A Matrix-Valued Inner Product for Matrix-Valued Signals and Matrix-Valued Lattices

XIANG-GEN XIA
Fellow, IEEE
Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA

Abstract: A matrix-valued inner product was proposed before to construct orthonormal matrix-valued wavelets for matrix-valued signals. It introduces a weaker orthogonality for matrix-valued signals than the orthogonality of all components in a matrix that is commonly used in orthogonal multiwavelet constructions. With the weaker orthogonality, it is easier to construct orthonormal matrix-valued wavelets. In this paper, we re-study the matrix-valued inner product more from the inner product viewpoint that is more fundamental and propose a new but equivalent norm for matrix-valued signals. We show that although it is not scalar-valued, it maintains most of the scalarvalued inner product properties. We introduce a new linear independence concept for matrix-valued signals and present some related properties. We then present the Gram-Schmidt orthonormalization procedure for a set of linearly independent matrix-valued signals. Finally we define matrix-valued lattices, where the newly introduced Gram-Schmidt orthogonalization might be applied.

Keywords: Matrix-valued inner product, matrix-valued signal space, nondegenerate matrix-valued signals, linearly independent matrix-valued signals, matrix-valued lattices, matrix-valued wavelets

Received: March 25, 2021. Revised: April 19, 2022. Accepted: July 16, 2022. Published: September 9, 2022.

1. Introduction

Matrix-valued (vector-valued) signals are everywhere these days, such as, videos, multi-spectral images, signals from multiarray multisensors, and high dimensional data. For these signals, there are correlations not only over the time but also across their matrix components. How to efficiently represent and process them plays an important and fundamental role in data science.

In [1], [2], a matrix-valued inner product was introduced for matrix-valued (or vector-valued) signals. It leads to a weaker orthogonality (called Orthogonality B) than the component-wise orthogonality (called Orthogonality A) in orthogonal multiwavelet constructions [11], [12]. The weaker orthogonality, i.e., Orthogonality B, provides an easier sufficient condition to construct orthonormal multiwavelets with Orthogonality B than the necessary and sufficient condition [12] to construct orthonormal multiwavelets with Orthogonality A. A connection between multiwavelets and matrix-valued/vector-valued wavelets can be found in [2]. After the works in [1], [2], there have been many studies on matrix-valued/vector-valued wavelets for matrix-valued/vector-valued signals in the literature, see, for example, [3]- [10].

On the other hand, the Orthogonality B induced from the matrix-valued inner product is stronger than the orthogonality, which is called Orthogonality C here, induced from the commonly used scalar-valued inner product for matrices. It is because, as we shall see later, Orthogonality B is basically the orthogonality between all row vectors of two matrix-valued signals, while Orthogonality C is the orthogonality of two long vectors of concatenated row vectors of two matrix-valued signals. It was proved in [2] that Orthogonality B or the matrix-valued inner product is able to completely decorrelate matrix-valued signals not only in time domain but also across the components inside matrix, i.e., it provides a complete Karhunen-Loève expansion for matrix-valued signals, while Orthogonality A or Orthogonality C may not do so. In other words, Orthogonality B induced from the matrix-valued inner product is the proper orthogonality for matrix-valued signals. This also means that the matrix-valued inner product is needed to study the decorrelation of matrix-valued signals and a conventional scalar-valued inner product may not be enough.

Since the main goal in [1], [2] was to construct orthonormal matrix-valued (vector-valued) wavelets, not much about the inner product or the orthogonality itself, which is more fundamental, was studied. In this paper, we study more propertis on the matrix-valued inner product and its induced Orthogonality B for matrix-valued signal space proposed in [1], [2]. We first define a different norm for matrix-valued signals than that defined in [2] and prove that these two norms are equivalent. The norm defined in this paper is consistent with the matrix-valued inner product similar to that for a scalar-valued inner product. We introduce a new linear independence concept for matrix-valued signals and present some related properties. We then present the Gram-Schmidt orthonormalization procedure for a set of linearly independent matrix-valued signals. We finally define matrix-valued lattices, where the newly introduced Gram-Schmidt orthogonalization might be applied. Due to the noncommuntativity of matrix multiplications, these concepts and properties for matrix-valued signals and/or inner product are not straightforward extensions of the conventional ones for scalar-valued signals and/or inner product.

The remainder of this paper is organized as follows. In Section II, we introduce matrix-valued signal space, matrix-valued inner product, and define a new norm for matrix-valued signals. We present some simple properties for the matrix-valued inner product and prove that the new norm proposed in this paper is equivalent to that used in [2]. In Section III, we

$$L^2(a,b;\mathbb{C}^{N\times N}) \triangleq \left\{ \mathbf{f}(t) = \begin{pmatrix} f_{11}(t) & f_{12}(t) & \cdots & f_{1N}(t) \\ f_{21}(t) & f_{22}(t) & \cdots & f_{2N}(t) \\ & & \cdots & \\ f_{N1}(t) & f_{N2}(t) & \cdots & f_{NN}(t) \end{pmatrix} : \; f_{kl} \in L^2(a,b), 1 \le k,l \le N \right\}. \quad (1)$$

first introduce the concepts of degenerate and nondegenerate matrix-valued signals and then introduce the concept of linear independence for matrix-valued signals. The newly introduced linear independence is different from but consistent with the conventional one for vectors. We also present some interesting properties on the linear independence and the orthogonality. We finally present the Gram-Schmidt orthonormalization procedure for a set of linearly independent matrix-valued signals, which has the similar form as the conventional one for vectors but not a straightforward generalization due to the noncommutativity of matrix multiplications and the matrix-valued inner product used in the procedure. In Section IV, we define matrix-valued lattices. In Section V, we conclude this paper.

2. Matrix-valued Signal Space and Matrix-valued Inner Product

We first introduce matrix-valued signal space studied in [1], [2]. Let $\mathbb{C}^{N\times N}$ denote all $N\times N$ matrices of complex-valued entries, and for $-\infty \le a < b \le \infty$, let $L^2(a,b)$ denote all the finite energy signals in the interval (a,b) and $L^2(a,b;\mathbb{C}^{N\times N})$ be defined in (1) at the top of this page. We call $L^2(a,b;\mathbb{C}^{N\times N})$ a matrix-valued signal space and $\mathbf{f}(t) \in L^2(a,b;\mathbb{C}^{N\times N})$, or simply $\mathbf{f} \in L^2(a,b;\mathbb{C}^{N\times N})$, a matrix-valued signal.

For any $A \in \mathbb{C}^{N\times N}$ and $\mathbf{f} \in L^2(a,b;\mathbb{C}^{N\times N})$, the products $A\mathbf{f}, \mathbf{f}A \in L^2(a,b;\mathbb{C}^{N\times N})$. This implies that the matrix-valued signal space $L^2(a,b;\mathbb{C}^{N\times N})$ is defined over $\mathbb{C}^{N\times N}$ and not simply over $\mathbb{C}$. For $\mathbf{f} \in L^2(a,b;\mathbb{C}^{N\times N})$, its integration $\int_a^b \mathbf{f}(t)dt$ is defined by the integrations of its components, i.e., $\int_a^b \mathbf{f}(t)dt = \left(\int_a^b f_{kl}(t)dt\right)$.

Let $\|\cdot\|_M$ denote a matrix norm on $\mathbb{C}^{N\times N}$, for example, the Frobenius norm,

$$\|A\|_M = \|A\|_F = \left(\sum_{k,l=1}^{N} |A_{kl}|^2\right)^{1/2},$$

where $A = (A_{kl})$. For each $\mathbf{f} \in L^2(a,b;\mathbb{C}^{N\times N})$, let $\|\mathbf{f}\|_M$ denote the norm of $\mathbf{f}$ associated with the matrix norm $\|\cdot\|_M$ as

$$\|\mathbf{f}\|_M \triangleq \left\| \int_a^b \mathbf{f}(t)\mathbf{f}^\dagger(t)dt \right\|_M^{1/2}, \quad (2)$$

where † denotes the complex conjugate transpose. Note that the norm $\|\mathbf{f}\|$ of $\mathbf{f}$ defined in [2] has the following form

$$\|\mathbf{f}\| \triangleq \left(\int_a^b \|\mathbf{f}(t)\|_M^2 dt \right)^{1/2}, \quad (3)$$

where $\|\mathbf{f}(t)\|_M$ is the matrix norm of matrix $\mathbf{f}(t)$ for a fixed t. We will show later that the above two norms $\|\mathbf{f}\|$ and $\|\mathbf{f}\|_M$ are equivalent in the sense that there exist two positive constants $C_1 > 0$ and $C_2 > 0$ such that

$$C_1\|\mathbf{f}\| \le \|\mathbf{f}\|_M \le C_2\|\mathbf{f}\|, \text{ for any } \mathbf{f} \in L^2(a,b;\mathbb{C}^{N\times N}). \quad (4)$$

We next define matrix-valued inner product for matrix-valued signals in $L^2(a,b;\mathbb{C}^{N\times N})$. For two matrix-valued signals $\mathbf{f},\mathbf{g} \in L^2(a,b;\mathbb{C}^{N\times N})$, their *matrix-valued inner product* (or simply inner product) $\langle \mathbf{f},\mathbf{g}\rangle$ is defined as the integration of the matrix product $\mathbf{f}(t)\mathbf{g}^\dagger(t)$, i.e.,

$$\langle \mathbf{f},\mathbf{g}\rangle \triangleq \int_a^b \mathbf{f}(t)\mathbf{g}^\dagger(t)dt. \quad (5)$$

With the definition (5), most properties of the conventional scalar-valued inner product hold for the above matrix-valued inner product. For instance, the following properties of the matrix-valued inner product are clear:

(i) $\langle \mathbf{f},\mathbf{g}\rangle = \langle \mathbf{g},\mathbf{f}\rangle^\dagger$.
(ii) $\langle \mathbf{f},\mathbf{f}\rangle = 0$ if and only if $\mathbf{f} = 0$.
(iii) $\|\mathbf{f}\|_M = \|\langle \mathbf{f},\mathbf{f}\rangle\|_M^{1/2}$.
(iv) For any $A,B \in \mathbb{C}^{N\times N}$, $\langle A\mathbf{f},B\mathbf{g}\rangle = A\langle \mathbf{f},\mathbf{g}\rangle B^\dagger$.

Note that Property (iii) may not hold for the norm (3) used in [2].

Two matrix-valued signals $\mathbf{f}$ and $\mathbf{g}$ in $L^2(a,b;\mathbb{C}^{N\times N})$ are called *orthogonal* if $\langle \mathbf{f},\mathbf{g}\rangle = 0$. A set of matrix-valued signals is called an *orthogonal set* if any two distinct matrix-valued signals in the set are orthogonal. A sequence $\mathbf{\Phi}_k(t) \in L^2(a,b;\mathbb{C}^{N\times N})$, $k \in \mathbb{Z}$, is called an *orthonormal set* in $L^2(a,b;\mathbb{C}^{N\times N})$ if

$$\langle \mathbf{\Phi}_k,\mathbf{\Phi}_l\rangle = \delta(k-l)I_N, \;\; k,l \in \mathbb{Z}, \quad (6)$$

where $\delta(k) = 1$ when $k = 0$ and $\delta(k) = 0$ when $k \ne 0$, I_N is the $N\times N$ identity matrix. Due to (i) above, the orthogonality/orthonormality between $\mathbf{f}$ and $\mathbf{g}$ is commutative, i.e., if $\mathbf{f}$ and $\mathbf{g}$ are orthogonal/orthonormal, then $\mathbf{g}$ and $\mathbf{f}$ are orthogonal/orthonormal too.

A sequence $\mathbf{\Phi}_k(t) \in L^2(a,b;\mathbb{C}^{N\times N})$, $k \in \mathbb{Z}$, is called an *orthonormal basis* for $L^2(a,b;\mathbb{C}^{N\times N})$ if it satisfies (6), and moreover, for any $\mathbf{f} \in L^2(a,b;\mathbb{C}^{N\times N})$ there exists a sequence of $N\times N$ constant matrices F_k, $k \in \mathbb{Z}$, such that

$$\mathbf{f}(t) = \sum_{k\in\mathbb{Z}} F_k\mathbf{\Phi}_k(t), \quad \text{for } t \in [a,b], \quad (7)$$

or simply

$$\mathbf{f} = \sum_{k\in\mathbb{Z}} F_k\mathbf{\Phi}_k,$$

where $F_k = \langle \mathbf{f},\mathbf{\Phi}_k\rangle$, the multiplication $F_k\mathbf{\Phi}_k(t)$ for each fixed t is the $N\times N$ matrix multiplication, and the convergence for the infinite summation is in the sense of the norm $\|\cdot\|_M$ defined by (2) for the matrix-valued signal space. The corresponding Parseval equality is

$$\langle \mathbf{f}, \mathbf{f} \rangle = \sum_{k \in \mathbb{Z}} F_k F_k^{\dagger}. \quad (8)$$

With the norm $\| \cdot \|_M$ in (2), it is clear that for any element $\mathbf{\Phi}_k$ in an orthonormal set in $L^2(a, b; \mathbb{C}^{N \times N})$, we have $\|\mathbf{\Phi}_k\|_M = 1$, which is consistent with the conventional relationship between vector norm and vector inner product. However, this property may not hold for the norm $\| \cdot \|$ in (3) used in [2]. We refer to [2] for the Karhunen-Loève expansion with an orthonormal basis for random processes of matrix-valued signals.

We next show the equivalence (4) of the two norms $\| \cdot \|_M$ in (2) and $\| \cdot \|$ in (3).

Proposition 1: The norms $\| \cdot \|_M$ in (2) and $\| \cdot \|$ in (3) are equivalent.

Proof: It is known that all matrix norms for constant matrices are equivalent. Hence, we show (4) only for the Frobenius norm, i.e., $\| \cdot \|_M = \| \cdot \|_F$. In this case,

$$\begin{aligned}
\|\mathbf{f}\|_M^4 &= \sum_{k,l=1}^{N} \left| \int_a^b \sum_{m=1}^{N} f_{km}(t) f_{lm}^*(t) dt \right|^2 \\
&= \sum_{k=1}^{N} \left| \sum_{m=1}^{N} \int_a^b |f_{km}(t)|^2 dt \right|^2 \\
&\quad + \sum_{k \neq l=1}^{N} \left| \sum_{m=1}^{N} \int_a^b f_{km}(t) f_{lm}^*(t) dt \right|^2 \\
&\geq \sum_{k=1}^{N} \left| \sum_{m=1}^{N} \int_a^b |f_{km}(t)|^2 dt \right|^2 \\
&\geq \frac{1}{N} \left(\int_a^b \sum_{k,l=1}^{N} |f_{kl}(t)|^2 dt \right)^2 \\
&= \frac{1}{N} \|\mathbf{f}\|^4.
\end{aligned}$$

Thus, we have

$$\frac{1}{N^{1/4}} \|\mathbf{f}\| \leq \|\mathbf{f}\|_M. \quad (9)$$

On the other hand,

$$\begin{aligned}
\|\mathbf{f}\|_M^4 &= \sum_{k,l=1}^{N} \left| \int_a^b \sum_{m=1}^{N} f_{km}(t) f_{lm}^*(t) dt \right|^2 \\
&\leq N \sum_{k,l=1}^{N} \sum_{m=1}^{N} \left| \int_a^b f_{km}(t) f_{lm}^*(t) dt \right|^2 \\
&\leq N \sum_{k,l=1}^{N} \sum_{m=1}^{N} \int_a^b |f_{km}(t)|^2 dt \int_a^b |f_{lm}(t)|^2 dt \\
&\leq \frac{N}{2} \sum_{k,l=1}^{N} \sum_{m=1}^{N} \left(\left(\int_a^b |f_{km}(t)|^2 dt \right)^2 \right. \\
&\quad \left. + \left(\int_a^b |f_{lm}(t)|^2 dt \right)^2 \right) \\
&\leq N^2 \sum_{k=1}^{N} \sum_{m=1}^{N} \left(\int_a^b |f_{km}(t)|^2 dt \right)^2 \\
&\leq N^2 \left(\int_a^b \sum_{k,m=1}^{N} |f_{km}(t)|^2 dt \right)^2 \\
&= N^2 \|\mathbf{f}\|^4.
\end{aligned}$$

This shows that

$$\|\mathbf{f}\|_M \leq \sqrt{N} \|\mathbf{f}\|. \quad (10)$$

Combining (9) and (10), the equivalence (4) with $C_1 = N^{-1/4}$ and $C_2 = N^{1/2}$ between the norms $\| \cdot \|_M$ in (2) and $\| \cdot \|$ in (3) is proved. **q.e.d.**

Due to the equivalence of the norm $\| \cdot \|_M$ proposed in this paper and the norm $\| \cdot \|$ used in [2], all the results on orthonormal matrix-valued wavelets obtained in [2] hold, when the norm $\| \cdot \|_M$ for matrix-valued signals in this paper is used.

As a remark, the conventional inner product for two matrices A and B is the scalar-valued inner product $tr(AB^{\dagger})$ where tr stands for the matrix trace. It is not hard to see that with this scalar-valued inner product, the orthogonality between two matrix-valued signals, which is called *Orthogonality C*, is the orthogonality of two long vectors of concatenated row vectors of two matrix-valued signals. As mentioned in Introduction, and it is also not hard to see from the above definition, the orthogonality (6) induced from the matrix-valued inner product in this paper for two matrix-valued signals is the orthogonality between any row vectors including the row vectors inside a matrix of the two matrix-valued signals, which is named *Orthogonality B* in [2]. Clearly Orthogonality B is stronger than Orthogonality C, while it is weaker than the component-wise orthogonality called *Orthogonality A* in [2], commonly used in multiwavelets [11], [12].

With Orthogonality A, a necessary and sufficient condition to construct orthonormal multiwavelets was given in [12] that is not easy to check. However, with Orthogonality B, an easy sufficient condition to construct orthonormal multiwavelets was obtained in [2]. Furthermore, it was shown in [2] that the matrix-valued inner product (5) and its induced Orthogonality B provide a complete decorrelation of matrix-valued signals along time and across matrix components, i.e., a complete Karhunen-Loève expansion for matrix-valued signals can be obtained. This may not be possible for Orthogonality A or Orthogonality C induced from a scalar-valued inner product [13], [14]. In other words, the matrix-valued inner product (5) is fundamental to study matrix-valued signals.

3. Linear Independence and Gram-schmidt Orthonormalization

Let us first introduce degenerate and linearly independent matrix-valued signals, and study their properties.

3.1 Degenerate and Linearly Independent Matrix-valued Signals

A matrix-valued signal $\mathbf{f}$ in $L^2(a,b;\mathbb{C}^{N\times N})$ is called *degenerate signal* if $\langle \mathbf{f},\mathbf{f}\rangle$ does not have full rank, otherwise it is called *nondegenerate signal*. A sequence of matrix-valued signals $\mathbf{f}_k$ in $L^2(a,b;\mathbb{C}^{N\times N})$, $k=1,2,...,K$, are called *linearly independent* if the following condition holds: if

$$\sum_{k=1}^{K} F_k \mathbf{f}_k \triangleq \mathbf{f} \tag{11}$$

for constant matrices $F_k \in \mathbb{C}^{N\times N}$, $k=1,2,...,K$, is degenerate, then the null space of matrix $F_k^\dagger$ includes the null space of matrix $\langle \mathbf{f},\mathbf{f}\rangle$ for every k, $k=1,2,...,K$. Clearly, the above linear independence returns to the conventional one when all the above matrices are diagonal. Furthermore, if $\mathbf{f}=0$ in (11), the above condition implies that all $F_k=0$, $k=1,2,...,K$, since in this case, the null space of $\langle \mathbf{f},\mathbf{f}\rangle$ is the whole space $\mathbb{C}^{N\times N}$. This concides with the condition of the conventional linear independence.

Proposition 2: If matrix-valued signals $\mathbf{f}_k$, $k=1,2,...,K$, are linearly independent, then, all signals $\mathbf{f}_k$, $k=1,2,...,K$, are nondegenerate.

Proof: Without loss of generality, assume $\mathbf{f}_1$ is degenerate. Let $F_1=I_N$ and $F_k=0$ for $k=2,3,...,K$. Then, we have that

$$\sum_{k=1}^{K} F_k \mathbf{f}_k = \mathbf{f}_1$$

is degenerate, while the null space of $F_1^\dagger$ is 0 only and does not include the null space of $\langle \mathbf{f}_1,\mathbf{f}_1\rangle$. In other words, $\mathbf{f}_k$, $k=1,2,...,K$, are not linearly independent. This contradicts the assumption in the proposition and therefore the proposition is proved. **q.e.d.**

As one can see, the above concept of degenerate signal is similar to that of 0 in the conventional linear dependence or independence.

Proposition 3: Let $G_k \in \mathbb{C}^{N\times N}$, $k=1,2,...,K$, be K constant matrices and at least one of them have full rank. If matrix-valued signals $\mathbf{f}_k$, $k=1,2,...,K$, are linearly independent, then $\sum_{k=1}^{K} G_k\mathbf{f}_k$ is nondegenerate.

Proof: Without loss of generality, let us assume G_1 has full rank. If $\sum_{k=1}^{K} G_k\mathbf{f}_k=\mathbf{g}$ is degenerate, then by the linear independence of $\mathbf{f}_k$, $k=1,2,...,K$, the null space of $G_1^\dagger$ cannot only contain 0, which contradicts the assumption that G_1 has full rank. **q.e.d.**

It is clear to see that Proposition 2 is a special case of Proposition 3.

Proposition 4: If matrix-valued signals $\mathbf{f}_k$, $k=1,2,...,K$, are linearly independent, then for any full rank constant matrices $G_k \in \mathbb{C}^{N\times N}$, $k=1,2,...,K$, matrix-valued signals $\mathbf{g}_k \triangleq G_k\mathbf{f}_k$, $k=1,2,...,K$, are also linearly independent.

Proof: For any constant matrices $F_k \in \mathbb{C}^{N\times N}$, $k=1,2,...,K$, if

$$\sum_{k=1}^{K} F_k\mathbf{g}_k = \sum_{k=1}^{K} F_kG_k\mathbf{f}_k = \mathbf{f}$$

is degenerate, then for each k, $1\le k\le K$, the null space of matrix $(F_kG_k)^\dagger = G_k^\dagger F_k^\dagger$ includes the null space of matrix $\langle \mathbf{f},\mathbf{f}\rangle$, since $\mathbf{f}_k$, $k=1,2,...,K$, are linearly independent. Because all matrices G_k, $k=1,2,...,K$, have full rank, for each k, $1\le k\le K$, the null spaces of $F_k^\dagger$ and $G_k^\dagger F_k^\dagger$ are the same, thus, the null space of $F_k^\dagger$ includes the null space of $\langle \mathbf{f},\mathbf{f}\rangle$ as well. This proves the proposition. **q.e.d.**

Similar to the conventional linear dependence of vectors, we have the following result for matrix-valued signals.

Proposition 5: For a matrix-valued signal $\mathbf{f} \in L^2(a,b;\mathbb{C}^{N\times N})$ and two constant matrices $A,B \in \mathbb{C}^{N\times N}$, matrix-valued signals $A\mathbf{f}$ and $B\mathbf{f}$ are linearly dependent.

Proof: If $A\mathbf{f}$ and $B\mathbf{f}$ are linearly independent, then, from Proposition 2 it is easy to see that matrices A and B all have full rank and $\mathbf{f}$ is nondegenerate. Then, we have

$$BA^{-1}A\mathbf{f} - B\mathbf{f} = 0,$$

which contradicts with the assumption of the linear independence of $A\mathbf{f}$ and $B\mathbf{f}$. This proves the proposition. **q.e.d.**

Although it is obvious for the conventional vectors, the result in Proposition 5 for matrix-valued signals may not be so, due to the matrix-valued coefficient multiplications as it can be seen from the above proof. We next consider more general linear combinations of linearly independent matrix-valued signals.

For $1\le p\le K$, let $S_1,...,S_p$ be a partition of the index set $\{1,2,...,K\}$ and each S_i has K_i elements, where $S_{i_1}\cap S_{i_2}=\emptyset$ for $1\le i_1\ne i_2\le p$, $\cup_{i=1}^{p} S_i=\{1,2,...,K\}$, and $1\le K_1,...,K_p\le K$ with $K_1+K_2+\cdots+K_p=K$.

Proposition 6: For each i, $1\le i\le p$, let $G_{k_i}\in\mathbb{C}^{N\times N}$, $k_i\in S_i$, be K_i constant matrices and at least one of them have full rank. If matrix-valued signals $\mathbf{f}_k$, $k=1,2,...,K$, are linearly independent, then the following p matrix-valued signals:

$$\sum_{k_i\in S_i} G_{k_i}\mathbf{f}_{k_i}, \quad \text{for } i=1,2,...,p,$$

are linearly independent.

Proof: Let $F_i\in\mathbb{C}^{N\times N}$, $i=1,2,...,p$, be constant matrices. Assume that

$$\sum_{i=1}^{p} F_i \sum_{k_i\in S_i} G_{k_i}\mathbf{f}_{k_i} = \mathbf{g}$$

is degenerate. Then,

$$\sum_{i=1}^{p}\sum_{k_i\in S_i} F_iG_{k_i}\mathbf{f}_{k_i} = \mathbf{g},$$

and by the linear independence of $\mathbf{f}_k$, $k = 1, 2, ..., K$, we know that the null space of $(F_i G_{k_i})^\dagger$ for every $k_i \in S_i$ and every $i = 1, ..., p$ contains the null space of matrix $\langle \mathbf{g}, \mathbf{g} \rangle$. From the condition in the proposition, without loss of generality, we may assume that $G_{k_{i,1}}$, for some $k_{i,1} \in S_i$, has full rank for $1 \le i \le p$. Thus, the null space of $(F_i G_{k_{i,1}})^\dagger$, or $G_{k_{i,1}}^\dagger F_i^\dagger$, contains the null space of $\langle \mathbf{g}, \mathbf{g} \rangle$ for $1 \le i \le p$. Since $G_{k_{i,1}}$ has full rank, the null space of $F_i^\dagger$ must contain the null space of $\langle \mathbf{g}, \mathbf{g} \rangle$ for $1 \le i \le p$. This proves the proposition. **q.e.d.**

Note that when $p = 1$ in Proposition 6, it returns to Proposition 3, and when $p = K$ in Proposition 6, it returns to Proposition 4.

Proposition 7: If $\mathbf{f}_k$, $k = 1, 2, ..., K$, form an orthonormal set in $L^2(a, b; \mathbb{C}^{N\times N})$, then, they must be linearly independent.

Proof: For constant matrices $F_k \in \mathbb{C}^{N\times N}$, $k = 1, 2, ..., K$, let

$$\sum_{k=1}^{K} F_k \mathbf{f}_k = \mathbf{f}.$$

Then, from the Parseval equality (8), we have

$$\sum_{k=1}^{K} F_k F_k^\dagger = \langle \mathbf{f}, \mathbf{f} \rangle.$$

Assume that for some vector $u \ne 0$, we have $\langle \mathbf{f}, \mathbf{f} \rangle u = 0$ but $F_{k_0}^\dagger u \ne 0$ for some k_0, $1 \le k_0 \le K$. Then,

$$0 < u^\dagger F_{k_0} F_{k_0}^\dagger u \le \sum_{k=1}^{K} u^\dagger F_k F_k^\dagger u = u^\dagger \langle \mathbf{f}, \mathbf{f} \rangle u = 0,$$

which leads to a contradiction. Thus, for each k, $1 \le k \le K$, the null space of $F_k^\dagger$ includes the null space of $\langle \mathbf{f}, \mathbf{f} \rangle$. This proves the linear independence of $\mathbf{f}_k$, $k = 1, 2, ..., K$. **q.e.d.**

Corollary 1: Assume $\mathbf{f}_k$, $k = 1, 2, ..., K$, are nondegenerate matrix-valued signals and form an orthogonal set in $L^2(a, b; \mathbb{C}^{N\times N})$. Then, $\mathbf{g}_k \triangleq \langle \mathbf{f}_k, \mathbf{f}_k \rangle^{-1/2} \mathbf{f}_k$, $k = 1, 2, ..., K$, form an orthonormal set in $L^2(a, b; \mathbb{C}^{N\times N})$, and $\mathbf{f}_k$, $k = 1, 2, ..., K$, are linearly independent.

Proof: Since all matrix-valued signals $\mathbf{f}_k$ are nondegenerate, matrices $\langle \mathbf{f}_k, \mathbf{f}_k \rangle$ all have full rank. From the property (iv) for the matrix-valued inner product and $\{\mathbf{f}_k\}$ is an orthogonal set, for every k, l, $1 \le k, l \le K$,

$$\langle \mathbf{g}_k, \mathbf{g}_l \rangle = \langle \mathbf{f}_k, \mathbf{f}_k \rangle^{-1/2} \langle \mathbf{f}_k, \mathbf{f}_l \rangle \langle \mathbf{f}_l, \mathbf{f}_l \rangle^{-1/2} = \delta(k-l) I_N.$$

Thus, $\mathbf{g}_k$, $k = 1, 2, ..., K$, form an orthonormal set in $L^2(a, b; \mathbb{C}^{N\times N})$.

Then, the linear independence of $\mathbf{f}_k = \langle \mathbf{f}_k, \mathbf{f}_k \rangle^{1/2} \mathbf{g}_k$, $k = 1, 2, ..., K$, immediately follows from Propositions 4 and 7. **q.e.d.**

The result in Corollary 1 is consistent with the conventional one for vectors, i.e., any orthogonal set of nonzero vectors must be linearly independent. However, there is a difference. In the above relationship between orthogonality and linear independence, matrix-valued signals need to be nondegenerate. Note that it is possible that a matrix-valued signal $\mathbf{f}$ in an orthogonal set in $L^2(a, b; \mathbb{C}^{N\times N})$ is degenerate, i.e., $\langle \mathbf{f}, \mathbf{f} \rangle$ may not necessarily have full rank, even though $\mathbf{f} \ne 0$. Thus, a general orthogonal set of matrix-valued signals may not have to be linearly independent. This does not occur for any orthogonal set of nonzero signals when a scalar-valued inner product is used.

3.2 Gram-schmidt Orthonormalization

We are now ready to present the Gram-Schmidt orthonormalization for a finite sequence of linearly independent matrix-valued signals. Let $\mathbf{f}_k \in L^2(a, b; \mathbb{C}^{N\times N})$, $k = 1, 2, ..., K$, be linearly independent. The Gram-Schmidt orthonormalization for this sequence is as follows, which is similar to, but not a straightforward extension of, the conventional one, due to the noncommutativity of matrix multiplications.

Since $\mathbf{f}_k \in L^2(a, b; \mathbb{C}^{N\times N})$, $k = 1, 2, ..., K$, are linearly independent, by Proposition 2, $\mathbf{f}_1$ is nondegenerate, i.e., matrix $\langle \mathbf{f}_1, \mathbf{f}_1 \rangle$ is invertible and positive definite. Let

$$\mathbf{g}_1 = \langle \mathbf{f}_1, \mathbf{f}_1 \rangle^{-1/2} \mathbf{f}_1. \tag{12}$$

Then, we have

$$\begin{aligned} \langle \mathbf{g}_1, \mathbf{g}_1 \rangle &= \langle \mathbf{f}_1, \mathbf{f}_1 \rangle^{-1/2} \int_a^b \mathbf{f}_1(t) \mathbf{f}_1^\dagger(t) dt \langle \mathbf{f}_1, \mathbf{f}_1 \rangle^{-1/2} \\ &= \langle \mathbf{f}_1, \mathbf{f}_1 \rangle^{-1/2} \langle \mathbf{f}_1, \mathbf{f}_1 \rangle \langle \mathbf{f}_1, \mathbf{f}_1 \rangle^{-1/2} = I_N. \end{aligned} \tag{13}$$

Let

$$\hat{\mathbf{g}}_2 = \mathbf{f}_2 - \langle \mathbf{f}_2, \mathbf{g}_1 \rangle \mathbf{g}_1, \tag{14}$$

$$\mathbf{g}_2 = \langle \hat{\mathbf{g}}_2, \hat{\mathbf{g}}_2 \rangle^{-1/2} \hat{\mathbf{g}}_2. \tag{15}$$

For (15) to be vaild, we need to show that $\hat{\mathbf{g}}_2$ in (14) is nondegenerate. In fact, if $\hat{\mathbf{g}}_2$ is degenerate, then, from (14) and (12), we have

$$\begin{aligned} \hat{\mathbf{g}}_2 &= \mathbf{f}_2 - \langle \mathbf{f}_2, \mathbf{g}_1 \rangle \langle \mathbf{f}_1, \mathbf{f}_1 \rangle^{-1/2} \mathbf{f}_1 \\ &= F_1 \mathbf{f}_1 + F_2 \mathbf{f}_2, \end{aligned}$$

where $F_2 = I_N$ and $F_1 = -\langle \mathbf{f}_2, \mathbf{g}_1 \rangle \langle \mathbf{f}_1, \mathbf{f}_1 \rangle^{-1/2}$. Similar to the proof of Proposition 2, this contradicts the assumption that $\mathbf{f}_1$ and $\mathbf{f}_2$ are linearly independent. Therefore, it proves that $\hat{\mathbf{g}}_2$ in (14) is nondegenerate and (15) is well-defined.

Let us then check the orthogonality between $\mathbf{g}_1$ and $\hat{\mathbf{g}}_2$. From (14) and (13), we have

$$\begin{aligned} \langle \hat{\mathbf{g}}_2, \mathbf{g}_1 \rangle &= \langle \mathbf{f}_2, \mathbf{g}_1 \rangle - \langle \mathbf{f}_2, \mathbf{g}_1 \rangle \langle \mathbf{g}_1, \mathbf{g}_1 \rangle \\ &= \langle \mathbf{f}_2, \mathbf{g}_1 \rangle - \langle \mathbf{f}_2, \mathbf{g}_1 \rangle = 0. \end{aligned}$$

From (15) and (13), we have that $\mathbf{g}_1$ and $\mathbf{g}_2$ form an orthonormal set.

Repeat the above process and for a general k, $2 \le k \le K$, we let

$$\hat{\mathbf{g}}_k = \mathbf{f}_k - \sum_{l=1}^{k-1} \langle \mathbf{f}_k, \mathbf{g}_l \rangle \mathbf{g}_l, \tag{16}$$

$$\mathbf{g}_k = \langle \hat{\mathbf{g}}_k, \hat{\mathbf{g}}_k \rangle^{-1/2} \hat{\mathbf{g}}_k. \tag{17}$$

With the same proof as the above $\mathbf{g}_1$ and $\mathbf{g}_2$, we have the following proposition.

Proposition 8: For a linearly independent set of matrix-valued signals $\mathbf{f}_k$, $k = 1, 2, ..., K$, let $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_K$ be constructed in (12) and (14)-(17). Then, $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_K$ form an orthonormal set.

As we can see, although the above Gram-Schmidt orthonormalization procedure for matrix-valued signals is similar to the conventional one for vectors, it is not a straightforward generalization due to 1) the noncommuntativity of matrix mulitpications and 2) the matrix-valued inner product used in the above procedure.

We also want to make a comment on the nondegenerate and linear independence for matrix-valued signals. The condition for nondegenerate matrix-valued signals is a weak condition. Unless the row vectors of functions are linearly dependent in the conventional sense, otherwise, a matrix-valued signal is usually nondegenerate. For a finite set of nondegenerate matrix-valued signals, they usually satisfy the condition for linear independence for matrix-valued signals defined above, i.e., they are usually linearly independent and therefore, they can be made to an orthonormal set by using the above Gram-Schmidt procedure.

Another comment on the linear independence for matrix-valued signals is that the definition in (11) is only for left multiplication of constant matrices F_k to matrix-valued signals $\mathbf{f}_k$. Similar definitions for linear independences of matrix-valued signals with right constant matrix multiplications and/or mixed left and right constant matrix multiplications may be possible. Although what is studied in this paper is for continuous-time matrix-valued signals, it can be easily generalized to discrete-time matrix-valued signals (sequences of finite or infinite length).

4. Matrix-valued Lattices

In this section, based on the matrix-valued signal space with the matrix-valued inner product, we introduce matrix-valued lattices.

We first introduce matrix-valued lattices. For convenience, in what follows we only consider the Frobenius norm for matrices, i.e., $\|\cdot\|_M = \|\cdot\|_F$, and real matrix-valued signal spaces, and also let $\mathcal{R}$ denote the real matrix-valued signal space $\mathcal{R} \triangleq L^2(a,b;\mathbb{R}^{N\times N})$. Let $\mathbb{Z}^{N\times N}$ denote all $N \times N$ matrices of integer entries.

For a finite many linearly independent real matrix-valued signals $\mathbf{f}_k \in \mathcal{R}$, $k = 1,2,...,K$, let $\mathcal{R}^K$ denote the matrix-valued signal space linearly expanded by them, i.e.,

$$\mathcal{R}^K = \left\{ \sum_{k=1}^{K} F_k \mathbf{f}_k : \ F_k \in \mathbb{R}^{N\times N}, \, k = 1,2,...,K \right\}. \quad (18)$$

From what was studied in the previous section, clearly, $\mathbf{f}_k$, $k = 1,2,...,K$, form a basis in $\mathcal{R}^K$. The *matrix-valued lattice* formed by this basis in $\mathcal{R}^K$ is defined as

$$\mathcal{L} = \left\{ \sum_{k=1}^{K} F_k \mathbf{f}_k : \ F_k \in \mathbb{Z}^{N\times N}, \, k = 1,2,...,K \right\}, \quad (19)$$

which is a subset/subgroup of $\mathcal{R}^K$. The basis $\mathbf{f}_k$, $k = 1,2,...,K$, is called a basis for the K dimensional matrix-valued lattice $\mathcal{L}$.

The fundamental region of this lattice $\mathcal{L}$ can be defined similar to the conventional lattice as follows. A set $\mathcal{F} \subset \mathcal{R}^K$ is called a fundamental region, if its translations $\mathbf{x} + \mathcal{F} = \{\mathbf{x}+\mathbf{f} : \ \mathbf{f} \in \mathcal{F}\}$ for $\mathbf{x} \in \mathcal{L}$ form a partition of $\mathcal{R}^K$. Since the basis elements $\mathbf{f}_k$, $k = 1,2,...,K$, are not constant real vectors as in the conventional lattices, it would not be convenient to define the determinant of the lattice. However, with the Gram-Schmidt orthonormalization developed in the previous section, we may define the determinant of the lattice directly as

$$\det(\mathcal{L}) = \prod_{k=1}^{K} \|\hat{\mathbf{f}}_k\|_F, \quad (20)$$

where $\hat{\mathbf{f}}_k$, $k = 1,2,...,K$, are from the following Gram-Schmidt orthogonalization of $\mathbf{f}_k$, $k = 1,2,...,K$, which is from the Gram-Schmidt orthonormalization in the previous section:

$$\hat{\mathbf{f}}_1 = \mathbf{f}_1,$$

$$\hat{\mathbf{f}}_k = \hat{\mathbf{g}}_k = \mathbf{f}_k - \sum_{l=1}^{k-1} \mu_{l,k} \hat{\mathbf{f}}_l, \quad (21)$$

where

$$\mu_{l,k} = \langle \mathbf{f}_k, \hat{\mathbf{f}}_l \rangle \langle \hat{\mathbf{f}}_l, \hat{\mathbf{f}}_l \rangle^{-1}, \ l = 1,2,...,k-1 \text{ and } k = 2,3,...,K. \quad (22)$$

It is clear to see that the spaces linearly spanned by $\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_K\}$ and $\{\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, ..., \hat{\mathbf{f}}_K\}$ are the same, i.e., $\mathcal{R}^K$ in (18), since they can be linearly (over $\mathbb{R}^{N\times N}$) represented by each other similar to the conventional vectors.

From the Gram-Schmidt orthogonalization (21), we have

$$\langle \mathbf{f}_k, \mathbf{f}_k \rangle = \langle \hat{\mathbf{f}}_k, \hat{\mathbf{f}}_k \rangle + \sum_{l=1}^{k-1} \mu_{l,k} \langle \hat{\mathbf{f}}_l, \hat{\mathbf{f}}_l \rangle \mu_{l,k}^{\dagger}, \quad (23)$$

for $k = 1,2,...,K$. Using Property (iii) in Section II, the identity (23) implies

$$\|\mathbf{f}_k\|_F^2 \le \|\hat{\mathbf{f}}_k\|_F^2 + \sum_{l=1}^{k-1} \|\mu_{l,k}\|_F^2 \cdot \|\hat{\mathbf{f}}_l\|_F^2, \quad (24)$$

for $k = 1,2,...,K$. From (23), it is also clear that $\|\mathbf{f}_k\|_F \ge \|\hat{\mathbf{f}}_k\|_F$, $k = 2,...,K$.

As a remark, we know that the conventional Gram-Schmidt orthogonalization plays an important role in the LLL algorithm for the conventional lattice basis reduction [20]. It is, however, not clear how the Gram-Schmidt orthogonalization for matrix-valued signals introduced above can be applied in matrix-valued lattice basis reduction.

5. Conclusion

In this paper, we re-studied the matrix-valued inner product previously proposed to construct orthonormal matrix-valued wavelets for matrix-valued signal analysis [1], [2] where not much on the matrix-valued inner product or its induced Orthogonality B, which is more fundamental, was studied. In order to study more on the matrix-valued inner product and its induced Orthogonality B, we first proposed a new norm for matrix-valued signals, which is more consistent with the matrix-valued inner product than that used in [2], and is similar to that with the conventional scalar-valued inner product. We showed that these two norms are equivalent, which means that with the newly proposed norm, all the results for contructing orthonormal matrix-valued wavelets obtained in [2] still hold. We then proposed the concepts of degenerate and nondegenerate matrix-valued signals and defined the linear independence for matrix-valued signals, which is different from but similar to the conventional linear independence for vectors. We also presented some properties on the linear independence and the orthogonality. We then presented the Gram-Schmidt orthonormalization procedure for a set of linearly independent matrix-valued signals. Although this procedure is similar to the conventional one for vectors, due to the noncommutativity of matrix multiplications and the matrix-valued inner product used in the procedure, it is not a straightforward generalization. We finally defined matrix-valued lattices, where the newly introduced Gram-Schmidt orthogonalization might be applied.

Since it was shown in [2] that the matrix-valued inner product and Orthogonality B provide a complete Karhunen-Loève expansion for matrix-valued signals, which a scalar-valued inner product may not do, it is believed that what was studied in this paper for matrix-valued inner product for matrix-valued signal space will have fundamental applications for high dimensional signal analysis in data science.

As a final note, after this paper was written, it has been found that the matrix-valued signal space with the matrix-valued inner product in this paper is related to Hilbert modules, see, for example, [15]- [19]. Interestingly, it was mentioned in [18] that there does not exist any general notion of "C^*-linear independence" due to the existence of zero-divisors. We believe that the linear independence for matrix-valued signals introduced in this paper is novel.

References

[1] X.-G. Xia and B. W. Suter, "Vector-valued wavelets and vector filter banks," *IEEE Trans. Signal Processing*, vol. 44, pp. 508-518, Mar. 1996.

[2] X.-G. Xia, "Orthonormal matrix valued wavelets and matrix Karhunen-Loève expansion," In *Wavelets, Multiwavelets, and Their Applications*, Contemporary Math., vol. 216, pp. 159-175, 1998. Also available at https://www.eecis.udel.edu/˜ xxia/matrix_wavelets.pdf

[3] A. T. Walden and A. Serroukh, "Wavelet analysis of matrix-valued time-series,"*Proc. R. Soc. Lond. A*, vol. 458, pp. 157-179, 2002.

[4] G. Bhatt, B. D. Johnson, and E. Weber, "Orthogonal wavelet frames and vector-valued wavelet transforms," *Applied and Computational Harmonic Analysis*, vol. 23, no. 2, pp. 215234, 2007.

[5] Q. Chen and Z. Shi, "Construction and properties of orthogonal matrix-valued wavelets and wavelet packets," *Chaos, Solutions, and Fractals*, vol. 37, pp. 75-86, 2008.

[6] L. Cui, B. Zhai, and T. Zhang, "Existence and design of biorthogonal matrix-valued wavelets," *Nonlinear Analysis: Real World Applications*, vol. 10, pp. 2679-2687, 2009.

[7] A. S. Antoln and R. A. Zalik, "Matrix-valued wavelets and multiresolution analysis," *J. Applied Functional Analysis*, vol. 7, no. 1-2, pp. 13-25, 2012.

[8] P. Ginzberg and A. T. Walden, "Matrix-valued and quaternion wavelets,"*IEEE Trans. Signal Processing*, vol. 61, pp. 1357-1367, Mar. 2013.

[9] M. Mittal, P. Manchanda, and A. H. Siddiqi, "Wavelets associated vector-valued nonuniform multiresolution analysis,"*Applicable Analysis*, vol. 93, no. 1, pp. 84-104, 2014.

[10] L. Cui, N. Zhu, Y. Wang, J. Sun, and Y. Cen, "Existence of matrix-valued multiresolution analysis-based matrix-valued tight wavelet frames," *Numerical Functional Analysis and Optimization*, vol. 37, no. 9, pp. 1089-1106, 2016.

[11] G. Strang and V. Strela, "Orthogonal multiwavelets with vanishing moments," *J. Optical Eng.*, vol. 33, pp. 2104-2107, Jul. 1994.

[12] G. Plonka, "Necessary and sufficient conditions for orthonormality of scaling vectors," In: G. Nürnberger, J. W. Schmidt, and G. Walz (eds.), *Multivariate Approximation and Splines*, ISNM International Series of Numerical Mathematics, vol 125. Birkhuser, Basel, 1997. https://doi.org/10.1007/978-3-0348-8871-4_17

[13] E. J. Kelly and W. L. Root, "A representation of vector-valued random processes," Group Rept. 55-21, revised, MIT, Lincoln Laboratory, Apr. 1960.

[14] H. Van Trees, *Detection, Estimation, and Modulation Theory I*, Wiley, 1968.

[15] I. Kaplansky, "Modules over operator algebras," *Amer. J. Math.*, vol. 75, pp. 839-853, 1953.

[16] A. Bultheel, "Inequalities in Hilbert modules of matrix-valued functions," *Proc. Amer. Math. Soc.*, vol. 85, no. 3, pp. 369-372, July 1982.

[17] E. C. Lance, "Hilbert C*-modules, A toolkit for operator algebraists," *London Math. Soc. Note Ser.*, vol. 210, Cambridge Univ. Press, 1995.

[18] M. Frank and D. R. Larson, "Modular frames for Hilbert C*-modules and symmetric approximation of frames," arXiv:math/0010115v1 [math.OA], Oct. 2000.

[19] M. Frank and D. R. Larson, "Frames in Hilbert C*-modules and C*-algebras," *J. Operator Theory*, vol. 48, no. 2, pp. 273-314, Fall 2002.

[20] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, no. 4, pp. 515-534, 1982.