2017

Adaptive Orthonormal Systems for Matrix-Valued Functions

Daniel Alpay
Chapman University, alpay@chapman.edu

Fabrizio Colombo
Politecnico di Milano

Tao Qian
University of Macau

Irene Sabadini
Politecnico di Milano

Tao Qian
University of Macau

Recommended Citation
D. Alpay, F. Colombo, I. Sabadini and T. Qian. Adaptive orthonormal systems for matrix-valued functions. Proceedings of the AMS, vol. 145, number 5 (2017), no. 5, 2089-2106.

Comments
First published in *Proceedings of the AMS* in volume 145, issue 5, in 2017, published by the American Mathematical Society. DOI: 10.1090/proc/13359

ADAPTIVE ORTHONORMAL SYSTEMS FOR MATRIX-VALUED FUNCTIONS

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

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. We show that it is possible, as in the scalar case, to decompose those functions as linear combinations of suitably modified matrix-valued Blaschke product, 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.

AMS Classification: 47A56, 41A20, 30H10.
Key words: Matrix-valued functions and Hardy spaces, matrix-valued Blaschke products, maximum selection principle, adaptive decomposition.

Contents

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

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

where the points $a_n \in \mathbb{D}$ are adaptively chosen according to the function to be decomposed, see [26]. 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 Takenaka–Malmquist system. It is a basis of the Hardy space $\mathbf{H}_2(\mathbb{D})$ and, more in general, of $\mathbf{H}_p(\mathbb{D})$, $1 \le p \le \infty$, if and only if (1.2) is satisfied. It is possible to show, see [26], that the points a_n can be chosen to decompose a given function f into basic functions, each of which has nonnegative analytic instantaneous frequency. A system (1.1) satisfying this property is called an adaptive rational orthonormal system. A signal that possesses a nonnegative analytic instantaneous frequency function is said to be mono-component. It can be real or complex-valued. If, in particular, taking $a_1 = 0$, then the boundary values of the modified Blaschke products B_n

The authors thank Macao Science and Technology Fund FDCT/098/2012/A3 and Macao Science and Technology Fund FDCT/099/2014/A2.

are mono-component for all $n \in \mathbb{N}$. We note that the condition (1.2) is not necessarily satisfied, and so the system is not necessarily complete in $\mathbf{H}_2(\mathbb{D})$. However, the convergence to f is fast.

As we said, the adaptive decomposition is designed in order to obtain a decomposition of a functions into mono-component signals. This method has been intensively studied in the past few years, see [23, 25, 26, 27, 28, 29]. It gives rise to an algorithm which is a variation of the greedy algorithm, see [22, 34, 30].

An algorithm to perform an adaptive decomposition, given f, 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 [26] $\mathcal{H} = \mathbf{H}_2(\mathbb{D})$ with its standard inner product, 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 $a_1 \in \mathbb{D}$

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

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. 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

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

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

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

and

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

The function f_k is called the k-th reduced remainder (see [25, (11) p. 850]). 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 [26, 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 [26] to the matrix-valued case. For $w \in \mathbb{D}$ we will use the notations e_w and b_w for the normalized Cauchy kernel and 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 $P = P^2 = P^*$.

Remark 1.1. In view of the polydisk setting, the operator-valued case will be considered in a later publication (see [2, 3, 30] for an approach to the polydisk setting using operator-valued functions).

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

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

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

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 [6, 7] for a study of these 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 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 $k_0 \in \{1, \ldots, p\}$, 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*

$$(1 - |w_0|^2)\,(\mathrm{Tr}\ [P_0F(w_0), F(w_0)]) \ \text{ is maximum.} \tag{2.1}$$

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

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

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

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

Let $\epsilon > 0$. In view of (1.7)-(1.8) there exists a $\mathbb{C}^{p\times q}$-valued polynomial P such that $\|F - P\| \le \epsilon$. We have

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

This ends the proof since $(1 - |w|^2)\|P(w)\|^2$ tends to 0 as w approaches the unit circle and since $\xi^*\xi$ is a rank k_0 orthogonal projection. $\square$

We write

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

Lemma 2.2. *Let*

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

It holds 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 originate with the work of Potapov [24] and can be defined (up to 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 so

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

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 [17] for further information on matrix-valued Blaschke products. We also note that (3.4) is a special case of (4.4), and that the proposition can be viewed as a special case of Proposition 4.1.

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

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

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

$$B(z)B(w)^*=I_p-P_0+P_0b_{w_0}(z)\overline{b_{w_0}(w)},$$

and so

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

Equation (3.4) follows in 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 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 $\mathbf{H}_2^{p\times q}$. 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 an unitary rational function and the dimension of its associated reproducing kernel space; see for instance [5, 6] 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 holds 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. (3.3)

$$B_{w_0,P_0}^{-1}(z)=I_p-P_0+P_0\frac{1}{b_{w_0}(z)}.$$

Write $P_0H(z)=(z-w_0)R(z)$, where R is $\mathbb{C}^{k\times q}$-valued and analytic in a neighborhood of the origin. 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 $\cap_{u=0}^{\infty}\ker CA^u=\{0\}$ 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.

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. The first claim follows from the series expansion

$$C(I_N - zA) = \sum_{u=0}^{\infty} z^u CA^u,$$

and from the observability of the pair (C, A).

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). ☐

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 [24]; see e.g. [6] for a proof.

Note that $q = 1$ in the next proposition.

Proposition 4.2. *Let B be a $\mathbb{C}^{p\times p}$-valued Blaschke product. There is a one-to-one correspondence between factorizations $B = B_1B_2$ of B into two Blaschke products (up to a right multiplicative unitary constant U for B_1 and the corresponding left multiplicative constant U^{-1} for B_2) and R_0-invariant subspaces of $\mathbf{H}_2^p \ominus B\mathbf{H}_2^p$.*

The following proposition uses Beurling-Lax theorem (see [20]). 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 4.3. *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, and so its orthogonal complement $(\mathcal{M}(F))^{\perp}$ is invariant by multiplication by z. The result follows then from the Beurling-Lax theorem. ☐

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

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

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, there holds 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 and P_0 formula (2.4) can be rewritten in a unique way as an orthogonal sum (orthogonal also with respect to the $[\cdot,\cdot]$ form)*

$$F(z)=M_0e_{w_0}(z)\sqrt{1-|w_0|^2}+B_{w_0,P_0}(z)F_1(z),\tag{5.1}$$

where $M_0\in\mathbb{C}^{p\times q}$ and $F_1\in\mathbf{H}_2^{p\times q}$. We have $M_0e_{w_0}\sqrt{1-|w_0|^2}\in\left(\mathbf{H}_2^{p\times q}\ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}\right)$ and $B_{w_0,P_0}F_1\in B_{w_0,P_0}\mathbf{H}_2^{p\times q}$. Finally,

$$[F,F]=(1-|w_0|^2)[P_0F(w_0),F(w_0)]+[F_1,F_1].\tag{5.2}$$

Proof. We have

$$F(z)=M_0e_{w_0}(z)\sqrt{1-|w_0|^2}+B_{w_0,P_0}(z)F_1(z),\tag{5.3}$$

where $M_0=P_0F(w_0)$ 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 Lemma 3.1,

$$P_0F(w_0)e_{w_0}\sqrt{1-|w_0|^2}\in\mathbf{H}_2^{p\times q}\ominus B_{w_0,P_0}\mathbf{H}_2^{p\times q}.$$

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. □

Note that the decomposition (2.4) is non-trivial if and only if $F\not\equiv 0_{p\times q}$.
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_2,

$$F_1(z)=P_1F(w_1)e_{w_1}(z)\sqrt{1-|w_1|^2}+B_{w_1,P_1}(z)F_2(z),\tag{5.5}$$

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 $M_1=P_1F(w_1)$)

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

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

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

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

$$\widetilde{B_0}(z) = P_0e_0(z) \quad \text{and} \quad \widetilde{B_k}(z) = P_ke_k(z)\overset{\curvearrowleft}{\prod_{u=0}^{k-1}} B_{w_u,P_u}, \quad k = 1, 2, \ldots \tag{5.8}$$

We have

$$F(z) = \sum_{k=0}^{N} M_k\widetilde{B_k}(z) + B_{w_N,P_N}(z)F_{N+1}(z).$$

Proposition 5.2. *If $B_{w_N,P_N}(z)F_{N+1}(z) = 0$, then the algorithm ends up after N steps. In such 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 [26, Theorem 2.2].

Theorem 5.3. *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*

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

where N_0 can be finite or infinite. In particular,

$$[F, F] = \sum_{k=0}^{N_0}[M_k, M_k], \tag{5.9}$$

where $M_k = P_kF_k(w_k)$, $k = 0, 1, \ldots,$ and

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

Proof. The proof follows the proof for the scalar case presented in [26, 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 sum of finite entries, and F is rational. We now suppose that $N_0 = \infty$. Let

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

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

STEP 1: *It holds that*

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

with

$$M_k=[F,\widetilde{B_k}]. \tag{5.11}$$

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

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

and the claim in the step follows.

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

$$[F,\widetilde{B_k}]=[R_k,\widetilde{B_k}]=[F_k,P_ke_k] \tag{5.12}$$

The first equality in (5.12) follows from

$$[\widetilde{B_k},\widetilde{B_u}]=0_{p\times p}\quad \text{for } \ u=0,\dots,k-1. \tag{5.13}$$

The second equality follows from

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

and from 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*

$$\text{Tr }[G,Pe_b]=\text{Tr }(PG(b))\neq 0.$$

In view of (5.13) the sum $\sum_{k=0}^{\infty}M_k\widetilde{B_k}$ converges in $\mathbf{H}_2^{p\times q}$ and so G is analytic in the open unit disk. The claim in the step follows thenfrom the analyticity of G inside the open unit disk.

STEP 4: *In the notation of the previous step, we have*

$$\sqrt{1-|b|^2}|\text{Tr}\,[PR_k(b)]|>\frac{|\text{Tr }[G,Pe_b]|}{2}. \tag{5.14}$$

In view of (5.10), and using the Cauchy-Schwarz inequality we see that there is $k_0\in\mathbb{N}$ such that for all $k\geq k_0$,

$$|\text{Tr}\,[\sum_{u=k}^{\infty}[M_v\widetilde{B_v},Pe_b]|<\frac{|\text{Tr }[G,Pe_b]|}{2}.$$

Hence,

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

so that $|\text{Tr}\,[R_k,Pe_b]|>\frac{|\text{Tr }[G,Pe_b]|}{2}$. By the reproducing kernel property this inequality can be rewritten as (5.14).

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

$$|\text{Tr}\,[PR_k(b)]|<(\text{Tr}\,P)^{1/2}(\text{Tr}\,PR_k(b)^*R_k(b)P)^{1/2}<(\text{Tr}\,P)^{1/2}(\text{Tr}\,PF_k(b)^*F_k(b)P)^{1/2},$$

and so

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

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

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

Equation (5.10) implies that $\lim_{k\to\infty} M_k = 0$. From (5.12) and (5.11) and the Cauchy-Schwarz inequality we have $\lim_{k\to\infty}[F_k, P_k e_k] = 0$, and so

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

and so

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

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

Remark 5.4. In the above arguments one could also use normalized factors of the form (3.2). It is needed to use them when one wishes the underlying Blaschke product to converge. See Remark 5.6.

Remark 5.5. If F is a rational general function, and the maximum selection principle is used at each step, then the only possibility that the algorithm stops after a finite N_0 steps is the case when $N_0 = 1$. In the case, the subspace $\mathcal{M}(F)$ of $\mathbf{H}_2^{p\times q}$ spanned by the functions $R_0^u FX$ where $u = 0, 1, \ldots$ and $X \in \mathbb{C}^{q\times q}$ is finite dimensional and R_0-invariant by construction. So it is of the form $\mathbf{H}_2^{p\times q} \ominus B\mathbf{H}_2^{p\times q}$ for some finite Blaschke product B. This does not mean that a backward-shift algorithm is then performed inside this space, and thus that it would end after a finite number of steps. In contrast, for a rational function, if we do not use the maximum selection principle but suitably select w_k and P_k, the algorithm can well stop after a finite N_0 steps, as concerned by Proposition 5.2.

Remark 5.6. We now consider the case where we take normalized Blaschke factors (see Remark 5.4). When the algorithm does not end in a finite number of steps, two cases occur depending on whether the infinite matrix-valued Blaschke product

$$\mathcal{B}(z) \stackrel{\text{def.}}{=} \prod_{n=0}^{\overset{\curvearrowleft}{\infty}} \mathcal{B}_{w_n,P_n}(z) = \lim_{N\to\infty}\mathcal{B}_{w_N,P_N}(z)\cdots\mathcal{B}_{w_1,P_1}(z)\mathcal{B}_{w_0,P_0}(z) \tag{5.15}$$

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 [26]). The infinite product (5.15) 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 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 $H_2^{p\times q}$ space. 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 $H_2^{p\times q}$ space, 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 with the condition $\sum_{n=0}^{\infty}(1-|a_n|) = \infty$.

Remark 5.7. Assume that the Blaschke product (5.15) 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

$$\mathcal{M}(F) \subset \mathbf{H}_2^{p\times q} \ominus \mathcal{B}\mathbf{H}_2^{p\times q}, \tag{5.16}$$

and F is not cyclic for R_0. Let Θ be the the $\mathbb{C}^{p\times\ell}$-valued function as in Theorem 4.3. The isometric inclusion (5.16) 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 [32], [21], [19], [4]) 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)

$$\mathrm{Tr}\,(A_0^T A_0) + \sum_{n=1}^{\infty} \mathrm{Tr}\,(A_n^T A_n + B_n^T B_n) < \infty.$$

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 $f(t)$

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

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 state (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 $F(e^{it}) = \sum_{n\in\mathbb{Z}} F_n e^{int}$ and let $F_+(e^{it}) = F_0 + \sum_{n=1}^{\infty} F_n e^{int}$. 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 to approximate real matrix-valued signals using the maximum selection principle algorithm presented in the previous sections.

7. Concluding remarks

The method developed in [25, 26] is extended here to the matrix-valued case. The results have impacts to 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 [33]), and Blaschke products can be defined; see [8]. 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, 16]. We note that in the later 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

([15]). More generally, one can consider complete Pick kernels, that is positive definite kernels whose inverse has one positive square, see [10, 11, 12, 18, 31].

References

[1] D. Alpay. *An advanced complex analysis problem book. Topological vector spaces, functional analysis, and Hilbert spaces of analytic functions*. Birkäuser Basel, 2015.

[2] D. Alpay, V. Bolotnikov, and L. Rodman. One-Sided Tangential Interpolation for Operator-Valued Hardy Functions on Polydisks. *Integral Equations and Operator Theory*, 35:253–270, 1999.

[3] D. Alpay, V. Bolotnikov, and L. Rodman. Two-sided tangential interpolation for Hilbert-Schmidt operator functions on polydisks. In *Recent advances in operator theory (Groningen, 1998)*, volume 124 of *Oper. Theory Adv. Appl.*, pages 21–62. Birkhäuser, Basel, 2001.

[4] D. Alpay, P. Dewilde, and H. Dym. On the existence and construction of solutions to the partial lossless inverse scattering problem with applications to estimation theory. *IEEE Trans. Inform. Theory*, 35:1184–1205, 1989.

[5] D. Alpay and H. Dym. On applications of reproducing kernel spaces to the Schur algorithm and rational J-unitary factorization. In I. Gohberg, editor, *I. Schur methods in operator theory and signal processing*, volume 18 of *Operator Theory: Advances and Applications*, pages 89–159. Birkhäuser Verlag, Basel, 1986.

[6] D. Alpay and I. Gohberg. *On orthogonal matrix polynomials*, volume 34 of *Operator Theory: Advances and Applications*, pages 25–46. Birkhäuser Verlag, Basel, 1988.

[7] D. Alpay and I. Gohberg. Unitary rational matrix functions. In I. Gohberg, editor, *Topics in interpolation theory of rational matrix-valued functions*, volume 33 of *Operator Theory: Advances and Applications*, pages 175–222. Birkhäuser Verlag, Basel, 1988.

[8] D. Alpay and H.T. Kaptanoğlu. Some finite-dimensional backward shift-invariant subspaces in the ball and a related interpolation problem. *Integral Equation and Operator Theory*, 42:1–21, 2002.

[9] W. Arveson. Subalgebras of C^*–algebras. III. Multivariable operator theory. *Acta Mathematica*, 181:159–228, 1998.

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

[11] J. Ball, V. Bolotnikov, and Q. Fang. Schur-class multipliers on the Arveson space: de Branges-Rovnyak reproducing kernel spaces and commutative transfer-function realizations. *J. Math. Anal. Appl.*, 341(1):519–539, 2008.

[12] J. Ball, T. Trent, and V. Vinnikov. Interpolation and commutant lifting for multipliers on reproducing kernel Hilbert spaces. In *Proceedings of Conference in honor of the 60–th birthday of M.A. Kaashoek*, volume 122 of *Operator Theory: Advances and Applications*, pages 89–138. Birkhauser, 2001.

[13] H. Bart, I. Gohberg, and M.A. Kaashoek. *Minimal factorization of matrix and operator functions*, volume 1 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 1979.

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

[15] M. Cowling and T. Qian. *A class of singular integralson the n-complex unit sphere*. Scientia Sinica (Series A), vol 42, No.12 (December 1999), 1233-1245.

[16] S. W. Drury. A generalization of von Neumann's inequality to the complex ball. *Proc. Amer. Math. Soc.*, 68(3):300–304, 1978.

[17] H. Dym. J contractive matrix functions, reproducing kernel Hilbert spaces and interpolation. CBMS Regional Conference Series in Mathematics, 71. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1989.

[18] D.C.V. Greene, S. Richter, and C. Sundberg. The structure of inner multipliers on spaces with complete Nevanlinna-Pick kernels. *J. Funct. Anal.*, 194(2):311–331, 2002.

[19] M. A. Kaashoek and J. Rovnyak. On the preceding paper by R. B. Leech. *Integral Equations Operator Theory*, 78(1):75–77, 2014.

[20] P. D. Lax and R. S. Phillips. *Scattering theory (revised edition)*, volume 26 of *Pure and Applied Mathematics*. Academic Press, New–York, 1989.

[21] R. B. Leech. Factorization of analytic functions and operator inequalities. *Integral Equations Operator Theory*, 78(1):71–73, 2014.

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

[23] W. Mi, T. Qian, and F. Wan. A fast adaptive model reduction method based on Takenaka-Malmquist systems. *Systems Control Lett.*, 61(1):223–230, 2012.

[24] V.P. Potapov. The multiplicative structure of J-contractive matrix–functions. *Trudy Moskow. Mat. Obs.*, 4:125–236, 1955. English translation in: American mathematical society translations (2), vol. 15, p. 131–243 (1960).

[25] T. Qian. Cyclic AFD algorithm for the best rational approximation. *Math. Methods Appl. Sci,*, 37(6):846–859, 2014.

[26] T. Qian and Y. B. Wang. Adaptive Decomposition Into Basic Signals of Non-negative Instantaneous Frequencies - A Variation and Realization of Greedy Algorithm. *Adv. Comput. Math.*, 34(3):279–293, 2011.

[27] T. Qian. Adaptive Fourier decompositions and rational approximations, part I: Theory. *Int. J. Wavelets Multiresolut. Inf. Process.*, 12(5):1461008, 13, 2014.

[28] T. Qian, H. Li, and M. Stessin. Comparison of adaptive mono-component decompositions. *Nonlinear Anal. Real World Appl.*, 14(2):1055–1074, 2013.

[29] T. Qian, Y.-B. Wang, and P. Dang. Adaptive decomposition into mono-components. *Adv. Adapt. Data Anal.*, 1(4):703–709, 2009.

[30] T. Qian and W. Sproessig, Two-Dimensional Adaptive Fourier Decomposition. Accepted to appear in *Math. Methods Appl. Sci.*, arXiv.org math arXiv:1406.2088.
[31] P. Quiggin. For which reproducing kernel Hilbert spaces is Pick's theorem true? *Integral Equations Operator Theory*, 16:244–266, 1993.
[32] M. Rosenblum and J. Rovnyak. *Hardy classes and operator theory*. Birkhäuser Verlag, Basel, 1985.
[33] W. Rudin. *Function theory in the unit ball of* $\mathbb{C}^n$. Springer–Verlag, 1980.
[34] V.N. Temlyakov. *Weak greedy algorithm*. Advances in Comout. Math., 12 (2000), 213-227.

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

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

(TQ) Department of Mathematics, Universityof Macau, Macau
E-mail address: fsttq@umac.mo

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