Orthonormal Matrix Valued Wavelets and Matrix Karhunen-Loève Expansion

Xiang-Gen Xia

Abstract. In this paper, we study orthonormal matrix valued wavelets for analyzing matrix (vector) valued signals based on matrix multiresolution analysis. We present a simple sufficient condition on the matrix filter $\mathbf{H}(\omega)$ that leads to orthonormal matrix valued wavelets. The sufficient condition is analogous to the one given by Mallat for scalar valued wavelets. The components at each column of matrix valued wavelets form multiwavelets for a scalar valued signal, where the orthonormality induced from the orthonormal matrix valued wavelets is weaker than the one in the current literature on orthonormal multiwavelets. With the new orthonormality, one is able to construct orthonormal matrix valued wavelets similar to the conventional multiresolution analysis based orthonormal wavelets. Moreover, we show that the new orthonormality provides a complete Karhunen-Loève decomposition for matrix valued signals.

1. Introduction

While wavelets and multiwavelets have been extensively studied lately for a scalar-valued signals, see for example [**1**]-[**17**], there are only a few researches, [**1**], on matrix (vector) valued wavelets for matrix (vector) valued signals. In practice, it is however often to encounter matrix (vector) valued signals, such as video images, multi-spectral images and color images. A significant difference between matrix (vector) valued signals and scalar valued signals is that there are correlations for a matrix (vector) valued signal not only in the time domain but also between its components (or the spatial domain) at a fixed time while there is correlation for a scalar valued signal only in the time domain. The aim of the construction of orthonormal matrix valued wavelets is to decorrelate a matrix (vector) valued signal in both the time and the spatial domains. As a side result, the components at each column of orthonormal matrix valued wavelets also form multiwavelets for scalar valued signals. We will see later that the orthonormality for the multiwavelets generated from orthonormal matrix valued wavelets is weaker than the orthonormality in current literature on orthonormal multiwavelets, [**4**]-[**15**]. In [**1**], orthonormal matrix (vector) multiresolution analysis was introduced for the purpose of constructing

---

1991 *Mathematics Subject Classification.* Primary 41A58, 46E40, 94A12; Secondary 46C50, 45B05, 46N30.

This work was supported in part by an initiative grant from the Department of Electrical Engineering, University of Delaware, the Air Force Office of Scientific Research (AFOSR) under Grant No. F49620-97-1-0253, and the National Science Foundation CAREER Program under Grant MIP-9703377.

orthonormal matrix valued wavelets. However, the theory in [**1**] is not complete in the continuous time case in the sense that there is not a simple sufficient condition on the matrix quadrature mirror filter (MQMF) $\mathbf{H}(\omega)$ that leads to orthonormal matrix valued wavelets.

In this paper, we first re-introduce matrix valued signal spaces and matrix valued multiresolution analysis studied in [**1**]. We then present a simple sufficient condition on the MQMF $\mathbf{H}(\omega)$ for constructing orthonormal matrix valued wavelets, which basically proves the conjecture proposed in [**1**]. A connection between orthonormal matrix valued wavelets and orthonormal multiwavelets in the current literature is studied. It can be seen that the orthonormality for the multiwavelets induced from the orthonormality of orthonormal matrix valued wavelets is weaker than the orthonormality for multiwavelets in the recent literature in the continuous time waveform case, see for example [**4**]-[**15**], while they are the same in the discrete time filterbank case. The weaker orthonormality in the continuous time case provides a weaker sufficient conditon for constructing multiwavelets with this weaker orthonormality.

In the second part of this paper, we show that the orthonormality studied in this paper for matrix valued signals gives a complete Karhunen-Loève decomposition for matrix valued signals, i.e., this orthonormality provides a complete decorrelation for a matrix valued signals in both the time and the spatial domains.

2. Matrix Valued Signal Space and Multiresolution Analysis

For convenience, we only study $N\times N$ matrix valued signals and wavelets. We introduce some notations first.

2.1. Matrix Valued Signal Space. Let

$$\mathbf{C}^{N\times N}=\{A:\ A \text{ is an } N\times N \text{ matrix with entries in the complex plane } \mathbf{C}\},$$

and

$$L^2(a,b;\mathbf{C}^{N\times N})\stackrel{\Delta}{=}\left\{\mathbf{f}(t)=(f_{k,l}(t))_{N\times N}:\ f_{k,l}(t)\in L^2(a,b), 1\leq k,l\leq N\right\}.$$

The signal space $L^2(a,b;\mathbf{C}^{N\times N})$ is called a matrix valued signal space. When $a=-\infty$ and $b=\infty$, $L^2(a,b;\mathbf{C}^{N\times N})$ is also denoted by $L^2(\mathbf{R},\mathbf{C}^{N\times N})$.

For any $A\in\mathbf{C}^{N\times N}$ and $\mathbf{f}\in L^2(a,b;\mathbf{C}^{N\times N})$, the products

$$A\mathbf{f},\ \mathbf{f}A\in L^2(a,b;\mathbf{C}^{N\times N}).$$

This implies that the matrix valued signal space $L^2(a,b;\mathbf{C}^{N\times N})$ is defined over $\mathbf{C}^{N\times N}$ and not simply over $\mathbf{C}$.

Let $\|\cdot\|_M$ denote a matrix norm on $\mathbf{C}^{N\times N}$. For each $\mathbf{f}\in L^2(a,b;\mathbf{C}^{N\times N})$, $\|\mathbf{f}\|$ denotes the norm of $\mathbf{f}$ associated with the matrix norm $\|\cdot\|_M$ as

$$\|\mathbf{f}\|\stackrel{\Delta}{=}\left(\int_a^b\|\mathbf{f}(t)\|_M^2dt\right)^{1/2}. \tag{2.1}$$

For $\mathbf{f}\in L^2(a,b;\mathbf{C}^{N\times N})$, its integration $\int\mathbf{f}(t)dt$ is defined by the integration of its components.

For two matrix valued signals $\mathbf{f},\mathbf{g}\in L^2(a,b;\mathbf{C}^{N\times N})$, $\langle\mathbf{f},\mathbf{g}\rangle$ denotes the integration of the matrix product $\mathbf{f}(t)\mathbf{g}^\dagger(t)$:

$$\langle\mathbf{f},\mathbf{g}\rangle\stackrel{\Delta}{=}\int_{\mathbf{R}}\mathbf{f}(t)\mathbf{g}^\dagger(t)dt, \tag{2.2}$$

where † denotes the conjugate transpose. For convenience, we still call the operation $\langle\, , \rangle$ in (2.2) *inner product* although it is not the inner product in the common sense. With the definition (2.2) it is clear that $\langle \mathbf{f}, \mathbf{g}\rangle = \langle \mathbf{g}, \mathbf{f}\rangle^\dagger$.

A sequence $\Phi_k(t) \in L^2(a, b; \mathbf{C}^{N\times N})$, $k \in \mathbf{Z}$, is called an *orthonormal set* in $L^2(a, b; \mathbf{C}^{N\times N})$ if

$$\langle \Phi_k, \Phi_l\rangle = \delta(k-l) I_N, \quad k, l \in \mathbf{Z}, \tag{2.3}$$

where $\delta(k) = 1$ when $k = 0$ and $\delta(k) = 0$ when $k \neq 0$ and I_N is the $N \times N$ identity matrix. A sequence $\Phi_k(t) \in L^2(a, b; \mathbf{C}^{N\times N})$, $k \in \mathbf{Z}$, is called an *orthonormal basis* for $L^2(a, b; \mathbf{C}^{N\times N})$ if it satisfies (2.3), and moreover, for any $\mathbf{f}(t) \in L^2(a, b; \mathbf{C}^{N\times N})$ there exists a sequence of $N \times N$ constant matrices F_k such that

$$\mathbf{f}(t) = \sum_{k\in\mathbf{Z}} F_k \Phi_k(t), \quad \text{for } t \in [a, b], \tag{2.4}$$

where the multiplication $F_k\Phi_k(t)$ 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 $\|\cdot\|$ defined by (2.1) for the matrix valued signal space.

2.2. Matrix Valued Multiresolution Analysis. We next define matrix valued multiresolution analysis, which is similar to the conventional multiresolution analysis.

A matrix valued multiresolution analysis (MMRA) of $L^2(\mathbf{R}, \mathbf{C}^{N\times N})$ is a nested sequence of closed subspaces $\mathbf{V}_j$, $j \in \mathbf{Z}$, of $L^2(\mathbf{R}, \mathbf{C}^{N\times N})$ such that

(i). $\mathbf{V}_j \subset \mathbf{V}_{j+1}$, $j \in \mathbf{Z}$,
(ii). $\cup_{j\in\mathbf{Z}} \mathbf{V}_j$ is dense in $L^2(\mathbf{R}, \mathbf{C}^{N\times N})$ and $\cap_{j\in\mathbf{Z}} \mathbf{V}_j = \{\mathbf{0}\}$, where $\mathbf{0}$ is the all zero matrix,
(iii). $\mathbf{f}(t) \in \mathbf{V}_j$ if and only if $\mathbf{f}(2t) \in \mathbf{V}_{j+1}$, $j \in \mathbf{Z}$,
(iv). There is a $\Phi \in \mathbf{V}_0$ such that its translations $\Phi(t-k)$, $k \in \mathbf{Z}$, form an *orthonormal* basis for $\mathbf{V}_0$.

The above definition for an MMRA is notationally similar to the one for the conventional multiresolution analysis (MRA). We call $\Phi(t)$ a *matrix valued scaling function* (or simply scaling function) for the MMRA $\{\mathbf{V}_j\}$. Since $\Phi(t) \in \mathbf{V}_0 \subset \mathbf{V}_1$, there exist constant $N \times N$ matrices H_k, $k \in \mathbf{Z}$, such that,

$$\Phi(t) = 2\sum_k H_k \Phi(2t-k). \tag{2.5}$$

Let

$$\mathbf{H}(\omega) = \sum_k H_k e^{-ik\omega}. \tag{2.6}$$

Then,

$$\hat{\Phi}(\omega) = \mathbf{H}(\frac{\omega}{2})\hat{\Phi}(\frac{\omega}{2}) = \mathbf{H}(\frac{\omega}{2})\mathbf{H}(\frac{\omega}{4})\cdots\hat{\Phi}(0), \tag{2.7}$$

where it is assumed that $\hat{\Phi}(\omega)$ is continuous at $\omega = 0$. This assumption is satisfied when $\mathbf{H}(\omega)$ has only finite terms and $\mathbf{H}(0) = I_N$. In this paper, for convenience we assume $\hat{\Phi}(0) = I_N$, which makes an important difference between matrix-valued wavelets and multiwavelets from the matrix scaling equation or refinable equation point of view. By this assumption,

$$\hat{\Phi}(\omega) = \mathbf{H}(\frac{\omega}{2})\mathbf{H}(\frac{\omega}{4})\cdots = \prod_{k=1}^{\infty} \mathbf{H}(\frac{\omega}{2^k}). \tag{2.8}$$

The equation (2.7) implies

$$\mathbf{H}(0) = I_N, \quad \text{or } \sum_k H_k = I_N. \tag{2.9}$$

It is not hard to see that the orthonormality of $\Phi(t-k)$, $k \in \mathbf{Z}$, (or the orthonormality of MMRA $\{\mathbf{V}_j\}$) is equivalent to

$$\sum_k \hat{\Phi}(\omega + 2\pi k)\hat{\Phi}^\dagger(\omega + 2\pi k) = 2\pi I_N, \ \forall \omega \in \mathbf{R}. \tag{2.10}$$

In terms of the filter $\mathbf{H}(\omega)$, the above orthonormality implies

$$\mathbf{H}(\omega)\mathbf{H}^\dagger(\omega) + \mathbf{H}(\omega+\pi)\mathbf{H}^\dagger(\omega+\pi) = I_N, \ \forall \omega \in \mathbf{R}. \tag{2.11}$$

The orthonormality (2.10) is in the continuous time domain for continuous-time waveforms while the one (2.11) is in the discrete time domain for discrete-time filterbanks.

Assume we have the above MMRA and $\mathbf{H}(\omega)$. We now want to construct its corresponding matrix valued wavelets that form an orthonormal basis for the whole matrix valued signal space $L^2(\mathbf{R}, \mathbf{C}^{N\times N})$.

Let $\mathbf{G}(\omega)$ satisfy

$$\mathbf{G}(\omega)\mathbf{H}^\dagger(\omega) + \mathbf{G}(\omega+\pi)\mathbf{H}^\dagger(\omega+\pi) = \mathbf{0}, \ \forall \omega \in \mathbf{R}, \tag{2.12}$$

and

$$\mathbf{G}(\omega)\mathbf{G}^\dagger(\omega) + \mathbf{G}(\omega+\pi)\mathbf{G}^\dagger(\omega+\pi) = I_N, \ \forall \omega \in \mathbf{R}. \tag{2.13}$$

Let

$$\hat{\Psi}(\omega) = \mathbf{G}(\frac{\omega}{2})\hat{\Phi}(\frac{\omega}{2}). \tag{2.14}$$

The following result was proved in [**1**].

THEOREM 2.1. *Let $\Psi(t)$ be the matrix valued function with its Fourier transform defined in (2.14). Then, its translations $\Psi(t-k)$, $k \in \mathbf{Z}$, form an orthonormal basis for $\mathbf{W}_0 \triangleq \mathbf{V}_1 \ominus \mathbf{V}_0$. Thus, $\Psi_{j,k}(t) \triangleq 2^{j/2}\Psi(2^j t - k)$, $j, k \in \mathbf{Z}$, form an orthonormal basis for the matrix valued signal space $L^2(\mathbf{R}, \mathbf{C}^{N\times N})$.*

The matrix filters $\mathbf{H}(\omega)$ and $\mathbf{G}(\omega)$ in (2.11)-(2.13) are called *matrix quadrature mirror filters* (MQMF). Given $\mathbf{H}(\omega)$, $\mathbf{G}(\omega)$ can be constructed by the following method.

Let $\hat{\mathbf{H}}(\omega) = (\mathbf{H}(\omega), \mathbf{H}(\omega+\pi))^\dagger$ and $\hat{\mathbf{G}}(\omega) = (\mathbf{G}(\omega), \mathbf{G}(\omega+\pi))^\dagger$. Then, the orthogonality (2.11)-(2.13) is equivalent to the paraunitariness of the $2N \times 2N$ matrix $(\hat{\mathbf{H}}(\omega), \hat{\mathbf{G}}(\omega))$. Let $\mathbf{H}_j(\omega)$ and $\mathbf{G}_j(\omega)$ for $j = 0, 1$ be the polyphase components of $\mathbf{H}(\omega)$ and $\mathbf{G}(\omega)$, respectively: $\mathbf{F}(\omega) = \mathbf{F}_0(2\omega) + e^{-i\omega}\mathbf{F}_1(2\omega)$, where $\mathbf{F}$ is $\mathbf{H}$ or $\mathbf{G}$. Then, the above paraunitariness is equivalent to the paraunitariness of the matrix $(\tilde{\mathbf{H}}(\omega), \tilde{\mathbf{G}}(\omega))$, where $\tilde{\mathbf{F}}(\omega) = (\mathbf{F}_0(\omega), \mathbf{F}_1(\omega))^\dagger$ for $\mathbf{F} = \mathbf{H}$ or $\mathbf{G}$. Thus, the construction of $\mathbf{G}(\omega)$ in (2.11)-(2.13) is equivalent to the completion of a $2N \times 2N$ paraunitary matrix given its first N orthogonal columns $\tilde{\mathbf{H}}(\omega)$. This completion can be obtained by employing the state-space description, see for example [**20**]-[**22**], where only the orthogonal completion of a constant orthogonal matrix is needed for the corresponding constant realization matrix.

In the next section, we want to construct orthonormal matrix valued scaling functions $\Phi(t)$ from the orthogonal filter $\mathbf{H}(\omega)$ in (2.11).

3. Construction of Matrix Valued Wavelets

It is known that the conventional scaling functions or MRA can be constructed from QMF $H(\omega)$ and necessary and sufficient conditions have been obtained, [**18**]-[**19**]. For matrix valued wavelets, we present the following results. We first present a lemma. In what follows, we are only interested in FIR MQMF $\mathbf{H}(\omega)$, i.e., $\mathbf{H}(\omega)$ is a polynomial matrix of $e^{-i\omega}$.

LEMMA 3.1. *Let $\mathbf{H}(\omega)$ satisfy (2.9) and (2.11). If there exist a constant $C>0$ and an integer K_0 such that for any $\omega \in (-2^K\pi, 2^K\pi)$ and any $K>K_0$,*

$$\|\prod_{l=1}^{K}\mathbf{H}(\frac{\omega}{2^l})\|_M \le C\|\prod_{l=1}^{\infty}\mathbf{H}(\frac{\omega}{2^l})\|_M, \tag{3.1}$$

then, the solution $\Phi(t)$ in the matrix dilation equation (2.5) is a matrix valued scaling function for an MMRA.

Proof. The assumption of the FIR property on $\mathbf{H}(\omega)$ leads to the finiteness of the right hand side of (3.1). To prove Lemma 3.1 we only need to prove the orthonormality of $\Phi(t-k)$, $k \in \mathbf{Z}$. The rest is similar to the conventional MRA theory, see for example [**19**].

For an integer $K>0$, let

$$\mu_K(\omega) = \prod_{l=1}^{K}\mathbf{H}(\frac{\omega}{2^l})\chi_{[-2^K\pi, 2^K\pi]}(\omega).$$

Then,

$$\begin{aligned}
&\int_{\mathbf{R}} \mu_K(\omega)\mu_K^\dagger(\omega)e^{-in\omega}d\omega \\
= &\int_{-2^K\pi}^{2^K\pi} \mathbf{H}(\frac{\omega}{2})\cdots\mathbf{H}(\frac{\omega}{2^K})\mathbf{H}^\dagger(\frac{\omega}{2^K})\cdots\mathbf{H}^\dagger(\frac{\omega}{2})e^{-in\omega}d\omega \\
= &\int_{0}^{2^{K+1}\pi} \mathbf{H}(\frac{\omega}{2})\cdots\mathbf{H}(\frac{\omega}{2^K})\mathbf{H}^\dagger(\frac{\omega}{2^K})\cdots\mathbf{H}^\dagger(\frac{\omega}{2})e^{-in\omega}d\omega \\
= &\int_{0}^{2^K\pi} \prod_{l=1}^{K-1}\mathbf{H}(\frac{\omega}{2^l})\left[\mathbf{H}(\frac{\omega}{2^K})\mathbf{H}^\dagger(\frac{\omega}{2^K}) + \mathbf{H}(\frac{\omega}{2^K}+\pi)\mathbf{H}^\dagger(\frac{\omega}{2^K}+\pi)\right] \\
&\cdot\left(\prod_{l=1}^{K-1}\mathbf{H}(\frac{\omega}{2^l})\right)^\dagger e^{-in\omega}d\omega \\
\text{by } \overset{(2.11)}{=} &\int_{\mathbf{R}} \mu_{K-1}(\omega)\mu_{K-1}^\dagger(\omega)e^{-in\omega}d\omega = \cdots = \int_0^{2\pi} e^{in\omega}d\omega I_N = 2\pi\delta(n)I_N.
\end{aligned}$$

It is clear that $\mu_K(\omega)$ converges to $\Phi(\omega)$ pointwisely in (2.8) since $\mathbf{H}(0)=I_N$ and $\mathbf{H}(\omega)$ is a polynomial matrix of $e^{-i\omega}$. By (3.1),

$$\|\mu_K(\omega)\mu_K^\dagger(\omega) - \hat{\Phi}(\omega)\hat{\Phi}^\dagger(\omega)\|_M \le (C+1)\|\hat{\Phi}(\omega)\hat{\Phi}^\dagger(\omega)\|_M, \ \forall \omega \in \mathbf{R}.$$

By the dominated convergence theorem, we have $\|\mu_K\mu_K^\dagger - \hat{\Phi}\hat{\Phi}^\dagger\| \to 0$ as $K\to\infty$. Therefore,

$$\int \Phi(t)\Phi^*(t-n)dt = \frac{1}{2\pi}\int_{\mathbf{R}} \hat{\Phi}(\omega)\hat{\Phi}^\dagger(\omega)e^{-in\omega}d\omega$$

$$= \frac{1}{2\pi}\lim_{K\to\infty}\int_{\mathbf{R}} \mu_K(\omega)\mu_K^\dagger(\omega)e^{-in\omega}d\omega = \delta(n)I_N.$$

This proves the orthonormality of $\Phi(t-k)$, $k\in\mathbf{Z}$. ♣

We next want to present a sufficient condition on $\mathbf{H}(\omega)$ so that (3.1) is satisfied.

LEMMA 3.2. *Let $\mathbf{H}(\omega)$ be a polynomial matrix of $e^{-i\omega}$ and $\mathbf{H}(0)=I_N$. Then, there exist an integer K_0 and a constant $C>0$ such that*

$$\|\prod_{l=1}^{K}\mathbf{H}(\frac{\omega}{2^l})\|_M \le C\|\prod_{l=1}^{\infty}\mathbf{H}(\frac{\omega}{2^l})\|_M,$$

for $\omega\in(-\pi,\pi)$ and $K>K_0$.

Proof. Since $\mathbf{H}(\omega)$ is a polynomial matrix of $e^{-i\omega}$ and $\hat{\Phi}(0)=I_N$, we have

$$\hat{\Phi}(\omega) = \prod_{k=1}^{\infty}\mathbf{H}(\frac{\omega}{2^k}),$$

and

$$\lim_{\omega\to 0}\|\hat{\Phi}(\omega)-I_N\|_M = 0.$$

Thus, there exists an integer $K_0>0$ such that, for $k>K_0$ and $|\omega|<\pi/2$,

$$\|\hat{\Phi}(\frac{\omega}{2^k}) - I_N\|_M \le \epsilon,$$

and

$$\|\hat{\Phi}^{-1}(\frac{\omega}{2^k})\|_M \le \frac{1}{\epsilon}, \quad \text{i.e., } \|\left(\prod_{l=k+1}^{\infty}\mathbf{H}(\frac{\omega}{2^l})\right)^{-1}\|_M \le \frac{1}{\epsilon},$$

where ϵ is a small positive constant.

Therefore, for $K>K_0$ and $|\omega|<\pi/2$,

$$\|\prod_{l=1}^{K}\mathbf{H}(\frac{\omega}{2^l})\|_M = \|\prod_{l=1}^{\infty}\mathbf{H}(\frac{\omega}{2^l})\left(\prod_{l=K+1}^{\infty}\mathbf{H}(\frac{\omega}{2^l})\right)^{-1}\|_M \le C\|\prod_{l=1}^{\infty}\mathbf{H}(\frac{\omega}{2^l})\|_M,$$

where $C=1/\epsilon$. ♣

LEMMA 3.3. *Let $\mathbf{H}(\omega)$ be a polynomial matrix of $e^{-i\omega}$ and $\mathbf{H}(0)=I_N$. If*

$$\inf_{|\omega|<\pi/2}|\lambda(\omega)| > 0$$

for any eigenvalue function $\lambda(\omega)$ of the polynomial matrix $\mathbf{H}(\omega)$ of variable $e^{-i\omega}$, then, there exists a constant $C>0$ such that, for any $\omega\in(-2^K\pi, 2^K\pi)$,

$$\|\prod_{l=1}^{K}\mathbf{H}(\frac{\omega}{2^l})\|_M \le C\|\prod_{l=1}^{\infty}\mathbf{H}(\frac{\omega}{2^l})\|_M.$$

Proof. For $\omega \in (-2^K\pi, 2^K\pi)$, if $k > K$, then $\omega/2^k \in (-\pi/2, \pi/2)$. By the proof of Lemma 3.2, for $\omega \in (-2^K\pi, 2^K\pi)$,

$$\| \left(\prod_{l=K+K_0+1}^{\infty} \mathbf{H}(\frac{\omega}{2^l}) \right)^{-1} \|_M \leq \frac{1}{\epsilon}.$$

Let us consider the case of $l \in \{K+1, K+2, ..., K+K_0\}$. Let $\delta > 0$ such that $\inf_{|\omega|<\pi/2} |\lambda(\omega)| \geq \delta$ for all eigenvalue functions of the polynomial matrix $\mathbf{H}(\omega)$. Then, $(\lambda(\omega))^{-1}$ is an eigenvalue function of the function matrix $(\mathbf{H}(\omega))^{-1}$ of variable $e^{-i\omega}$ for $|\omega| < \pi/2$. Thus, there exists positive constant C_1, which only depends on δ, such that, for $|\omega| < \pi/2$,

$$\|(\mathbf{H}(\omega))^{-1}\|_M \leq C_1.$$

Therefore, for any $\omega \in (-2^K\pi, 2^K\pi)$,

$$\begin{aligned} \| \prod_{l=1}^{K} \mathbf{H}(\frac{\omega}{2^l}) \|_M &= \| \prod_{l=1}^{\infty} \mathbf{H}(\frac{\omega}{2^l}) \left(\prod_{l=K+K_0+1}^{\infty} \mathbf{H}(\frac{\omega}{2^l}) \right)^{-1} \left(\prod_{l=K+1}^{K+K_0} \mathbf{H}(\frac{\omega}{2^l}) \right)^{-1} \|_M \\ &\leq C \| \prod_{l=1}^{\infty} \mathbf{H}(\frac{\omega}{2^l}) \|_M, \end{aligned}$$

where $C = C_1^{K_0}/\epsilon$. ♣

By combining the above three lemmas, we have proved the following result.

THEOREM 3.4. *Let $\mathbf{H}(\omega)$ be a polynomial matrix of $e^{-i\omega}$ and satisfy (2.9) and (2.11). If*

$$\inf_{|\omega|<\pi/2} |\lambda(\omega)| > 0$$

for any eigenvalue function $\lambda(\omega)$ of the polynomial matrix $\mathbf{H}(\omega)$ of variable $e^{-i\omega}$, then, the solution $\Phi(t)$ in the matrix dilation equation (2.5) is a matrix valued scaling function for an MMRA, and therefore $\Psi_{j,k}(t)$, $j, k \in \mathbf{Z}$, form an orthonormal basis for the matrix valued signal space $L^2(\mathbf{R}, \mathbf{C}^{N\times N})$.

Notice that the above sufficient condition is analogous of the one given by Mallat [**18**]. With the above sufficient condition, it is not hard to construct nontrivial families of orthonormal matrix valued wavelets. The following is an example.

It is not hard to show that, if $\mathbf{H}(\omega) = \frac{1}{2}(I_N + e^{i\omega}\mathbf{E}(2\omega))$ and $\mathbf{E}(\omega)$ is paraunitary, i.e., $\mathbf{E}(\omega)\mathbf{E}^\dagger(\omega) = I_N$, then $\mathbf{G}(\omega) = e^{-i\omega}\mathbf{H}^\dagger(\omega+\pi)$ and $\mathbf{H}(\omega)$ form a pair of MQMF satisfying (2.11)-(2.13). Such property for $\mathbf{H}(\omega)$ is called the sampling property in [**1**]. Let $\mathbf{E}(\omega) = \mathbf{U}(\omega)\mathrm{diag}(e^{-ik_1\omega}, \cdots, e^{-ik_N\omega})\mathbf{U}^\dagger(\omega)$ for $k_j = 0$ or 1, where $\mathbf{U}(\omega)$ is an arbitrary paraunitary polynomial matrix and $\mathbf{U}(0) = I_N$. Then, it is not hard to see that the above $\mathbf{H}(\omega)$ and $\mathbf{G}(\omega)$ satisfy (2.11)-(2.13) and the sufficient condition in Theorem 3.4.

4. Connection to Multiwavelets

Let $(\Phi(t))_{lk}$, $(\Psi(t))_{lk}$ and $(\mathbf{V}_j)_{lk}$ be the components at the lth column and kth row of $\Phi(t)$, $\Psi(t)$ and $\mathbf{V}_j$, respectively, $l, k = 1, 2, ..., N$ and $j \in \mathbf{Z}$. Then,

$$(\mathbf{V}_j)_{lk} \subset (\mathbf{V}_{j+1})_{lk}, \quad \text{and } f(t) \in (\mathbf{V}_j)_{lk} \Longleftrightarrow f(2t) \in (\mathbf{V}_{j+1})_{lk},$$

and

$$\cap_{j\in\mathbf{Z}} (\mathbf{V}_j)_{lk} = \{0\}, \quad \text{and } \cup_{j\in\mathbf{Z}} (\mathbf{V}_j)_{lk} \text{ is dense in } L^2(\mathbf{R}).$$

Moreover, for any $f_{lk} \in (\mathbf{V}_0)_{lk}$, there exist constants $a_{k_1,m,l,k}$ such that

$$f_{lk}(t) = \sum_{k_1 \in \mathbf{Z}} \sum_{m=1}^{N} a_{k_1,m,l,k} (\Phi(t-k_1))_{mk}, \ t \in \mathbf{R}. \tag{4.1}$$

And, for any $f \in L^2(\mathbf{R})$, there exist constants $a_{j,k_1,l,k}$ such that

$$f(t) = \sum_{j,k_1 \in \mathbf{Z}} \sum_{l=1}^{N} a_{j,k_1,l,k} (\Psi_{jk_1}(t))_{lk}, \ t \in \mathbf{R}, \tag{4.2}$$

where k is any integer with $1 \leq k \leq N$. This implies the following proposition.

THEOREM 4.1. *Let $\Phi(t)$ be a matrix valued scaling function of an MMRA $\{\mathbf{V}_j\}$ and $\Psi(t)$ be its an associated matrix valued wavelet function. Then, for any fixed k, $1 \leq k \leq N$, the functions $(\Phi(t))_{lk}$, $l = 1,2,...,N$, form multiscaling functions and $(\Psi(t))_{lk}$, $l = 1,2,...,N$, form multiwavelets. Moreover, for each pair (l,k), the spaces $(\mathbf{V}_j)_{lk}$, $j \in \mathbf{Z}$, form a multiresolution analysis of multiplicity r_k where r_k is the maximum number of linearly independent functions of $(\Phi(t))_{lk}$, $l = 1,2,...,N$.*

For more about multiresolution analysis of multiplicity r, see [**2**]-[**3**]. We next want to study the orthonormality of the column multiscaling functions induced from the orthonormality for matrix valued scaling functions, which is

$$\sum_{m=1}^{N} \int (\Phi(t-\tau_1))_{lm} (\Phi^*(t-\tau_2))_{km} dt = \delta(\tau_1 - \tau_2)\delta(l-k). \tag{4.3}$$

Or,

$$\int (\Phi(t-\tau_1))_{lk} (\Phi^*(t-\tau_2))_{kk} dt + \sum_{m=1, m\neq k}^{N} \int (\Phi(t-\tau_1))_{lm} (\Phi^*(t-\tau_2))_{km} dt \\ = \delta(\tau_1 - \tau_2)\delta(l-k). \tag{4.4}$$

Consider the multiscaling functions from the kth column $(\Phi(t))_{lk}(t)$, $1 \leq l \leq N$, of $\Phi(t)$. The conventional orthogonality studied in the current literature for multiwavelets is

$$\int (\Phi(t-\tau_1))_{l_1 k} (\Phi^*(t-\tau_2))_{l_2 k} dt = \delta(\tau_1 - \tau_2)\delta(l_1 - l_2). \tag{4.5}$$

We call the orthogonality (4.5) *Orthogonality A*, and the orthogonality (4.4) *Orthogonality B*, for multiscaling functions $(\Phi(t))_{lk}(t)$, $1 \leq l \leq N$. One can see that the second term in the left hand side of (4.4), Orthogonality B, is the flexibility term over (4.5), Orthogonality A.

LEMMA 4.2. *The conventional Orthogonality A for all column vectors of a matrix valued scaling function implies Orthogonality B induced from the orthogonality for matrix valued scaling functions.*

Proof. To prove (4.4), we only need to prove (4.3), which is

$$\sum_{m=1}^{N} \int (\Phi(t-\tau_1))_{lm} (\Phi^*(t-\tau_2))_{km} dt \overset{(4.5)}{=} \sum_{m=1}^{N} \delta(\tau_1-\tau_2)\delta(l-k) = N\delta(\tau_1-\tau_2)\delta(l-k).$$

♣

Comparing Orthogonality A in (4.5) and Orthogonality B in (4.4) or (4.3), one can see that the former requires the orthogonality for each individual component in a vector while the later only needs the orthogonality for the vector itself. This implies that Orthogonality B is weaker than Orthogonality A. On the other hand, these two orthogonalities imply the same orthogonality (2.11) for the discrete matrix filterbank $\mathbf{H}(\omega)$.

We now consider a subspace of $L^2(\mathbf{R}, \mathbf{C}^{N\times N})$:

$$L^2(\mathbf{R}, \mathbf{C}^N) = \{\mathbf{f} = (f_{k,l}(t))_{N\times N} \in L^2(\mathbf{R}, \mathbf{C}^{N\times N}) : \ f_{k,l}(t) = 0 \text{ for } 2 \le l \le N\},$$

which is isomorphic to the $N \times 1$ vector valued signal space. We may define its corresponding MAR, scaling functions, wavelet functions similarly. In this case, $\Phi(t) = ((\Phi(t))_{kl})_{N\times N}$ with $(\Phi(t))_{kl} = 0$ for $2 \le l \le N$. Clearly, Orthogonality A and Orthogonality B are equivalent in this case. In other words, Orthogonality A only corresponds to Orthogonality B in a proper subspace of the matrix valued signal space.

With Orthogonality A, necessary and sufficient conditions on $\mathbf{H}(\omega)$ that leads to orthogonal multiwavelets have been obtained, see for example [**15**]. Since the stronger Orthogonality A is used, the necessary and sufficient condition on $\mathbf{H}(\omega)$ is not easy to check or use. However, with the weaker Orthogonality B, the condition on $\mathbf{H}(\omega)$ in Theorem 3.4 is much easier to check so that one is able to use it to construct families of nontrivial orthogonal(B) multiwavelets as studied in Section 3. The basic idea doing this is to embed an $N \times 1$ vector into an $N \times N$ matrix and then use the matrix orthogonality. Another way to interpret this idea is that we lift a one dimensional vector into a two dimensional matrix with additional freedoms to play with, which makes the construction easier. One now might want to ask whether this new Orthogonality B is physically meaningful. The answer is *yes* because it provides a complete decorrelation for matrix valued signals as we shall study in the next section.

5. Matrix Karhunen Loève Expansion

In this section, we show that Orthogonality B provides a complete decorrelation for matrix valued random processes.

5.1. Matrix KL Expansion: Definition. Let $\mathbf{X}(t)$, $t \in [a, b]$ with $-\infty < a < b < \infty$, be a matrix valued random process with finite second moments, i.e.,

$$E(\mathbf{X}^\dagger(t)\mathbf{X}(t)) \in \mathbf{C}^{N\times N},$$

and each path $\mathbf{X}(t) \in L^2(\mathbf{R}; \mathbf{C}^{N\times N})$. Let

$$\mathbf{R}(s,t) \triangleq E(\mathbf{X}^\dagger(s)\mathbf{X}(t)), \ \ s,t \in [a,b]. \tag{5.1}$$

If there exist $\Phi_n(t) \in L^2(a, b; \mathbf{C}^{N\times N})$, $\Lambda_n \in \mathbf{C}^{N\times N}$, $n = 1, 2, ...$, such that

$$\int_a^b \Phi_n(s)\mathbf{R}(s,t)ds = \Lambda_n\Phi_n(t), \ \ n = 1, 2, ..., \ t \in [a, b], \tag{5.2}$$

$$\langle \Phi_n, \Phi_m\rangle = \delta(m-n)I_N, \ \ m, n = 1, 2, ..., \tag{5.3}$$

and

$$\mathbf{X}(t) = \sum_{n=1}^{\infty} \langle \mathbf{X}, \Phi_n\rangle\Phi_n(t), \ \ t \in [a, b], \tag{5.4}$$

then, the expansion of $\mathbf{X}(t)$ in (5.4) is called the *matrix Karhunen-Loève expansion* of $\mathbf{X}(t)$. If the matrix Karhunen-Loève (MKL) expansion of $\mathbf{X}(t)$ exists, then $\mathbf{X}(t)$ is decorrelated into a matrix valued random sequence $\mathbf{Y}_n \triangleq \langle \Phi_n, \mathbf{X} \rangle$ as

$$E(\mathbf{Y}_n \mathbf{Y}_m^\dagger) = \delta(n-m)\Lambda_n, \quad m, n = 1, 2, ... \tag{5.5}$$

The random sequence $\mathbf{Y}_n$, $n = 0, 1, 2, ...$, is called the matrix Karhunen-Loève transform of $\mathbf{X}(t)$.

Notice that when $N = 1$, the above MKL expansions/transforms are reduced to the conventional KL expansions/transforms. The object of this section is to study the MKL expansion of $\mathbf{X}(t)$.

Two special cases were studied in [**23**]-[**24**]. In one, the constant matrix Λ_n in (5.2) was replaced by a scalar value and in the other, $\Phi_n(t)$ in (5.2) was replaced by a scalar-valued function. As mentioned in §3.7 in [**24**], only a few cases satisfy these assumptions, and therefore they are not complete. The main reason for not using the product of two matrices at the right hand side in (5.2) is due to the difficulty of handling the noncommutativity of matrix products.

5.2. The Generalized Hilbert-Schmidt and Mercer's Theorems. Without loss of generality, in what follows we assume $a = 0$ and $b = T > 0$. Let $\mathbf{K}(s,t)$, $s, t \in [0, T]$, be a matrix valued function of two variables in $L^2(0, T; \mathbf{C}^{N\times N})$. In other words, for each $s \in [0, T]$, $K(s, \cdot) \in L^2(0, T; \mathbf{C}^{N\times N})$, and for each $t \in [0, T]$, $K(\cdot, t) \in L^2(0, T; \mathbf{C}^{N\times N})$, and

$$\int_0^T \int_0^T \|\mathbf{K}(s,t)\|_M^2 ds dt < \infty. \tag{5.6}$$

If $\mathbf{K}(s,t)$ satisfies the above conditions, then $\mathbf{K}(s,t)$ is called a matrix Fredholm integral operator. It is clear that a matrix Fredholm integral operator $\mathbf{K}(s,t)$ maps $L^2(0, T; \mathbf{C}^{N\times N})$ into itself:

$$(K\mathbf{f})(t) \triangleq \int_0^T \mathbf{f}(s)\mathbf{K}(s,t)ds \in L^2(0, T; \mathbf{C}^{N\times N}).$$

Let $\Phi(t) \in L^2(0, T; \mathbf{C}^{N\times N})$ with $\langle \Phi, \Phi \rangle = I_N$, and $\Lambda \in \mathbf{C}^{N\times N}$. If the following identity holds:

$$\int_0^T \Phi(s)\mathbf{K}(s,t)ds = \Lambda\Phi(t), \quad t \in [0, T], \tag{5.7}$$

then, $\Phi(t)$ and Λ are called *eigen-matrix-functions* and *eigen-matrix-values* of the operator $\mathbf{K}(s,t)$, respectively.

Notice that the property $\langle \Phi, \Phi \rangle = I_N$ is required in the above definitions of eigen-matrix-functions and eigen-matrix-values, which is different from the scalar-valued case. In the scalar-valued case, if $\phi(t)$ is an eigenfunction associated with an eigenvalue λ for a scalar Fredholm integral operator, then $a\phi(t)$ for any constant $a \neq 0$ is also an eigenfunction associated with λ. It is not known, however, whether the following statement is true: If $\Phi(t)$ is an eigen-matrix-function associated with an eigen-matrix-value Λ for a matrix Fredholm integral operator $\mathbf{K}(s,t)$, then $A\Phi(t)$ or $\Phi(t)A$ for an $N\times N$ matrix $A \in \mathbf{C}^{N\times N}$ is also an eigen-matrix-function associated with Λ for the operator $\mathbf{K}(s,t)$. The difficulty is due to the noncommutativity of matrix multiplications.

A matrix Fredholm integral operator $\mathbf{K}(s,t)$ is called *Hermitian* if $\mathbf{K}(s,t) = \mathbf{K}^\dagger(t,s)$ for $s,t \in [0,T]$. If $\mathbf{K}(s,t)$ is Hermitian and Λ is its eigen-matrix-value, then $\Lambda = \Lambda^\dagger$, i.e., Λ is also Hermitian. This is because

$$\langle \Phi, \mathbf{K}\Phi \rangle = \Lambda = (\langle \Phi, \mathbf{K}\Phi \rangle)^\dagger = \Lambda^\dagger.$$

We associate each matrix Fredholm integral operator $\mathbf{K}(s,t)$ on $[0,T]\times[0,T]$ with the following scalar Fredholm integral operator $K(s,t)$ on $[0,NT]\times[0,NT]$:

$$K(s,t) \triangleq K_{k,l}(s-(k-1)T, t-(l-1)T), \tag{5.8}$$

if $(s,t) \in ((k-1)T, kT] \times ((l-1)T, lT]$, $k,l = 1,2,...,N$, where $K_{k,l}(s,t)$ is the component function of $\mathbf{K}(s,t)$ at the kth row and the lth column. The property (5.6) implies the following properties for $K(s,t)$:

$$\int_0^{NT}\int_0^{NT} |K(s,t)|^2 dtds < \infty, \tag{5.9}$$

and if $\mathbf{K}(s,t)$ is Hermitian then $K(s,t)$ is also Hermitian, i.e., $K(s,t) = K^*(t,s)$, where * means the complex conjugate.

We now have the following generalized Hilbert-Schmidt theorem.

THEOREM 5.1. *Let $\mathbf{K}(s,t)$, $s,t \in [0,T]$, be a Hermitian matrix Fredholm integral operator and $K(s,t)$, $s,t \in [0,NT]$, be its associated scalar Fredholm integral operator. Let $\lambda_1, \lambda_2, ...,$ all be eigenvalues (including multiples) of $K(s,t)$ with $|\lambda_1| \geq |\lambda_2| \geq \cdots$. Then, an $N\times N$ matrix Λ is an eigen-matrix-value of the operator $\mathbf{K}(s,t)$ if and only if*

$$\Lambda = U\, diag(\lambda_{n_1}, \cdots, \lambda_{n_N}) U^\dagger, \tag{5.10}$$

where U is a certain $N\times N$ unitary matrix, and $n_1, ..., n_N$ are positive integers with $n_1 < n_2 < \cdots < n_N$. Moreover, if the operator $K(s,t)$ doesn't have zero eigenvalue, i.e., $|\lambda_n| > 0$, $n = 1,2,...,$ then, the eigen-matrix-functions $\Phi_n(t)$ corresponding to the eigen-matrix-values $\Lambda_n \triangleq diag(\lambda_{(n-1)N+1}, \cdots, \lambda_{nN})$, $n = 1,2,...,$ form an orthonormal basis for the matrix valued signal space $L^2(0,T;\mathbf{C}^{N\times N})$.

Proof: From the definition of an eigen-matrix-value in (5.7), $U^\dagger \Lambda U$ is an eigen-matrix-value of $\mathbf{K}(s,t)$ if Λ is an eigen-matrix-value of $\mathbf{K}(s,t)$ and U is an $N\times N$ unitary matrix. Thus, to prove Λ in (5.10) is an eigen-matrix-value of $\mathbf{K}(s,t)$, we only need to prove the diagonal matrix diag$(\lambda_{n_1}, \cdots, \lambda_{n_N})$ is an eigen-matrix-value of $\mathbf{K}(s,t)$. In fact, without loss of generality, we only need to prove Λ_n is an eigen-matrix-value of $\mathbf{K}(s,t)$ for any integer $n > 1$.

Let $\phi_n(t)$, $t \in [0,NT]$, be the eigenfunctions of $K(s,t)$ corresponding to λ_n, $n = 1,2,...,$ i.e., $\phi_n(t)$, $n=1,2,...,$ form an orthonormal set of $L^2(0,NT;\mathbf{C})$, and

$$\int_0^{NT} \phi_n(s) K(s,t) ds = \lambda_n \phi_n(t), \;\; t \in [0,NT]. \tag{5.11}$$

Then, equation (5.11) can be rewritten as

$$\int_0^T \sum_{k=0}^{N-1} \phi_n(s+kT) K(s+kT, t) ds = \lambda_n \phi_n(t), \;\; t \in [0,NT]. \tag{5.12}$$

Let $\phi_{k,n}(s) \triangleq \phi_n(s+kT),\ \ s\in[0,T],\ k=0,1,...,N-1$. Then,

$$\int_0^T \sum_{k=0}^{N-1} \phi_{k,n}(s) K(s+kT,t)ds = \lambda_n \phi_{l,n}(t-lT), \tag{5.13}$$

for $t\in(lT,(l+1)T]$, $l=0,1,...,N-1$. Let

$$\Phi_n(s) \triangleq \left(\begin{array}{cccc} \phi_{0,(n-1)N+1}(s) & \phi_{0,(n-1)N+2}(s) & \cdots & \phi_{0,nN}(s) \\ \phi_{1,(n-1)N+1}(s) & \phi_{1,(n-1)N+2}(s) & \cdots & \phi_{1,nN}(s) \\ & & \cdots & \\ \phi_{N-1,(n-1)N+1}(s) & \phi_{N-1,(n-1)N+2}(s) & \cdots & \phi_{N-1,nN}(s) \end{array} \right). \tag{5.14}$$

By (5.8), (5.13) can be rewritten as

$$\int_0^T \Phi_n(s)\mathbf{K}(s,t)dt = \Lambda_n\Phi_n(t),\ \ n=1,2,...,\ \ t\in[0,T]. \tag{5.15}$$

By the orthonormality of $\phi_n(s)$, $t\in[0,NT]$, it is not hard to see that

$$\langle \Phi_m, \Phi_n\rangle = \delta(m-n)I_N,\ \ m,n=1,2,.... \tag{5.16}$$

Therefore, we have proved that Λ_n, $n=1,2,...$, are eigen-matrix-values of the operator $\mathbf{K}(s,t)$.

Conversely, let Λ be an eigen-matrix-value of the operator $\mathbf{K}(s,t)$. By the previous discussion we know that Λ is Hermitian. Thus, there exists a unitary matrix U such that $\Lambda = U\mathrm{diag}(\alpha_1,\cdots,\alpha_N)U^\dagger$ with $|\alpha_1|\geq\cdots\geq|\alpha_N|$. By definition (5.7) of an eigen-matrix-value, $\mathrm{diag}(\alpha_1,\cdots,\alpha_N)$ is also an eigen-matrix-value of $\mathbf{K}(s,t)$, i.e., there is $\Phi(t)\in L^2(0,T;\mathbf{C}^{N\times N})$ with $\langle\Phi,\Phi\rangle = I_N$ such that

$$\int_0^T \Phi(s)\mathbf{K}(s,t)ds = \mathrm{diag}(\alpha_1,\cdots,\alpha_N)\Phi(t),\ \ t\in[0,T]. \tag{5.17}$$

Assume $\phi_{m,n}(s)$ is the mth row and the nth column component function of $\Phi(s)$. Let $\phi_n(s)=\phi_{m,n}(s-(m-1)T)$ if $s\in((m-1)T,mT]$ for $m,n=1,2,...,N$. By (5.8) and (5.17), the function $\phi_n(s)$ is an eigenfunction of the operator $K(s,t)$ with its corresponding eigenvalue α_n, $n=1,2,...,N$. Thus, $\alpha_k=\lambda_{n_k}$ for some k with $n_1<n_2<\cdots<n_N$. This proves (5.10).

When $K(s,t)$ has no zero eigenvalue, by the scalar Hilbert-Schmidt Theorem (see [**25**]), the eigenfunctions $\phi_n(t)$, $n=1,2,...$, form an orthonormal basis for $L^2(0,NT;\mathbf{C}^{N\times N})$. Therefore, any $f(t)\in L^2(0,NT;\mathbf{C})$ can be represented as

$$f(t)=\sum_{n=1}^{\infty}\langle f,\phi_n\rangle\phi_n(t),\ \ t\in[0,NT]. \tag{5.18}$$

Similarly, (5.18) can be rewritten as

$$\mathbf{f}(t)=\sum_{n=1}^{\infty}\int_0^T \mathbf{f}(s)(\phi_{0,n}(s),\cdots,\phi_{N-1,n}(s))^\dagger ds(\phi_{0,n}(t),\cdots,\phi_{N-1,n}(t)),\ t\in[0,T],$$

for any $N\times 1$ vector-valued $\mathbf{f}\in L^2(0,T;\mathbf{C}^N)$. By regrouping the above summation, we have

$$\mathbf{f}(t)=\sum_{n=1}^{\infty}\int_0^T \mathbf{f}(s)\Phi_n^\dagger(s)\Phi_n(t)ds,\ t\in[0,T],\ \mathbf{f}\in L^2(0,T;\mathbf{C}^N). \tag{5.19}$$

Extending $\mathbf{f}(t) \in L^2(0,T;\mathbf{C}^N)$ to $\mathbf{f}(t) \in L^2(0,T;\mathbf{C}^{N\times N})$, we have

$$\mathbf{f}(t) = \sum_{n=1}^{\infty} \langle \mathbf{f}, \Phi_n \rangle \Phi_n(t), \;\; t \in [0,T], \; \mathbf{f}(t) \in L^2(0,T;\mathbf{C}^{N\times N}). \tag{5.20}$$

This proves that the sequence $\Phi_n(t)$, $n = 1, 2, ...$, forms an orthonormal basis for $L^2(0,T;\mathbf{C}^{N\times N})$. ♣

From the above proof, the eigen-matrix-function $\Phi_n(t)$ in Theorem 5.1 associated with the eigen-matrix-value Λ_n in Theorem 5.1 is formulated by (5.14), for $n = 1, 2, ...$ We next want to generalize Mercer's Theorem. A matrix Fredholm integral operator $\mathbf{K}(s,t)$ is called *positive* if the $N \times N$ matrix $\langle \mathbf{f}, \mathbf{K}\mathbf{f} \rangle$ for any $\mathbf{f}(t) \in L^2(0,T;\mathbf{C}^{N\times N})$ is nonnegative definite, i.e., $\mathbf{x}^\dagger \langle \mathbf{f}, \mathbf{K}\mathbf{f} \rangle \mathbf{x} \geq 0$ for any $\mathbf{x} \in \mathbf{C}^N$.

LEMMA 5.2. *A matrix Fredholm integral operator $\mathbf{K}(s,t)$ is positive if and only if its associated scalar Fredholm integral operator $K(s,t)$ is positive.*

Proof: Writing $\langle f, Kf \rangle$ up, similar to the proof of Theorem 5.1, we have

$$\int_0^{NT} \int_0^{NT} f^*(s) K^*(s,t) f(t) ds dt = \int_0^T \int_0^T \mathbf{f}(t) \mathbf{K}^\dagger(s,t) \mathbf{f}^\dagger(s) dt ds, \tag{5.21}$$

where $\mathbf{f}(t) \in L^2(0,T;\mathbf{C}^N)$. On the other hand,

$$\mathbf{x}^\dagger \int_0^T \int_0^T \mathbf{f}(t) \mathbf{K}^\dagger(s,t) \mathbf{f}^\dagger(s) dt ds \mathbf{x} = \int_0^T \int_0^T (\mathbf{x}^\dagger \mathbf{f}(t)) \mathbf{K}^\dagger(s,t) (\mathbf{x}^\dagger \mathbf{f}(s))^\dagger dt ds, \tag{5.22}$$

where $\mathbf{x} \in \mathbf{C}^N$ and $\mathbf{f}(t) \in L^2(0,T;\mathbf{C}^{N\times N})$. Since

$$L^2(0,T;\mathbf{C}^N) = \{\mathbf{f}(t)\mathbf{x} : \; \mathbf{x} \in \mathbf{C}^N, \, \mathbf{f} \in L^2(0,T;\mathbf{C}^{N\times N})\},$$

the values in (5.21) are nonnegative for all $\mathbf{f}(t) \in L^2(0,T;\mathbf{C}^N)$ is equivalent to that the values in (5.22) are nonnegative for all $\mathbf{x} \in \mathbf{C}^N$ and all $\mathbf{f}(t) \in L^2(0,T;\mathbf{C}^{N\times N})$. This proves Lemma 5.2. ♣

we have the following generalized form of Mercer's Theorem.

THEOREM 5.3. *Let $\mathbf{K}(s,t)$ be a Hermitian matrix Fredholm integral operator. If $\mathbf{K}(s,t)$ is positive and its associated scalar Fredholm integral operator $K(s,t)$ is continuous in $[0,NT] \times [0,NT]$, then*

$$\mathbf{K}(s,t) = \sum_{n=1}^{\infty} \Phi_n^\dagger(s) \Lambda_n \Phi_n(t), \;\; s, t \in [0,T], \tag{5.23}$$

where $\Phi_n(t)$ and Λ_n are the same as in Theorem 5.1 and the convergence of the infinite summation is uniform.

Proof: By Lemma 5.2, the operator $K(s,t)$ is also positive. By Mercer's theorem for the operator $K(s,t)$ (see [**25**]),

$$K(s,t) = \sum_{n=1}^{\infty} \phi_n^*(s) \phi_n(t) \lambda_n, \;\; s, t \in [0,NT],$$

where ϕ_n, λ_n are eigenfunctions and eigenvalues of $K(s,t)$ and the convergence is uniform. Regrouping the above summation and using the same technique in the proof of Theorem 5.1, (5.23) can be proved. ♣

5.3. Matrix KL Expansions for Continuous-Time Matrix Valued Signals. We now come back to the MKL expansions for continuous-time matrix valued signals.

Let $\mathbf{R}(s,t)$ be the correlation matrix function defined by (5.1) of a matrix valued random process $\mathbf{X}(t)$ with $a=0$ and $b=T$. Assume $\mathbf{R}(s,t) \in L^2(0,T;\mathbf{C}^{N\times N})$. Then $\mathbf{R}(s,t)$ is a Hermitian matrix Fredholm integral operator on $L^2(0,T;\mathbf{C}^{N\times N})$; moreover $\mathbf{R}(s,t)$ is positive. Therefore, we can apply the generalized Hilbert-Schmidt Theorem and the generalized Mercer's Theorem.

Let $R(s,t)$ be the associated scalar Fredholm integral operator of the operator $\mathbf{R}(s,t)$, that is defined by (5.8). Let $\phi_n(t)$, λ_n, $n=1,2,...$, all be eigenfunctions and eigenvalues (including multiples) of the operator $R(s,t)$ with

$$\int_0^{NT} \phi_n(s)R(s,t)ds = \lambda_n\phi_n(t),\ t\in[0,NT],\ n=1,2,..., \tag{5.24}$$

and

$$\int_0^{NT} \phi_m(t)\phi_n^*(t)dt = \delta(m-n),\ \ m,n=1,2,..., \tag{5.25}$$

where $|\lambda_1| \geq |\lambda_2| \geq \cdots$. Since the operator $\mathbf{R}(s,t)$ is positive, by Lemma 5.2, the operator $R(s,t)$ is also positive. Thus, $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$.

Let

$$\Lambda_n \triangleq \mathrm{diag}(\lambda_{(n-1)N+1},\cdots,\lambda_{nN}),\ \ n=1,2,..., \tag{5.26}$$

and, for $t\in[0,T]$, $n=1,2,...$, and $\Phi_n(t)$ defined by (5.14). Then, by Theorem 5.1, its proof and (5.25), $\Phi_n(t)$ is an eigen-matrix-function of the operator $\mathbf{R}(s,t)$ corresponding to the eigen-matrix value Λ_n in (5.26) for $n=1,2,....$ This gives the following first condition on signals so that their MKL expansions exist.

THEOREM 5.4. *Let $\mathbf{X}(t)$, $t\in[0,T]$, be a random process with its correlation matrix function $\mathbf{R}(s,t)\in L^2(0,T;\mathbf{C}^{N\times N})$. If $\lambda_n>0$, $n=1,2,...$, then, for each path of $\mathbf{X}(t)$,*

$$\mathbf{X}(t) = \sum_{n=1}^{\infty}\langle \mathbf{X},\Phi_n\rangle\Phi_n(t),\ \ t\in[0,T], \tag{5.27}$$

i.e., the MKL expansion of $\mathbf{X}(t)$ exists in the sense (5.2)-(5.4).

The second condition is given by the following theorem.

THEOREM 5.5. *Let $\mathbf{X}(t)$, $t\in[0,T]$, be a random process with its correlation matrix function $\mathbf{R}(s,t)\in L^2(0,T;\mathbf{C}^{N\times N})$. If its associated scalar Fredholm integral operator $R(s,t)$ is continuous in $[0,NT]\times[0,NT]$, then the MKL expansion of $\mathbf{X}(t)$ exists:*

$$\mathbf{X}(t) = \sum_{n=1}^{\infty}\langle \mathbf{X},\Phi_n\rangle\Phi_n(t),\ \ t\in[0,T], \tag{5.28}$$

where the convergence is in the mean square sense.

The proofs of the above two theorems are straightforward by using the results in Section 5.2.

From Theorems 5.4-5.5, it seems that the MKL expansions of $\mathbf{X}(t)$ depend on the definition of the associated scalar Fredholm integral operator $R(s,t)$ of $\mathbf{R}(s,t)$. One might ask, when the existence of the MKL expansion of $\mathbf{X}(t)$ in the sense of (5.2)-(5.4) is assumed, whether the MKL expansion of $\mathbf{X}(t)$ changes if the way to define $R(s,t)$ in (5.8) changes. The answer is *NO*. In other words, the MKL expansions (5.27) and (5.28) in Theorems 5.4-5.5 are necessary.

THEOREM 5.6. *Let $\mathbf{X}(t)$, $t \in [0,T]$, be a random process with its correlation matrix function $\mathbf{R}(s,t) \in L^2(0,T;\mathbf{C}^{N\times N})$. If the MKL expansion of $\mathbf{X}(t)$ exists in the sense of (5.2)-(5.4), then the MKL expansion of $\mathbf{X}(t)$ can always be written as*

$$\mathbf{X}(t) = \sum_{n=1}^{\infty} \langle \mathbf{X}, \Phi_n \rangle \Phi_n(t), \ \ t \in [0,T], \tag{5.29}$$

where $\Phi_n(t)$, $n = 1, 2...$, are defined in (5.14).

Proof: By (5.2)-(5.4), there exist $\Phi_n^{'}(t) \in L^2(0,T;\mathbf{C}^{N\times N})$ and $\Lambda_n^{'} \in \mathbf{C}^{N\times N}$, $n = 1, 2, ...,$ such that

$$\int_0^T \Phi_n^{'}(s)\mathbf{R}(s,t)ds = \Lambda_n^{'}\Phi_n^{'}(t), \ \ n = 1, 2, ..., \ t \in [0,T],$$

$$\langle \Phi_n^{'}, \Phi_m^{'} \rangle = \delta(n-m)I_N, \ \ m, n = 1, 2, ...,$$

and

$$\mathbf{X}(t) = \sum_{n=1}^{\infty} \langle \mathbf{X}, \Phi_n^{'} \rangle \Phi_n^{'}(t), \ \ t \in [0,T]. \tag{5.30}$$

Thus, $\Phi_n^{'}(t)$ is an eigen-matrix-function of the operator $\mathbf{R}(s,t)$ corresponding to the eigen-matrix-value $\Lambda_n^{'}$ for $n = 1, 2, ...$. By Theorem 5.1, there exist unitary matrices U_n such that $\Lambda_n = U_n^{\dagger}\Lambda_n^{'}U_n$ for $n = 1, 2, ...,$ where the order of the eigenvalues λ_n is rearranged if necessary. Moreover, Λ_n is an eigen-matrix-value of $\mathbf{R}(s,t)$ with its eigen-matrix-function $U_n\Phi_n^{'}(t)$, $n = 1, 2, ...$. Then, similar to the proof of Theorem 5.1, one can show that $\Phi_n(t) = U_n\Phi_n^{'}(t)$, $n = 1, 2, ...$. By (5.30),

$$\mathbf{X}(t) = \sum_{n=1}^{\infty} \langle \mathbf{X}, U_n^{\dagger}\Phi_n \rangle U_n^{\dagger}\Phi_n(t) = \sum_{n=1}^{\infty} \langle \mathbf{X}, \Phi_n \rangle \Phi_n(t).$$

This proves (5.29). ♣

From Thoerems 5.1-5.6, one can clearly see that a matrix valued random process $\mathbf{X}(t)$ is completely decorrelated in the both time and the spatial domains using Orthogonality B.

6. Conclusion

In this paper, we studied orthonormal matrix valued multiresolution analysis and wavelets. A simple sufficient condition on the matrix filter $\mathbf{H}(\omega)$ that leads to orthonormal matrix valued wavelets is presented, which is analogous to the one given by Mallat in [**18**] for scalar valued wavelets. This sufficient condition enables us to construct families of nontrivial orthonormal matrix valued wavelets. With orthonormal matrix valued wavelets, one is able to construct multiwavelets with a different orthonormality (called Orthogonality B in this paper) from the one people currently use (called Orthogonality A in this paper). It was shown that Orthogonality B is weaker than Orthogonality A. We believe that this weaker orthogonality makes the sufficient condition simple. The main idea behind it is that one dimensional vectors are lifted to two dimensional matrices, and therefore more freedoms are avaliable. It was also shown that Orthogonality B provides a complete Karhunen-Loève expansion, i.e., a complete decorrelation, for matrix valued signals.

Acknowledgement

The author would like to thank Mr. Quangcai Zhou for providing Lemmas 3.2-3.3 and their proofs. He also would like to thank for the referees' useful comments and suggestions.

References

[1] X.-G. Xia and B. W. Suter, *Vector-valued wavelets and vector filter banks* IEEE Trans. on Signal Processing **44** (1996), 508–518.
[2] T. N. T. Goodman and S. L. Lee, *Wavelets of multiplicity r* Trans. Amer. Math. Soc.**342** (1994), 307-324.
[3] L. Hervé, *Multi-resolution analysis of multiplicity d: applications to dyadic interpolation* Applied and Computational Harmonic Analysis **1** (1994), 299–315.
[4] J. S. Geronimo, D. P. Hardin and P. R. Massopust, *Fractal functions and wavelet expansions based on several scaling functions* J. Approx. Theory **78** (1994), 373–401.
[5] G. Donovan, J. S. Geronimo, and D. P. Hardin, *Intertwining multiresolution analysis and the construction of piecewise polynomial wavelets* Preprint (1994).
[6] G. Strang and V. Strela, *Orthogonal multiwavelets with vanishing moments* J. Optical Eng. **33** (1994), 2104–2107.
[7] G. Strang and V. Strela, *Short wavelets and matrix dilation equations* IEEE Trans. on Signal Processing **43** (1995), 108–115.
[8] M. Vetterli and G. Strang, *Time-varying filter banks and multiwavelets* Sixth Digital Signal Processing Workshop (1994), Yosemite.
[9] C. Heil and D. Colella, *Matrix refinement equations: existence and uniqueness* J. Fourier Anal. Appl. **2** (1996), 363-377.
[10] X.-G. Xia, J. S. Geronimo, D. P. Hardin, and B. W. Suter, *Design of prefilters for discrete multiwavelet transforms* IEEE Trans. on Signal Processing **44** (1996), 25–35.
[11] V. Strela, P. N. Heller, G. Strang, P. Topiwala, C. Heil, *The application of multiwavelet filter banks to image processing* IEEE Trans. on Image Processing, to appear.
[12] W. Lawton, S. L. Lee, and Z. Shen, *An algorithm for matrix extension and wavelet construction* Math. Comp. **65** (1996), 723–737.
[13] J. Z. Wang, *Stability and linear independence associated with scaling vectors* SIAM J. Math. Anal., to appear.
[14] J. Lian, *Orthogonality criteria for multi-scaling functions* Preprint (1996).
[15] G. Plonka, *Necessary and sufficient conditions for orthonormality of scaling vectors* Preprint (1997).
[16] A. Aldroubi, *Oblique and biorthogonal multi-wavelet bases with fast-filtering algorithms* SPIE Proceedings **2569** (1995), San Diego, 15–26.
[17] P. Rieder, J. Götze, J. A. Nossek, *Multiwavelet transforms based on several scaling functions* Proceedings of IEEE Int. Symp. on Time-Freq. and Time-Scale Anal. (1994).
[18] S. Mallat, *Multiresolution approximations and wavelet orthonormal bases of* $L^2(\mathbf{R})$ Trans. Amer. Math. Soc. **315** (1989), 69–87.
[19] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
[20] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, NJ: Prentice Hall, 1993.
[21] Z. Doğanata, P. P. Vaidyanathan and T. Q. Nguyen, *General synthesis procedures for FIR lossless transfer matrices, for perfect-reconstruction multirate filter bank applications* IEEE Trans. on Acoust. Speech and Signal Proc. **36** (1988), 1561–1574.
[22] X.-G. Xia and B. W. Suter, *FIR paraunitary filter banks given several analysis filters: factorization and construction* IEEE Trans. on Signal Processing **44** (1996), 720–723.
[23] E. J. Kelly and W. L. Root, *A representation of vector-valued random processes* Group Rept. 55-21, revised, MIT, Lincoln Laboratory, April 22, 1960.
[24] H. Van Trees, *Detection, Estimation, and Modulation Theory I*, Wiley, 1968.
[25] F. Riesz and B.Sz. Nagy, *Functional Analysis*, New York, Ungar, 1955.

Department of Electrical Engineering, University of Delaware, Newark, DE 19716
E-mail address: xxia@ee.udel.edu