\mathcal{U}_i = \{0\}$ when $i \in \mathcal{I}_{non}$.

5 Strong reducibility and Schur's Lemma

We now consider strongly coupled reducibility. Our goal is a version of Schur's lemma for families that are not strongly reducible in the coupled sense, with a conclusion similar to that of Theorem 4.2: each X_i is either zero or nonsingular. The next example shows that for such a conclusion, we need some restrictions on A_{ij} and B_{ij}.

Example 5.1. Let $n = m = 2$ (so $\mathcal{V} = \mathcal{W} = \mathbb{F}^2$), and $K = 2$. Put

$$A_{11} = A_{22} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad A_{21} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad A_{12} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$B_{11} = B_{22} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B_{21} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad B_{12} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

In terms of the matrices A, B:

$$A = \left(\begin{array}{cc|cc} 0 & 1 & a & b \\ 0 & 0 & c & d \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \qquad B = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline a & b & 0 & 1 \\ c & d & 0 & 0 \end{array}\right).$$

Let $\mathcal{U}$ be the subspace spanned by $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. One may easily check that $\mathcal{A}$ is properly reducible in the coupled sense with $\mathcal{U}_1 = \mathcal{U}$ and $\mathcal{U}_2 = \{0\}$, while $\mathcal{B}$ is properly reducible in the coupled sense with $\mathcal{U}_1 = \{0\}$ and $\mathcal{U}_2 = \mathcal{U}$. However, if $c \neq 0$, then neither $\mathcal{A}$ nor $\mathcal{B}$ is strongly reducible in the coupled sense. The reason is that $\mathcal{U}$ is the only nonzero, proper invariant subspace for the diagonal blocks, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, of A and B, so $\mathcal{U}_1 = \mathcal{U}_2 = \mathcal{U}$ is the only possible choice for nonzero, proper subspaces $\mathcal{U}_1$ and $\mathcal{U}_2$. If $c \neq 0$, then $A_{12}(\mathcal{U}) \not\subset \mathcal{U}$, and $B_{21}(\mathcal{U}) \not\subset \mathcal{U}$. So if $c \neq 0$, neither $\mathcal{A}$ nor $\mathcal{B}$ is strongly reducible in the coupled sense. Set $X_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $X_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $X = X_1 \oplus X_2$. One may check that $AX = XB = 0$ and hence $A_{ij}X_j = X_iB_{ij}$ for $i, j = 1, 2$. The point is that the matrix X_1 is neither zero nor nonsingular.

Our theorem for coupled pairs $\mathcal{A}$, $\mathcal{B}$ that are not strongly reducible will be for the case when $n_i = n$ and $m_i = m$ for all i. It will have a hypothesis about graphs related to $\mathcal{A}$ and $\mathcal{B}$; roughly speaking, this hypothesis will tell us there are "enough" nonsingular A_{ij}'s and B_{ij}'s. Although our main result assumes $\mathcal{A}$ is a family of $n \times n$ matrices and $\mathcal{B}$ is a family of $m \times m$ matrices, we define the graphs for families with matrices of any size.

Recall that a matrix is said to have *full column rank* if the columns are linearly independent; thus, the rank of the matrix equals the number of columns. If A is a $p \times q$ matrix with full column rank, and $\mathcal{U}$ is a subspace of $\mathbb{F}^q$, then $A(\mathcal{U})$ has the same dimension as $\mathcal{U}$.

Consider $\mathcal{A} = \{A_{ij}\}_{i,j \in I}$, and subspaces $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$, satisfying $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ for all i, j. Let d_i be the dimension of $\mathcal{U}_i$. If A_{ij} has full column rank, then $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ tells us $d_j \leq d_i$. If A_{ji} also has full column rank, then we also have $d_i \leq d_j$, and hence $d_i = d_j$. When A_{ij} and A_{ji} both have full column rank, $n_j \leq n_i$ and $n_i \leq n_j$, so $n_i = n_j$; hence, A_{ij} and A_{ji} are actually square, nonsingular matrices. If all of the A_{ij}'s have full column rank, then all of the n_i's have the same value, n, and all of the subspaces $\mathcal{U}_j$ have the same dimension, d. However, we need not assume *all* of the matrices A_{ij} are nonsingular in order to show the $\mathcal{U}_j$'s all have the same dimension. To explore this further, we introduce a directed graph in which directed edges correspond to the A_{ij}'s of full column rank.

Definition 5.1. Let $\mathcal{A} = \{A_{ij}\}_{i,j \in I}$, with A_{ij} of size $n_i \times n_j$. The *directed graph (digraph) of* $\mathcal{A}$, denoted $\mathcal{D}(\mathcal{A})$, is the graph on vertices $\{v_i\}_{i \in \mathcal{I}}$, such that there is a directed edge (v_i, v_j) from v_i to v_j if and only if A_{ij} has full column rank.

For a finite index set, $\mathcal{I} = \{1, \ldots, K\}$, there are K vertices. If $n_i = 1$ for all i, our $\mathcal{D}(\mathcal{A})$ is just the usual directed graph associated with a $K \times K$ matrix.

More generally, there is a vertex for each $i \in \mathcal{I}$, so there could be infinitely many vertices. We use the same definition for directed walk as for graphs with a finite number of vertices. A *directed walk* is a finite sequence of vertices, $v_{i_1}, v_{i_2}, \ldots, v_{i_p}$, such that $(v_{i_j}, v_{i_{(j+1)}})$ is a directed edge for $1 \leq j \leq (p-1)$. In this case, we write $v_{i_1} \to v_{i_2} \to \cdots \to v_{i_p}$. Vertices v and w in a directed graph $\mathcal{D}$ are said to be *strongly connected* if there is a directed walk from v to w and a directed walk from w to v. We say $\mathcal{D}$ is strongly connected if each pair of vertices of $\mathcal{D}$ is strongly connected.

Proposition 5.1. Let $\mathcal{A} = \{A_{ij}\}_{i,j \in \mathcal{I}}$ and suppose the subspaces $\{U_i\}_{i \in \mathcal{I}}$ satisfy $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$, for all i, j. Then the following hold.

1. If there is a directed walk from v_i to v_j in $\mathcal{D}(\mathcal{A})$, then $n_j \leq n_i$, and $dim(\mathcal{U}_j) \leq dim(\mathcal{U}_i)$.

2. If the vertices v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{A})$, then $n_i = n_j$, and $dim(\mathcal{U}_j) = dim(\mathcal{U}_i)$

3. If $\mathcal{D}(\mathcal{A})$ is strongly connected, all of the n_i's are equal, and all of the subspaces $\mathcal{U}_i$ have the same dimension.

Proof. Let $d_i = dim(\mathcal{U}_i)$. If (v_i, v_j) is a directed edge of $\mathcal{D}(\mathcal{A})$, then A_{ij} has full column rank, so, as we have already observed, $n_j \leq n_i$ and $d_j \leq d_i$.

More generally, suppose $v_i = v_{i_1} \to v_{i_2} \to \cdots \to v_{i_p} = v_j$ is a directed walk from v_i to v_j in $\mathcal{D}(\mathcal{A})$. Working from right to left, we have

$$n_j = n_{i_p} \leq n_{i_{p-1}} \leq \cdots \leq n_{i_2} \leq n_{i_1} = n_i$$

and

$$d_j = d_{i_p} \leq d_{i_{p-1}} \leq \cdots \leq d_{i_2} \leq d_{i_1} = d_i.$$

Hence, $n_j \leq n_i$ and $d_j \leq d_i$.

If v_i and v_j are strongly connected, there is a directed walk from v_i to v_j and a directed walk from v_j to v_i. So $n_i = n_j$ and $d_i = d_j$.

If $\mathcal{D}(\mathcal{A})$ is strongly connected, then, for all i, j, we have $d_i = d_j$ and $n_i = n_j$, so all of the subspaces $\mathcal{U}_j$ have the same dimension and all of the n_i's have the same value. $\square$

As an example, suppose $\mathcal{I} = \{1, \ldots, K\}$ and $A_{12}, A_{23}, \ldots A_{K-1,K}, A_{K1}$ all have full column rank. Then $\mathcal{D}(\mathcal{A})$ contains the directed cycle

$$v_1 \to v_2 \to \cdots \to v_{K-1} \to v_K \to v_1,$$

and is strongly connected.

If $\mathcal{D}(\mathcal{A})$ is not strongly connected, the strong components identify sets of n_i's which must be equal, and sets of subspaces $\mathcal{U}_i$ which must have the same dimension. For each strong component, $\mathcal{C}$, of $\mathcal{D}(\mathcal{A})$, all n_i's corresponding to vertices of $\mathcal{C}$ must be equal, and all subspaces $\mathcal{U}_i$ corresponding to vertices of $\mathcal{C}$ must have the same dimension. For a finite $\mathcal{I}$, we can use the strong components to put the $N \times N$ matrix A into a block triangular form in which none of the A_{ij}'s below the diagonal blocks has full column rank. (See [3], section 3.2.)

For the proofs of coupled versions of Schur's Lemma, the subspaces $\mathcal{U}_i$ of interest are the kernels and ranges of the matrices $\{X_i\}_{i \in \mathcal{I}}$.

Proposition 5.2. Let $\mathcal{A} = \{A_{ij}\}_{i,j \in \mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j \in \mathcal{I}}$. Let X_i be $n_i \times m_i$, and suppose $A_{ij} X_j = X_i B_{ij}$ for all $i, j \in \mathcal{I}$. Then the following hold.

1. If v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{A})$, then $range(X_i)$ and $range(X_j)$ have the same dimension, i.e., X_i and X_j have the same rank.

2. If v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{B})$, then $ker(X_i)$ and $ker(X_j)$ have the same dimension, i.e., X_i and X_j have the same nullity.

3. If $\mathcal{D}(\mathcal{A})$ is strongly connected, all of the n_i's have the same value, n, and all of the X_i's have the same rank.

4. If $\mathcal{D}(\mathcal{B})$ is strongly connected, all of the m_i's have the same value, m, and all of the X_i's have the same nullity, d.
5. If v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{B})$, then X_i and X_j have the same rank.
6. If $\mathcal{D}(\mathcal{B})$ is strongly connected, all of the X_i's have the same rank.

Proof. The first four parts follow from Lemma 4.1 and Proposition 5.1. For part 5, suppose v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{B})$. Then $m_i = m_j$, so the matrices X_i and X_j have the same number of columns. From part 2, we know X_i and X_j have the same nullity. The rank plus nullity theorem then tells us X_i and X_j have the same rank. Part 6 is an immediate consequence of part 5. $\square$

We now have a version of Schur's lemma for families $\mathcal{A}$, $\mathcal{B}$ when neither is strongly reducible in the coupled sense.

Theorem 5.1. Assume neither $\mathcal{A} = \{A_{ij}\}_{i,j\in I}$ nor $\mathcal{B} = \{B_{ij}\}_{i,j\in I}$ is strongly reducible in the coupled sense. Assume also that $\mathcal{D}(\mathcal{A})$ and $\mathcal{D}(\mathcal{B})$ are strongly connected. Let X_i be $n \times m$ for all $i \in \mathcal{I}$, and suppose $A_{ij}X_j = X_iB_{ij}$ for all $i, j \in \mathcal{I}$. Then either $X_i = 0$ for all i, or X_i is nonsingular for all i. In the latter case we must have $m = n$. If $\mathcal{A} = \mathcal{B}$ and is a family of complex matrices, then there is some scalar α such that $X_i = \alpha I_n$ for all i.

Proof. Note first that since $\mathcal{D}(\mathcal{A})$ and $\mathcal{D}(\mathcal{B})$ are both strongly connected, the A_{ij}'s are all square matrices of the same size, n, and the B_{ij}'s are all square matrices of the same size, m.

By Proposition 5.2, the subspaces $ker(X_i)$, for $i \in \mathcal{I}$, all have the same dimension, d. Since $\mathcal{B}$ is not strongly reducible in the coupled sense, either $d = 0$ or $d = m$. If $d = m$, then $X_i = 0$ for all i and we are done.

Assume then that $d = 0$. Proposition 5.2 tells us the subspaces $range(X_i)$ all have the same dimension, r. Since $\mathcal{A}$ is not strongly reducible in the coupled sense either $r = 0$ or $r = n$. If $r = 0$, then $X_i = 0$ for all j. If $r = n$, then, since we also have $d = 0$, the X_i's are nonsingular; we then have $m = n$.

If $\mathcal{A} = \mathcal{B}$, we have $A_{ij}X_j = X_iA_{ij}$ for all i, j. Fix p and let λ be an eigenvalue of X_p with corresponding eigenspace $\mathcal{U}_p(\lambda)$; note the subspace $\mathcal{U}_p(\lambda)$ is nonzero, because λ is an eigenvalue of X_p. From part 3 of Lemma 4.1, we have $A_{ij}(\mathcal{U}_j(\lambda)) \subseteq \mathcal{U}_i(\lambda)$ for all i, j. Since $\mathcal{D}(\mathcal{A})$ is strongly connected, Proposition 5.1 tells us the spaces $\mathcal{U}_i(\lambda)$ all have the same dimension; call it f. Since $\mathcal{U}_p(\lambda)$ is nonzero, we know $f > 0$. Hence, since $\mathcal{A}$ is not strongly reducible in the coupled sense, we must have $f = n$, and $X_i = \lambda I_n$ for all i. □

Remark 5.1. Earlier work [6] gives a proof, using block matrix computation, for the case where all A_{ij}'s and B_{ij}'s are assumed to be nonsingular.

The proof of Theorem 5.1 uses the assumption that both $\mathcal{D}(\mathcal{A})$ and $\mathcal{D}(\mathcal{B})$ are strongly connected in two ways: to establish that $n_i = n$ and $m_i = m$ for all i, and to show that the relevant subspaces (kernels and ranges of the X_i's) have the same dimension. We now develop another version of Theorem 5.1, in which we weaken the hypothesis about the graphs, but then need to explicitly assume that $n_i = n$ and $m_i = m$ for all i. The key point for this second version is that X_i and X_j have the same rank whenever v_i and v_j are strongly connected in *either* of the digraphs $\mathcal{D}(\mathcal{A})$ or $\mathcal{D}(\mathcal{B})$.

We use $\mathcal{A}$ and $\mathcal{B}$ to define an undirected graph, $\mathcal{G}(\mathcal{A}, \mathcal{B})$, as follows.

Definition 5.2. The undirected graph, $\mathcal{G}(\mathcal{A}, \mathcal{B})$, is the graph on vertices $\{v_i\}_{i\in\mathcal{I}}$, such that $\{v_i, v_j\}$ is an (undirected) edge of $\mathcal{G}(\mathcal{A}, \mathcal{B})$ if and only if the vertices v_i and v_j are either strongly connected in $\mathcal{D}(\mathcal{A})$, or in $\mathcal{D}(\mathcal{B})$ (or both). We call this the *linked graph of* $\mathcal{A}$ *and* $\mathcal{B}$.

Proposition 5.3. Let $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in\mathcal{I}}$. Let X_i be $n_i \times m_i$ and suppose $A_{ij}X_j = X_iB_{ij}$ for all $i, j \in \mathcal{I}$. If v_i and v_j are connected in $\mathcal{G}(\mathcal{A}, \mathcal{B})$ then X_i and X_j have the same rank. If $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is connected, then all of the matrices X_i have the same rank.

Proof. Suppose v_i and v_j are connected in $\mathcal{G}(\mathcal{A}, \mathcal{B})$. Then there is a sequence of vertices, $v_i = v_{i_1}, v_{i_2}, v_{i_3}, \ldots, v_{i_{p-1}}, v_{i_p} = v_j$, such that $\{v_{i_k}, v_{i_{k+1}}\}$ is an edge of $\mathcal{G}(\mathcal{A}, \mathcal{B})$ for $k = 1, \ldots, p-1$. This means v_{i_k} and $v_{i_{k+1}}$ are either strongly connected in $\mathcal{D}(\mathcal{A})$ or strongly connected in $\mathcal{D}(\mathcal{B})$, or both. Therefore, $rank(X_{i_k}) = rank(X_{i_{k+1}})$ for $k = 1, \ldots, p-1$, and $rank(X_i) = rank(X_j)$. □

If either $\mathcal{D}(\mathcal{A})$ or $\mathcal{D}(\mathcal{B})$ is strongly connected, then $\mathcal{G}(\mathcal{A}, \mathcal{B})$ will be connected. However, $\mathcal{G}(\mathcal{A}, \mathcal{B})$ can be a connected graph even if neither of the digraphs $\mathcal{D}(\mathcal{A})$ or $\mathcal{D}(\mathcal{B})$ is strongly connected. For example, suppose $K = 3$ and

$$A = \begin{pmatrix} 0 & A_{12} & 0 \\ A_{21} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 & B_{13} \\ 0 & 0 & 0 \\ B_{31} & 0 & 0 \end{pmatrix},$$

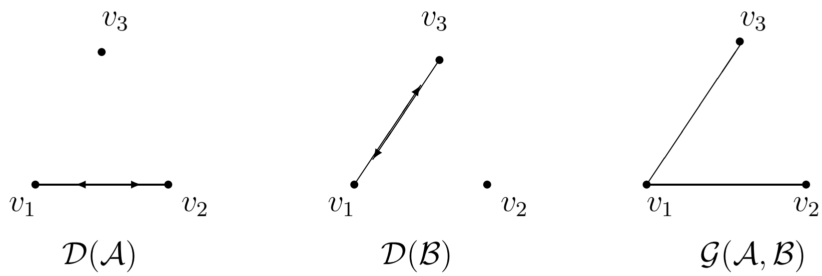

Figure 1: $\mathcal{D}(\mathcal{A})$, $\mathcal{D}(\mathcal{B})$ and $\mathcal{G}(\mathcal{A},\, \mathcal{B})$

where A_{12}, A_{21}, B_{13} and B_{31} are all nonsingular. Neither $\mathcal{D}(\mathcal{A})$ nor $\mathcal{D}(\mathcal{B})$ is connected, but $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is connected. (See Figure 1.)

As an example, suppose $\mathcal{I}_1$ and $\mathcal{I}_2$ are nonempty, disjoint subsets of $\mathcal{I}$ such that $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$. Partition the vertices of $\mathcal{G}(\mathcal{A},\, \mathcal{B})$ into two sets corresponding to $\mathcal{I}_1$ and $\mathcal{I}_2$, setting

$$\mathcal{S} = \{v_i \mid i \in \mathcal{I}_1\} \quad \text{and} \quad \mathcal{T} = \{v_i \mid i \in \mathcal{I}_2\}.$$

Suppose $rank(A_{ij}) < n_j$ and $rank(B_{ij}) < m_j$, whenever $i \in \mathcal{I}_1$ and $j \in \mathcal{I}_2$. Then neither $\mathcal{D}(\mathcal{A})$ nor $\mathcal{D}(\mathcal{B})$ has any directed edges from vertices in $\mathcal{S}$ to vertices in $\mathcal{T}$. The linked graph $\mathcal{G}(\mathcal{A}, \mathcal{B})$ then has no edges from vertices in $\mathcal{S}$ to vertices in $\mathcal{T}$ and hence is not connected.

Now suppose that, whenever $i \in \mathcal{I}_1$ and $j \in \mathcal{I}_2$, we have $rank(A_{ij}) < n_j$ and $rank(B_{ji}) < m_i$, (note the reversal of subscripts on B_{ji}). In this case, $\mathcal{D}(\mathcal{A})$ has no directed edges from vertices in $\mathcal{S}$ to vertices in $\mathcal{T}$, while $\mathcal{D}(\mathcal{B})$ has no directed edges from vertices in $\mathcal{T}$ to vertices in $\mathcal{S}$. Consequently, if $v \in \mathcal{S}$ and $w \in \mathcal{T}$, then the pair v, w is not strongly connected in either $\mathcal{D}(\mathcal{A})$ or $\mathcal{D}(\mathcal{B})$. Hence, $\mathcal{G}(\mathcal{A}, \mathcal{B})$ has no edges between vertices in $\mathcal{S}$ and vertices in $\mathcal{T}$; thus $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is not connected.

The following variation of Theorem 5.1 uses this linked graph, $\mathcal{G}(\mathcal{A},\, \mathcal{B})$.

Theorem 5.2. Assume neither $\mathcal{A} = \{A_{ij}\}_{i,j \in I}$ nor $\mathcal{B} = \{B_{ij}\}_{i,j \in I}$ is strongly reducible in the coupled sense. Assume also that $n_i = n$ and $m_i = m$ for all i, and that $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is connected. Let X_i be $n \times m$, and suppose $A_{ij} X_j = X_i B_{ij}$ for all $i, j \in \mathcal{I}$. Then either $X_i = 0$ for all i, or X_i is nonsingular for all i. In the latter case we must have $m = n$. If $\mathcal{A} = \mathcal{B}$ and is a family of complex matrices, then there is some scalar α such that $X_i = \alpha I_n$ for all i.

Proof. By Proposition 5.3, the subspaces $range(X_i)$, for i in $\mathcal{I}$, all have the same dimension, r. Since all of the X_i's have the same number of columns, the rank plus nullity theorem tells us the subspaces $ker(X_i)$, for $i \in \mathcal{I}$, must also all have the same dimension, d. The remainder of the proof is the same as that for Theorem 5.1. $\square$

Comparing Theorems 4.2, 5.1, and 5.2, the simplest version is part 1 of Theorem 4.2. It is the closest to the usual Schur's Lemma. However, the hypothesis that $\mathcal{A}$, $\mathcal{B}$ be irreducible in the coupled sense is more restrictive than the hypothesis of part 2 of Theorem 4.2. The conclusion of part 2 has more options for the X_i's than part 1. Theorems 5.1 and 5.2 apply to the larger class of pairs, $\mathcal{A}$, $\mathcal{B}$, which are not strongly reducible in the coupled sense, but have additional restrictions about the connectivity of the graphs $\mathcal{D}(\mathcal{A})$, $\mathcal{D}(\mathcal{B})$, and $\mathcal{G}(\mathcal{A}, \mathcal{B})$ and the equality of the n_i's and m_i's.

6 Normality and coupled normality

We now consider a refinement of Schur's Lemma for irreducible sets of normal matrices. This is closely related to Lemma A.4 of [11]. We obtain corresponding results for sets $\mathcal{A}$, $\mathcal{B}$ satisfying a "coupled normality" condition. For this section we work over the field of complex numbers. We use * to denote the transpose conjugate of a matrix. If $\mathcal{U}$ is a subspace of $\mathcal{V}$, we use $\mathcal{U}^\perp$ for the orthogonal complement of $\mathcal{U}$. We will need the following facts.

Proposition 6.1. Let A be a normal matrix; let S be nonsingular and let $B = S^{-1}AS$. Then the following are equivalent.

1. The matrix B is normal.
2. $S^{-1}A^*S = B^*$.
3. The matrix SS^* commutes with A.
4. The matrix SS^* commutes with A^*.
5. The matrix S^*S commutes with B.
6. The matrix S^*S commutes with B^*.

Proof. The equivalence of 2, 3 and 4 is easily shown. Using $B = S^{-1}AS$,

$$\begin{aligned} S^{-1}A^*S = B^* \quad &\Longleftrightarrow \quad S^{-1}A^*S = S^*A^*S^{-*} \\ &\Longleftrightarrow \quad A^*SS^* = SS^*A^* \\ &\Longleftrightarrow \quad SS^*A = ASS^*, \end{aligned}$$

where the third line comes from taking the transpose conjugate of the equation in the second line. A similar calculation, starting with $A = SBS^{-1}$, shows 2, 5 and 6 are equivalent:

$$\begin{aligned} SB^*S^{-1} = A^* \quad &\iff \quad SB^*S^{-1} = S^{-*}B^*S^* \\ &\iff \quad S^*SB^* = B^*S^*S \\ &\iff \quad BS^*S = S^*SB. \end{aligned}$$

The fact that 2 implies 1 is also easy. If $S^{-1}A^*S = B^*$, use $AA^* = A^*A$ to get

$$BB^* = (S^{-1}AS)(S^{-1}A^*S) = (S^{-1}A^*S)(S^{-1}AS) = B^*B.$$

The only part needing any work at all is to show 1 implies 2. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A and let D be the diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Since A is normal, $A = U^*DU$ for some unitary matrix U, and $A^* = U^*\overline{D}U$, where the bar denotes complex conjugation. Note $\overline{\lambda}_1, \dots, \overline{\lambda}_n$ are the eigenvalues of A^*. Let $p(x)$ be a polynomial such that $p(\lambda_i) = \overline{\lambda}_i$ for each eigenvalue λ_i. Then $\overline{D} = p(D)$, and

$$A^* = U^*p(D)U = p(U^*DU) = p(A).$$

Since B is similar to A, the matrix B also has eigenvalues $\lambda_1, \dots, \lambda_n$ and B^* has eigenvalues $\overline{\lambda}_1, \dots, \overline{\lambda}_n$. If B is normal, $B = V^*DV$ for some unitary matrix V. Hence,

$$B^* = V^*\overline{D}V = V^*p(D)V = p(V^*DV) = p(B).$$

But $p(B) = p(S^{-1}AS) = S^{-1}p(A)S = S^{-1}A^*S$, so $B^* = S^{-1}A^*S$. □

This gives an easy proof of the following.

Theorem 6.1. Suppose $\{A_i\}_{i\in\mathcal{I}}$ and $\{B_i\}_{i\in\mathcal{I}}$ are irreducible families of normal matrices, and S is a nonsingular matrix such that $S^{-1}A_iS = B_i$ for all $i \in \mathcal{I}$. Then S is a scalar multiple of a unitary matrix.

Proof. By the preceding proposition, SS^* commutes with each A_i. Since $\{A_i\}_{i\in\mathcal{I}}$ is an irreducible family, SS^* must be a scalar matrix. Since S is nonsingular, the Hermitian matrix SS^* is positive definite; hence $SS^* = \alpha I$ where α is a positive real number. Set $U = \frac{1}{\sqrt{\alpha}}S$. Then $UU^* = \frac{1}{\alpha}SS^* = I$. So U is unitary and $S = \sqrt{\alpha}U$. □

Remark 6.1. This argument is essentially the proof of Lemma A.4 of [11], which says that if two irreducible representations of a $*$-algebra of square matrices are equivalent, then they are similar via a unitary similarity. Let $\mathcal{S}$ be the algebra generated by $\{A_i\}_{i\in\mathcal{I}}$ and let $\mathcal{T}$ be the algebra generated by $\{B_i\}_{i\in\mathcal{I}}$. For any normal matrix, N, the matrix N^* is a polynomial in N, so the algebras $\mathcal{S}$ and $\mathcal{T}$ are $*$-algebras, (which means that whenever A is in the algebra, so is A^*). Let S be a nonsingular matrix such that $S^{-1}A_iS = B_i$ for all i. Proposition 6.1 tells us $S^{-1}A_i^*S = B_i^*$ for all i, so S may be extended to an isomorphism of the $*$-algebras $\mathcal{S}$ and $\mathcal{T}$ in the usual way.

We now introduce the idea of coupled normality.

Definition 6.1. The family $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ is *normal in the coupled sense* if for all $i,j \in \mathcal{I}$ we have $A_{ij}^*A_{ij} = A_{ji}A_{ji}^*$.

If $\mathcal{A}$ is normal in the coupled sense, setting $i = j$ gives $A_{ii}^*A_{ii} = A_{ii}A_{ii}^*$, so A_{ii} is normal for all i. Note also that if $A_{ji} = A_{ij}^*$ for all i,j, then $\mathcal{A}$ is coupled normal. When $\mathcal{I} = \{1,\ldots,K\}$, the condition $A_{ji} = A_{ij}^*$ for all i,j holds when A is a Hermitian matrix. In the JISA model, A is a covariance matrix, and hence is a real, symmetric matrix, so it is Hermitian.

Recall that, for any matrix G, the four matrices G, G^*, GG^*, and G^*G all have the same rank. Hence, when $\mathcal{A}$ is normal in the coupled sense, the matrices A_{ij} and A_{ji} have the same rank. In particular, note that A_{ij} is nonsingular if and only if A_{ji} is nonsingular.

Let C be a $q\times p$ matrix, let D be a $p\times q$ matrix, and let M be the $(p+q)\times(p+q)$ matrix

$$M = \begin{pmatrix} 0 & D \\ C & 0 \end{pmatrix},$$

where the zero blocks are $p\times p$ and $q\times q$. Then

$$MM^* = \begin{pmatrix} DD^* & 0 \\ 0 & CC^* \end{pmatrix} \quad \text{and} \quad M^*M = \begin{pmatrix} C^*C & 0 \\ 0 & D^*D \end{pmatrix}.$$

Hence, M is normal if and only if $C^*C = DD^*$ and $D^*D = CC^*$. The connection with coupled normality is this: if we set $M_{ij} = \begin{pmatrix} 0 & A_{ij} \\ A_{ji} & 0 \end{pmatrix}$, then $\mathcal{A} = \{A_{ij}\}_{i\in\mathcal{I}}$ is normal in the coupled sense if and only if M_{ij} is normal for all $i,j \in \mathcal{I}$.

Suppose $\mathcal{A}$ is normal in the coupled sense and the subspaces $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$ satisfy $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ for all i,j. Let d_i be the dimension of $\mathcal{U}_i$. We use the fact that M_{ij} is normal to show that $A_{ij}(\mathcal{U}_j^\perp) \subseteq \mathcal{U}_i^\perp$ for all i,j.

Proposition 6.2. Let C, D be matrices of sizes $q\times p$ and $p\times q$, respectively, such that $C^*C = DD^*$ and $D^*D = CC^*$. Suppose there are subspaces $\mathcal{U}$ of $\mathbb{C}^p$, and $\mathcal{W}$ of $\mathbb{C}^q$, such that $C(\mathcal{U}) \subseteq \mathcal{W}$ and $D(\mathcal{W}) \subseteq \mathcal{U}$. Then $C(\mathcal{U}^\perp) \subseteq \mathcal{W}^\perp$ and $D(\mathcal{W}^\perp) \subseteq \mathcal{U}^\perp$.

Proof. Let $M = \begin{pmatrix} 0 & D \\ C & 0 \end{pmatrix}$. For any $\mathbf{x} \in \mathbb{C}^p$ and $\mathbf{y} \in \mathbb{C}^q$,

$$M \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} D\mathbf{y} \\ C\mathbf{x} \end{pmatrix}.$$

If $\mathbf{x} \in \mathcal{U}$ and $\mathbf{y} \in \mathcal{W}$, then $D\mathbf{y} \in \mathcal{U}$ and $C\mathbf{x} \in \mathcal{W}$. So $\mathcal{U} \oplus \mathcal{W}$, (which is a subspace of $\mathbb{C}^p \oplus \mathbb{C}^q$), is invariant under M. Since M is normal, the orthogonal complement of $\mathcal{U} \oplus \mathcal{W}$ in $\mathbb{C}^p \oplus \mathbb{C}^q$ must also be invariant under M. Hence, $\mathcal{U}^\perp \oplus \mathcal{W}^\perp$ is invariant under M. This means that, for $\mathbf{x} \in \mathcal{U}^\perp$ and $\mathbf{y} \in \mathcal{W}^\perp$, we have $D\mathbf{y} \in \mathcal{U}^\perp$ and $C\mathbf{x} \in \mathcal{W}^\perp$. So $C(\mathcal{U}^\perp) \subseteq \mathcal{W}^\perp$ and $D(\mathcal{W}^\perp) \subseteq \mathcal{U}^\perp$. □

Apply Proposition 6.2 to the normal matrix $M_{ij} = \begin{pmatrix} 0 & A_{ij} \\ A_{ji} & 0 \end{pmatrix}$, to get $A_{ij}(\mathcal{U}_j^\perp) \subseteq \mathcal{U}_i^\perp$ for all i, j. Hence, if $\mathcal{A}$ is normal in the coupled sense, and is reducible in the coupled sense, then it is fully reducible in the coupled sense, because we can form T_j using a basis for $\mathcal{U}_j$ for the first d_j columns and a basis for $\mathcal{U}_j^\perp$ for the remaining $n - d_j$ columns. If we use orthonormal bases for $\mathcal{U}_j$ and $\mathcal{U}_j^\perp$, then T_j will be unitary. Hence $\mathcal{A}$ is fully reducible in the coupled sense with a coupled unitary similarity.

We will give three versions of Theorem 6.1 for $\mathcal{A}$, $\mathcal{B}$ which are normal in the coupled sense, corresponding to the three types of reducibility. The proofs depend on the following proposition. The first two statements are a "coupled" version of Proposition 6.1. Part 4 uses the digraphs $\mathcal{D}(\mathcal{A})$ and $\mathcal{D}(\mathcal{B})$.

Proposition 6.3. Assume the families $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in\mathcal{I}}$, where A_{ij} and B_{ij} are complex matrices, are normal in the coupled sense. Suppose $A_{ij}S_j = S_iB_{ij}$ for all i, j, where S_i is $n_i \times m_i$. For any $i \in \mathcal{I}$, and any scalar α, define

$$\mathcal{U}_i(\alpha) = \{\mathbf{v} \mid S_iS_i^*\mathbf{v} = \alpha\mathbf{v}\} \quad \text{and} \quad \mathcal{Y}_i(\alpha) = \{\mathbf{w} \mid S_i^*S_i\mathbf{w} = \alpha\mathbf{w}\}.$$

Then the following hold.

1. If S_i is nonsingular then $S_iS_i^*$ commutes with A_{ii}, and $S_i^*S_i$ commutes with B_{ii}.

2. If S_i and S_j are both nonsingular,

$$S_i S_i^* A_{ij} = A_{ij} S_j S_j^* \quad \text{and} \quad S_i^* S_i B_{ij} = B_{ij} S_j^* S_j.$$

3. If S_i and S_j are both nonsingular,

$$A_{ij}(\mathcal{U}_j(\alpha)) \subseteq \mathcal{U}_i(\alpha) \quad \text{and} \quad B_{ij}(\mathcal{Y}_j(\alpha)) \subseteq \mathcal{Y}_i(\alpha).$$

   If A_{ij} is also nonsingular, then $dim(\mathcal{U}_i(\alpha)) = dim(\mathcal{U}_j(\alpha))$.

   If B_{ij} is also nonsingular, then $dim(\mathcal{Y}_i(\alpha)) = dim(\mathcal{Y}_j(\alpha))$.

4. Assume S_i is nonsingular for all $i \in \mathcal{I}$. Then the following hold.

   If v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{A})$, then $\mathcal{U}_i(\alpha)$ and $\mathcal{U}_j(\alpha)$ have the same dimension.

   If v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{B})$, then $\mathcal{Y}_i(\alpha)$ and $\mathcal{Y}_j(\alpha)$ have the same dimension.

5. If $\alpha \neq 0$ then $\mathcal{U}_i(\alpha)$ and $\mathcal{Y}_i(\alpha)$ have the same dimension.

6. Assume S_i is nonsingular for all $i \in \mathcal{I}$. Then if v_i and v_j are connected in $\mathcal{G}(\mathcal{A}, \mathcal{B})$, and $\alpha \neq 0$, we have $dim(\mathcal{U}_i(\alpha)) = dim(\mathcal{U}_j(\alpha))$.

Proof. Suppose S_i is nonsingular. Since A_{ii} and B_{ii} are both normal, and $S_i^{-1} A_{ii} S_i = B_{ii}$, Proposition 6.1 tells us that $S_i S_i^*$ commutes with A_{ii} and $S_i^* S_i$ commutes with B_{ii}.

Now suppose $i \neq j$, and S_i and S_j are both nonsingular. Set $M_{ij} = \begin{pmatrix} 0 & A_{ij} \\ A_{ji} & 0 \end{pmatrix}$. Then

$$\begin{aligned}
M_{ij} \begin{pmatrix} S_i & 0 \\ 0 & S_j \end{pmatrix} &= \begin{pmatrix} 0 & A_{ij} \\ A_{ji} & 0 \end{pmatrix} \begin{pmatrix} S_i & 0 \\ 0 & S_j \end{pmatrix} \\
&= \begin{pmatrix} 0 & A_{ij} S_j \\ A_{ji} S_i & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & S_i B_{ij} \\ S_j B_{ji} & 0 \end{pmatrix} \\
&= \begin{pmatrix} S_i & 0 \\ 0 & S_j \end{pmatrix} \begin{pmatrix} 0 & B_{ij} \\ B_{ji} & 0 \end{pmatrix}.
\end{aligned}$$

So,

$$\begin{pmatrix} S_i & 0 \\ 0 & S_j \end{pmatrix}^{-1} M_{ij} \begin{pmatrix} S_i & 0 \\ 0 & S_j \end{pmatrix} = \begin{pmatrix} 0 & B_{ij} \\ B_{ji} & 0 \end{pmatrix}.$$

Since $\mathcal{A}$ and $\mathcal{B}$ are both normal in the coupled sense, M_{ij} and $\begin{pmatrix} 0 & B_{ij} \\ B_{ji} & 0 \end{pmatrix}$ are both normal. Set $S = \begin{pmatrix} S_i & 0 \\ 0 & S_j \end{pmatrix}$. Proposition 6.1 tells us that SS^* commutes with M_{ij}. Hence,

$$\begin{pmatrix} S_iS_i^* & 0 \\ 0 & S_jS_j^* \end{pmatrix} \begin{pmatrix} 0 & A_{ij} \\ A_{ji} & 0 \end{pmatrix} = \begin{pmatrix} 0 & A_{ij} \\ A_{ji} & 0 \end{pmatrix} \begin{pmatrix} S_iS_i^* & 0 \\ 0 & S_jS_j^* \end{pmatrix},$$

and $S_iS_i^*A_{ij} = A_{ij}S_jS_j^*$. Use the fact that S^*S commutes with $\begin{pmatrix} 0 & B_{ij} \\ B_{ji} & 0 \end{pmatrix}$ to show $S_i^*S_iB_{ij} = B_{ij}S_j^*S_j$ for all i, j.

For part 3, assume S_i and S_j are nonsingular. Let $\mathbf{v} \in \mathcal{U}_j(\alpha)$. By part 2, $S_iS_i^*(A_{ij}\mathbf{v}) = A_{ij}(S_jS_j^*\mathbf{v}) = \alpha(A_{ij}\mathbf{v})$. This shows $A_{ij}(\mathcal{U}_j(\alpha)) \subseteq \mathcal{U}_i(\alpha)$. If A_{ij} is nonsingular, $dim(A_{ij}(\mathcal{U}_j(\alpha))) = dim(\mathcal{U}_j(\alpha))$, so $dim(\mathcal{U}_j(\alpha)) \leq dim(\mathcal{U}_i(\alpha))$. Since $\mathcal{A}$ is coupled normal, A_{ji} is also nonsingular, giving the reverse inequality, so $\dim(\mathcal{U}_j(\alpha)) = \dim(\mathcal{U}_i(\alpha))$. The corresponding facts for $\mathcal{B}$ come from the same argument, using $S_i^*S_iB_{ij} = B_{ij}S_j^*S_j$.

For part 4, assume v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{A})$. Proposition 5.1, together with part 3, gives $\dim(\mathcal{U}_i(\alpha)) = \dim(\mathcal{U}_i(\alpha))$. The same argument applies when v_i and v_j are strongly connected in $\mathcal{D}(\mathcal{B})$

Part 5 comes from the fact that $S_j^*S_j$ and $S_jS_j^*$ have the same nonzero eigenvalues with the same multiplicities.

For part 6, suppose v_i and v_j are connected in $\mathcal{G}(\mathcal{A}, \mathcal{B})$. Then there is a sequence of vertices, $v_i = v_{i_1}, v_{i_2}, v_{i_3}, \ldots, v_{i_{p-1}}, v_{i_p} = v_j$, such that $\{v_{i_k}, v_{i_{k+1}}\}$ is an edge of $\mathcal{G}(\mathcal{A}, \mathcal{B})$ for $k = 1, \ldots, p-1$. This means v_{i_k} and $v_{i_{k+1}}$ are either strongly connected in $\mathcal{D}(\mathcal{A})$ or strongly connected in $\mathcal{D}(\mathcal{B})$ (or both). If v_{i_k} and $v_{i_{k+1}}$ are strongly connected in $\mathcal{D}(\mathcal{A})$, then $dim(\mathcal{U}_{i_k}(\alpha)) = dim(\mathcal{U}_{i_{k+1}}(\alpha))$ by part 4. If v_{i_k} and $v_{i_{k+1}}$ are strongly connected in $\mathcal{D}(\mathcal{B})$, then part 4 tells us $dim(\mathcal{Y}_{i_k}(\alpha)) = dim(\mathcal{Y}_{i_{k+1}}(\alpha))$. But, since α is nonzero, $\mathcal{U}_i(\alpha)$ and $\mathcal{Y}_i(\alpha)$ have the same dimension. So, in either case, $dim(\mathcal{U}_{i_k}(\alpha)) = dim(\mathcal{U}_{i_{k+1}}(\alpha))$ for $1 \leq k \leq p-1$, and hence $dim(\mathcal{U}_i(\alpha)) = dim(\mathcal{U}_j(\alpha))$. □

With these preliminaries completed, we state and prove a version of Schur's Lemma for $\mathcal{A}$, $\mathcal{B}$ that are normal in the coupled sense. The three cases correspond to the three types of coupled reducibility.

Theorem 6.2. Let $\mathcal{A} = \{A_{ij}\}_{i,j\in\mathcal{I}}$ and $\mathcal{B} = \{B_{ij}\}_{i,j\in\mathcal{I}}$ where A_{ij} is $n_i \times n_j$ and B_{ij} is $m_i \times m_j$. Assume $\mathcal{A}$ and $\mathcal{B}$ are normal in the coupled sense. Suppose S_i is $n_i \times m_i$ and $A_{ij}S_j = S_iB_{ij}$ for all i, j.

1. If $\mathcal{A}$ and $\mathcal{B}$ are both irreducible in the coupled sense, then either $S_i = 0$ for all i, or there is a scalar α such that every S_i is α times a unitary matrix; i.e., $S_i = \alpha U_i$, where U_i is unitary. In the latter case, $m_i = n_i$ for all i. Furthermore, if $\mathcal{A} = \mathcal{B}$, then there is a scalar β such that $S_i = \beta I_{n_i}$ for all i.

2. If neither $\mathcal{A}$ nor $\mathcal{B}$ is properly reducible in the coupled sense, then, for each i, either $S_i = 0$ or S_i is a scalar multiple of a unitary matrix. In the latter case, $m_i = n_i$. Furthermore, if $\mathcal{A} = \mathcal{B}$, then every S_i is a scalar matrix.

3. Suppose neither $\mathcal{A}$ nor $\mathcal{B}$ is strongly reducible in the coupled sense. Assume also that $n_i = n$ and $m_i = m$ for all $i \in \mathcal{I}$, and that the graph $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is connected. Then either $S_i = 0$ for all i, or there is a scalar α such that each S_i is a α times a unitary matrix; i.e., $S_i = \alpha U_i$, where U_i is unitary. In the latter case we must have $m = n$. Furthermore, if $\mathcal{A} = \mathcal{B}$, then there is some scalar β such that $S_i = \beta I_n$ for all i.

Proof. The proofs are similar to those of Theorems 4.2 and 5.1.

Suppose $\mathcal{A}$ and $\mathcal{B}$ are both irreducible in the coupled sense. Part 1 of Theorem 4.2 tells us that, either $S_i = 0$ for all i, or S_i is nonsingular for all i. In the latter case we must have $m_i = n_i$ for all i. Suppose S_i is nonsingular for all i. Fix p and let λ be an eigenvalue of $S_pS_p^*$. Proposition 6.3 gives $A_{ij}(\mathcal{U}_j(\lambda)) \subseteq \mathcal{U}_i(\lambda)$ for all i, j. Since $\mathcal{A}$ is irreducible in the coupled sense, either all of the subspaces $\mathcal{U}_i(\lambda)$ are zero, or $\mathcal{U}_i(\lambda) = \mathcal{V}_i$ for all $i \in \mathcal{I}$. Since λ is an eigenvalue of $S_pS_p^*$, the space $\mathcal{U}_p(\lambda)$ is nonzero. Therefore, $\mathcal{U}_i(\lambda) = \mathcal{V}_i$ for all i, and $S_iS_i^* = \lambda I_{n_i}$ for all i. Since $S_iS_i^*$ is positive definite, λ is a positive real number and $U_i = \frac{1}{\sqrt{\lambda}}S_i$ is a unitary matrix.

For the second version, assume neither $\mathcal{A}$ nor $\mathcal{B}$ is properly reducible in the coupled sense. From part 2 of Theorem 4.2, we know that, for each i, either $S_i = 0$ or S_i is nonsingular. If S_i is nonsingular we must have $m_i = n_i$. Suppose S_p is nonsingular for some p. Let λ_p be an eigenvalue of $S_pS_p^*$. Since S_p is nonsingular, $\lambda_p \neq 0$. Let $\mathcal{N}$ denote the set of all q such that S_q is nonsingular. Consider the statement

$$A_{ij}(\mathcal{U}_j(\lambda_p)) \subseteq \mathcal{U}_i(\lambda_p). \tag{6}$$

If i, j are both in $\mathcal{N}$, then S_i, S_j are both nonsingular and Proposition 6.3 tells us (6) holds. If $j \notin \mathcal{N}$, then $S_j = 0$, and hence, since λ_p is nonzero, $\mathcal{U}_j(\lambda_p) = \{0\}$, so (6) holds. Finally, if $i \notin \mathcal{N}$ but $j \in \mathcal{N}$, then $S_i = 0$ and S_j is nonsingular. In this case, $A_{ij}S_j = S_iB_{ij}$ tells us $A_{ij} = 0$, and (6) holds. Hence, (6) holds for all i, j. Since $\mathcal{A}$ is not strongly reducible in the coupled sense, there are only two possibilities for each $\mathcal{U}_i(\lambda)$: it is either zero or the whole space $\mathcal{V}_i$. Since λ_p is an eigenvalue of $S_pS_p^*$, we know $\mathcal{U}_p(\lambda_p)$ is nonzero; therefore it must be the whole space and $S_pS_p^* = \lambda_p I_{n_p}$. Since $S_pS_p^*$ is positive definite, λ_p is a positive real number and $U_p = \frac{1}{\sqrt{\lambda_p}}S_p$ is a unitary matrix.

Finally, consider the third version, where we assume neither $\mathcal{A}$ nor $\mathcal{B}$ is strongly reducible in the coupled sense and $\mathcal{G}(\mathcal{A}, \mathcal{B})$ is connected. From Theorem 5.2, either $S_i = 0$ for all i, or S_i is nonsingular for all i. In the latter case, $m = n$.

Suppose S_i is nonsingular for all i. Fix p and let λ be an eigenvalue of $S_pS_p^*$. Since S_p is nonsingular, $\lambda \neq 0$. From Proposition 6.3, we have $A_{ij}(\mathcal{U}_j(\lambda)) \subseteq \mathcal{U}_i(\lambda)$, for all i, j, and the subspaces $\mathcal{U}_i(\lambda)$ all have the same dimension. Let f be the dimension of these subspaces. Since λ is an eigenvalue of $S_pS_p^*$, the eigenspace $\mathcal{U}_p(\lambda)$ is nonzero. Hence, $f > 0$. Since $\mathcal{A}$ is not strongly reducible in the coupled sense we must have $f = n$. Therefore $S_iS_i^* = \lambda I_n$ for all i, the number λ must be a positive real number and $U_i = \frac{1}{\sqrt{\lambda}}S_i$ is a unitary matrix. $\square$

7 Appendix

We construct examples to establish the claims made in Section 3.

Let $\mathcal{I}$ be the index set; let $\{n_i\}_{i\in\mathcal{I}}$ be a family of positive integers. If $n_i = 1$, set $N_i = (0)$. If $n_i \geq 2$, let N_i be the $n_i \times n_i$ matrix with a 1 in each superdiagonal entry and zeroes elsewhere. This is the standard nilpotent matrix used in the blocks of the Jordan canonical form. For any $\mathbf{x} \in \mathbb{F}^{n_i}$,

$$N_i\mathbf{x} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_i-1} \\ x_{n_i} \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_{n_i} \\ 0 \end{pmatrix}.$$

Multiplying $\mathbf{x}$ on the left by N_i moves the coordinates up one position and puts a 0 in the last entry. Let $\mathbf{e}_j^i$ denote the vector with n_i coordinates that has a 1 in entry j and zeroes in all other positions. Thus, $\mathbf{e}_1^i, \ldots, \mathbf{e}_{n_i}^i$ are the unit coordinate vectors for $\mathbb{F}^{n_i}$. Then $N_i \mathbf{e}_j^i = \mathbf{e}_{j-1}^i$. Henceforth, we omit the superscript i on $\mathbf{e}_j$, as the number of coordinates will be clear from the context. For example, if we write $A_{ij}\mathbf{v}$, then it is understood that $\mathbf{v}$ has n_j coordinates.

Here is the key fact used in the examples.

Proposition 7.1. For $n \geq 2$, let N be the $n \times n$ matrix with a 1 in each superdiagonal entry and zeroes elsewhere. Suppose $\mathcal{U}$ is a nonzero, proper invariant subspace of N. Then $\mathbf{e}_1 \in \mathcal{U}$ and $\mathbf{e}_n \notin \mathcal{U}$.

Proof. Let $\mathbf{x}$ be a nonzero vector in $\mathcal{U}$, and let x_k be the last nonzero coordinate of $\mathbf{x}$, i.e., $x_{k+1} = \cdots = x_n = 0$. Then $N^{k-1}\mathbf{x} = x_k \mathbf{e}_1$, so $\mathbf{e}_1 \in \mathcal{U}$.

For the second part, note that $N^{n-1}\mathbf{e}_n, N^{n-2}\mathbf{e}_n, \ldots, N\mathbf{e}_n, \mathbf{e}_n$ are the unit coordinate vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Hence, if $\mathbf{e}_n \in \mathcal{U}$, then $\mathcal{U}$ is the whole space $\mathcal{V}$. Since $\mathcal{U}$ is a proper subspace of $\mathcal{V}$, the vector $\mathbf{e}_n$ cannot be in $\mathcal{U}$. $\square$

Remark 7.1. Let $\mathcal{Y}_j$ be the j-dimensional subspace spanned by $\mathbf{e}_1, \ldots, \mathbf{e}_j$, i.e., the set of all vectors with zeroes in the last $n-j$ entries. A similar argument shows that the nonzero invariant subspaces of N are the subspaces $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$.

We now construct some examples.

Example 7.1. Assume $|\mathcal{I}| \geq 2$ and that $n_p \geq 2$ for some $p \in \mathcal{I}$. Define $\mathcal{A}$ as follows.

1. $A_{ii} = N_i$ for all $i \in \mathcal{I}$.

2. If $j \neq p$, set $A_{pj} = 0$.

3. If $i \neq p$ let A_{ip} be any matrix which has $\mathbf{e}_{n_i}$ in the first column.

4. If $i \neq p$, and $j \neq p$, and $i \neq j$, then A_{ij} can be any $n_i \times n_j$ matrix.

Set $\mathcal{U}_i = \mathcal{V}_i$ for $i \neq p$, and let $\mathcal{U}_p$ be the line spanned by $\mathbf{e}_1$. Since $n_p \geq 2$, the subspace $\mathcal{U}_p$ is a nonzero, proper subspace of $\mathcal{V}_p$. One can easily check that the subspaces $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$ properly reduce $\mathcal{A}$.

We now show $\mathcal{A}$ is not strongly reducible in the coupled sense. Suppose there were nonzero, proper subspaces $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$ that reduced $\mathcal{A}$. (Note we must then have $n_i \geq 2$ for all i.) Each $\mathcal{U}_i$ is a nonzero, proper invariant subspace of N_i, so $\mathbf{e}_1 \in \mathcal{U}_i$ and $\mathbf{e}_{n_i} \notin \mathcal{U}_i$. Choose $i \neq p$. Then A_{ip} has $\mathbf{e}_{n_i}$ in its first column, so $A_{ip}\mathbf{e}_1 = \mathbf{e}_{n_i}$. But $\mathbf{e}_1 \in \mathcal{U}_p$ and $\mathbf{e}_{n_i} \notin \mathcal{U}_i$, so $A_{ip}(\mathcal{U}_p) \not\subseteq \mathcal{U}_i$. Hence, we have a contradiction, and $\mathcal{A}$ is not strongly reducible in the coupled sense.

Example 7.1 shows $\mathcal{A}$ can be properly reducible in the coupled sense without being strongly reducible. Thus, for any field $\mathbb{F}$, when $|\mathcal{I}| \geq 2$ and $n_i \geq 2$ for at least one i, we have $StrRed(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset PropRed(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$.

The next example shows that if $|\mathcal{I}| \geq 4$, and $n_i \geq 2$ for at least one value of i, we have $PropRed(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}}) \subset Red(\mathbb{F}, \{n_i\}_{i\in\mathcal{I}})$.

Example 7.2. Assume $|\mathcal{I}| \geq 4$ and that $n_p \geq 2$ for some $p \in \mathcal{I}$. Choose any $q \in \mathcal{I}$, with $q \neq p$, and define $\mathcal{A}$ as follows.

1. $A_{ii} = N_i$ for all $i \in \mathcal{I}$.
2. For all i with $i \neq p$ and $i \neq q$, set $A_{ip} = 0$ and $A_{iq} = 0$.
3. For all other choices of i, j with $i \neq j$, let A_{ij} be any matrix with $\mathbf{e}_{n_1}$ in the first column.

We illustrate for $\mathcal{I} = \{1, 2, \ldots, K\}$, with $p = 1$ and $q = 2$.

$$A = \left(\begin{array}{cc|cccc} N_1 & * & * & * & \cdots & * \\ * & N_2 & * & * & \cdots & * \\ \hline 0 & 0 & N_3 & * & \cdots & * \\ 0 & 0 & * & N_4 & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & * & \cdots & N_K \end{array}\right),$$

where each asterisk ($*$) represents an $n_i \times n_j$ matrix with $\mathbf{e}_{n_i}$ in the first column.

Set $\mathcal{U}_p = \mathcal{V}_p$, and $\mathcal{U}_q = \mathcal{V}_q$. For all other values of i, set $\mathcal{U}_i = 0$. One can check that $\mathcal{A}$ is coupled reducible via $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$.

We now show $\mathcal{A}$ is not properly reducible. Suppose $\mathcal{A}$ could be properly reduced by subspaces $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$. At least one $\mathcal{U}_i$ must be a nonzero, proper subspace; we first show this holds for at most one value of i. Suppose $\mathcal{U}_i$ and $\mathcal{U}_j$ were both nonzero, proper subspaces, with $i \neq j$. We must then have $n_i \geq 2$ and $n_j \geq 2$. Since $\mathcal{U}_i$ is a nonzero, proper invariant subspace of N_i, and $\mathcal{U}_j$ is a nonzero, proper invariant subspace of N_j, we know $\mathbf{e}_1$ is in both $\mathcal{U}_i$ and $\mathcal{U}_j$, and $\mathbf{e}_{n_i} \notin \mathcal{U}_i$ and $\mathbf{e}_{n_j} \notin \mathcal{U}_j$. If $i = p$ and $j = q$, use the matrix A_{pq}. Since $A_{pq}\mathbf{e}_1 = \mathbf{e}_{n_p}$, we see $A_{pq}(\mathcal{U}_q) \not\subseteq \mathcal{U}_p$. The same argument, using A_{qp}, applies when $i = q$ and $j = p$. Suppose then that at least one of i, j is different from p and q. Without loss of generality, assume $j \notin \{p, q\}$. Then use A_{ij}, which has $\mathbf{e}_{n_i}$ in its first column. So $A_{ij}(\mathcal{U}_j) \not\subseteq \mathcal{U}_i$.

So, at most one $\mathcal{U}_i$ is a proper, nonzero subspace; each of the other subspaces is either the whole space $\mathcal{V}_i$ or the zero subspace. Assume $\mathcal{U}_i$ is the nonzero, proper subspace; note $n_i \geq 2$. We claim we can then choose $j \neq i$ so that A_{ij} has $\mathbf{e}_{n_i}$ in the first column and A_{ji} has $\mathbf{e}_{n_j}$ in the first column. If $i = p$, choose $j = q$, and if $i = q$, choose $j = p$. If $i \neq p$ and $i \neq q$, choose any j which is different from i, p and q. (This is where we use the fact that $|\mathcal{I}| \geq 4$.) The subspace $\mathcal{U}_j$ is either the full space $\mathcal{V}_j$, or it is the zero subspace. If $\mathcal{U}_j = \mathcal{V}_j$, then $A_{ij}(\mathcal{V}_j)$ contains $A_{ij}\mathbf{e}_1 = \mathbf{e}_{n_i}$, which is not in $\mathcal{U}_i$. So $A_{ij}(\mathcal{U}_j) \not\subseteq \mathcal{U}_i$. If $\mathcal{U}_j = \{0\}$, then, since $\mathbf{e}_1 \in \mathcal{U}_i$, we have $A_{ji}\mathbf{e}_1 = \mathbf{e}_{n_j} \in A_{ij}(\mathcal{U}_i)$. So $A_{ji}(\mathcal{U}_i) \not\subseteq \mathcal{U}_j$. Hence, $\mathcal{A}$ is not properly reducible in the coupled sense.

In the example above, we needed $|\mathcal{I}| \geq 4$. What can we say when $K = 2$ or $K = 3$? In these cases, the field $\mathbb{F}$ must be considered. The reason is, that for $\mathcal{A}$ to be properly reducible in the coupled sense, at least one A_{ii} must have a nonzero, proper invariant subspace. If $\mathbb{F}$ is algebraically closed and $n \geq 2$, then any $n \times n$ matrix over $\mathbb{F}$ has an eigenvalue in $\mathbb{F}$, and the line spanned by a corresponding eigenvector is a nonzero, proper invariant subspace. But if $\mathbb{F}$ is not algebraically closed, there may be $n \times n$ matrices over $\mathbb{F}$ which have no proper invariant subspaces. We shall give an example for the real numbers later, but first we show that if $\mathbb{F}$ is an algebraically closed field, then $PropRed(\mathbb{F}, n, 2) = Red(\mathbb{F}, n, 2)$ and $PropRed(\mathbb{F}, n, 3) = Red(\mathbb{F}, n, 3)$ for all $n \geq 2$.

We use the following lemma to deal with the cases $K = 2$ and $K = 3$.

Lemma 7.1. Let $\mathcal{A} = \{A_{ij}\}_{i,j \in \mathcal{I}}$ where A_{ij} is $n_i \times n_j$. Suppose $\mathcal{A}$ is coupled reducible with $\{\mathcal{U}_i\}_{i \in \mathcal{I}}$ satisfying one of the following.

1. $\mathcal{U}_p = \mathcal{V}_p$ for exactly one index value p, and $\mathcal{U}_i = \{0\}$ when $i \neq p$.

2. $\mathcal{U}_p = \{0\}$ for exactly one index value p and $\mathcal{U}_i = \mathcal{V}_i$ when $i \neq p$.

Suppose $\mathcal{W}_p$ is a nonzero, proper invariant subspace of A_{pp}. Then $\mathcal{A}$ is properly reducible by coupled similarity via the subspaces obtained by replacing $\mathcal{U}_p$ by $\mathcal{W}_p$, and leaving the other $\mathcal{U}_i$'s unchanged.

Proof. Since $\mathcal{U}_p$ is the only subspace that is changed, we continue to have $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$ whenever i and j are both different from p. Also, $\mathcal{W}_p$ is chosen to satisfy $A_{pp}(\mathcal{W}_p) \subseteq \mathcal{W}_p$. It remains to consider A_{ip} and A_{pi} for $i \neq p$.

In case 1, we have $\mathcal{U}_i = \{0\}$ for $i \neq p$, so $A_{pi}(\mathcal{U}_i) = \{0\} \subseteq \mathcal{W}_p$. We also have $A_{ip}(\mathcal{U}_p) \subseteq \mathcal{U}_i = \{0\}$. Since $\mathcal{U}_p = \mathcal{V}_p$, we must have $A_{ip}(\mathcal{W}_p) = \{0\} = \mathcal{U}_i$.

In case 2, we have $\mathcal{U}_i = \mathcal{V}_i$ for $i \neq p$, and $\mathcal{U}_p = \{0\}$. So $A_{pi}(\mathcal{U}_i) = \{0\} \subseteq \mathcal{W}_p$. We also have $A_{ip}(\mathcal{W}_p) \subseteq \mathcal{V}_i = \mathcal{U}_i$ for $i \neq p$. □

Now suppose $\mathbb{F}$ is algebraically closed, and $n \geq 2$. Any $n \times n$ matrix over $\mathbb{F}$ has a nonzero, proper invariant subspace. For $K = 2$, Lemma 7.1 immediately tells us that $\mathcal{A}$ is coupled reducible if and only if it is properly reducible, i.e., $PropRed(\mathbb{F}, n, 2) = Red(\mathbb{F}, n, 2)$. For the case $K = 3$, suppose $\mathcal{A}$ is reduced by $\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_3$. If none of the $\mathcal{U}_i$'s is a nonzero proper subspace, then each is either $\mathcal{V}$ or 0, so either two of them are $\mathcal{V}$, with the third being zero, or vice versa, two of them are zero, with the third being $\mathcal{V}$. Lemma 7.1 then tells us $\mathcal{A}$ is properly reducible. Hence, for algebraically closed $\mathbb{F}$ and $n \geq 2$ we have $PropRed(\mathbb{F}, n, 3) = Red(\mathbb{F}, n, 3)$.

If $\mathbb{F}$ is not algebraically closed, then a matrix over $\mathbb{F}$ need not have a proper invariant subspace. Consider the case $\mathbb{F} = \mathbb{R}$, the field of real numbers. Let A be a real $n \times n$ matrix, where $n \geq 2$. The eigenvalues of A are in $\mathbb{C}$, and the non-real eigenvalues occur in conjugate pairs. If λ is a real eigenvalue of A then there is a corresponding real eigenvector, $\mathbf{v}$, and the line spanned by $\mathbf{v}$ is a proper, nonzero invariant subspace of A. For a pair of complex conjugate, non-real eigenvalues, $\lambda, \overline{\lambda}$, there is a corresponding two dimensional invariant subspace.

Consider the following example for $K \geq 2$ and 2×2 real matrices.

Example 7.3. Choose an angle θ with $0 < \theta < \pi$. For $1 \leq i \leq K$, set

$$A_{ii} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

This is the matrix for rotation of the plane $\mathbb{R}^2$ by angle θ. Since no line through the origin is mapped to itself by this rotation, this map has no nonzero, proper invariant subspace. Hence, for any choice of the A_{ij}'s when $i \neq j$, the set $\mathcal{A}$ is not properly reducible in the coupled sense. It is, however, possible to find A_{ij}'s such that $\mathcal{A}$ is reducible in the coupled sense. Choose a positive integer s with $1 \leq s < K$ and set $A_{ij} = 0$ whenever $i > s$ and $j \leq s$. Set $\mathcal{U}_i = \mathbb{R}^2$ for $1 \leq i \leq s$ and $\mathcal{U}_i = \{0\}$ for $s + 1 \leq i \leq K$. It is easy to check that the subspaces $\mathcal{U}_1, \ldots, \mathcal{U}_K$ reduce $\mathcal{A}$. For, when i and j are both less than or equal to s, we have $\mathcal{U}_i = \mathcal{U}_j = \mathbb{R}^2$, and hence $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$. If i and j are both greater than s, then $\mathcal{U}_i = \mathcal{U}_j = \{0\}$, so $A_{ij}(\mathcal{U}_j) \subseteq \mathcal{U}_i$. If $i > r$ and $j \leq s$, then $A_{ij} = 0$; hence $A_{ij}(\mathcal{U}_j) = \{0\} \subseteq \mathcal{U}_i$. Finally, if $i \leq s$ and $j > s$, then $\mathcal{U}_j = \{0\}$ so $A_{ij}(\mathcal{U}_j) = \{0\} \subseteq \mathcal{U}_i$. So $\mathcal{A}$ is reducible in the coupled sense, but not properly reducible.

So for $K \geq 2$, we have $PropRed(\mathbb{R}, 2, K) \subset Red(\mathbb{R}, 2, K)$. From Example 7.2, we already knew this for $K \geq 4$; the new information is that $PropRed(\mathbb{R}, 2, 2) \subset Red(\mathbb{R}, 2, 2)$ and $PropRed(\mathbb{R}, 2, 3) \subset Red(\mathbb{R}, 2, 3)$.

However, for $n \geq 3$, any $n \times n$ real matrix has a nonzero proper invariant subspace. Lemma 7.1 then gives $PropRed(\mathbb{R}, n, 2) = Red(\mathbb{R}, n, 2)$ and $PropRed(\mathbb{R}, n, 3) = Red(\mathbb{R}, n, 3)$ when $n \geq 3$.

Acknowledgements. D. Lahat thanks Dr. Jean-François Cardoso for insightful discussions during her Ph.D., which set the foundations for and motivated this work.

References

[1] T. Adalı, D. Lahat, and C. Jutten. Data fusion: Benefits of fully exploiting diversity. Tutorial presented at *EUSIPCO*, Aug. 2016.

[2] Michael Artin. *Algebra*, 2nd edition, Pearson Prentice Hall, New Jersey, 2011.

[3] Brualdi and Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, 1991.

[4] D. Lahat. A data fusion perspective on source separation. Talk presented at the *Hausdorff School: Low-rank Tensor Techniques in Numerical Analysis and Optimization*, Apr. 2016.

[5] D. Lahat and C. Jutten, *Joint blind source separation of multidimensional components: Model and algorithm*, in *Proc. EUSIPCO*, Lisbon, Portugal, Sep. 2014, pp. 1417–1421.

[6] D. Lahat and C. Jutten, A generalization to Schurs lemma with an application to joint independent subspace analysis, Technical

Report hal-01247899, GIPSA-Lab, Grenoble, France, Dec. 2015, https://hal.archives-ouvertes.fr/hal-01247899.

[7] D. Lahat and C. Jutten. On the uniqueness of coupled matrix block diagonalization in the joint analysis of multiple datasets (talk). In *SIAM Conference on Applied Linear Algebra*, Atlanta, GA, USA, Oct. 2015, http://meetings.siam.org/sess/dsp_talk.cfm?p=72077.

[8] D. Lahat and C. Jutten, Joint independent subspace analysis using second-order statistics, *IEEE Trans. Signal Process.*, 64 (2016), pp. 4891–4904, https://doi.org/10.1109/tsp.2016.

[9] D. Lahat and C. Jutten. Joint independent subspace analysis by coupled block decomposition: Non-identifiable cases. In *Proc. ICASSP*, Calgary, Canada, Apr. 2018.

[10] D. Lahat and C. Jutten, *Joint independent subspace analysis: Uniqueness and identifiability* *IEEE Trans. Signal Process.*, 2018, to appear, DOI 10.1109/TSP.2018.2880714

[11] K. Murota, Y. Kanno, M. Kojima, and S. Kojima, A numerical algorithm for block-diagonal decomposition of matrix *-algebras with application to semidefinite programming, *Japan J. Indust. Appl. Math.*, 27 (2010), pp. 125–160.

[12] I. Schur, Neue Begründung der Theorie der Gruppencharaktere, *Sitzungsberichte der Königlich-Preussischen Akademie der Wissenschaften zu Berlin*, Berlin, Germany, Jan.–Jun. 1905, pp. 406–432, https://archive.org/details/sitzungsberichte1905deutsch.

[13] J. J. Sylvester. Sur l'équation en matrices px=xq. *Comptes Rendus*, 99: 67–71, 1884.