MIMO BIORTHOGONAL PARTNERS AND APPLICATIONS*

Bojan Vrcelj and P. P. Vaidyanathan
Contact Author: P. P. Vaidyanathan, Department of Electrical Engineering 136-93,
California Institute of Technology, Pasadena, CA 91125 USA,
Phone: (626) 395-4681 E-mail: ppvnath@systems.caltech.edu

October 30, 2001

EDICS number: 2-MWAV

ABSTRACT

Multiple Input Multiple Output (MIMO) biorthogonal partners arise in many different contexts, one of them being multiwavelet theory. They also play a central role in the theory of MIMO channel equalization, especially with fractionally spaced equalizers. In this paper we first derive some theoretical properties of MIMO biorthogonal partners. We develop conditions for the existence of MIMO biorthogonal partners and conditions under which FIR solutions are possible. In the process of constructing FIR MIMO biorthogonal partners we exploit the non-uniqueness of the solution. This will lead to the design of flexible *fractionally spaced* MIMO zero-forcing equalizers. The additional flexibility in design makes these equalizers more robust to channel noise. Finally, other situations where MIMO biorthogonal partners occur will also be considered, such as prefiltering in multiwavelet theory and deriving the vector version of the least squares signal projection problem.

*Work supported in part by the National Science Foundation under grant MIP 0703755.

1 Introduction

Digital filters $H(z)$ and $F(z)$ are called biorthogonal partners of each other with respect to an integer M if their cascade $H(z)F(z)$ obeys the Nyquist(M) property [20]. The theory of biorthogonal partners was developed recently in [20] for the simple, single input single output (SISO) case. Multiple input multiple output (MIMO) biorthogonal partners are defined using a similar approach [24, 23]. However, in the MIMO case the "biorthogonal partner" relation is not symmetric, so we distinguish between a *left biorthogonal partner* (**LBP**) and a *right biorthogonal partner* (**RBP**). In this paper we first derive some theoretical properties of MIMO biorthogonal partners. Many of these properties are extensions to the vector case of some known results from the case of scalar signals [20]. However, some of the properties take a different form in the case of vector signals and, furthermore, lead to some new applications. One of the applications of MIMO biorthogonal partners that will be explored in this paper is the equalization of vector digital communication channels. Specifically, we will be interested in zero-forcing fractionally-spaced MIMO equalizers. Fractionally-spaced equalizers (FSE) demonstrate many advantages over symbol-spaced equalizers (SSE), such as the existence of an FIR solution and reduced sensitivity to the shift in sampling instances [11]. Moreover, the FSE turns out to be a particular form of a biorthogonal partner of the equivalent channel transfer matrix and therefore we can resort to the theory of MIMO biorthogonal partners in the process of designing fractionally spaced equalizers.

1.1 Paper Outline and Relation to Past Work

In Sec. 2 we introduce the precise definition of MIMO biorthogonal partners. We derive a general closed form expression for a MIMO transfer function $\mathbf{H}(z)$ to be a biorthogonal partner of $\mathbf{F}(z)$. We also derive a set of necessary and sufficient conditions on $\mathbf{F}(z)$ which allow for the existence of its MIMO biorthogonal partner.

In Sec. 3 we consider FIR MIMO biorthogonal partners in greater detail. A set of necessary and sufficient conditions for the existence of FIR MIMO biorthogonal partners will be derived. We will also concentrate on the fact that the FIR MIMO biorthogonal partner (if it exists) is not unique. In Sec. 4 we will exploit this non-uniqueness in order to reduce the noise power at the output of fractionally spaced equalizers for vector channels. Finally, we will address the performance of our algorithms through simple examples of fractionally spaced equalizers (**FSE**) for vector channels.

In Sec. 5 we deal with other possible applications of MIMO biorthogonal partners. In particular, we will review their role in the least squares approximation of vector signals. In the context of this paper the least squares problem will be limited to that of finding the approximation for a vector signal $\mathbf{x}(n)$ within a certain signal model. In the scalar case the idea originated in the context of spline interpolation [3], where it was suggested that the signal corrupted by noise could be approximated within the model of oversampled splines. We show that in the vector signal case the solution to this problem involves a particular form of MIMO biorthogonal partners. This work is also closely related to the concept of oblique projections studied intensively in [1] and [2]. Finally, we will consider the relation between biorthogonal partners and multiwavelets. In one of the pioneering works on the subject [26], the use of prefiltering for multiwavelet transform was introduced. In this contribution we consider the prefiltering problem in the light of biorthogonal partners and draw the connection between the two. Portions of this paper have been presented at the ICASSP 2001 and ICC 2001 conferences [23, 24].

1.2 Notations

If not stated otherwise, all notations are as in [19]. We use the notation $[x(n)]_{\downarrow M}$ and $[X(z)]_{\downarrow M}$ to denote the decimated version $x(Mn)$ and its z-transform. The expanded version

$$\begin{cases} x(n/M) & \text{for } n = \text{mul of } M, \\ 0 & \text{otherwise} \end{cases}$$

is similarly denoted by $[x(n)]_{\uparrow M}$, and its z-transform $X(z^M)$ denoted by $[X(z)]_{\uparrow M}$. In a block diagram, the scalar decimation and expansion operations will be denoted by encircled symbols $\downarrow M$ and $\uparrow M$ respectively. In the case of vectors and matrices, the decimation and expansion are performed on each element separately. The corresponding vector sequence decimation/expansion symbols are placed in square boxes as in Fig. 2.

The polyphase decomposition [19] is also valid in the matrix case. Thus for example if $\mathbf{F}(z)$ is a matrix transfer function, then it can be written in the Type-2 polyphase form as

$$\mathbf{F}(z) = \sum_{k=0}^{M-1} z^k \mathbf{F}_k(z^M). \tag{1}$$

If not mentioned otherwise, all the matrices in this paper are rectangular. It is implicit that their dimensions are such that the matrix products in question are well defined and that the product matrix has the appropriate size.

2 MIMO Biorthogonal Partners: Definition and Properties

We start the discussion in this section by defining the notion of a MIMO biorthogonal partner.

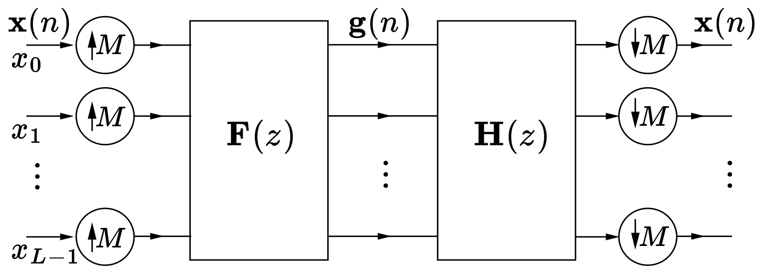

Figure 1: Block diagram interpretation of a left biorthogonal partner.

Definition 1. MIMO Biorthogonal partners. A MIMO transfer function $\mathbf{H}(z)$ is said to be a left biorthogonal partner (LBP) of $\mathbf{F}(z)$ with respect to an integer M if

$$[\mathbf{H}(z)\mathbf{F}(z)]_{\downarrow M} = \mathbf{I} \tag{2}$$

Similarly, a MIMO transfer function $\mathbf{H}(z)$ is said to be a right biorthogonal partner (RBP) of $\mathbf{F}(z)$ with respect to an integer M if $[\mathbf{F}(z)\mathbf{H}(z)]_{\downarrow M} = \mathbf{I}$.

The interpretation of the first part of the above definition is shown in Fig. 1. Recall that the multirate system in Fig. 1 is just an LTI system with transfer function $[\mathbf{H}(z)\mathbf{F}(z)]_{\downarrow M}$, which under the condition (2) becomes the identity. It can be seen that if $\mathbf{H}(z)$ is a LBP of $\mathbf{F}(z)$ this implies that $\mathbf{F}(z)$ is a RBP of $\mathbf{H}(z)$, but it does not imply that $\mathbf{H}(z)$ is also a RBP of $\mathbf{F}(z)$. The latter would happen if, for example, the two matrices commuted. The other important point to make here is that if M is changed, the two filters might not remain partners. However, we will often omit the term "with respect to M", since it will usually be understood from the context.

In the following we concentrate on the issues of existence and the general form of MIMO biorthogonal partners. The first result gives the most general form of a biorthogonal partner. In the subsequent discussion, the question of uniqueness of biorthogonal partners will also be addressed. The second result states necessary and sufficient conditions on a transfer matrix $\mathbf{F}(z)$ and integer M such that there exists a biorthogonal partner of $\mathbf{F}(z)$ with respect to M.

2.1 General Expression

We first derive a general expression for $\mathbf{H}(z)$ in terms of $\mathbf{F}(z)$ in Fig. 1. The theorem has two parts, one for left biorthogonal partners and the other for right biorthogonal partners. It is very intuitive that whatever holds for LBPs should also hold for RBPs (in a slightly modified form), and this comes into play in the proof of the Theorem 1.

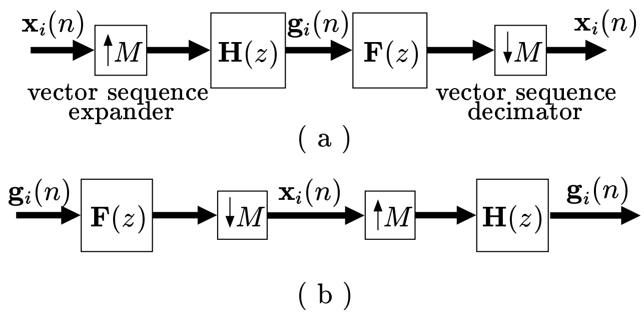

Figure 2: Pertaining to the proof of Theorem 1.

Theorem 1. General form of biorthogonal partner.

1. A MIMO transfer function $\mathbf{H}(z)$ is a LBP of $\mathbf{F}(z)$ if and only if it can be written in the form

$$\mathbf{H}(z) = \left([\mathbf{G}(z)\mathbf{F}(z)]_{\downarrow M\uparrow M}\right)^{-1}\mathbf{G}(z) \tag{3}$$

   for some MIMO transfer function $\mathbf{G}(z)$.

2. A MIMO transfer function $\mathbf{H}(z)$ is a RBP of $\mathbf{F}(z)$ if and only if it can be written in the form

$$\mathbf{H}(z) = \mathbf{G}(z)\left([\mathbf{F}(z)\mathbf{G}(z)]_{\downarrow M\uparrow M}\right)^{-1} \tag{4}$$

   for some MIMO transfer function $\mathbf{G}(z)$.

Proof. First we will prove the "if part" of the statement one. Given $\mathbf{H}(z)$ as in (3), we have

$$[\mathbf{H}(z)\mathbf{F}(z)]_{\downarrow M} = [\left([\mathbf{G}(z)\mathbf{F}(z)]_{\downarrow M\uparrow M}\right)^{-1}\mathbf{G}(z)\mathbf{F}(z)]_{\downarrow M} = \left([\mathbf{G}(z)\mathbf{F}(z)]_{\downarrow M}\right)^{-1}[\mathbf{G}(z)\mathbf{F}(z)]_{\downarrow M} = \mathbf{I}$$

The "if part" of the second statement follows in the same manner. Now we will prove the "only if part" of the **second** statement. For this, first consider Fig. 2(a). Here $\mathbf{x}_i(n)$ is an arbitrary vector sequence and $\mathbf{g}_i(n)$ is the corresponding output of $\mathbf{H}(z)$. By assumption $\mathbf{H}(z)$ is a RBP of $\mathbf{F}(z)$ and from the definition we have that the output of the system has to be $\mathbf{x}_i(n)$ again. However, this also means that the signal $\mathbf{g}_i(n)$ when input to the system in Fig. 2(b) comes out as $\mathbf{g}_i(n)$. Thus we have

$$\mathbf{H}(z)[\mathbf{F}(z)\mathbf{G}_i(z)]_{\downarrow M\uparrow M} = \mathbf{G}_i(z). \tag{5}$$

This equality holds for any $\mathbf{G}_i(z)$ obtained as in Fig. 2(a). We repeat the procedure sufficient number of times, each time taking $\mathbf{X}_n(z)$ to be linearly independent from the previous vectors $\mathbf{X}_1(z)$, $\mathbf{X}_2(z)$, ... $\mathbf{X}_{n-1}(z)$. Collecting those vectors as columns in a matrix $\mathbf{X}(z)$, and the corresponding vectors $\mathbf{G}_i(z)$ in a matrix $\mathbf{G}(z)$, we have the following

$$\mathbf{H}(z)[\mathbf{F}(z)\mathbf{G}(z)]_{\downarrow M\uparrow M} = \mathbf{G}(z)$$

which after solving for $\mathbf{H}(z)$ gives

$$\mathbf{H}(z) = \mathbf{G}(z) \left([\mathbf{F}(z)\mathbf{G}(z)]_{\downarrow M \uparrow M} \right)^{-1} \tag{6}$$

and this concludes the proof of (4). Notice that $[\mathbf{F}(z)\mathbf{G}(z)]_{\downarrow M\uparrow M} = [\mathbf{X}(z)]_{\uparrow M}$ so that by choosing the sequences $\mathbf{x}_i(n)$ carefully we can ensure that the matrix inversion in (6) is valid. Now we move on to prove the "only if part" of the first statement. For this we notice that if $\mathbf{H}(z)$ is a LBP of $\mathbf{F}(z)$, then $\mathbf{H}^T(z)$ is a RBP of $\mathbf{F}^T(z)$, with the superscript T denoting the transpose of a matrix. Thus from (4) we have

$$\mathbf{H}^T(z) = \mathbf{G}^T(z) \left([\mathbf{F}^T(z)\mathbf{G}^T(z)]_{\downarrow M\uparrow M} \right)^{-1},$$

for some matrix $\mathbf{G}^T(z)$. Finally, taking the transpose of both sides we arrive at (3) and this concludes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

In the proof of Theorem 1 we used the idea of "transposing the result" for RBP in order to prove a similar result for LBP. The same trick could also be used for the remaining results in the paper. That is why we will consider only left biorthogonal partners in the following; very similar results hold for right biorthogonal partners.

Clearly, from the equations (3) and (4) we have that MIMO biorthogonal partners are in general **not unique**. Any stable transfer matrix $\mathbf{G}(e^{j\omega})$ such that $\det\left([\mathbf{G}(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}\right)$ is nonzero for all ω gives rise to a stable LBP of $\mathbf{F}(z)$. The similar conclusion holds for right biorthogonal partners. Here are some special cases of interest.

Example 1. In the square case, if $\left|\det\left[\mathbf{F}(e^{j\omega})\right]\right| > 0$ for all ω then $\mathbf{H}(z) = \mathbf{F}^{-1}(z)$ is a theoretically stable biorthogonal partner (both LBP and RBP) of $\mathbf{F}(z)$. It can be obtained from (3) or (4) with the choice $\mathbf{G}(z) = \mathbf{F}^{-1}(z)$. This is conceptually the simplest biorthogonal partner.

Example 2. If the construction of biorthogonal partners from Example 1 does not work for a particular $\mathbf{F}(z)$, we can try the following. Suppose that $\det\left[[\mathbf{F}(e^{j\omega})]_{\downarrow M}\right]$ is nonzero for all ω. This condition is easily verified to be looser than the one in the previous example. Then, substituting $\mathbf{G}(z) = \mathbf{I}$ in (3) or (4) we get a biorthogonal partner $\mathbf{H}(z) = \left([\mathbf{F}(z)]_{\downarrow M\uparrow M}\right)^{-1}$.

Example 3. To get yet another solution for a LBP, consider the matrix filter

$$\mathbf{H}(z) = \left([\tilde{\mathbf{F}}(z)\mathbf{F}(z)_{\downarrow M\uparrow M} \right)^{-1} \tilde{\mathbf{F}}(z),$$

where $\tilde{\mathbf{F}}(z) = \mathbf{F}^{\dagger}(1/z^*)$. This solution is obtained from (3) with $\mathbf{G}(z) = \tilde{\mathbf{F}}(z)$, and is valid as long as $\det\left(\mathbf{F}^{\dagger}(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}\right)$ is nonzero on the unit circle. In the rest of the paper this solution will play a significant role, because it occurs in several different contexts.

2.2 Existence

In the following, we look into the problem of the existence of biorthogonal partners more closely. We present a necessary and sufficient condition on a MIMO transfer function $\mathbf{F}(z)$ for the existence of its MIMO biorthogonal partner $\mathbf{H}(z)$. Throughout this paper by "existence of a biorthogonal partner" we actually mean "existence of a stable biorthogonal partner".

Theorem 2. Existence of LBP. A MIMO transfer function $\mathbf{F}(z)$ with the Type-2 polyphase form as in (1) has a LBP if and only if the following implication holds for each ω in $0 \le \omega < 2\pi$

$$\mathbf{C}^T(e^{j\omega})[\mathbf{F}_0^T(e^{j\omega}) \;\; \mathbf{F}_1^T(e^{j\omega}) \;\; \cdots \;\; \mathbf{F}_{M-1}^T(e^{j\omega})] = \mathbf{0} \quad \Rightarrow \quad \mathbf{C}(e^{j\omega}) = \mathbf{0}.$$

Therefore, for any fixed ω there cannot exist a nonzero common annihilating vector $\mathbf{C}(e^{j\omega})$ for all the M polyphase components of $\mathbf{F}(e^{j\omega})$. Note that in order for $\mathbf{F}(z)$ to have an *inverse* we need to have $\det[\mathbf{F}(e^{j\omega})] \neq 0$, for all ω, and that condition is stricter than the one in Theorem 2.

Proof. We start by proving the forward part of the theorem, i.e. supposing $\mathbf{H}(z)$ is a stable LBP of $\mathbf{F}(z)$, we need to show that there cannot exist a nonzero common annihilating vector $\mathbf{C}(e^{j\omega})$. By the supposition we have that $[\mathbf{H}(z)\mathbf{F}(z)]_{\downarrow M} = \mathbf{I}$ and that implies that there cannot exist a nonzero vector $\mathbf{C}(z)$ such that $\mathbf{F}(z)\mathbf{C}(z^M) = \mathbf{0}$. Indeed, if we assume there exists such nonzero vector $\mathbf{C}(z)$, we end up with the following contradiction

$$\mathbf{0} = [\mathbf{H}(z)\mathbf{F}(z)\mathbf{C}(z^M)]_{\downarrow M} = \mathbf{C}(z).$$

Rewriting $\mathbf{F}(z)$ in the Type-2 polyphase form (1) we then have that there cannot exist a nonzero vector $\mathbf{C}(z)$ such that

$$\sum_{k=0}^{M-1} z^k \mathbf{F}_k(z^M)\mathbf{C}(z^M) = \mathbf{0}$$

or equivalently, such that

$$\mathbf{F}_k(z)\mathbf{C}(z) = \mathbf{0} \quad \forall k, \quad 0 \le k \le M-1.$$

Therefore, if there exists a stable LBP of $\mathbf{F}(z)$, then there cannot exist a common nonzero annihilating vector $\mathbf{C}(e^{j\omega})$ for all the M polyphase components $\mathbf{F}_k(e^{j\omega})$.

Now we proceed to prove the converse. To that end we suppose that for no ω does there exist a common nonzero vector $\mathbf{C}(e^{j\omega})$ that annihilates all the M polyphase components $\mathbf{F}_k(e^{j\omega})$. That is to say, given any $\omega \in [0, 2\pi)$ such that $\mathbf{C}(e^{j\omega}) \neq \mathbf{0}$, there exists k such that $\mathbf{F}_k(e^{j\omega})\mathbf{C}(e^{j\omega}) \neq \mathbf{0}$. This implies that the following matrix $\mathbf{S}(\omega)$ is positive definite for all ω

$$\mathbf{S}(\omega) = \sum_{k=0}^{M-1} \mathbf{F}_k^\dagger(e^{j\omega})\mathbf{F}_k(e^{j\omega}). \tag{7}$$

To justify this, observe that for any vector $\mathbf{C}(e^{j\omega})$ and $\mathbf{S}(\omega)$ as in (7), the entity $\mathbf{C}^\dagger(e^{j\omega})\mathbf{S}(\omega)\mathbf{C}(e^{j\omega})$ is a summation of nonnegative terms. Moreover, as asserted previously, for any nonzero choice of $\mathbf{C}(e^{j\omega})$ at least one of those terms is strictly positive, so that the overall result is positive. Now we observe that the matrix $\mathbf{S}(\omega)$ defined by (7) can be rewritten as $\mathbf{S}(\omega) = [\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}$ and from the previous discussion we have that

$$\det\left([\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}\right) > 0. \tag{8}$$

The final conclusion is that if there does not exist a common nonzero annihilating vector $\mathbf{C}(e^{j\omega})$ for all the M polyphase components $\mathbf{F}_k(e^{j\omega})$ then $\mathbf{F}(z)$ has a stable LBP. In particular, one such LBP is obtained as in Example 3 and is given by

$$\mathbf{H}(z) = \left([\tilde{\mathbf{F}}(z)\mathbf{F}(z)_{\downarrow M}\right)^{-1}_{\uparrow M}\tilde{\mathbf{F}}(z). \tag{9}$$

This conclues the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

In the following we will see that the LBP given by (9) has some other interesting properties. The next corollary asserts that if $\mathbf{F}(z)$ has any LBP, the choice (9) will be a valid one.

Corollary 1. A MIMO transfer function $\mathbf{F}(z)$ has a left biorthogonal partner if and only if $\mathbf{S}(\omega) = [\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}$ is a positive definite matrix for all ω in the range $[0, 2\pi)$.

Proof. If $[\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}$ is a positive definite matrix for all ω then (8) holds and thus (9) is a valid choice for LBP. Conversely, suppose that there exists a stable LBP of $\mathbf{F}(z)$. Consider $\mathbf{S}(\omega)$, which is obviously a positive semi-definite matrix for all ω. Writing $\mathbf{S}(\omega)$ as in (7) and recalling from Theorem 1 that the polyphase components $\mathbf{F}_k(e^{j\omega})$ cannot have a common annihilating vector we finally conclude that $\mathbf{S}(\omega)$ has to be positive definite, which concludes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

3 Existence of FIR LBP

In Theorem 2 and Corollary 1 we saw the necessary and sufficient conditions for a transfer matrix $\mathbf{F}(z)$ to have a biorthogonal partner. In practice the situation of most significance is when $\mathbf{F}(z)$ is a rational function of z. A question of considerable interest is the following: under what conditions does a rational function $\mathbf{F}(z)$ have an FIR biorthogonal partner $\mathbf{H}(z)$? In fact it suffices to pose the previous question for any **FIR** filter $\mathbf{F}(z)$, which is evident by the following reasoning. Let $\mathbf{F}_r(z)$ be an arbitrary rational transfer matrix and let $D(z)$ be the least common multiple of the polynomials appearing in the denominators of the rational entries of $\mathbf{F}_r(z)$. Then we can write $\mathbf{F}_r(z) = \mathbf{F}(z)/D(z)$, where $\mathbf{F}(z)$ is an FIR matrix. If there exists an FIR biorthogonal partner $\mathbf{H}(z)$ of $\mathbf{F}(z)$, then $\mathbf{H}_r(z) = \mathbf{H}(z)D(z)$ is the corresponding FIR biorthogonal partner of $\mathbf{F}_r(z)$.

In view of all this, we begin the discussion in this section by finding the conditions for the existence of an FIR biorthogonal partner of an FIR transfer matrix. To this end we need to revisit the notion of *greatest right common divisors* (**grcd**) of polynomial matrices [19, 7]. In the linear systems literature, grcd's are most commonly defined for square matrices. In this setting, we will extend this definition to the case of rectangular matrices. In principle, we can define the grcd of a $p_1 \times r$ polynomial matrix $\mathbf{A}(z)$ and a $p_2 \times r$ polynomial matrix $\mathbf{B}(z)$ to be any $m \times r$ polynomial matrix $\mathbf{R}(z)$ such that:

1. $\mathbf{R}(z)$ is a common right divisor of $\mathbf{A}(z)$ and $\mathbf{B}(z)$, i.e. there exist polynomial matrices $\mathbf{A}_1(z)$ and $\mathbf{B}_1(z)$ such that $\mathbf{A}(z) = \mathbf{A}_1(z)\mathbf{R}(z)$ and $\mathbf{B}(z) = \mathbf{B}_1(z)\mathbf{R}(z)$;

2. If $\mathbf{R}_1(z)$ is another $m_1 \times r$ common right divisor of $\mathbf{A}(z)$ and $\mathbf{B}(z)$, then $\mathbf{R}_1(z)$ is a right divisor of $\mathbf{R}(z)$, i.e. there exists a $m \times m_1$ polynomial matrix $\mathbf{T}(z)$ such that $\mathbf{R}(z) = \mathbf{T}(z)\mathbf{R}_1(z)$.

However, for the purpose of this paper it is enough to consider only square grcd's $\mathbf{R}(z)$, so from now on by grcd we shall mean *square* grcd. Now we can state the following result.

Theorem 3. Existence of FIR LBP. Suppose $\mathbf{F}(z)$ is a causal and FIR $p \times r$ matrix, with the Type-2 polyphase form as in (1). Then there exists a causal FIR $r \times p$ matrix $\mathbf{H}(z)$ such that $[\mathbf{H}(z)\mathbf{F}(z)]_{\downarrow M} = \mathbf{I}$ if and only if grcd$[\mathbf{F}_0(z), \mathbf{F}_1(z), \ldots \mathbf{F}_{M-1}(z)]$ is a unimodular[1] matrix $\mathbf{R}(z)$.

Before proceeding to the proof of Theorem 3, several comments are due. Given an arbitrary MIMO transfer function, the grcd-condition is almost always satisfied. For example let

$$\mathbf{F}(z) = \begin{bmatrix} 3 + 2z^{-1} + z^{-2} & 2 + 3z^{-1} + z^{-2} \\ 1 + 3z^{-2} & 2 + z^{-1} + 3z^{-2} \end{bmatrix}.$$

The trivial biorthogonal partner (as in Example 1) is IIR in this case, since $\det[\mathbf{F}(z)] = 4 + 4z^{-1} + 6z^{-2} - 2z^{-3}$. However, it can be verified that the grcd of the two polyphase components of $\mathbf{F}(z)$ is unimodular, with one solution being (for construction of a grcd, see [7])

$$\mathbf{R}(z) = -2 \begin{bmatrix} 4 & 2 \\ 0 & 1 \end{bmatrix}.$$

Therefore, an FIR LBP for $M = 2$ indeed exists and one possibility is (see the proof of Theorem 3)

$$\mathbf{H}(z) = \frac{1}{16} \begin{bmatrix} 8 + 4z^{-1} + 3z^{-2} & -8 - z^{-2} \\ -4 - 8z^{-1} - 6z^{-2} & 12 + 2z^{-2} \end{bmatrix}.$$

In the statement of Theorem 3 we have not assumed anything about the integers p and r - the dimensions of $\mathbf{F}(z)$. It will soon become clear that the necessary relation between them is given by

$$r \leq 2p. \tag{10}$$

[1] A square polynomial matrix is said to be unimodular if its determinant is a nonzero constant.

Also, the constraint on $\mathbf{F}(z)$ and its LBP to be causal is unnecessary; it can be avoided if we allow the determinant of $\mathbf{R}(z)$ to be of the form cz^k, with $k \in \mathbb{Z}$, rather than just a constant.

Proof of Theorem 3. First we consider the case $M = 2$. If $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$ are right coprime (which is equivalent to saying that $\mathbf{R}(z)$, i.e. the grcd$[\mathbf{F}_0(z), \mathbf{F}_1(z)]$ is unimodular) then there exist polynomial matrices $\mathbf{H}_0(z)$ and $\mathbf{H}_1(z)$ such that

$$\mathbf{H}_0(z)\mathbf{F}_0(z) + \mathbf{H}_1(z)\mathbf{F}_1(z) = \mathbf{I}. \tag{11}$$

This follows from the *simple Bezout identity* [7], extended to the rectangular case. (Although the extension is straightforward, we summarize these results in the Appendix for convenience.) In fact, from the construction for a grcd [7] it follows that there exists a unimodular matrix $\mathbf{U}(z)$ such that

$$\underbrace{\begin{array}{cc} & \begin{array}{cc} p & p \end{array} \\ \begin{array}{c} r \\ 2p-r \end{array} & \begin{bmatrix} \mathbf{U}_{11}(z) & \mathbf{U}_{12}(z) \\ \mathbf{U}_{21}(z) & \mathbf{U}_{22}(z) \end{bmatrix} \end{array}}_{\mathbf{U}(z)} \begin{array}{cc} r & \\ \begin{bmatrix} \mathbf{F}_0(z) \\ \mathbf{F}_1(z) \end{bmatrix} & \begin{array}{c} p \\ p \end{array} \end{array} = \begin{array}{cc} r & \\ \begin{bmatrix} \mathbf{R}(z) \\ \mathbf{0} \end{bmatrix} & \begin{array}{c} r \\ 2p-r \end{array} \end{array} \tag{12}$$

with indicated sizes of the building blocks. From (12) it is easy to see that we can choose

$$\mathbf{H}_0(z) = \mathbf{R}^{-1}(z)\mathbf{U}_{11}(z), \quad \mathbf{H}_1(z) = \mathbf{R}^{-1}(z)\mathbf{U}_{12}(z) \tag{13}$$

and that these are really polynomial (actually causal FIR) matrices since $\mathbf{R}(z)$ is unimodular.

So far we have considered the $M = 2$ case, but the extension to arbitrary M follows readily by applying the rule

$$\text{grcd}_{0\le k\le M-1}[\mathbf{F}_k(z)] = \text{grcd}[\mathbf{F}_{M-1}(z), \text{grcd}_{0\le k\le M-2}[\mathbf{F}_k(z)]]. \tag{14}$$

Now, suppose by contradiction that $\mathbf{F}(z)$ has a causal FIR LBP $\mathbf{H}(z)$, but that

$$\text{grcd}[\mathbf{F}_0(z), \mathbf{F}_1(z), \ldots \mathbf{F}_{M-1}(z)] = \mathbf{C}(z)$$

is not unimodular. Writing $\mathbf{H}(z)$ in the Type-1 polyphase form [19] we have

$$\mathbf{I} = [\mathbf{H}(z)\mathbf{F}(z)]_{\downarrow M} = \sum_{k=0}^{M-1} \mathbf{H}_k(z)\mathbf{F}_k(z) = \left(\sum_{k=0}^{M-1} \mathbf{H}_k(z)\hat{\mathbf{F}}_k(z)\right)\mathbf{C}(z)$$

and it follows that

$$\sum_{k=0}^{M-1} \mathbf{H}_k(z)\hat{\mathbf{F}}_k(z) = \mathbf{C}^{-1}(z).$$

The left hand side of the above equation is a causal FIR matrix (since all $\mathbf{H}_k(z)$ and $\hat{\mathbf{F}}_k(z)$ are causal FIR), but the right hand side is not. This contradiction concludes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

Notice that (12) readily implies that $r \leq 2p$ in order for this particular construction to work. To see that (10) has to hold for *any* FIR LBP to exist, observe that (11) can be rewritten as

$$\left[\begin{array}{cc} \mathbf{H}_0(z) & \mathbf{H}_1(z) \end{array}\right] \underbrace{\left[\begin{array}{c} \mathbf{F}_0(z) \\ \mathbf{F}_1(z) \end{array}\right]}_{\mathbf{P}(z)} = \mathbf{I}.$$

If $r > 2p$, the matrix $\mathbf{P}(z)$ above becomes "fat", i.e. has no left inverse, thus in this case there is no FIR LBP of $\mathbf{F}(z)$. We return to this relation in Sec. 5, when we talk about multiwavelet theory.

It is important to notice here that, if it exists, FIR LBP is not unique. There are two reasons for this. Firstly, the grcd of two matrices is unique only up to a premultiplication by a unimodular matrix. Secondly, there are many unimodular matrices $\mathbf{U}(z)$ that satisfy (12) and each of them provides a valid solution. The issue of parametrization of these solutions will be treated in the following section. Also, notice that in the successive applications of the construction (12), as implied by the right-hand side of (14), grcd's of rectangular $p \times r$ matrices $\mathbf{F}_i(z)$ and square $r \times r$ matrices $\mathbf{R}(z)$ are computed. The result will again be a $r \times r$ matrix, and the necessary condition now becomes $r \leq p + r$, which is always satisfied. The sizes of the building blocks $\mathbf{U}_{ij}(z)$ from (12) will also need to be adjusted accordingly.

4 Application in Channel Equalization

In the following we will consider the case where an FIR LBP is used as a MIMO channel equalizer. We will show that the flexibility in the choice of $\mathbf{H}(z)$ can be exploited in order to reduce the undesirable amplification of the channel noise. But, before proceeding to these results, we give a brief overview of some equalization techniques.

The discrete-time equivalent of a MIMO digital communication system with symbol-spaced equalizer (**SSE**) [11] is shown in Fig. 3(a). The symbol rate at the input $\mathbf{x}(n)$ is $1/T$. Notice that the equalizer $\mathbf{H}_2(z)$ works at the same rate (thus the name symbol-spaced equalizer). The discrete versions of the pulse shaping filter and the channel, $\mathbf{G}_2(z)$ and $\mathbf{C}_2(z)$ respectively, are obtained by sampling the corresponding continuous-time impulse responses also at the rate $1/T$. We will refer to their cascade $\mathbf{F}_2(z) = \mathbf{C}_2(z)\mathbf{G}_2(z)$ as the equivalent channel for the SSE case. Therefore, as for the signal $\mathbf{x}(n)$, the system from Fig. 3(a) can be represented as a cascade of the equivalent channel $\mathbf{F}_2(z)$ and a SSE $\mathbf{H}_2(z)$. An ideal equalizer (or a zero-forcing equalizer [11]) $\mathbf{H}_2(z)$ is then obtained as a left inverse of the equivalent channel $\mathbf{F}_2(z)$.

From this discussion, several drawbacks of symbol-spaced equalizers are apparent. The MIMO transfer function $\mathbf{F}_2(z)$ does not have a left inverse if it is a fat matrix. Even if the matrix is not

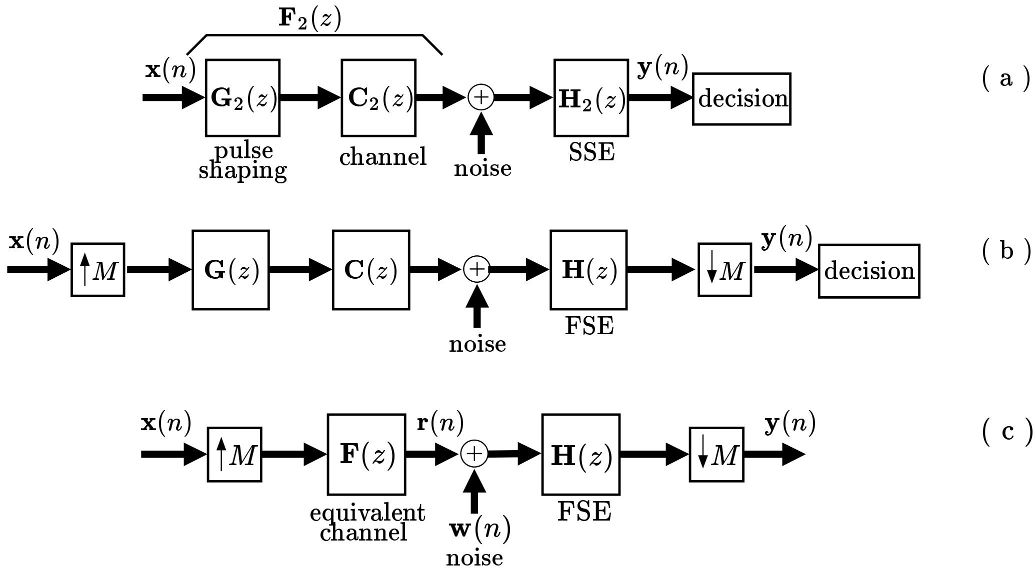

Figure 3: (a) Discrete-time equivalent of a digital communication system with SSE; the equivalent channel is $\mathbf{F}_2(z) = \mathbf{C}_2(z)\mathbf{G}_2(z)$. (b) Digital communication system from (a), now equalized with FSE $\mathbf{H}(z)$. (c) Further simplification of the system from (b); the equivalent channel is $\mathbf{F}(z) = \mathbf{C}(z)\mathbf{G}(z)$.

fat, its invertibility will depend on the rank. Furthermore, if $\mathbf{F}_2(z)$ is invertible, its inverse is most probably IIR, which often amplifies the noise at the receiver. Finally, it has been observed that the ISI suppression achieved by this equalizer is very sensitive to the phase of the sampling at the receiver [11, 17]. For all these reasons, a popular alternative is to use a so called fractionally spaced equalizer (**FSE**). It can be shown to be far less sensitive to the sampling phase [17], it can be used with fat channel transfer functions, and it often allows for FIR solutions while SSE does not.

The idea behind a FSE is to let the equalizer work at a higher rate. Because of this additional redundancy, FSEs are both more flexible and more robust than SSEs. In the continuous-time communication system, FSE is realized by sampling the received waveform at M times the symbol rate, and feeding such oversampled signal to the equalizer, which now operates at the rate M/T. In discrete-time this is modeled as shown in Fig. 3(b). The discrete transfer functions $\mathbf{G}(z)$ and $\mathbf{C}(z)$ are obtained after sampling the corresponding continuous-time impulse responses at the rate M/T. Thus, the equivalent channel $\mathbf{F}(z)$ in this case is such that $\mathbf{F}_2(z) = [\mathbf{F}(z)]_{\downarrow M}$ and the simplified scheme is shown in Fig. 3(c). Note that the noise also needs to be modified, but this is not the main point of discussion here. We recall from Sec. 3 that a zero-forcing FSE $\mathbf{H}(z)$ in Fig. 3(c) is nothing but a LBP of the channel matrix $\mathbf{F}(z)$. In this section we will exploit the nonuniqueness of this biorthogonal partner with the aim of minimizing the noise power at the receiver.

Optimization of MIMO systems of the type shown in Fig. 3(a) has been considered by several authors in many different contexts (e.g. [10], [12], [27]). The authors in [27] derive the optimal transmitter *and* receiver for a given channel in the sense of minimizing the overall mean squared error. This MMSE solution clearly outperforms any zero-forcing equalizer, however the price is paid in terms of complexity: the solution in [27] involves ideal filtering. Here we have taken a simplistic approach of decoupling the problems of ISI and noise suppression. Moreover, the system shown in Fig. 3(b) brings in an additional element of freedom, which will be exploited in this section.

4.1 Optimizing LBP for Channel Equalization

The size of the channel $\mathbf{F}(z)$ will be assumed to be $p \times r$, with $r \leq 2p$. For simplicity we will first assume that the oversampling ratio M is equal to 2 (see Fig. 4(a)). In this case the system can be redrawn as in Fig. 4(b). Here $\mathbf{w}_0(n)$ and $\mathbf{w}_1(n)$ are the corresponding polyphase components of the noise vector sequence $\mathbf{w}(n)$ from Fig. 4(a), while $\mathbf{F}_0(z)$, $\mathbf{F}_1(z)$ and $\hat{\mathbf{H}}_0(z)$, $\hat{\mathbf{H}}_1(z)$ are the polyphase components of $\mathbf{F}(z)$ and $\hat{\mathbf{H}}(z)$, respectively. Recall that if the conditions of Theorem 2 are satisfied, then $\mathbf{H}_0(z)$ and $\mathbf{H}_1(z)$ as in (13) lead to one possible solution for $\hat{\mathbf{H}}(z)$. However, from (12) we see that another class of solutions is given by

$$\hat{\mathbf{H}}_0(z) = \mathbf{H}_0(z) + \mathbf{A}(z)\mathbf{U}_{21}(z), \quad \hat{\mathbf{H}}_1(z) = \mathbf{H}_1(z) + \mathbf{A}(z)\mathbf{U}_{22}(z) \tag{15}$$

for an arbitrary $r \times 2p - r$ matrix $\mathbf{A}(z)$ and matrices $\mathbf{U}_{ij}(z)$ defined in (12). Our goal here is to design $\mathbf{A}(z)$ such that the noise component of $\mathbf{y}(n)$ in Fig. 4(a) is minimized. For that purpose, we consider the noise model shown in Fig. 4(b). Let us define the following

$$\mathbf{e}(n) = \begin{bmatrix} \mathbf{w}_0(n) \\ \mathbf{w}_1(n) \end{bmatrix}, \quad \mathbf{B}(z) = \begin{bmatrix} \mathbf{H}_0(z) & \mathbf{H}_1(z) \\ \mathbf{U}_{21}(z) & \mathbf{U}_{22}(z) \end{bmatrix}. \tag{16}$$

Then the equivalent of the system in Fig. 4(b) is shown in Fig. 4(c). Our task now becomes that of finding the matrix $\mathbf{A}(z) = \sum_{i=0}^{N_A-1} \mathbf{A}_i z^{-i}$ such that the norm of

$$\hat{\mathbf{e}}(n) = \mathbf{u}(n) + \sum_{i=0}^{N_A-1} \mathbf{A_i}\mathbf{v}(n-i)$$

is less than the corresponding norm when any other polynomial matrix $\bar{\mathbf{A}}(z)$ of the same or lower order is used. That turns out to be equivalent to the problem of finding the best linear estimator of order $N_A - 1$ for the vector process $\mathbf{u}(n)$ given the observations $\mathbf{v}$(n). The solution to this problem is well-known and is based on the orthogonality principle. Let us define the $r \times N_A(2p - r)$ matrix $\mathcal{A}$ (corresponding to the optimal solution) to be

$$\mathcal{A} \stackrel{\text{def}}{=} [\mathbf{A}_0 \;\; \mathbf{A}_1 \ldots \mathbf{A}_{N_A-1}] \tag{17}$$

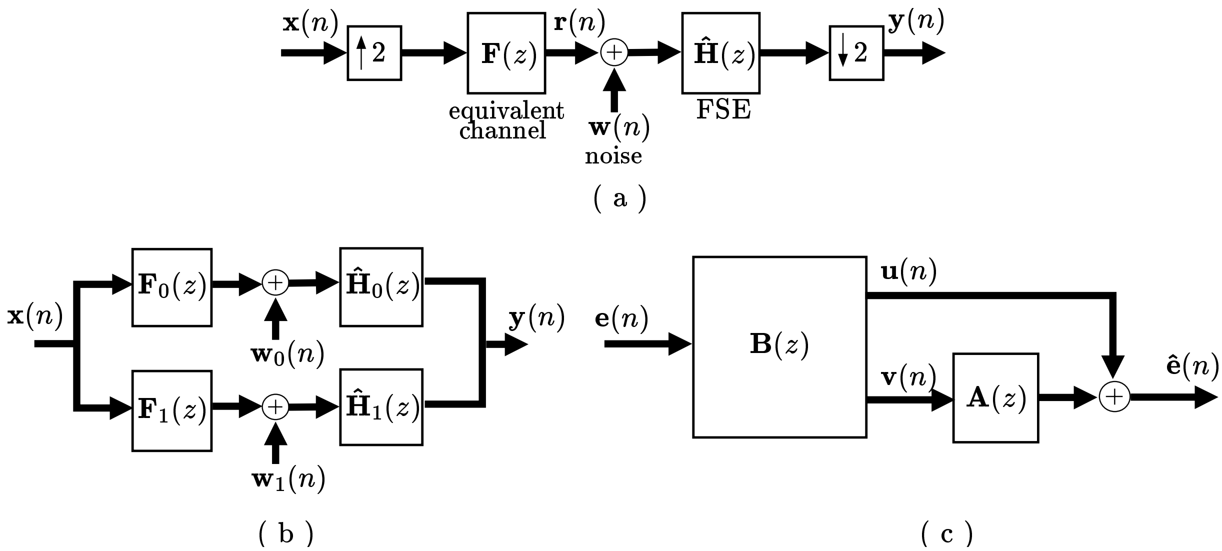

Figure 4: Block diagram interpretation of the construction of FSE for $M = 2$. (a) Discrete-time equivalent communication channel with FSE, (b) equivalent of (a) obtained using noble identities [19], and (c) equivalent model for noise.

and the $N_A(2p-r)\times 1$ vector sequence $\mathcal{V}(n)$ to be

$$\mathcal{V}(n) \stackrel{\text{def}}{=} \left[\mathbf{v}^T(n)\ \ \mathbf{v}^T(n-1)\ldots\mathbf{v}^T(n-N_A+1)\right]^T$$

By the orthogonality principle we then have[2] $E[\mathbf{u}\mathcal{V}^\dagger] + E[\mathcal{A}\mathcal{V}\mathcal{V}^\dagger] = \mathbf{0}$, so that the solution for the optimum estimator $\mathcal{A}$ becomes

$$\mathcal{A} = -E[\mathbf{u}\mathcal{V}^\dagger]\mathbf{R}_{\mathcal{V}\mathcal{V}}{}^{-1} \tag{18}$$

with $\mathbf{R}_{\mathcal{V}\mathcal{V}}$ denoting the autocorrelation of $\mathcal{V}$. This is nothing but a standard Wiener solution [6].

Now we need to express (18) in terms of the statistics of the input noise $\mathbf{e}(n)$. Define $\bar{\mathbf{H}}(z)$ ($r\times 2p$ matrix) to consist of the first r rows of the $2p\times 2p$ matrix $\mathbf{B}(z)$ defined in (16) and similarly $\bar{\mathbf{U}}(z)$ ($2p-r\times 2p$ matrix) to consist of the last $2p-r$ rows of $\mathbf{B}(z)$. If N_B-1 is the order of $\mathbf{B}(z)$, we define $\bar{\mathbf{H}}_i$ and $\bar{\mathbf{U}}_i$ as

$$\bar{\mathbf{H}}(z) = \sum_{i=0}^{N_B-1}\bar{\mathbf{H}}_i z^{-i} \quad \text{and} \quad \bar{\mathbf{U}}(z) = \sum_{i=0}^{N_B-1}\bar{\mathbf{U}}_i z^{-i}.$$

Now, the $2p(N_A+N_B-1)\times 1$ vector $\mathcal{E}(n)$, the $r\times 2p(N_A+N_B-1)$ matrix $\mathcal{H}$ and the $N_A(2p-r)\times 2p(N_A+N_B-1)$ matrix $\mathcal{U}$ are defined as follows

$$\mathcal{E}(n) \stackrel{\text{def}}{=} \left[\mathbf{e}^T(n)\ \ \mathbf{e}^T(n-1)\ldots\mathbf{e}^T(n-N_A-N_B+1)\right]^T$$

[2]Symbol $E[\cdot]$ denotes the expected value.

$$\mathcal{H} \stackrel{\text{def}}{=} \left[\bar{\mathbf{H}}_0 \;\; \bar{\mathbf{H}}_1 \ldots \bar{\mathbf{H}}_{N_B-1} \;\; \mathbf{0} \ldots \mathbf{0}\right]$$

$$\mathcal{U} \stackrel{\text{def}}{=} \left[\begin{array}{cccccc} \bar{\mathbf{U}}_0 & \ldots & \bar{\mathbf{U}}_{N_B-1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{U}}_0 & \ldots & \bar{\mathbf{U}}_{N_B-1} & \ldots & \mathbf{0} \\ \vdots & & \ddots & & \ddots & \\ \mathbf{0} & \ldots & \mathbf{0} & \bar{\mathbf{U}}_0 & \ldots & \bar{\mathbf{U}}_{N_B-1} \end{array}\right].$$

Then the following holds

$$\mathbf{u}(n) = \mathcal{H}\mathcal{E}(n), \quad \mathcal{V}(n) = \mathcal{U}\mathcal{E}(n). \tag{19}$$

Finally, all we need to do is substitute (19) in (18) and arrive at the final solution for $\mathcal{A}$

$$\mathcal{A} = -\mathcal{H}\mathbf{R}_{\mathcal{E}\mathcal{E}}\mathcal{U}^\dagger \left(\mathcal{U}\mathbf{R}_{\mathcal{E}\mathcal{E}}\mathcal{U}^\dagger\right)^{-1}. \tag{20}$$

Notice that the final solution depends only on the statistics of the input noise ($\mathbf{R}_{\mathcal{E}\mathcal{E}}$) and the elements of the previously determined matrix $\mathbf{B}(z)$. Also notice that the solution (20) provides constant matrices $\mathbf{A}_i$ (as in (17)) and the linear estimator $\mathbf{A}(z)$ is given by $\mathbf{A}(z) = \sum_{i=0}^{N_A-1}\mathbf{A}_i z^{-i}$. It should also be noted that the solution (15) is not the most general one; it is possible that there exists another FIR LBP $\mathbf{H}'(z)$ which will outperform any $\hat{\mathbf{H}}(z)$ of the same order obtained via (15). The linear estimator $\mathbf{A}(z)$ derived in this section is unique given the form of the solution (15) and the initial values $\mathbf{H}_0(z)$, $\mathbf{H}_1(z)$. However, (15) presents only *one possible* form of the solution. The problem of the general solution is further treated in [22].

4.2 Case of General M

So far in this section we have only considered the $M = 2$ case, which leads to the solutions as in (12), (13), and (15). In the following we consider the case when $M > 2$. The changes that need to be made with respect to the $M = 2$ case are twofold. Firstly, the procedure for finding the initial values of polyphase components $\mathbf{H}_k(z)$ (given by (12) and (13) when $M = 2$) has to be extended to accommodate for larger values of M. Secondly, the introduced redundancy for LBP optimization given by (15) also needs to be modified.

We deal with the initial values $\mathbf{H}_k(z)$ first. The way of extending the construction (12) and (13) is suggested by (14); apart from (12), $M-2$ additional equations also need to be satisfied:

$$\begin{array}{c} \\ r \\ p \end{array}\underbrace{\begin{array}{c} \begin{array}{cc} \;\;r & \qquad p \end{array} \\ \left[\begin{array}{cc} \mathbf{U}_{11}^{(i)}(z) & \mathbf{U}_{12}^{(i)}(z) \\ \mathbf{U}_{21}^{(i)}(z) & \mathbf{U}_{22}^{(i)}(z) \end{array}\right] \end{array}}_{\mathbf{U}^{(i)}(z)} \left[\begin{array}{c} \mathbf{R}_{i-1}(z) \\ \mathbf{F}_{i+1}(z) \end{array}\right]\begin{array}{c} r \\ p \end{array} = \begin{array}{c} r \\ \left[\begin{array}{c} \mathbf{R}_i(z) \\ \mathbf{0} \end{array}\right] \end{array}\begin{array}{c} r \\ p \end{array}, \quad \text{for} \;\; 1 \le i \le M-2. \tag{21}$$

Here $\mathbf{R}_i(z) = \text{grcd}\left[\mathbf{F}_0(z), \mathbf{F}_1(z), \ldots \mathbf{F}_{i+1}(z)\right]$. If we denote $\mathbf{R}(z) = \text{grcd}[\mathbf{F}_0(z), \mathbf{F}_1(z), \ldots \mathbf{F}_{M-1}(z)]$ (as we did in Sec. 3), then the polyphase components $\mathbf{H}_k(z)$ can be found from (21) as $\mathbf{H}_k(z) =$

$\mathbf{R}^{-1}(z)\mathbf{V}_k(z)$. Here the $r \times p$ matrices $\mathbf{V}_k(z)$ are given by

$$\mathbf{V}_0(z) = \left(\prod_{i=0}^{M-2} \mathbf{U}_{11}^{(M-2-i)}(z)\right); \quad \mathbf{V}_{M-1}(z) = \mathbf{U}_{12}^{(M-2)}(z);$$

$$\mathbf{V}_k(z) = \left(\prod_{i=0}^{M-2-k} \mathbf{U}_{11}^{(M-2-i)}(z)\right) \mathbf{U}_{12}^{(k-1)}(z), \quad \text{for} \;\; 1 \leq k \leq M-2. \tag{22}$$

Notice that when $M = 2$ the above equations (22) reduce to $\mathbf{V}_0(z) = \mathbf{U}_{11}(z)$ and $\mathbf{V}_1(z) = \mathbf{U}_{12}(z)$, which is in compliance with previously established result (13).

Now we move on to find the equivalent of (15) for $M > 2$. If we want to keep the same structure of having only one matrix $\mathbf{A}(z)$ to represent the degrees of freedom, then the solution equivalent to the one in (15) is going to be

$$\hat{\mathbf{H}}_k(z) = \mathbf{H}_k(z) + \mathbf{A}(z)\mathbf{W}_k(z), \quad \text{for} \;\; 0 \leq k \leq M-1 \tag{23}$$

with the $p \times p$ matrices $\mathbf{W}_k(z)$ satisfying $\sum_{k=0}^{M-1} \mathbf{W}_k(z)\mathbf{F}_k(z) = \mathbf{0}$. As it turns out, the matrices $\mathbf{W}_k(z)$ can be obtained from (21) in a fashion similar to the one when we obtained $\mathbf{U}_{21}(z)$ and $\mathbf{U}_{22}(z)$ from (12). They are given by

$$\mathbf{W}_0(z) = \mathbf{U}_{21}^{(M-2)}(z) \left(\prod_{i=0}^{M-3} \mathbf{U}_{11}^{(M-3-i)}(z)\right); \quad \mathbf{W}_{M-1}(z) = \mathbf{U}_{22}^{(M-2)}(z);$$

$$\mathbf{W}_k(z) = \mathbf{U}_{21}^{(M-2)}(z) \left(\prod_{i=0}^{M-3-k} \mathbf{U}_{11}^{(M-3-i)}(z)\right) \mathbf{U}_{12}^{(k-1)}(z), \quad \text{for} \;\; 1 \leq k \leq M-2. \tag{24}$$

Now we can repeat the whole procedure of finding the optimal estimator $\mathbf{A}(z)$, along the same lines as in (16)-(20), the only difference being in the dimensions of the vector $\mathbf{e}(n)$ and the matrix $\mathbf{B}(z)$.

Finally, we note that it is possible to construct solutions similar to the one in (23), but with more degrees of freedom. For example, we may consider any pair of polyphase components $\mathbf{F}_i(z)$ and $\mathbf{F}_j(z)$ $(i \neq j)$ and find the corresponding matrices $\mathbf{W}_{ij,0}(z)$ and $\mathbf{W}_{ij,1}(z)$ such that

$$\mathbf{W}_{ij,0}(z)\mathbf{F}_i(z) + \mathbf{W}_{ij,1}(z)\mathbf{F}_j(z) = \mathbf{0}.$$

The solution to this problem is again based on (12). Now pick any $r \times 2p - r$ matrix $\mathbf{A}_{ij}(z)$. It follows that if $\mathbf{H}_i(z)$ and $\mathbf{H}_j(z)$ are valid polyphase components of a LBP corresponding to $\mathbf{F}_i(z)$ and $\mathbf{F}_j(z)$, then the following solutions are also valid:

$$\hat{\mathbf{H}}_i(z) = \mathbf{H}_i(z) + \mathbf{A}_{ij}(z)\mathbf{W}_{ij,0}(z); \quad \hat{\mathbf{H}}_j(z) = \mathbf{H}_j(z) + \mathbf{A}_{ij}(z)\mathbf{W}_{ij,1}(z).$$

Repeating the above reasoning for all the pairs $(i,\, j)$ we can generate proper LBPs with much more degrees of freedom, given by the matrices $\mathbf{A}_{ij}(z)$. However it becomes very difficult to optimize all these parameters, since the matrices $\mathbf{A}_{ij}(z)$ cannot be interpreted as linear estimators.

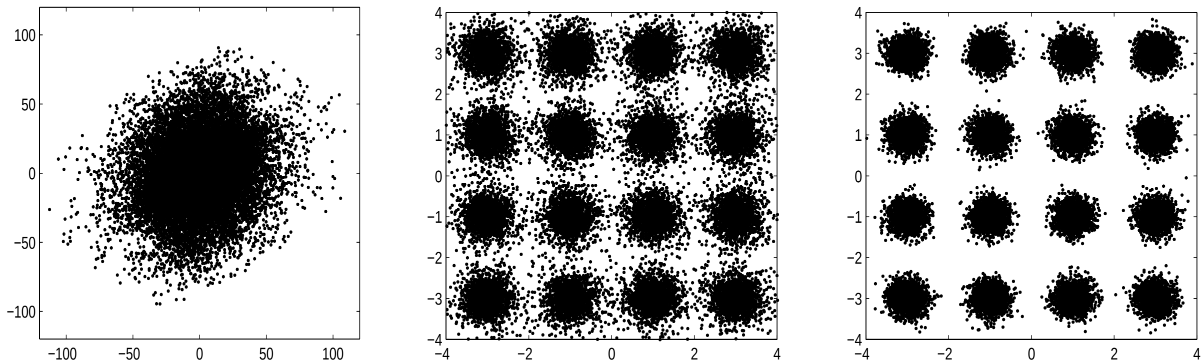

Figure 5: MIMO equalization of a square 3×3 channel: (left) plain old SSE, (middle) direct derivation of FSE, and (right) FSE optimized as in Sec. 4.1, with $\mathbf{A}(z)$ of order 4.

4.3 Experimental Results

In this section we present the results of numerical simulations. Three different methods for the MIMO channel equalization are compared: SSE, FIR FSE as described previously in Sec. 3, and the generalized solution for FIR FSE as in (15) with $\mathbf{A}(z)$ chosen optimally for the given noise statistics. A block diagram of the digital communication systems with SSE and FSE were shown in Fig. 3(a) and 3(c), respectively.

In the following we present the results of equalizing two MIMO channels: a square 3×3 channel $\mathbf{F}_{sq}(z)$ and a rectangular, fat 2×3 channel $\mathbf{F}_{rec}(z)$. For simplicity $\mathbf{F}_{rec}(z)$ was chosen such that it consists of the first two rows of $\mathbf{F}_{sq}(z)$; in other words

$$\mathbf{F}_{rec}(z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{F}_{sq}(z).$$

In both examples we chose $M = 2$. The MIMO channel $\mathbf{F}_{sq}(z)$ was characterized by a 3×3 matrix polynomial of order 6 and the corresponding coefficients can be found at [28]. The constellation was chosen to be 16-QAM. The noise on the channel was taken to be white. The signal to noise ratio used in the experiments was obtained as (see Fig. 4(a))

$$\text{SNR} = 20 \log_{10} \frac{||\mathbf{r}||_2}{||\mathbf{w}||_2}.$$

The transfer function $\mathbf{F}_2(z)$ of the equivalent channel in the SSE case (see Fig. 3(a)), was chosen in such way that the inverse matrix $\mathbf{H}_2(z)$ is stable, but with two poles very close to the unit circle. At the same time, the conventional FIR solution for the FSE in this case exists and can be found at [28]. In the absence of noise, SSE was performing as well as both the FSEs, i.e. all the symbols were received intact, but in the presence of noise the received symbols were virtually unintelligible.

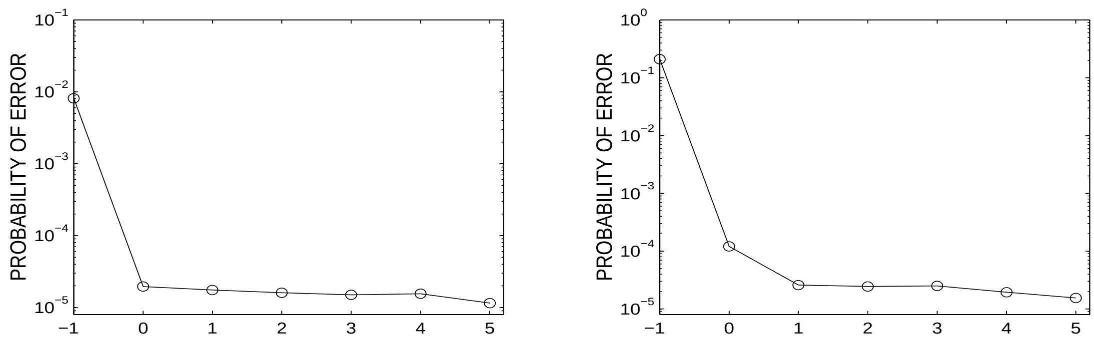

Figure 6: Probability of error as a function of the estimator order: (left) square 3×3 channel, and (right) rectangular 2×3 channel - see the text.

To demonstrate this, in the left part of Fig. 5 we show the output of SSE with SNR = 18dB. At the same noise level a simple FIR FSE (middle of Fig. 5) performed much better, with less than one percent of the symbols being misinterpreted. Finally, when we used the improved FIR FSE with only a third order estimator $\mathbf{A}_{sq}(z)$ (the coefficients can also be found at [28]), almost all the symbols were detected correctly and the eye-diagram (right part of Fig. 5) shows clear separation.

While it is intuitive that the higher order estimators $\mathbf{A}_{sq}(z)$ should perform better than the lower order ones, it is still of importance to quantitatively evaluate this improvement in performance. Notice that the procedure for obtaining linear estimators as in Sec. 4.1 applies to any desired order $N_A - 1$. The only difference for higher orders is that the matrix $\mathcal{A}$ in (17) becomes larger, and in turn so do the matrices $\mathcal{H}$ and $\mathcal{U}$ in (20). What we have noticed in our examples (taking the noise to be white) is that, with increasing N_A, the performance of the estimator does not improve much after a certain point (see Fig. 6). This is because all the terms in matrices $\mathbf{A}_k$ in (17) tend to decay very fast for large values of k and their influence on the equalizer performance diminishes in the similar fashion.

In the left part of Fig. 6 we plotted the estimated probability of error at the detector in the 3×3 channel case as a function of the order of estimator $N_A - 1$. The probability of error is obtained as the average value of the error probabilities in each of the three channels. The first measurement, for $N_A - 1 = -1$ corresponds to the case where there is no optimization of the equalizer, i.e. where $\mathbf{A}_{sq}(z)$ is a zero matrix. The probability of error in this case with SNR = 18dB was equal to 0.82 percent. Interestingly, the probability of error can be reduced by several orders of magnitude by employing just the zeroth order (constant) matrix $\mathbf{A}_{sq}(z)$. Only two out of 10^5 symbols were misinterpreted in this case. However, there is not much improvement as the order increases.

Very similar findings were true in the rectangular case as well. Although the left inverse does not exist in this case, it is still possible to equalize the 2×3 channel $\mathbf{F}_{rec}(z)$ with oversampling just by two [28]. However, the default equalizer obtained by the LBP construction as in Sec. 3 does not perform very well. As shown in the right part of Fig. 6, for no estimator correction $\mathbf{A}_{rec}(z)$ employed, the error rate was more than than 20 percent for SNR = 32dB. However, even the constant correction $\mathbf{A}_{rec}(z)$ resulted in the dramatic decrease in the probability of error to 10^{-4}, while the higher order corrections kept the error probability below 4×10^{-5}. This, together with the previous, square, example stands to show that exploiting the redundancy of LBP even to the smallest extent can prove to be very fruitful.

5 Some Further Applications of Biorthogonal Partners

In this section we will consider two other situations where we encounter MIMO biorthogonal partners. In both these instances, the solutions are already well-known and our intention here is to make a connection to the biorthogonal partner theory described in this paper.

5.1 Least Squares Signal Approximation

First we will address the problem of least squares signal approximation for vector signals. In the scalar case, a similar problem is very common in multiresolution theory [9] as well as in the spline approximation theory [18], [20]. This topic has been treated extensively in the mathematics literature, in the more general setting of *oblique projections* [1, 2]. The article by Aldroubi and Unser [2] is especially insightful and is closely connected to the material in this section.

Suppose we are given the signal model as shown in Fig. 7(a). The vector signal $\mathbf{y}(n)$ is obtained by upsampling the vector sequence $\mathbf{c}(n)$ and passing the result through the matrix transfer function $\mathbf{F}(z)$. Now, given a vector signal $\mathbf{x}(n)$, suppose we want to approximate it by a signal $\mathbf{y}(n)$ admitting the described model

$$\mathbf{y}(n) = \sum_{k \in \mathbb{Z}} \mathbf{F}(n - kM)\mathbf{c}(k), \tag{25}$$

or in the z-domain

$$\mathbf{Y}(z) = \mathbf{F}(z)\mathbf{C}(z^M). \tag{26}$$

It turns out that the optimum vector sequence $\mathbf{c}(n)$ can be determined as in Fig. 7(b). The prefilter $\mathbf{H}(z)$ turns out to be a particular form of a MIMO biorthogonal partner of $\mathbf{F}(z)$. In the following we refer to this as the least squares problem.

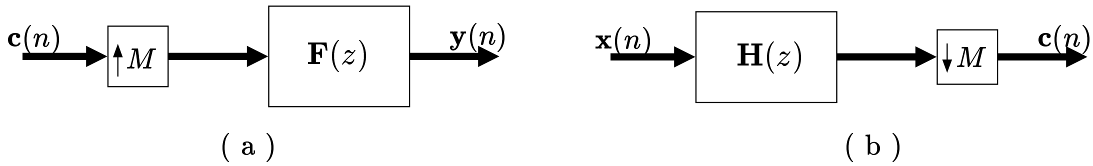

Figure 7: Least squares signal modeling: (a) signal model and (b) least squares solution (see text).

A very similar problem arises in multiwavelet theory (see [26] and also the next section). Consider the two-band multiwavelet transform. The space V_0 is spanned by N scaling functions and their integer shifts. Similarly, the space W_0 is spanned by N wavelets and their integer shifts. Those two spaces together form a finer resolution space V_1. Suppose we have a signal $x_1(n)$ belonging to the space V_1 and we want to find a coarser signal $x_0(n)$ from V_0 such that the distance (in the ℓ_2 sense) from the signal $x_1(n)$ is minimized. This problem can be formulated as a vector valued least squares problem, so the solution is again given by Fig. 7.

We first state the vector valued least squares problem in the general form. Consider the space $\mathcal{F}$ of all vector signals $\mathbf{y}(n)$ satisfying the model (25), where $\mathbf{c}(n)$ is an arbitrary ℓ_2 vector sequence.[3] This situation is depicted in Fig. 7(a). Here $\mathbf{F}(z)$ is a given MIMO transfer function. The problem is as follows. Given any vector signal $\mathbf{x}(n)$, we want to find the corresponding **projection** in $\mathcal{F}$, i.e. the vector signal $\mathbf{y}(n) \in \mathcal{F}$ such that

$$\sum_n \|\mathbf{y}(n) - \mathbf{x}(n)\|^2 \tag{27}$$

is minimized. Here $\|\cdot\|$ denotes the vector norm in ℓ_2. The following theorem describes the algorithm by which this is achieved and the corresponding corollary will address the uniqueness of the proposed solution.

Theorem 4. Solution to least squares problem. Given a MIMO transfer function $\mathbf{F}(z)$ and assuming that $\mathbf{S}(e^{j\omega}) = [\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}$ is a positive definite matrix for all ω, we define the (orthogonal) **projection filter** by

$$\mathbf{H}(z) = \left([\tilde{\mathbf{F}}(z)\mathbf{F}(z)_{\downarrow M}\right)^{-1}_{\uparrow M} \tilde{\mathbf{F}}(z). \tag{28}$$

If we pass the vector signal $\mathbf{x}(n)$ through the projection filter and decimate the outputs by M we get the optimal driving sequence $\mathbf{c}(n)$ (see Fig. 7(b)). This $\mathbf{c}(n)$ can be used to find the least squares approximation $\mathbf{y}(n)$ as in Fig. 7(a).

[3]This means that all the scalar sequences corresponding to the vector entries are square summable.

Notice that the projection filter is equal to the generic LBP given by (9). The positive-definiteness condition is necessary only to ensure the stability of $\mathbf{H}(z)$.

Proof. The error (27) that needs to be minimized can be rewritten in the frequency domain

$$\sum_n \|\mathbf{y}(n)-\mathbf{x}(n)\|^2 = \int_0^{2\pi} \|\mathbf{Y}(e^{j\omega})-\mathbf{X}(e^{j\omega})\|^2 \frac{d\omega}{2\pi} = \int_0^{2\pi} \underbrace{\|\mathbf{F}(e^{j\omega})\mathbf{C}(e^{j\omega M})-\mathbf{X}(e^{j\omega})\|^2}_{\mathcal{E}(\omega)} \frac{d\omega}{2\pi}.$$

Note that $\mathbf{C}(e^{j\omega M})$ appearing in the integrand is periodic with period $2\pi/M$, and thus can be chosen independently only in the range $0 \le \omega \le 2\pi/M$. That is why the integrand can be rewritten as

$$\mathcal{E}(\omega) = \sum_{k=0}^{M-1} \|\mathbf{F}(e^{j(\omega+\frac{2\pi k}{M})})\mathbf{C}(e^{j\omega M}) - \mathbf{X}(e^{j(\omega+\frac{2\pi k}{M})})\|^2.$$

For each ω in $0 \le \omega \le 2\pi/M$ we can choose $\mathbf{C}(e^{j\omega M})$ such that the nonnegative integrand $\mathcal{E}(\omega)$ is minimized and that would in turn minimize the projection error (27). Define the vector $\mathbf{a}(\omega)$ and the matrix $\mathbf{B}(\omega)$ as

$$\mathbf{a}(\omega) = [\mathbf{X}^T(e^{j\omega}) \;\; \mathbf{X}^T(e^{j(\omega+\frac{2\pi}{M})}) \ldots \mathbf{X}^T(e^{j(\omega+\frac{2\pi(M-1)}{M})})]^T$$

$$\mathbf{B}(\omega) = [\mathbf{F}^T(e^{j\omega}) \;\; \mathbf{F}^T(e^{j(\omega+\frac{2\pi}{M})}) \ldots \mathbf{F}^T(e^{j(\omega+\frac{2\pi(M-1)}{M})})]^T.$$

The problem now reduces to that of minimizing

$$\begin{aligned}
\mathcal{E}(\omega) &= \|\mathbf{B}(\omega)\mathbf{C}(e^{j\omega M}) - \mathbf{a}(\omega)\|^2 \\
&= [\mathbf{C}^\dagger(e^{j\omega M}) - \mathbf{a}^\dagger(\omega)\mathbf{B}(\omega)\mathbf{S}^{-1}(\omega)]\mathbf{S}(\omega)[\mathbf{C}(e^{j\omega M}) - \mathbf{S}^{-1}(\omega)\mathbf{B}^\dagger(\omega)\mathbf{a}(\omega)] \\
&\quad + \mathbf{a}^\dagger(\omega)\mathbf{a}(\omega) - \mathbf{a}^\dagger(\omega)\mathbf{B}(\omega)\mathbf{S}^{-1}(\omega)\mathbf{B}^\dagger(\omega)\mathbf{a}(\omega)
\end{aligned} \tag{29}$$

where $\mathbf{S}(\omega) = \mathbf{B}^\dagger(\omega)\mathbf{B}(\omega)$. The form (29) was obtained by the "completion of squares". Consider the right hand side of the last equality in (29). It consists of two parts; the first part depends on the choice of $\mathbf{C}(e^{j\omega M})$ and the second part does not. Since the first part is always nonnegative, we should choose $\mathbf{C}(e^{j\omega M})$ such that it becomes zero. Note that the matrix $\mathbf{S}(\omega) = \mathbf{B}^\dagger(\omega)\mathbf{B}(\omega)$ is positive definite, which follows from the assumption $[\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M} > 0$. Therefore, the only way to make the first part zero is to choose $\mathbf{C}(e^{j\omega M}) = \left(\mathbf{B}^\dagger(\omega)\mathbf{B}(\omega)\right)^{-1}\mathbf{B}^\dagger(\omega)\mathbf{a}(\omega)$. In order to rewrite this solution in terms of multirate building blocks, we note [19] that for any transfer function $\mathbf{A}(e^{j\omega})$, $[\mathbf{A}(e^{j\omega})]_{\downarrow M} = \frac{1}{M}\sum_{k=0}^{M-1}\mathbf{A}(e^{j\frac{\omega+2\pi k}{M}})$. Therefore,

$$\mathbf{B}^\dagger(\omega)\mathbf{B}(\omega) = \sum_{k=0}^{M-1} \mathbf{F}^\dagger(e^{j(\omega+\frac{2\pi k}{M})})\mathbf{F}(e^{j(\omega+\frac{2\pi k}{M})}) = M[\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M \uparrow M},$$

$$\mathbf{B}^\dagger(\omega)\mathbf{a}(\omega) = \sum_{k=0}^{M-1} \mathbf{F}^\dagger(e^{j(\omega+\frac{2\pi k}{M})})\mathbf{X}(e^{j(\omega+\frac{2\pi k}{M})}) = M[\mathbf{F}^\dagger(e^{j\omega})\mathbf{X}(e^{j\omega})]_{\downarrow M\uparrow M}.$$

The optimal $\mathbf{C}(e^{j\omega M})$ is therefore

$$\mathbf{C}(e^{j\omega M}) = \left[\left([\mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega})]_{\downarrow M}\right)^{-1}_{\uparrow M} \mathbf{F}^\dagger(e^{j\omega})\mathbf{X}(e^{j\omega})\right]_{\downarrow M\uparrow M}.$$

Thus we have $\mathbf{C}(z) = [\mathbf{H}(z)\mathbf{X}(z)]_{\downarrow M}$, with $\mathbf{H}(z)$ given by (28). This concludes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

The next corollary states that the least squares solution proposed by Theorem 2 is unique. While the proposed proof provides an elegant argument, the result of Corollary 2 also follows from the uniqueness of the orthogonal projection onto a closed subspace [2].

Corollary 2. Uniqueness of projection filter. Consider Fig. 7. For fixed $\mathbf{F}(z)$ satisfying the condition of Theorem 2 and $\mathbf{x}(n) \in \ell_2$, the least squares approximation $\mathbf{y}(n)$ is unique. Next, suppose the prefilter $\mathbf{H}(z)$ in Fig. 7(b) is such that the output of $\mathbf{F}(z)$ (Fig. 7(a)) is the least squares approximation of $\mathbf{x}(n)$ for any choice of the ℓ_2 input $\mathbf{x}(n)$. Then $\mathbf{H}(z)$ is unique and is therefore given by (28).

Proof. The uniqueness of $\mathbf{c}(n)$ and thus $\mathbf{y}(n)$ follows from the proof of Theorem 2. Next, let two different prefilters $\mathbf{H}(z)$ and $\mathbf{H}_1(z)$ both be optimal for all $\mathbf{x}(n) \in \ell_2$. Thus by the uniqueness of $\mathbf{c}(n)$ we have that

$$[(\mathbf{H}(z) - \mathbf{H}_1(z))\mathbf{X}(z)]_{\downarrow M} = \mathbf{0}$$

for any choice of $\mathbf{X}(z)$. The choice $\mathbf{X}(z) = z^k\mathbf{e}_i$, where $\mathbf{e}_i$ is the ith unit vector (i.e. the ith column of the identity matrix), implies that the kth polyphase component of the ith column of $\mathbf{H}(z)-\mathbf{H}_1(z)$ is zero. This holds for all i and k, so the conclusion is that all the polyphase components of all columns of $\mathbf{H}(z) - \mathbf{H}_1(z)$ are zero, and thus $\mathbf{H}(z) = \mathbf{H}_1(z)$, i.e. the prefilter is indeed unique.

$\bigtriangledown\bigtriangledown\bigtriangledown$

5.2 Multiwavelets and Prefiltering

Multiwavelet theory emerged recently as the extension of wavelet theory to the case where there is more than one scaling function and mother wavelet. It has been shown [16] that multiwavelets have some advantages over the conventional wavelets, especially in data compression. In this section we provide the connection between MIMO biorthogonal partners and prefilters employed in multiwavelet theory. To that end we first give a brief overview of some of the results in this area. For a more thorough and comprehensive exposition to multiwavelets, reader is referred to the works by Geronimo et al. [5, 4], Xia et al. [26, 25], Vetterli and Strang [21, 8, 15], and Selesnick [14, 13].

Consider the set of N scaling functions $\{\phi_n(t)\}$, $0 \le n \le N-1$ and the corresponding set of N mother wavelets $\{\psi_n(t)\}$. The scaling functions are chosen in such way that their integer shifts $\{\phi_n(t-k)\}$, together with the shifts of the dilated versions $\{\phi_n(2^j t - k)\}$ (integer j is called the *scale*) span a sequence of nested subspaces of L_2. These subspaces $V_j \subset V_{j+1}$, $j \in \mathbb{Z}$ form a multiresolution analysis [5] of L_2. Some of the desirable properties of the scaling functions are linear phase, orthogonality and compact support. In the scalar case (for $N=1$) these properties occur simultaneously only in the Haar basis, while in the multiwavelet case ($N>1$) many such examples are known [5, 15, 14]. In the following we will assume that the scaling functions are orthogonal and compactly supported. Let $x_c(t)$ (subscript c stands for *continuous*) be a continuous-time signal contained in V_0. Then it can be written as

$$\begin{aligned} x_c(t) &= \sum_{n=0}^{N-1} \sum_{k\in\mathbb{Z}} c_{0,n}(k)\phi_n(t-k) \\ &= \sum_{n=0}^{N-1} \sum_{k\in\mathbb{Z}} c_{J_0,n}(k) 2^{J_0/2}\phi_n(2^{J_0}t-k) + \sum_{n=0}^{N-1} \sum_{J_0\le j<0} \sum_{k\in\mathbb{Z}} d_{j,n}(k) 2^{j/2}\psi_n(2^j t-k), \end{aligned} \tag{30}$$

where $J_0, j < 0$. This decomposition algorithm suggests possibilities for a tree-like signal decomposition in terms of the coefficients at coarser scales. In fact, given the coefficients $\{c_{0,n}(k)\}$ the corresponding coefficients $\{c_{j,n}(k)\}$ and $\{d_{j,n}(k)\}$, $j<0$ can be found using the concept of vector-valued wavelets and filterbanks (see [26] and references therein). Let us denote the vectors of stacked coefficients at scale j

$$\mathbf{c}_j(k) = [c_{j,0}(k) \;\; c_{j,1}(k) \;\; \cdots \;\; c_{j,N-1}(k)]^T, \;\; \mathbf{d}_j(k) = [d_{j,0}(k) \;\; d_{j,1}(k) \;\; \cdots \;\; d_{j,N-1}(k)]^T.$$

Starting from the dilation equation on scaling functions and wavelets, it can be shown that

$$\mathbf{c}_{j-1}(k) = \sum_{m\in\mathbb{Z}} \mathbf{H}_m \mathbf{c}_j(2k-m) \;\; \text{and} \;\; \mathbf{d}_{j-1}(k) = \sum_{m\in\mathbb{Z}} \mathbf{G}_m \mathbf{c}_j(2k-m)$$

for some appropriate matrix polynomials $\mathbf{H}(z) = \sum_m \mathbf{H}_m z^{-m}$ and $\mathbf{G}(z) = \sum_m \mathbf{G}_m z^{-m}$. In other words, this gives rise to the vector-valued filterbank pyramid for decomposition, as shown in Fig. 8. From the orthogonality condition on the scaling functions it follows [26] that the reconstruction part is as shown in the lower part of Fig. 8.

The initial step in the multiwavelet decomposition is the so-called **prefiltering** [26, 25] and is not shown in Fig. 8. In order to understand the nature of this operation, consider again the first equality in (30) and rewrite it at instances $t = \frac{n}{M}$ (the significance of the factor M will be clear shortly)

$$x(n) \stackrel{\text{def}}{=} x_c(\frac{n}{M}) = \sum_{k\in\mathbb{Z}} \boldsymbol{\phi}_c^T(\frac{n}{M}-k)\mathbf{c}_0(k), \tag{31}$$

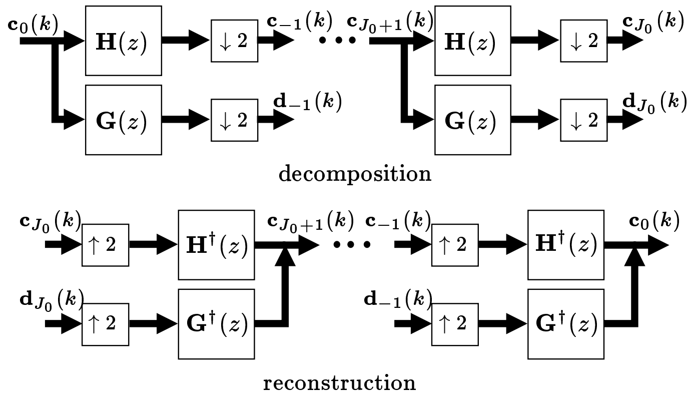

Figure 8: Vector-valued wavelet pyramid: decomposition to coarser scales and reconstruction.

where we introduced the column vector $\boldsymbol{\phi}_c(t) = [\phi_0(t) \;\; \phi_1(t) \;\; \cdots \;\; \phi_{N-1}(t)]^T$. Now, let us define the following

$$\boldsymbol{\phi}_M(n) \stackrel{\text{def}}{=} \boldsymbol{\phi}_c(\frac{n}{M}), \quad \text{with} \quad \boldsymbol{\Phi}_M(z) = \sum_{k\in\mathbb{Z}} \boldsymbol{\phi}_M(n) z^{-n}.$$

Then we can rewrite (31) as

$$x(n) = \sum_{k\in\mathbb{Z}} \boldsymbol{\phi}_M^T(n-kM)\mathbf{c}_0(k), \quad \text{or} \quad X(z) = \boldsymbol{\Phi}_M^T(z)\mathbf{C}_0(z^M). \tag{32}$$

This is depicted in Fig. 9(a). Notice here that $x(n)$ is a sequence of samples of $x_c(t)$ oversampled by M. The purpose of prefiltering is to obtain the multiwavelet coefficients $\mathbf{c}_0(k)$ at scale 0 from the signal samples $x(n)$. This prefiltering is unnecessary in the case of the so called interpolating multiwavelets [13]. The scaling functions there are chosen such that the multiwavelet coefficients are equal to signal samples. Balanced multiwavelets [8, 14] present similar attempts at circumventing the prefiltering operation, however they are in general only approximate methods. Therefore, in the most general multiwavelet setting the prefiltering operation is unavoidable. To understand how this operation is performed, we define the scalar filter $F_n(z)$ to be the nth entry in the column vector $\boldsymbol{\Phi}_M(z)$. In other words we have

$$F_n(z) = \sum_{k\in\mathbb{Z}} \phi_n(\frac{k}{M}) z^{-k}.$$

Now we can redraw the signal model from Fig. 9(a) as in the left half of Fig. 9(b). Notice how this figure resembles the transmultiplexer [19], where the composite signal $x(k)$ is obtained as the sum of signals $c_{0,n}(k)$ modulated by filters $F_n(z)$, for $n = 0, 1, \ldots, N-1$. The recovery of partial signals $c_{0,n}(k)$ is then achieved as in the right part of Fig. 9(b), for appropriately chosen filters $H_n(z)$.

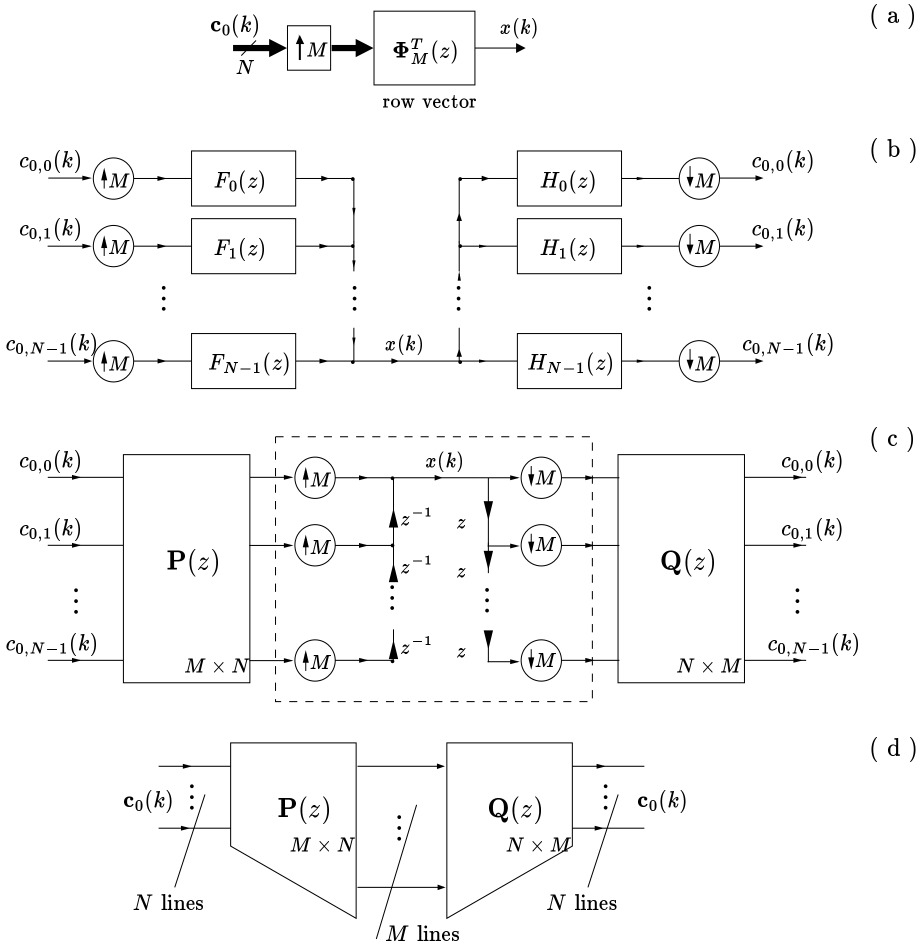

Figure 9: Prefiltering for multiwavelet transform. (a) Signal model. (b) Equivalent drawing of (a) together with the prefiltering part. (c) Equivalent drawing using polyphase matrices. (d) Final form of the traditional method for prefiltering by left-inverting the polyphase matrix. See text.

In order to find those *prefilters* $H_n(z)$, we present both filterbanks from Fig. 9(b) in terms of their polyphase matrices [19], which is shown in Fig. 9(c). Those matrices are defined by

$$\mathbf{P}(z) = \begin{bmatrix} F_{0,0}(z) & F_{0,1}(z) & \cdots & F_{0,N-1}(z) \\ F_{1,0}(z) & F_{1,1}(z) & \cdots & F_{1,N-1}(z) \\ & & & \\ F_{M-1,0}(z) & F_{M-1,1}(z) & \cdots & F_{M-1,N-1}(z) \end{bmatrix}, \quad \text{and} \tag{33}$$

$$\mathbf{Q}(z) = \begin{bmatrix} H_{0,0}(z) & H_{0,1}(z) & \cdots & H_{0,M-1}(z) \\ H_{1,0}(z) & H_{1,1}(z) & \cdots & H_{1,M-1}(z) \\ & & & \\ H_{N-1,0}(z) & H_{N-1,1}(z) & \cdots & H_{N-1,M-1}(z) \end{bmatrix}. \tag{34}$$

The entries in the above matrices $\mathbf{P}(z)$ and $\mathbf{Q}(z)$ are nothing but the Type-1 and Type-2 polyphase components of order M of filters $F_n(z)$ and $H_n(z)$, respectively given by

$$F_n(z) = \sum_{k=0}^{M-1} F_{k,n}(z^M)z^{-k}, \quad H_n(z) = \sum_{k=0}^{M-1} H_{n,k}(z^M)z^{k}.$$

Now we see that given the samples $x_c(\frac{n}{M})$, we can obtain the corresponding multiwavelet coefficients $\mathbf{c}_0(n)$ at scale 0 as shown in the right half of Fig. 9(c). Notice here that $\mathbf{P}(z)$ is a known $M \times N$ matrix and $\mathbf{Q}(z)$ is an unknown $N \times M$ matrix, which describes the prefiltering operation. Next we note that the middle part of Fig. 9(c) (enclosed in a dashed box) is equivalent to identity, since it is a cascade of unblocking and blocking operators. Therefore, this is equivalently redrawn as in Fig. 9(d). The matrix $\mathbf{Q}(z)$ is obtained as a **left inverse** of $\mathbf{P}(z)$ (see Fig. 9(d)), i.e. $\mathbf{Q}(z)\mathbf{P}(z) = \mathbf{I}$. From here it immediately follows that the minimum amount of oversampling $M \geq N$ is necessary, since we need $\mathbf{P}(e^{j\omega})$ to have rank N on the unit circle.

In the following we show the connection between the polynomial matrix inversion problem (as motivated in the multiwavelet setting) and the problem of constructing MIMO biorthogonal partners. As it turns out, these two problems are completely equivalent. For simplicity here we assume that M is an even integer, i.e. $M = 2L$. Then, we can redraw Fig. 9(b) as in Fig. 10(a). Now we consider the middle part of Fig. 10(a) (enclosed in a dashed box) and present it using the polyphase matrices of filters $F_n(z)$ and $H_n(z)$, only now with respect to $L = \frac{M}{2}$. This is shown in Fig. 10(b), with

$$\bar{\mathbf{P}}(z) = \begin{bmatrix} \bar{F}_{0,0}(z) & \bar{F}_{0,1}(z) & \cdots & \bar{F}_{0,N-1}(z) \\ \bar{F}_{1,0}(z) & \bar{F}_{1,1}(z) & \cdots & \bar{F}_{1,N-1}(z) \\ & & & \\ \bar{F}_{L-1,0}(z) & \bar{F}_{L-1,1}(z) & \cdots & \bar{F}_{L-1,N-1}(z) \end{bmatrix}, \quad \text{and} \tag{35}$$

$$\bar{\mathbf{Q}}(z) = \begin{bmatrix} \bar{H}_{0,0}(z) & \bar{H}_{0,1}(z) & \cdots & \bar{H}_{0,L-1}(z) \\ \bar{H}_{1,0}(z) & \bar{H}_{1,1}(z) & \cdots & \bar{H}_{1,L-1}(z) \\ & & & \\ \bar{H}_{N-1,0}(z) & \bar{H}_{N-1,1}(z) & \cdots & \bar{H}_{N-1,L-1}(z) \end{bmatrix}, \tag{36}$$

and the entries in the above matrices satisfying

$$F_n(z) = \sum_{k=0}^{L-1} \bar{F}_{k,n}(z^L)z^{-k}, \quad H_n(z) = \sum_{k=0}^{L-1} \bar{H}_{n,k}(z^L)z^{k}.$$

Comparing Fig. 10(b) with Fig. 1, we see that $\bar{\mathbf{Q}}(z)$ is found as a left biorthogonal partner of $\bar{\mathbf{P}}(z)$ with respect to 2, or $[\bar{\mathbf{Q}}(z)\bar{\mathbf{P}}(z)]_{\downarrow 2} = \mathbf{I}$. Therefore, we conclude that the problem of finding a left inverse $\mathbf{Q}(z)$ as in Fig. 9(d) is completely equivalent to finding a LBP $\bar{\mathbf{Q}}(z)$ as in Fig. 10(b). Moreover, comparing (35) and (36) to (33) and (34), we can easily verify the following relation

$$\begin{bmatrix} \bar{\mathbf{P}}_0(z) \\ \bar{\mathbf{P}}_1(z) \end{bmatrix} = \mathbf{P}(z), \quad \begin{bmatrix} \bar{\mathbf{Q}}_0(z) & \bar{\mathbf{Q}}_1(z) \end{bmatrix} = \mathbf{Q}(z), \quad \text{with} \tag{37}$$

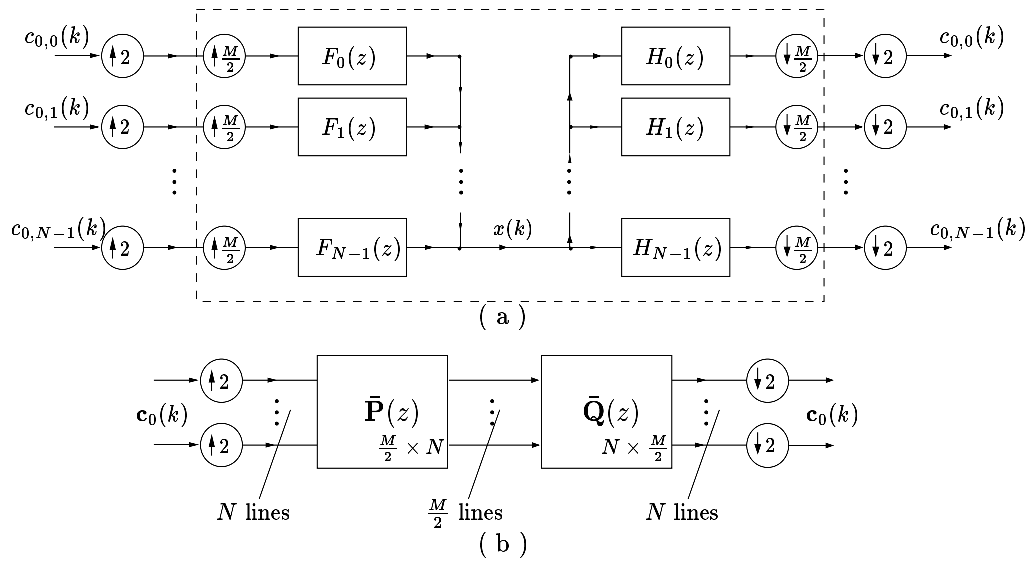

Figure 10: Biorthogonal partners in prefiltering for multiwavelet transform. (a) Equivalent of Fig. 9(b) for even M. (b) Equivalent of Fig. 9(d) for even M. See text.

$$\bar{\mathbf{P}}(z) = \bar{\mathbf{P}}_0(z^2) + z^{-1}\bar{\mathbf{P}}_1(z^2), \quad \text{and} \quad \bar{\mathbf{Q}}(z) = \bar{\mathbf{Q}}_0(z^2) + z\bar{\mathbf{Q}}_1(z^2). \tag{38}$$

From the relation between the solutions $\mathbf{Q}(z)$ and $\bar{\mathbf{Q}}(z)$, it is evident that many properties of one solution immediately apply to the other as well. For example, if one matrix is rational/FIR, the other will also be rational/FIR. The previous findings are summarized in the following lemma.

Lemma 1. Consider the continuous-time signal $x_c(t) \in V_0$, given by its samples $x_c(\frac{k}{M})$. Let the space V_0 be spanned by a set of N scaling functions $\{\phi_n(t)\}$ and their integer shifts; furthermore, let $M \geq N$ be an even integer and let $\mathbf{P}(z)$ be the corresponding polyphase matrix defined by (33). Then there will exist (an FIR) prefilter $\mathbf{Q}(z)$ for the corresponding multiwavelet transform if and only if there exists (an FIR) LBP $\bar{\mathbf{Q}}(z)$ of the $\frac{M}{2} \times N$ polyphase matrix $\bar{\mathbf{P}}(z)$ given by (35). The relation between them is given by (37) and (38).

Significance of Lemma 1. First notice that the equivalence between the matrix inversion problem and that of finding a LBP holds for matrix transfer functions in general, and is not restricted to polyphase matrices of multiscaling functions. Therefore, we can use one method to solve the other problem, and vice versa. This approach sometimes proves to be beneficial. Moreover, we can use the results that are well understood in one setting and apply them in the other setting. As an example, recall the discussion after the proof of Theorem 3. In order to prove the necessary relation (10) between the dimensions of the polynomial matrix $\mathbf{F}(z)$, we used the well-known fact that fat matrices do not have left inverses. Finally, the construction of grcd's and

MIMO biorthogonal partners have been studied extensively in Sec. 3 and Sec. 4 and all the results shown there can be applied in the prefilter construction problem as introduced in Sec. 5.2. In particular, the non-uniqueness of a left FIR inverse (thus a prefilter matrix) can be exploited in a similar fashion as described in Sec 4. This may become useful in many applications where prefilters are required to have certain properties [8, 26, 25], and is an interesting topic for further research.

6 Concluding Remarks

Multiple input multiple output (MIMO) biorthogonal partners arise naturally in many signal processing applications including MIMO channel equalization and multiwavelet theory. The theory of MIMO biorthogonal partners is to some extent a natural extension of the SISO case as considered in [20]. However, in the vector case we are presented with additional degrees of freedom which may be useful in some applications. In this paper we have considered many theoretical aspects of MIMO biorthogonal partners and pointed out several applications.

7 Appendix

In the appendix we provide the statement and the proof of the simple Bezout identity for rectangular matrices. This result occurs frequently in linear systems theory (see [7] for example), but is usually stated for the case where one of the matrices is square. In this context (cf. Sec. 3) both matrices are of the same size and are allowed to be rectangular.

Theorem 5. Rectangular $p \times r$ polynomial matrices $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$ are right coprime if and only if there exist polynomial matrices $\mathbf{H}_0(z)$ and $\mathbf{H}_1(z)$ such that

$$\mathbf{H}_0(z)\mathbf{F}_0(z) + \mathbf{H}_1(z)\mathbf{F}_1(z) = \mathbf{I}.$$

Before proving the simple Bezout identity for rectangular matrices, we need to introduce the following results.

Lemma 2. Given any two $p \times r$ polynomial matrices $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$, there exists a unimodular matrix $\mathbf{U}(z)$, such that (12) holds true for some $\mathbf{R}(z)$. This $\mathbf{R}(z)$ is a grcd of $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$. Furthermore, if $\mathbf{R}_0(z)$ is any grcd of $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$, it can be written as

$$\mathbf{W}_0(z)\mathbf{F}_0(z) + \mathbf{W}_1(z)\mathbf{F}_1(z) = \mathbf{R}_0(z), \tag{39}$$

for some polynomial matrices $\mathbf{W}_0(z)$ and $\mathbf{W}_1(z)$.

Proof of Lemma 2. The first statement about the existence of a unimodular matrix $\mathbf{U}(z)$ is proved in the general case (see Theorem 6.3-2. of [7]). Now, since $\mathbf{U}(z)$ is unimodular, its inverse is a polynomial matrix; call it $\mathbf{V}(z)$. Therefore, from (12) we have

$$\begin{bmatrix} \mathbf{F}_0(z) \\ \mathbf{F}_1(z) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{V}_{11}(z) & \mathbf{V}_{12}(z) \\ \mathbf{V}_{21}(z) & \mathbf{V}_{22}(z) \end{bmatrix}}_{\mathbf{V}(z)} \begin{bmatrix} \mathbf{R}(z) \\ \mathbf{0} \end{bmatrix},$$

which implies

$$\mathbf{F}_0(z) = \mathbf{V}_{11}(z)\mathbf{R}(z); \quad \mathbf{F}_1(z) = \mathbf{V}_{21}(z)\mathbf{R}(z).$$

Thus $\mathbf{R}(z)$ is a common right divisor of $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$. If $\bar{\mathbf{R}}(z)$ is any other common right divisor, we can write similarly

$$\mathbf{F}_0(z) = \mathbf{X}(z)\bar{\mathbf{R}}(z); \quad \mathbf{F}_1(z) = \mathbf{Y}(z)\bar{\mathbf{R}}(z) \tag{40}$$

with the appropriately chosen polynomial matrices $\mathbf{X}(z)$ and $\mathbf{Y}(z)$. Now, reading the first matrix equality from (12) we get

$$\mathbf{U}_{11}(z)\mathbf{F}_0(z) + \mathbf{U}_{12}(z)\mathbf{F}_1(z) = \mathbf{R}(z), \tag{41}$$

which after substituting (40) leads to

$$(\mathbf{U}_{11}(z)\mathbf{X}(z) + \mathbf{U}_{12}(z)\mathbf{Y}(z))\bar{\mathbf{R}}(z) = \mathbf{R}(z),$$

and this proves that $\mathbf{R}$(z) is indeed a grcd of $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$. Finally, we use the fact that any two grcd's are identical up to a premultiplication by a unimodular matrix [7]. Therefore, an arbitrary grcd $\mathbf{R}_0(z)$ can be written as $\mathbf{R}_0(z) = \mathbf{T}(z)\mathbf{R}(z)$. Applying this to (41) leaves us with

$$\underbrace{\mathbf{T}(z)\mathbf{U}_{11}(z)}_{\mathbf{W}_0(z)}\mathbf{F}_0(z) + \underbrace{\mathbf{T}(z)\mathbf{U}_{12}(z)}_{\mathbf{W}_1(z)}\mathbf{F}_1(z) = \mathbf{R}_0(z),$$

which concludes the proof of Lemma 2. $\bigtriangledown\bigtriangledown\bigtriangledown$

Proof of Theorem 5. First we show the "if part". From Lemma 2 we have that if $\mathbf{R}(z) = \text{grcd}[\mathbf{F}_0(z), \mathbf{F}_1(z)]$, then there exist polynomial matrices $\bar{\mathbf{H}}_0(z)$ and $\bar{\mathbf{H}}_1(z)$ such that

$$\bar{\mathbf{H}}_0(z)\mathbf{F}_0(z) + \bar{\mathbf{H}}_1(z)\mathbf{F}_1(z) = \mathbf{R}(z). \tag{42}$$

If $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$ are right coprime, then $\mathbf{H}_0(z) \stackrel{\text{def}}{=} \mathbf{R}^{-1}(z)\bar{\mathbf{H}}_0(z)$ and $\mathbf{H}_1(z) \stackrel{\text{def}}{=} \mathbf{R}^{-1}(z)\bar{\mathbf{H}}_1(z)$ are also polynomial matrices. Thus, premultiplying both sides of (42) by $\mathbf{R}^{-1}(z)$, the "if part" follows.

Conversely, suppose there exist polynomial matrices $\mathbf{H}_0(z)$ and $\mathbf{H}_1(z)$ such that (11) holds. Let $\mathbf{R}(z)$ be any grcd of $\mathbf{F}_0(z)$ and $\mathbf{F}_1(z)$. It follows that $\mathbf{F}_0(z) = \bar{\mathbf{F}}_0(z)\mathbf{R}(z)$ and $\mathbf{F}_1(z) = \bar{\mathbf{F}}_1(z)\mathbf{R}(z)$ for appropriate polynomial matrices $\bar{\mathbf{F}}_0(z)$ and $\bar{\mathbf{F}}_1(z)$. Thus we have

$$(\mathbf{H}_0(z)\bar{\mathbf{F}}_0(z) + \mathbf{H}_1(z)\bar{\mathbf{F}}_1(z))\mathbf{R}(z) = \mathbf{I}, \quad \text{or} \quad \mathbf{R}^{-1}(z) = \mathbf{H}_0(z)\bar{\mathbf{F}}_0(z) + \mathbf{H}_1(z)\bar{\mathbf{F}}_1(z).$$

Since the right-hand side of the above equation is a polynomial matrix, this shows that $\mathbf{R}(z)$ is a unimodular matrix, which concludes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

References

[1] A. Aldroubi, "Oblique projections in atomic spaces," *Proc. Amer. Math. Soc.*, vol. 124, pp. 2051-2060, 1996.

[2] A. Aldroubi and M. Unser, "Oblique projections in discrete signal subspaces of ℓ_2 and the wavelet transform," Proc. SPIE, vol. 2303, *Wavelet applications in signal and image processing, II*, pp. 36-45, San Diego, CA, 1994.

[3] C. de Boor, *A Practical Guide to Splines*. New York: Springer-Verlag 1978.

[4] G. C. Donovan, J. S. Geronimo and D. P. Hardin, "Orthogonal polynomials and the construction of piecewise polynomial smooth wavelets," *SIAM J. Math. Anal.*, vol. 30 (5), pp. 1029-1056, 1999.

[5] J. S. Geronimo, D. P. Hardin and P. R. Massopust, "Fractal functions and wavelet expansions based on several scaling functions," *J. Approx. Theory*, vol. 78, pp. 373-401, 1994.

[6] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, Inc., Upper Saddle River, N.J., 1996.

[7] T. Kailath, *Linear Systems*. Prentice Hall, Inc., Englewood Cliffs, N.J., 1980.

[8] J. Lebrun and M. Vetterli, "Balanced multiwavelets theory and design," *IEEE Trans. Signal Processing*, vol. 46 (4), pp. 1119-1125, Apr. 1998.

[9] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, London, 1998.

[10] H. S. Malvar and D. H. Staelin, "Optimal pre- and postfilters for multichannel signal processing," *IEEE Trans. Acoust., Speech and Signal Process.*, vol. 36 (2), pp. 287-289, Feb. 1988.

[11] J. G. Proakis, *Digital Communications,* McGraw-Hill, New York, 1995.

[12] A. Scaglione, G. B. Giannakis and S. Barbarossa, "Redundant filter bank precoders and equalizers Part I: Unification and optimal design," *IEEE Trans. Signal Proc.*, vol. 47 (7), pp. 1988-2006, July 1999.

[13] I. W. Selesnick, "Interpolating multiwavelet bases and the sampling theorem," *IEEE Trans. Signal Processing*, vol. 47 (6), pp. 1615-1621, June 1999.

[14] I. W. Selesnick, "Multiwavelet bases with extra approximation properties," *IEEE Trans. Signal Processing*, vol. 46 (11), pp. 1898-2908, Nov. 1998.

[15] G. Strang and V. Strela, "Short wavelets and matrix dilation equation," *IEEE Trans. Signal Processing*, vol. 43 (1), pp. 108-115, Jan. 1995.

[16] V. Strela, P. N. Heller, G. Strang, P. Topiwala, and C. Heil "The application of multiwavelet filterbanks to image processing," *IEEE Trans. Image Processing*, vol. 8(4), Apr. 1999.

[17] J. R. Treichler, I. Fijalkow and C. R. Johnson, Jr., "Fractionally spaced equalizers: how long should they really be?," *IEEE Signal Processing Magazine*, pp. 65-81, May 1996.

[18] M. Unser, A. Aldroubi and M. Eden, "B-spline signal processing: Part I - Theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 821-833, Feb. 1993.

[19] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Prentice Hall, Inc., Englewood Cliffs, N.J.,1993.

[20] P. P. Vaidyanathan and B. Vrcelj, "Biorthogonal partners and applications," *IEEE Trans. Signal Processing*, vol. 49(5), pp. 1013-1028, May 2001.

[21] M. Vetterli and G. Strang, "Time-varying filter banks and multiwavelets," *Sixth Digital Signal Processing Workshop, IEEE*, pp. 223-226, 1994.

[22] B. Vrcelj and P. P. Vaidyanathan, "On the general form of FIR MIMO biorthogonal partners," *Proc. of the 35th Asilomar Conf. on Sign., Syst. and Comp.*, Pacific Grove, CA, Nov. 2001.

[23] B. Vrcelj and P. P. Vaidyanathan, "Results on vector biorthogonal partners," *Proceedings ICASSP*, Salt Lake City, Utah, May 2001.

[24] B. Vrcelj and P. P. Vaidyanathan, "Theory of MIMO biorthogonal partners and their application in channel equalization," *Proceedings ICC*, Helsinki, Finland, Jun. 2001.

[25] X.-G. Xia, "A new prefilter design for discrete multiwavelet transform," *IEEE Trans. Signal Processing*, vol. 46 (6), pp. 1558-1570, June 1998.

[26] X.-G. Xia, J. S. Geronimo, D. P. Hardin and B. W. Suter, "Design of prefilters for discrete multiwavelet transforms," *IEEE Trans. Signal Processing*, vol. 44, pp. 25-35, Jan. 1996.

[27] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Comm.*, vol. 42(12), pp. 3221-3231, Dec. 1994.

[28] `http://www.systems.caltech.edu/bojan/papers/mimo_jrnl.html`