PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 145, Number 5, May 2017, Pages 2089–2106
http://dx.doi.org/10.1090/proc/13359
Article electronically published on January 6, 2017

ADAPTIVE ORTHONORMAL SYSTEMS FOR MATRIX-VALUED FUNCTIONS

DANIEL ALPAY, FABRIZIO COLOMBO, TAO QIAN, AND IRENE SABADINI

(Communicated by Stephan Ramon Garcia)

Abstract. In this paper we consider functions in the Hardy space $\mathbf{H}_2^{p\times q}$ defined in the unit disc of matrix-valued functions. We show that it is possible, as in the scalar case, to decompose those functions as linear combinations of suitably modified matrix-valued Blaschke products, in an adaptive way. The procedure is based on a generalization to the matrix-valued case of the maximum selection principle which involves not only selections of suitable points in the unit disc but also suitable orthogonal projections. We show that the maximum selection principle gives rise to a convergent algorithm. Finally, we discuss the case of real-valued signals.

1. Introduction

Functions in the Hardy space $\mathbf{H}_2(\mathbb{D})$ of the open unit disc $\mathbb{D}$ can be decomposed into linear combinations of functions which are modified Blaschke products

$$B_n(z) = \frac{\sqrt{1-|a_n|^2}}{1-z\overline{a_n}} \prod_{k=1}^{n-1} \frac{z-a_k}{1-z\overline{a_k}}, \qquad n = 1, 2, \ldots \tag{1.1}$$

where the points $a_n \in \mathbb{D}$ are adaptively chosen according to the function to be decomposed (see [32]), and thus they may be repeated. It is important to note that these points do not necessarily satisfy the so-called hyperbolic non-separability condition

$$\sum_{n=1}^{\infty} 1 - |a_n| = \infty. \tag{1.2}$$

The system (1.1), which is orthonormal, is called the Takenaka–Malmquist system. It is a basis of the Hardy space $\mathbf{H}_2(\mathbb{D})$ and, more generally, of $\mathbf{H}_p(\mathbb{D})$, $1 \leq p < \infty$, if and only if (1.2) is satisfied.

It is shown in [32] that the points a_n can be suitably chosen in order to decompose a given function f into basic functions, each of which has non-negative analytic instantaneous frequency. Given a real-valued function f, interpreted as a signal, we can define the function $Af = f + iHf$, called an analytic signal associated with f, where Hf is the Hilbert transform of f. Its analytic instantaneous frequency is the phase derivative $\varphi'(t)$ of Af, written in the form $Af(t) = \rho(t)\exp(i\varphi(t))$,

Received by the editors May 1, 2016 and, in revised form, June 30, 2016.
2010 *Mathematics Subject Classification.* Primary 47A56, 41A20, 30H10.
Key words and phrases. Matrix-valued functions and Hardy spaces, matrix-valued Blaschke products, maximum selection principle, adaptive decomposition.

when the derivative exists as a measurable function. A signal that possesses a non-negative analytic instantaneous frequency function is said to be a mono-component, and it can be real- or complex-valued. If, in particular, $a_1 = 0$, then the boundary values of the modified Blaschke products B_n, $n \in \mathbb{N}$, are all mono-components. The theory of mono-components is built upon the characteristic property of boundary functions of Moebius transforms: their phase derivatives are the corresponding Poisson kernels as density functions of harmonic measures (see [19, Section I.3]). We note that with adaptive approximation the condition (1.2) is not necessarily satisfied, and thus the system is not necessarily complete in $\mathbf{H}_2(\mathbb{D})$. However, the convergence to f is fast.

This and related methods have been intensively studied in the past few years; see [26, 28–30, 32, 33]. It gives rise to algorithms which are variations and improvements of the standard greedy algorithms; see [25, 31, 37]. In a recent paper of Coifman et al. a related fast mono-component decomposition is studied in depth [15]. Below we discuss an algorithm arising in this framework, which can be generalized to the matrix-valued case. For matrix-valued functions we do not have a counterpart for the notion of mono-component signal. However, as we will show, we have the counterparts of all the necessary ingredients to successfully perform the algorithm.

An algorithm to perform an adaptive decomposition can be assigned as follows. One considers a so-called dictionary $\mathcal{D}$, being a family of elements of unit norm whose span is dense in the Hilbert space $\mathcal{H}$. Given $f \in \mathcal{H}$ we select $u_1, \ldots, u_n \in \mathcal{D}$ such that

$$f = \sum_{k=1}^{\infty} \langle f_k, u_k \rangle u_k$$

where the functions f_k are defined inductively, starting from $f_1 = f$ and setting

$$f_k = f - \sum_{\ell=1}^{k-1} \langle f_\ell, u_\ell \rangle u_\ell,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathcal{H}$.

In the paper [32], $\mathcal{H} = \mathbf{H}_2(\mathbb{D})$ with its standard inner product, and the dictionary consists of the normalized Szegö kernels

$$\mathcal{D} = \left\{ e_a(z) = \frac{\sqrt{1 - |a|^2}}{1 - z\overline{a}}, \quad a \in \mathbb{D} \right\}.$$

Note that the reproducing kernel property of e_a in $\mathbf{H}_2(\mathbb{D})$ yields

$$\langle f, e_a \rangle = \sqrt{1 - |a|^2} f(a).$$

Let $f \in \mathbf{H}_2(\mathbb{D})$ and set $f_1 = f$. For any $a_1 \in \mathbb{D}$ let

$$f(z) = \langle f_1, e_{a_1} \rangle e_{a_1}(z) + f_2(z) \frac{z - a_1}{1 - z\overline{a_1}} \tag{1.3}$$

where

$$f_2(z) = \frac{f_1(z) - \langle f_1, e_{a_1} \rangle e_{a_1}(z)}{\frac{z - a_1}{1 - z\overline{a_1}}}.$$

One can show that $f_2 \in \mathbf{H}_2(\mathbb{D})$ and so the procedure can be repeated. We recall that the backward-shift is defined as

$$R_0 f(z) = \begin{cases} \dfrac{f(z) - f(0)}{z}, & z \neq 0, \\ f'(0), & z = 0, \end{cases}$$

The transformation mapping f_1 to f_2 when $a_1 = 0$ coincides with $R_0 f$, so, in the general case, the transformation from f_1 to f_2 is called generalized backward-shift. The two summands in (1.3) are orthogonal, thus

$$\|f\|^2 = |\langle f_1, e_{a_1} \rangle|^2 + \|f_2\|^2.$$

The maximal selection principle asserts that it is possible to choose $a_1 \in \mathbb{D}$ such that

$$(1 - |a_1|^2)|f_1(a_1)|^2 = \max\{|\langle f_1, e_a \rangle|^2 = (1 - |a|^2)|f_1(a)|^2, \quad a \in \mathbb{D}\}.$$

The procedure can be iterated, and after n steps one has

$$f(z) = \sum_{k=1}^{n} \langle f_k, e_{a_k} \rangle B_k(z) + f_{n+1}(z) \prod_{k=1}^{n} \frac{z - a_k}{1 - z\overline{a_k}},$$

where

$$(1 - |a_k|^2)|f_1(a_k)|^2 = \max\{|\langle f_k, e_a \rangle|^2 = (1 - |a|^2)|f_k(a)|^2, \quad a \in \mathbb{D}\}, \quad k = 1, \ldots, n,$$

and

$$f_k(z) = \frac{f_{k-1}(z) - \langle f_{k-1}, e_{a_{k-1}} \rangle e_{a_{k-1}}(z)}{\frac{z - a_{k-1}}{1 - z\overline{a_{k-1}}}}, \quad k = 2, \ldots, n. \tag{1.4}$$

The function f_k is called the k-th reduced remainder (see [28, (11), p. 850]). As we shall see, its matrix-valued counterpart is given by (5.7). One can easily show the relations

$$\langle f_k, e_{a_k} \rangle = \langle g_k, B_k \rangle = \langle f, B_k \rangle, \tag{1.5}$$

where g_k is the k-th standard remainder, defined through

$$g_k = f - \sum_{l=1}^{k-1} \langle f, B_l \rangle B_l.$$

As before, the orthogonality of the summands and the fact that B_k is unimodular on $\partial\mathbb{D}$ give

$$\|f(z) - \sum_{k=1}^{n} \langle f_k, e_{a_k} \rangle B_k(z)\|^2 = \|f(z)\|^2 - \sum_{k=1}^{n} |\langle f_k, e_{a_k} \rangle|^2 = \|f_{n+1}\|^2.$$

Since it can be shown that $\|f_{n+1}\| \to 0$ as $n \to \infty$ (see [32, Theorem 2.2]), we have the formula

$$f(z) = \sum_{k=1}^{\infty} \langle f_k, e_{a_k} \rangle B_k(z),$$

called adaptive Fourier decomposition, abbreviated as AFD.

In this paper, we extend some of the results of [32] to the matrix-valued case. For $w \in \mathbb{D}$ we will use the notation e_w and b_w for the normalized Cauchy kernel and the Blaschke factor at the point w respectively, that is:

$$e_w(z) = \frac{\sqrt{1-|w|^2}}{1-z\overline{w}} \quad \text{and} \quad b_w(z) = \frac{z-w}{1-z\overline{w}}. \tag{1.6}$$

The Szegö dictionary now consists of the $\mathbb{C}^{p\times p}$-valued functions Pe_w, where w belongs to the open unit disk $\mathbb{D}$ and $P \in \mathbb{C}^{p\times p}$ is any orthogonal projection that satisfies $P = P^2 = P^*$.

We denote by $\mathbf{H}_2^{p\times q}$ the space of $p \times q$ matrices with entries in $\mathbf{H}_2(\mathbb{D})$. When $q = 1$ we write $\mathbf{H}_2^p$ rather than $\mathbf{H}_2^{p\times q}$. The set $\mathbf{H}_2^{p\times q}$ can be considered as a $\mathbb{C}^{q\times q}$ right Hilbert module; see [22].

A function $F \in \mathbf{H}_2^{p\times q}$ if and only if it can be written as

$$F(z) = \sum_{n=0}^{\infty} F_n z^n, \tag{1.7}$$

where $F_\ell \in \mathbb{C}^{p\times q}$, $\ell = 1, 2, \ldots$, are such that

$$\sum_{n=0}^{\infty} \mathrm{Tr}\ (F_n^* F_n) < \infty. \tag{1.8}$$

Let G be another element of $\mathbf{H}_2^{p\times q}$, with power series expansion $G(z) = \sum_{n=1}^{\infty} G_n z^n$ at the origin. We set

$$[F, G] = \sum_{n=0}^{\infty} G_n^* F_n \ \in\ \mathbb{C}^{q\times q} \tag{1.9}$$

and

$$\|F\|^2 = \mathrm{Tr}\,[F, F] = \sum_{n=0}^{\infty} \mathrm{Tr}\ (F_n^* F_n).$$

We note that (1.9) can be rewritten as

$$[F, G] = \lim_{\substack{r\to 1\\ r\in(0,1)}} \frac{1}{2\pi} \int_0^{2\pi} G(re^{it})^* F(re^{it}) dt \tag{1.10}$$

and so we also have

$$\sum_{n=0}^{\infty} \mathrm{Tr}\ (G_n^* F_n) = \lim_{\substack{r\to 1\\ r\in(0,1)}} \frac{1}{2\pi} \int_0^{2\pi} \mathrm{Tr}\, G(re^{it})^* F(re^{it}) dt. \tag{1.11}$$

Most, if not all, the material of Sections 3 and 4 is classical. Some proofs are provided for the convenience of the reader. We refer to [5, 6] for a study of rational matrix-valued Blaschke products using state space methods.

An important condition in the algorithm is whether F is a cyclic vector for the backward-shift operator R_0, that is, whether the closed linear span $\mathcal{M}(F)$ of the functions

$$R_0^n F X, \quad n = 0, 1, 2, \ldots, \ \text{ and } \ X \in \mathbb{C}^{q\times q}$$

is strictly included in $\mathbf{H}_2^{p\times q}$ or not.

2. The maximum selection principle

In this section we show that the maximum selection principle holds also in the matrix-valued case. It allows one to adaptively choose a sequence of points together with orthogonal projections for any given function in the Hardy space. We note that this selection principle does not exclude the possibility that the obtained sequence of points contains elements repeating more than once.

Proposition 2.1. *Let $k_0 \in \{1,\dots,p\}$ and let $F \in \mathbf{H}_2^{p\times q}$. There exists $w_0 \in \mathbb{D}$ and an orthogonal projection P_0 of rank k_0 such that*

$$(1-|w|^2)\left(\text{Tr }[PF(w),F(w)]\right) \le (1-|w_0|^2)\left(\text{Tr }[P_0F(w_0),F(w_0)]\right), \tag{2.1}$$

for every choice of $w \in \mathbb{D}$ and every orthogonal projection P of rank k_0.

Proof. We first recall that for $f \in \mathbf{H}_2(\mathbb{D})$ (that is, $p = q = 1$), with power series $f(z) = \sum_{n=0}^{\infty} f_n z^n$, and for $w \in \mathbb{D}$, we have

$$\sqrt{1-|w|^2}|f(w)| = |[f,e_w]| \le \|f\|. \tag{2.2}$$

Let $F = (f_{ij}) \in \mathbf{H}_2^{p\times q}$, where the $f_{ij} \in \mathbf{H}_2(\mathbb{D})$ ($i = 1,\dots,p$ and $j = 1,\dots,q$), and $\xi \in \mathbb{C}^{k_0\times p}$ is such that $\xi\xi^* = I_{k_0}$. Then,

$$\begin{aligned}\text{Tr }[\xi F(w),\xi F(w)] &= \text{Tr }F(w)^*\xi^*\xi F(w)\\ &\le \text{Tr }F(w)^*F(w)\\ &= \sum_{i=1}^{p}\sum_{j=1}^{q}|f_{ij}(w)|^2.\end{aligned}$$

Using (2.2) for every f_{ij}, we obtain

$$(1-|w|^2)\left(\text{Tr }[\xi F(w),\xi F(w)]\right) \le \sum_{i=1}^{p}\sum_{j=1}^{q}\|f_{ij}\|^2 = \|F\|^2. \tag{2.3}$$

Let $\epsilon > 0$. In view of (1.7)-(1.8) there exists a $\mathbb{C}^{p\times q}$-valued polynomial Q such that $\|F - Q\| \le \epsilon$. We have (since $\xi^*\xi$ is a rank k_0 orthogonal projection)

$$\begin{aligned}&(1-|w|^2)\left(\text{Tr }[\xi F(w),\xi F(w)]\right)\\ &\qquad = (1-|w|^2)\left(\text{Tr }[\xi(F-Q)(w)+\xi Q(w),\xi(F-Q)(w)+\xi Q(w)]\right)\\ &\qquad = (1-|w|^2)\|\xi(F-Q)(w)+\xi Q(w)\|^2\\ &\qquad \le (1-|w|^2)\left(\|\xi(F-Q)(w)\|+\|\xi Q(w)\|\right)^2\\ &\qquad \le 2(1-|w|^2)\|\xi(F-Q)(w)\|^2 + 2(1-|w|^2)\|\xi Q(w)\|^2\\ &\qquad \le 2\|F-Q\|^2 + 2(1-|w|^2)\|\xi Q(w)\|^2 \quad \text{(where we have used (2.3))}\\ &\qquad \le 2\epsilon^2 + 2(1-|w|^2)\|Q(w)\|^2.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, $(1-|w|^2)\|Q(w)\|^2$ tends to 0 as w approaches the unit circle. We know from (2.3) that $(1-|w|^2)\left(\text{Tr }[\xi F(w),\xi F(w)]\right)$ is uniformly bounded in $\mathbb{D}$ and thus it admits a finite supremum $(1-|w_0|^2)\left(\text{Tr }[\xi_0F(w_0),\xi_0F(w_0)]\right)$, which is in fact a maximum, for w_0 in a compact subset of $\mathbb{D}$ (since the function goes to zero on the boundary of $\mathbb{D}$) and since the projections belong to a compact set. This ends the proof by setting $P_0 = \xi_0^*\xi_0$. $\square$

Let $w_0 \in \mathbb{D}$ and let $P_0 \in \mathbb{C}^{p\times p}$ be an orthogonal projection. We can write

$$F(z) = P_0F(w_0)e_{w_0}(z)\sqrt{1-|w_0|^2} + F(z) - P_0F(w_0)e_{w_0}(z)\sqrt{1-|w_0|^2}. \tag{2.4}$$

Note that the decomposition (2.4) is non-trivial if and only if $F \not\equiv 0_{p\times q}$.

Lemma 2.2. *Let $w_0 \in \mathbb{D}$, let $P_0 \in \mathbb{C}^{p\times p}$ be an orthogonal projection and let*

$$H(z) = F(z) - P_0F(w_0)e_{w_0}(z)\sqrt{1-|w_0|^2},$$
$$H_0(z) = P_0F(w_0)e_{w_0}(z)\sqrt{1-|w_0|^2}.$$

We have that

$$P_0H(w_0) = 0 \tag{2.5}$$

and

$$[F,F] = [H_0,H_0] + [H,H]. \tag{2.6}$$

Proof. First we have (2.5) since

$$P_0H(w_0) = P_0F(w_0) - P_0F(w_0)e_{w_0}(w_0)\sqrt{1-|w_0|^2} = 0.$$

Using (2.5) we have

$$[H, P_0F(w_0)e_{w_0}(z)\sqrt{1-|w_0|^2}] = F(w_0)^*P_0H(w_0)(1-|w_0|^2) = 0.$$

So, $[H,H_0] = 0$ and

$$[F,F] = [H_0+H, H_0+H] = [H_0,H_0] + [H,H].$$

□

To proceed and take care of the condition (2.5) (that is, in the scalar case, to divide by a Blaschke factor) we first need to define matrix-valued Blaschke factors. This is done in the next section.

3. Matrix-valued Blaschke factors

Matrix-valued Blaschke factors originated with the work of Potapov [27] and can be defined (up to a right multiplicative constant) as

$$B_{w_0,P_0}(z) = I_p - P_0 + P_0b_{w_0}(z), \tag{3.1}$$

where for $w_0 \in \mathbb{D}$, b_{w_0} is defined as in (1.6), and $P_0 \in \mathbb{C}^{p\times p}$ is any orthogonal projection. The *degree* $\deg B_{w_0,P_0}$ of the Blaschke factor is by definition the rank of the projection P_0. When considering infinite products, it will be more convenient to consider for $w_0 \neq 0$ the Blaschke factor

$$\mathcal{B}_{w_0,P_0}(z) = I_p - P_0 + P_0\frac{|w_0|}{w_0}\frac{w_0-z}{1-z\overline{w_0}}. \tag{3.2}$$

Note that

$$B^{-1}_{w_0,P_0}(z) = I_p - P_0 + P_0\frac{1}{b_{w_0}(z)} \tag{3.3}$$

and

$$\mathcal{B}_{w_0,P_0}(z) = B_{w_0,P_0}(z)U \quad \text{with} \quad U = I_p - P_0 - \frac{|w_0|}{w_0}P_0.$$

In (3.5) in the following proposition, deg refers to the McMillan degree of a matrix-valued rational function. We refer e.g. to [13] for the definition and properties of the McMillan degree and to [18] for further information on matrix-valued Blaschke products.

Proposition 3.1. *Let B_{w_0,P_0} be defined by* (3.1)*. Then*

$$K_{B_{w_0,P_0}}(z,w) \stackrel{\text{def.}}{=} \frac{I_p - B_{w_0,P_0}(z)B_{w_0,P_0}(w)^*}{1-z\overline{w}} = \frac{(1-|w_0|^2)}{(1-z\overline{w_0})(1-w_0\overline{w})}P_0. \tag{3.4}$$

and

$$\mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q} = \left\{ \frac{P_0V}{1-z\overline{w_0}},\ V \in \mathbb{C}^{p\times q} \right\}$$

is the reproducing kernel Hilbert space with reproducing kernel $K_{B_{w_0,P_0}}(z,w)$, meaning that the function $z \mapsto K_{B_{w_0,P_0}}(z,w)X$ belongs to $\mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}$ for every $X \in \mathbb{C}^{p\times q}$ and

$$[F(\cdot), K_{B_{w_0,P_0}}(\cdot,w)X] = [P_0F(w_0), X].$$

Finally (and with $q=1$)

$$\deg B_{w_0,P_0} = \dim \mathbf{H}_2^p \ominus B_{w_0,P_0}\mathbf{H}_2^p. \tag{3.5}$$

Proof. We include the proof for completeness. In the proof we write $B(z)$ rather than B_{w_0,P_0} to ease the notation. Since $P_0(I_p - P_0) = 0$ we have

$$B(z)B(w)^* = I_p - P_0 + P_0 b_{w_0}(z)\overline{b_{w_0}(w)},$$

and so

$$I_p - B(z)B(w)^* = P_0(1 - b_{w_0}(z)\overline{b_{w_0}(w)}).$$

Equation (3.4) follows since

$$\frac{1 - b_{w_0}(z)\overline{b_{w_0}(w)}}{1-z\overline{w}} = \frac{(1-|w_0|^2)}{(1-z\overline{w_0})(1-w_0\overline{w})}.$$

It follows that the $\mathbb{C}^{p\times p}$-valued function $K_{B_{w_0,P_0}}(z,w)$ is positive definite (in the sense of reproducing kernels theory) in the open unit disk and that the associated reproducing kernel Hilbert space $\mathcal{H}(K_{Bw_0,P_0})$ of $\mathbb{C}^{p\times q}$-valued functions is exactly the set of functions of the form $z \mapsto \frac{P_0V}{1-z\overline{w_0}}$ when V varies in $\mathbb{C}^{p\times q}$. Equation (3.4) also implies that the space $\mathcal{H}(K_{Bw_0,P_0})$ is isometrically included in $\mathbf{H}_2^{p\times q}$. The fact that

$$\mathcal{H}(K_{Bw_0,P_0}) = \mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}$$

follows then from the kernel decomposition

$$\frac{I_p}{1-z\overline{w}} = \frac{I_p - B(z)B(w)^*}{1-z\overline{w}} + \frac{B(z)B(w)^*}{1-z\overline{w}}.$$

The last claim follows from the identification of the McMillan degree of a unitary rational function and the dimension of its associated reproducing kernel space; see for instance [4, 5] for the latter. ☐

We note that in Proposition 3.1 one can replace B_{w_0,P_0} by $\mathcal{B}_{w_0,P_0}$. It is easily seen that

$$K_{B_{w_0,P_0}}(z,w) = K_{\mathcal{B}_{w_0,P_0}}(z,w).$$

Lemma 3.2. *Let $H \in \mathbf{H}_2^{p\times q}$ be such that $P_0H(w_0) = 0_{p\times q}$. Then*

$$G = B_{w_0,P_0}^{-1} H \in \mathbf{H}_2^{p\times q}$$

and

$$[H, H] = [G, G]. \tag{3.6}$$

Proof. Write $P_0H(z) = (z - w_0)R(z)$, where R is $\mathbb{C}^{k\times q}$-valued and analytic in a neighborhood of w_0. Using (3.3) we have for $z \neq w_0$,

$$B_{w_0,P_0}^{-1}(z)H(z) = P_0H(z)\frac{1}{b_{w_0}(z)} = R(z)(1 - z\overline{w_0}),$$

and the point w_0 is a removable singularity of P_0H. Hence, $B_{w_0,P_0}^{-1}(z)H(z)$ has a removable singularity at w_0. Furthermore, since $B_{w_0,P_0}(e^{it})^* B_{w_0,P_0}(e^{it}) = I_p$, and using (1.10), we have (3.6). □

4. Backward-shift invariant subspaces

We define for $\alpha \in \mathbb{C}$ the resolvent-like operator

$$R_\alpha f(z) = \begin{cases} \dfrac{f(z) - f(\alpha)}{z - \alpha}, & z \neq \alpha, \\ f'(\alpha), & z = \alpha, \end{cases}$$

where the (possibly vector-valued) function f is analytic in a neighborhood of α.

A finite dimensional space $\mathcal{M}$ of $\mathbb{C}^{p\times q}$-valued functions analytic in a neighborhood of the origin is R_0-invariant if and only if there exists a pair of matrices $(C, A) \in \mathbb{C}^{p\times N} \times \mathbb{C}^{N\times N}$ which is observable, meaning $\bigcap_{u=0}^{\infty} \ker CA^u = \{0\}$ and

$$\mathcal{M} = \left\{F(z) = C(I_N - zA)^{-1}X, \quad X \in \mathbb{C}^{N\times q}\right\}.$$

The following proposition is a particular case of the Beurling-Lax theorem in the finite dimensional setting (see [8] for the unit disc case and [23] for the half-plane case).

Proposition 4.1. *Let $(C, A) \in \mathbb{C}^{p\times N} \times \mathbb{C}^{N\times N}$ be an observable pair of matrices, and let $\mathcal{M}$ denote the span of the functions of the form $F(z) = C(I_N - zA)^{-1}X$, where X runs through $\mathbb{C}^{N\times q}$. Then $\mathcal{M} \subset \mathbf{H}_2^{p\times q}$ if and only if $\rho(A) < 1$. When this is the case, we have $\mathcal{M} = \mathbf{H}_2^{p\times q} \ominus B\mathbf{H}_2^{p\times q}$, that is,*

$$\mathcal{M}^\perp = B\mathbf{H}_2^{p\times q},$$

where B is a finite Blaschke product, defined up to a unitary right constant, by the formula

$$B(z) = I_p - (1 - z)C(I_N - zA)^{-1}\mathbf{P}^{-1}(I_N - A)^{-*}C^*, \tag{4.1}$$

with

$$\mathbf{P} = \sum_{u=0}^{\infty} A^{*u}C^*CA^u. \tag{4.2}$$

Proof. By hypothesis, the pair (C, A) is observable. We assume $\mathcal{M} \subset \mathbf{H}_2^{p\times q}$. Let λ be an eigenvalue of A, with eigenvector ξ. We have $CA^n\xi = \lambda^n C\xi$, and the observability implies $C\xi \neq 0$. Thus

$$F(z)\xi = C(I - zA)^{-1}\xi = \frac{C\xi}{1 - \lambda z} \neq 0.$$

So $F(z)$ belongs to $\mathbf{H}_2^{p\times q}$ if and only if $\lambda \in \mathbb{D}$, that is, if and only if $\rho(A) < 1$.

To prove the second claim we remark that (4.2) indeed converges since $\rho(A) < 1$ and that the matrix P is the unique solution of the Stein equation

$$\mathbf{P} - A^*\mathbf{P}A = C^*C. \tag{4.3}$$

The second claim follows then from the identity

$$C(I_N - zA)^{-1}\mathbf{P}^{-1}(I_N - wA)^{-*}C^* = \frac{I_p - B(z)B(w)^*}{1 - z\overline{w}}, \tag{4.4}$$

which is proved by a direct computation, taking into account (4.3) and upon setting $w = 1$. $\square$

We also note that (3.4) is a special case of (4.4) and that Proposition 3.1 can be viewed as a special case of Proposition 4.1.

Using the above theorem or using state space methods, one can prove that a finite Blasckhe product is a finite product of degree one Blaschke factors. This result originates with the work of Potapov [27]; see e.g. [5] for a proof.

The following proposition can be proved using the Beurling-Lax theorem or using the analysis in [3] for backward-shift invariant subspaces. In the statement a $\mathbb{C}^{p\times\ell}$-valued inner function is an analytic $\mathbb{C}^{p\times\ell}$-valued function Θ such that the operator of multiplication by Θ is an isometry from $\mathbf{H}_2^{\ell\times q}$ into $\mathbf{H}_2^{p\times q}$.

Proposition 4.2. *Let $F \in \mathbf{H}_2^{p\times q}$ and assume that the closed linear span $\mathcal{M}(F)$ of the functions*

$$R_0^n FX, \quad n = 0, 1, 2, \ldots, \ \text{and} \ X \in \mathbb{C}^{q\times q}$$

is strictly included in $\mathbf{H}_2^{p\times q}$. Then there exists a $\mathbb{C}^{p\times\ell}$-valued inner function Θ such that

$$\mathcal{M}(F) = \mathbf{H}_2^{p\times q} \ominus \Theta\mathbf{H}_2^{\ell\times q}. \tag{4.5}$$

Proof. The space $\mathcal{M}(F)$ is R_0 invariant, so its orthogonal complement $(\mathcal{M}(F))^\perp$ is invariant by multiplication by z. The result follows then from the Beurling-Lax theorem. $\square$

Note that Θ need not be square; for instance, if $p = 2$, we can have

$$\Theta(z) = \begin{pmatrix} 0 \\ b(z) \end{pmatrix},$$

where b is a Blaschke product. Then,

$$\mathcal{M}(F) = \left\{ \begin{pmatrix} f \\ g \end{pmatrix}, \ f \in \mathbf{H}_2(\mathbb{D}) \text{ and } g \in \mathbf{H}_2(\mathbb{D}) \ominus b\mathbf{H}_2(\mathbb{D}) \right\}.$$

Still for $p = 2$, the case

$$\Theta(z) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ b(z) \end{pmatrix},$$

where

$$b(z) = \prod_{n=1}^{N} \frac{z - w_n}{1 - z\overline{w_n}} = c_0 + \sum_{n=1}^{N} \frac{c_n}{1 - z\overline{w_n}} \quad \text{for uniquely defined } c_0, \ldots, c_N \in \mathbb{C}$$

when the zeros of the Blaschke product are all different from 0 and simple, leads to

$$\mathcal{M}(F) = \left\{ \begin{pmatrix} \overline{c_0} f(z) + \sum_{n=1}^{N} \overline{c_n} \frac{zg(z) - w_n g(w_n)}{z - w_n} \\ g(z) \end{pmatrix}, \ g \in \mathbf{H}_2 \right\},$$

where we have used (for instance) [1, Exercise 8.3.1] to compute the first component.

These examples suggest a classification of the functions $F \in \mathbf{H}_2^{p\times q}$ depending on the value ℓ and the precise structure of Θ.

5. The algorithm

For any w_0 in the unit disc and any projection P_0, $\mathbf{H}_2^{p\times q}$ admits the orthogonal decomposition

$$\mathbf{H}_2^{p\times q} = \left(\mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}\right) \oplus B_{w_0,P_0}\mathbf{H}_2^{p\times q},$$

as is explained in the following lemma.

Lemma 5.1. *For any given $w_0 \in \mathbb{D}$ and any orthogonal projection $P_0 \in \mathbb{C}^{p\times p}$, formula (2.4) can be rewritten in a unique way as an orthogonal sum (orthogonal also with respect to the $[\cdot,\cdot]$ form defined in (1.9))*

$$F(z) = M_0 e_{w_0}(z)\sqrt{1 - |w_0|^2} + B_{w_0,P_0}(z)F_1(z), \tag{5.1}$$

where $M_0 \in \mathbb{C}^{p\times q}$ and $F_1 \in \mathbf{H}_2^{p\times q}$. We have

$$M_0 e_{w_0}\sqrt{1-|w_0|^2} \in \left(\mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}\right)$$

and $B_{w_0,P_0}F_1 \in B_{w_0,P_0}\mathbf{H}_2^{p\times q}$.

Finally,

$$[F, F] = (1 - |w_0|^2)[P_0 F(w_0), F(w_0)] + [F_1, F_1]. \tag{5.2}$$

Proof. We have

$$F(z) = M_0 e_{w_0}(z)\sqrt{1 - |w_0|^2} + B_{w_0,P_0}(z)F_1(z), \tag{5.3}$$

where $M_0 = P_0 F(w_0)$ and $F_1 = B_{w_0,P_0}^{-1}\left(F - P_0F(w_0)e_{w_0}\sqrt{1 - |w_0|^2}\right) \in \mathbf{H}_2^{p\times q}$. By Proposition 3.1,

$$P_0 F(w_0)e_{w_0}\sqrt{1 - |w_0|^2} \in \mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}.$$

Furthermore,

$$[P_0F(w_0)e_{w_0}\sqrt{1 - |w_0|^2}, B_{w_0,P_0}(z)F_1] = 0_{q\times q}, \tag{5.4}$$

and so (5.2) holds. □

Assume that in (5.1) $F_1 \not\equiv 0$. We can then reiterate and, after fixing $k_1 \in \{1, \ldots, p\}$, get a decomposition of the form (5.1) for F_1,

$$F_1(z) = P_1 F(w_1)e_{w_1}(z)\sqrt{1 - |w_1|^2} + B_{w_1,P_1}(z)F_2(z), \tag{5.5}$$

where w_1 is any complex number in the disc, and P_1 is any orthogonal projection of rank k_1. Thus F admits the orthogonal (also with respect to the $[\cdot,\cdot]$ form) decomposition (with $M_1 = P_1F(w_1)$)

$$\begin{aligned}F(z) &= M_0e_{w_0}(z)\sqrt{1-|w_0|^2}\\ &\quad + B_{w_0,P_0}(z)M_1e_{w_1}(z)\sqrt{1-|w_1|^2} + B_{w_0,P_0}(z)B_{w_1,P_1}(z)F_2(z)\end{aligned} \tag{5.6}$$

along the decomposition

$$\mathbf{H}_2^{p\times q} = \left(\mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}\right) \oplus B_{w_0,P_0}\left(\mathbf{H}_2^{p\times q} \ominus B_{w_1,P_1}\mathbf{H}_2^{p\times q}\right) \oplus B_{w_0,P_0}B_{w_1,P_1}\mathbf{H}_2^{p\times q}$$

of $\mathbf{H}_2^{p\times q}$. Note that

$$\left(\mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}\right) \oplus B_{w_0,P_0}\left(\mathbf{H}_2^{p\times q} \ominus B_{w_1,P_1}\mathbf{H}_2^{p\times q}\right) = \mathbf{H}_2^{p\times q} \ominus B_{w_0,P_0}B_{w_1,P_1}\mathbf{H}_2^{p\times q}.$$

Iterating the algorithm we get a family $F_0 = F, F_1, F_2, \ldots$ of functions in $\mathbf{H}_2^{p\times q}$ such that

$$F_k(z) = \left(B_{w_{k-1},P_{k-1}}(z)\right)^{-1}\left(F_{k-1}(z) - M_{k-1}e_{w_{k-1}}(z)\sqrt{1-|w_{k-1}|^2}\right), \ k=1,2,\ldots, \tag{5.7}$$

where at each stage one takes a projection such that $P_kF_k \not\equiv 0$. If there is no such projection it means that the algorithm ends at the given step.

The function F_k is called the k-th reduced remainder and is the matrix-valued analogue of (1.4). Let

$$\widetilde{B_0}(z) = P_0e_0(z) \quad \text{and} \quad \widetilde{B_k}(z) = \left(\prod_{u=0}^{k-1} B_{w_u,P_u}(z)\right) P_ke_{w_k}(z), \quad k = 1,2,\ldots; \tag{5.8}$$

setting $\widetilde{M_k} = M_k\sqrt{1-|w_k|^2} = P_kF(w_k)\sqrt{1-|w_k|^2}$, we have

$$F(z) = \sum_{k=0}^{N} \widetilde{B_k}(z)\widetilde{M_k} + \left(\prod_{u=0}^{N} B_{w_u,P_u}(z)\right) F_{N+1}(z). \tag{5.9}$$

Proposition 5.2. *If $F_{N+1}(z) = 0$, then the algorithm ends after N steps. In this case F is rational.*

Proof. Indeed, if the algorithm finishes after a finite number of steps, there is a finite Blaschke product B such that $F \in \mathbf{H}_2^{p\times q} \ominus B\mathbf{H}_2^{p\times q}$, and the elements of the latter space are rational functions. □

If our selections of w_k and P_k follow the maximum selection principle (that is, because of the choices of the point and the projection at each stage) we have the following result, which is the matrix-version of [32, Theorem 2.2]. The result is proved for $q = p$; the general case $q \neq p$ can be proved by taking the norm instead of the trace.

Theorem 5.3. *Let $F \in \mathbf{H}_2^{p\times p}$. Suppose that at each step one selects w_k and P_k according to the maximum selection principle. Then, the algorithm* (5.7) *converges, meaning that*

$$\lim_{N\to N_0} \mathrm{Tr}\,[F(z) - \sum_{k=0}^{N} \widetilde{B_k}(z)\widetilde{M_k},\ F(z) - \sum_{k=0}^{N} \widetilde{B_k}(z)\widetilde{M_k}] = 0,$$

where N_0 can be finite or infinite. In particular,

$$[F,F] = \sum_{k=0}^{N_0} [\widetilde{M_k}, \widetilde{M_k}], \tag{5.10}$$

where $\widetilde{M_k} = P_k F(w_k)\sqrt{1-|w_k|^2}$, $k = 0, 1, \ldots,$ *and*

$$\mathrm{Tr}\,[F,F] = \sum_{k=0}^{N_0} \mathrm{Tr}\,[\widetilde{M_k}, \widetilde{M_k}]. \tag{5.11}$$

Proof. The proof follows the proof for the scalar case presented in [32, Theorem 2.2]. Before proceeding, it is important to recall that the maximum (2.1) is computed on all projections of given rank and all points in $\mathbb{D}$. The case $N_0 < \infty$ means that the algorithm ceases after a finite number of steps. We then obtain a decomposition of F into a finite sum of rational matrices, and F is rational. We now suppose that $N_0 = \infty$. Let

$$G = F - \sum_{k=0}^{\infty} \widetilde{B_k}\widetilde{M_k} \not\equiv 0.$$

We proceed in a number of steps to get a contradiction.

Step 1. Using (5.9), we have that

$$[F,F] = \sum_{k=0}^{N} [\widetilde{M_k}, \widetilde{M_k}] + [F_{N+1}, F_{N+1}]$$

and

$$\widetilde{M_k} = [F, \widetilde{B_k}]. \tag{5.12}$$

It follows from the unitarity of the Blaschke factors B_{w_u,P_u} on the unit circle that

$$[\widetilde{B_k}(z), \widetilde{B_\ell}(z)] = \begin{cases} 0_{p\times p} & \text{if } k \neq \ell, \\ P_k & \text{if } k = \ell, \end{cases}$$

and the claim in the step follows.

Step 2. Let $R_k = F - \sum_{u=0}^{k-1} \widetilde{B_u}\widetilde{M_u}$. We have

$$[F, \widetilde{B_k}] = [R_k, \widetilde{B_k}] = [F_k, P_k e_{w_k}]. \tag{5.13}$$

The first equality in (5.13) follows from

$$[\widetilde{B_k}, \widetilde{B_u}] = 0_{p\times p} \quad \text{for } u = 0, \ldots, k-1. \tag{5.14}$$

The second equality follows using (5.9),

$$[R_k, \widetilde{B_k}] = [\left(\prod_{u=0}^{k-1} B_{w_u,P_u}\right) F_k, P_k e_{w_k} \left(\prod_{u=0}^{k-1} B_{w_u,P_u}\right)],$$

and the unitarity of the factors B_{w_u,P_u} on the unit circle.

Step 3. There exist a projection P, which we assume of rank one, and a point $b \in \mathbb{D}$ such that

$$\mathrm{Tr}\ [G, Pe_b] = \mathrm{Tr}\ (PG(b)\sqrt{1-|b|^2}) \neq 0.$$

In view of (5.14) the sum $\sum_{k=0}^{\infty}\widetilde{B_k}\widetilde{M_k}$ converges in $\mathbf{H}_2^{p\times p}$, and so G is analytic in the open unit disk. The claim in the step follows then from the analyticity of G inside the open unit disk and the fact that $G\not\equiv 0$.

Step 4. In the notation of the previous step, we have that there exists $k_0\in\mathbb{N}$ such that for all $k\geq k_0$,

$$\sqrt{1-|b|^2}|\mathrm{Tr}\,[PR_k(b)]|>\frac{|\mathrm{Tr}\ [G,Pe_b]|}{2}. \tag{5.15}$$

In view of (5.11) and using the Cauchy-Schwarz inequality we see that there exists $k_0\in\mathbb{N}$ such that for $k\geq k_0$,

$$|\mathrm{Tr}\sum_{u=k}^{\infty}[\widetilde{B_u}\widetilde{M_u},Pe_b]|<\frac{|\mathrm{Tr}\ [G,Pe_b]|}{2}.$$

Hence,

$$\begin{aligned}|\mathrm{Tr}\,[R_k,Pe_b]|+\frac{|\mathrm{Tr}\ [G,Pe_b]|}{2}&>|\mathrm{Tr}\,[R_k,Pe_b]|+|\mathrm{Tr}\,[\sum_{u=k}^{\infty}[\widetilde{B_u}\widetilde{M_u},Pe_b]|\\&\geq|\mathrm{Tr}\,[G,Pe_b]|,\end{aligned}$$

so that $|\mathrm{Tr}\,[R_k,Pe_b]|>\frac{|\mathrm{Tr}\ [G,Pe_b]|}{2}$. By the reproducing kernel property this inequality can be rewritten as (5.15).

Step 5. We conclude the proof.

By the Cauchy-Schwarz inequality, and since $B_{w_n,P_n}(b)^*B_{w_n,P_n}(b)\leq I_p$, we have

$$\begin{aligned}|\mathrm{Tr}\,[PR_k(b)]|&<(\mathrm{Tr}\,P)^{1/2}(\mathrm{Tr}\,(PR_k(b)^*R_k(b)P))^{1/2}\\&<(\mathrm{Tr}\,P)^{1/2}(\mathrm{Tr}\,(PF_k(b)^*F_k(b)P))^{1/2},\end{aligned}$$

and so

$$\sqrt{1-|b|^2}(\mathrm{Tr}\,P)^{1/2}(\mathrm{Tr}\,(PF_k(b)^*F_k(b)P))^{1/2}>\sqrt{1-|b|^2}|\mathrm{Tr}\,[PR_k(b)]|>\frac{|\mathrm{Tr}\ [G,Pe_b]|}{2}.$$

Since P has rank 1, it has trace equal to 1 and we have

$$\sqrt{1-|b|^2}(\mathrm{Tr}\,(PF_k(b)^*F_k(b)P))^{1/2}>\frac{|\mathrm{Tr}\ [G,Pe_b]|}{2}. \tag{5.16}$$

Equation (5.11) implies that $\lim_{k\to\infty}\widetilde{M_k}=0$. From (5.13) and (5.12) and the Cauchy-Schwarz inequality we have $\lim_{k\to\infty}[F_k,P_ke_{w_k}]=0$, and so

$$\lim_{k\to\infty}\sqrt{1-|w_k|^2}P_kF_k(w_k)=0_{p\times p},$$

and

$$\lim_{k\to\infty}(1-|w_k|^2)\mathrm{Tr}\,[P_kF_k(w_k),F_k(w_k)]=0,$$

and hence (5.16) is a contradiction to the maximum selection condition (2.1), since the maximum (2.1) is computed on all projections of given rank and all points in $\mathbb{D}$. □

Remark 5.4. In the above arguments one could also use normalized factors of the form (3.2). They are needed when one wishes the underlying Blaschke product to converge. In this case, when the algorithm does not end in a finite number of steps, two cases occur depending on whether the infinite matrix-valued Blaschke product

$$\mathcal{B}(z) \overset{\text{def.}}{=} \prod_{n=0}^{\infty} \mathcal{B}_{w_n,P_n}(z) = \lim_{N\to\infty} B_{w_0,P_0}(z)\mathcal{B}_{w_1,P_1}(z)\cdots B_{w_N,P_N}(z) \tag{5.17}$$

converges or not. The first case can be achieved by requiring the numbers a_n to satisfy $\sum_{n=0}^{\infty}(1-|a_n|)<\infty$ (see [32]). The infinite product (5.17) then converges for all $z\in\mathbb{D}$ (the proof is as in the scalar case; see for instance [14, TG IX.82] for infinite products in a normed algebra) and $F\in\mathbf{H}_2^{p\times q}\ominus\mathcal{B}\mathbf{H}_2^{p\times q}$. The second case then occurs when $\sum_{n=0}^{\infty}(1-|a_n|)=\infty$. In such a case an infinite Blaschke product cannot be defined, but instead, the shift invariant space reduces to zero, and the backward-shift invariant space coincides with the whole Hardy space $\mathbf{H}_2^{p\times q}$. The proof of this fact is based on the Beurling-Lax theorem. In fact, if the backward-shift invariant space does not coincide with the Hardy space $\mathbf{H}_2^{p\times q}$, then its orthogonal complement is a non-trivial shift invariant space. By the Beurling-Lax theorem the latter is of the form $\mathcal{B}\mathbf{H}_2^{p\times q}$, where $\mathcal{B}$ is the Blaschke product generated by the $w_k's$ and the $P_k's$. But this contradicts the condition $\sum_{n=0}^{\infty}(1-|a_n|)=\infty$.

Remark 5.5. Assume that the Blaschke product (5.17) converges. Then $F\in\mathbf{H}_2^{p\times q}\ominus\mathcal{B}\mathbf{H}_2^{p\times q}$. But this latter space is R_0-invariant, and so the subspace $\mathcal{M}(F)$ (see Proposition 4.2 for its definition) is such that

$$\mathcal{M}(F)\subset\mathbf{H}_2^{p\times q}\ominus\mathcal{B}\mathbf{H}_2^{p\times q}, \tag{5.18}$$

and F is not cyclic for R_0. Let Θ be the $\mathbb{C}^{p\times\ell}$-valued function as in Theorem 4.2. The isometric inclusion (5.18) implies that the kernel

$$\frac{\Theta(z)\Theta(w)^*-\mathcal{B}(z)\mathcal{B}(w)^*}{1-z\overline{w}}=\frac{I_p-\mathcal{B}(z)\mathcal{B}(w)^*}{1-z\overline{w}}-\frac{I_p-\Theta(z)\Theta(w)^*}{1-z\overline{w}}$$

is positive definite in $\mathbb{D}$. Leech's factorization theorem (see [2, 21, 24, 35]) implies that there is a $\mathbb{C}^{\ell\times p}$-valued function Θ_1 analytic and contractive in $\mathbb{D}$ and such that $\mathcal{B}(z)=\Theta(z)\Theta_1(z)$. Since $\mathcal{B}$ takes unitary values almost everywhere on the unit circle it follows that $\ell=p$.

6. The case of real signals

We begin with two definitions. Let $A=(a_{jk})_{\substack{j=1,\ldots,p\\k=1,\ldots,q}}\in\mathbb{C}^{p\times q}$. We say that A is real if the entries of A are real, that is, $A=\overline{A}$, where $\overline{A}$ is the matrix with (j,k)-entry $\overline{a_{jk}}$.

A matrix-valued real signal of finite energy is a function of the form

$$f(t)=A_0+\sum_{n=1}^{\infty}A_n\cos(nt)+B_n\sin(nt),$$

where the matrices A_n and B_n belong to $\mathbb{R}^{p\times q}$ and such that (with A^T denoting the transpose of the matrix A)

$$\operatorname{Tr}(A_0^TA_0)+\sum_{n=1}^{\infty}\operatorname{Tr}(A_n^TA_n+B_n^TB_n)<\infty.$$

Since

$$f(t)=A_0+\sum_{n=1}^{\infty}A_n\frac{e^{int}+e^{-int}}{2}+B_n\frac{e^{int}-e^{-int}}{2i}$$

we can rewrite $f(t)$ as

$$f(t)=F(e^{it})=\sum_{n\in\mathbb{Z}}F_ne^{int},$$

with

$$F_n=\begin{cases}A_0\,, & \text{if}\quad n=0,\\ \frac{A_n-iB_n}{2}\,, & \text{if}\quad n=1,2,\ldots,\\ \frac{A_n+iB_n}{2}\,, & \text{if}\quad n=-1,-2,\ldots.\end{cases}$$

Note that $F_{-n}=\overline{F_n}$. With these computations at hand we can prove (in the statement $\mathbb{T}$ denotes the unit circle):

Proposition 6.1. *Let $F\in\mathbf{L}_2^{p\times q}(\mathbb{T})$, with power series $F(e^{it})=\sum_{n\in\mathbb{Z}}F_ne^{int}$ and let $F_+(e^{it})=F_0+\sum_{n=1}^{\infty}F_ne^{int}$. Then, $F_+\in\mathbf{H}_2^{p\times q}$ and*

$$F(e^{it})=F_+(e^{it})+\overline{F_+(e^{it})}-F_0. \tag{6.1}$$

Proof. Let $F_-(e^{it})=\sum_{n=1}^{\infty}F_{-n}e^{-int}$. Since the Fourier coefficients are real we can write

$$\begin{aligned}\overline{F_+(e^{it})}&=F_0+\sum_{n=1}^{\infty}\overline{F_n}e^{-int}\\&=F_0+\sum_{n=1}^{\infty}F_{-n}e^{-int}\\&=F_0+F_-(e^{it}),\end{aligned}$$

and so (6.1) holds. □

The preceding result allows us to approximate real matrix-valued signals using the maximum selection principle algorithm presented in the previous sections.

7. Concluding remarks

The method developed in [28, 32] is extended here to the matrix-valued case. The results have impact on rational approximation and interpolation of matrix-valued functions. In a sequel to the present paper we may consider the case of the ball $\mathbb{B}_N$ of $\mathbb{C}^N$. Then the counterpart of Blaschke elementary factors exists (see [36]), and Blaschke products can be defined; see [7]. One has then to consider the Drury-Arveson space of the ball, that is, the reproducing kernel Hilbert space of functions analytic in $\mathbb{B}_N$ with reproducing kernel $\frac{1}{1-\sum_{j=1}^{N}z_j\overline{w_j}}$ rather than the Hardy space of the ball, whose reproducing kernel is $\frac{1}{(1-\sum_{j=1}^{N}z_j\overline{w_j})^N}$; see [9, 17]. We note that in the latter mentioned reproducing kernel Hilbert space, viz., the Hardy $\mathbf{H}_2$ space inside the ball, there exists the H^∞-functional calculus of the radial Dirac operator $\sum_{k=1}^{N}z_k\frac{\partial}{\partial z_k}$ or, equivalently, the singular integral operator algebra generalizing the Hilbert transformation on the sphere ([16]). More generally, one can consider complete Pick kernels, that is, positive definite kernels whose inverse has one positive square; see [10–12, 20, 34].

Acknowledgments

The authors are grateful to the anonymous referee, whose comments improved the readability of the paper. They are also grateful for the grant Macao Government FDCT 098.

References

[1] Daniel Alpay, *An advanced complex analysis problem book*, Topological vector spaces, functional analysis, and Hilbert spaces of analytic functions, Birkhäuser/Springer, Cham, 2015. MR3410523

[2] Daniel Alpay, Patrick Dewilde, and Harry Dym, *On the existence and construction of solutions to the partial lossless inverse scattering problem with applications to estimation theory*, IEEE Trans. Inform. Theory **35** (1989), no. 6, 1184–1205, DOI 10.1109/18.45275. MR1036623

[3] Daniel Alpay, Aad Dijksma, James Rovnyak, and Hendrik de Snoo, *Schur functions, operator colligations, and reproducing kernel Pontryagin spaces*, Operator Theory: Advances and Applications, vol. 96, Birkhäuser Verlag, Basel, 1997. MR1465432

[4] Daniel Alpay and Harry Dym, *On applications of reproducing kernel spaces to the Schur algorithm and rational J unitary factorization*, I. Schur methods in operator theory and signal processing, Oper. Theory Adv. Appl., vol. 18, Birkhäuser, Basel, 1986, pp. 89–159, DOI 10.1007/978-3-0348-5483-2_5. MR902603

[5] Daniel Alpay and Israel Gohberg, *On orthogonal matrix polynomials*, Orthogonal matrix-valued polynomials and applications (Tel Aviv, 1987–88), Oper. Theory Adv. Appl., vol. 34, Birkhäuser, Basel, 1988, pp. 25–46, DOI 10.1007/978-3-0348-5472-6_2. MR1021059

[6] Daniel Alpay and Israel Gohberg, *Unitary rational matrix functions*, Topics in interpolation theory of rational matrix-valued functions, Oper. Theory Adv. Appl., vol. 33, Birkhäuser, Basel, 1988, pp. 175–222, DOI 10.1007/978-3-0348-5469-6_5. MR960699

[7] Daniel Alpay and H. Turgay Kaptanoğlu, *Some finite-dimensional backward-shift-invariant subspaces in the ball and a related interpolation problem*, Integral Equations Operator Theory **42** (2002), no. 1, 1–21, DOI 10.1007/BF01203020. MR1866874

[8] Damir Z. Arov and Harry Dym, *J-contractive matrix valued functions and related topics*, Encyclopedia of Mathematics and its Applications, vol. 116, Cambridge University Press, Cambridge, 2008. MR2474532

[9] William Arveson, *Subalgebras of C^*-algebras. III. Multivariable operator theory*, Acta Math. **181** (1998), no. 2, 159–228, DOI 10.1007/BF02392585. MR1668582

[10] Joseph A. Ball and Vladimir Bolotnikov, *Nevanlinna-Pick interpolation for Schur-Agler class functions on domains with matrix polynomial defining function in $\mathbb{C}^n$*, New York J. Math. **11** (2005), 247–290 (electronic). MR2154356

[11] Joseph A. Ball, Vladimir Bolotnikov, and Quanlei Fang, *Schur-class multipliers on the Arveson space: de Branges-Rovnyak reproducing kernel spaces and commutative transfer-function realizations*, J. Math. Anal. Appl. **341** (2008), no. 1, 519–539, DOI 10.1016/j.jmaa.2007.10.033. MR2394102

[12] Joseph A. Ball, Tavan T. Trent, and Victor Vinnikov, *Interpolation and commutant lifting for multipliers on reproducing kernel Hilbert spaces*, Operator theory and analysis (Amsterdam, 1997), Oper. Theory Adv. Appl., vol. 122, Birkhäuser, Basel, 2001, pp. 89–138. MR1846055

[13] Harm Bart, Israel Gohberg, and Marinus A. Kaashoek, *Minimal factorization of matrix and operator functions*, Operator Theory: Advances and Applications, vol. 1, Birkhäuser Verlag, Basel-Boston, Mass., 1979. MR560504

[14] N. Bourbaki, *Topologie générale*, Chapitre 5 à 10, Diffusion C.C.L.S., Paris, 1974.

[15] R. R. Coifman and S. Steinerberger, *Nonlinear phase unwinding of functions*, preprint, arXiv:1508.01241.

[16] Michael Cowling and Tao Qian, *A class of singular integrals on the n-complex unit sphere*, Sci. China Ser. A **42** (1999), no. 12, 1233–1245, DOI 10.1007/BF02876023. MR1749933

[17] S. W. Drury, *A generalization of von Neumann's inequality to the complex ball*, Proc. Amer. Math. Soc. **68** (1978), no. 3, 300–304, DOI 10.2307/2043109. MR480362

[18] Harry Dym, *J contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*, CBMS Regional Conference Series in Mathematics, vol. 71, published for the Conference Board of the Mathematical Sciences, Washington, DC, by the American Mathematical Society, Providence, RI, 1989. MR1004239

[19] John B. Garnett, *Bounded analytic functions*, Pure and Applied Mathematics, vol. 96, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981. MR628971

[20] Devin C. V. Greene, Stefan Richter, and Carl Sundberg, *The structure of inner multipliers on spaces with complete Nevanlinna-Pick kernels*, J. Funct. Anal. **194** (2002), no. 2, 311–331, DOI 10.1006/jfan.2002.3928. MR1934606

[21] M. A. Kaashoek and J. Rovnyak, *On the preceding paper by R. B. Leech*, Integral Equations Operator Theory **78** (2014), no. 1, 75–77, DOI 10.1007/s00020-013-2108-7. MR3147403

[22] C. Lance, *Hilbert C^*-modules: A toolkit for operator algebraists*, London Math. Soc. Lecture Notes, 210, 1994.

[23] Peter D. Lax and Ralph S. Phillips, *Scattering theory*, 2nd ed., with appendices by Cathleen S. Morawetz and Georg Schmidt, Pure and Applied Mathematics, vol. 26, Academic Press, Inc., Boston, MA, 1989. MR1037774

[24] Robert B. Leech, *Factorization of analytic functions and operator inequalities*, Integral Equations Operator Theory **78** (2014), no. 1, 71–73, DOI 10.1007/s00020-013-2107-8. MR3147402

[25] S. Mallat and Z. Zhang, *Matching pursuit with time-frequency dictionaries*, IEEE Trans. Signal Proc. (1993), 3397–3415.

[26] Wen Mi, Tao Qian, and Feng Wan, *A fast adaptive model reduction method based on Takenaka-Malmquist systems*, Systems Control Lett. **61** (2012), no. 1, 223–230, DOI 10.1016/j.sysconle.2011.10.016. MR2878709

[27] V. P. Potapov, *The multiplicative structure of J-contractive matrix functions* (Russian), Trudy Moskov. Mat. Obšč. **4** (1955), 125–236. MR0076882

[28] Tao Qian, *Cyclic AFD algorithm for the best rational approximation*, Math. Methods Appl. Sci. **37** (2014), no. 6, 846–859, DOI 10.1002/mma.2843. MR3188530

[29] Tao Qian, *Adaptive Fourier decompositions and rational approximations, part I: Theory*, Int. J. Wavelets Multiresolut. Inf. Process. **12** (2014), no. 5, 1461008, 13, DOI 10.1142/S0219691314610086. MR3257806

[30] Tao Qian, Hong Li, and Michael Stessin, *Comparison of adaptive mono-component decompositions*, Nonlinear Anal. Real World Appl. **14** (2013), no. 2, 1055–1074, DOI 10.1016/j.nonrwa.2012.08.017. MR2991133

[31] Tao Qian, *Two-dimensional adaptive Fourier decomposition*, Math. Methods Appl. Sci. **39** (2016), no. 10, 2431–2448, DOI 10.1002/mma.3649. MR3512759

[32] Tao Qian and Yan-Bo Wang, *Adaptive Fourier series—a variation of greedy algorithm*, Adv. Comput. Math. **34** (2011), no. 3, 279–293, DOI 10.1007/s10444-010-9153-4. MR2776445

[33] Tao Qian, Yan-Bo Wang, and Pei Dang, *Adaptive decomposition into mono-components*, Adv. Adapt. Data Anal. **1** (2009), no. 4, 703–709, DOI 10.1142/S1793536909000278. MR2669206

[34] Peter Quiggin, *For which reproducing kernel Hilbert spaces is Pick's theorem true?*, Integral Equations Operator Theory **16** (1993), no. 2, 244–266, DOI 10.1007/BF01358955. MR1205001

[35] Marvin Rosenblum and James Rovnyak, *Hardy classes and operator theory*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1985. MR822228

[36] Walter Rudin, *Function theory in the unit ball of $\mathbf{C}^n$*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science], vol. 241, Springer-Verlag, New York-Berlin, 1980. MR601594

[37] V. N. Temlyakov, *Weak greedy algorithms*, Adv. Comput. Math. **12** (2000), no. 2-3, 213–227, DOI 10.1023/A:1018917218956. MR1745113

Department of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
E-mail address: dany@math.bgu.ac.il

Dipartimento di Matematica, Politecnico di Milano, Via E. Bonardi, 9, 20133 Milano, Italy
E-mail address: fabrizio.colombo@polimi.it

Department of Mathematics, University of Macau, Avenida da Universidade, Taipa, Macau
E-mail address: fsttq@umac.mo

Dipartimento di Matematica, Politecnico di Milano, Via E. Bonardi, 9, 20133 Milano, Italy
E-mail address: irene.sabadini@polimi.it