Remarks on certain new methods for blind identification of FIR channels

P. P. Vaidyanathan and Borching Su

Dept. Electrical Engr., Caltech, Pasadena, CA
ppvnath@systems.caltech.edu borching@systems.caltech.edu

Abstract. This paper discusses a number of issues pertaining to blind identification of channels. The basics of blind identification are first discussed and a method called Vandermonde method is presented which is based on elementary linear system principles. Then some remarks are made about precoders with paraunitary antipodal preprocessors. It is argued that such preprocessors usually destroy signal richness which is a necessary feature in blind identification systems.[1]

1. INTRODUCTION

Our discussions in this paper will center around the transmultiplexer structure shown in Fig. 1. The signals $s_k(n)$ are symbol streams such as PAM or QAM signals, [4]. These could be symbols generated by different users who wish to transmit messages over the channel. Or they could be different independent parts of the signals generated by one user [11]. The symbol streams $s_k(n)$ are passed through the interpolation filters or transmitter filters $F_k(z)$ to produce the signals

$$x_k(n) = \sum_i s_k(i) f_k(n - iP)$$

The sum $x(n)$ of the signals $x_k(n)$ is then transmitted over a common channel. The channel is described by a linear time invariant filter $C(z)$ followed by additive noise. At the receiver end, the filters $H_k(z)$ have the task of separating the signals and reducing them to the original rates by P-fold decimation. Since the M signals are mutiplexed into one channel, it is necessary to have $P \geq M$. When $P > M$ we have a *redundant transmultiplexer*.

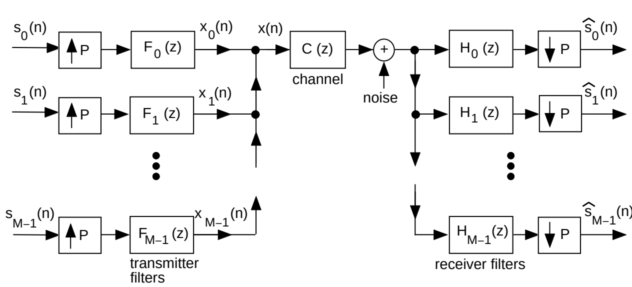

Fig. 1. The transmultiplexer system ($P \geq M$).

[1]Work supported in part by the NSF grant CCF-0428326 and Caltech's Moore Fellowship.

Figure 2 shows the polyphase representation for this system [2], [10]. It is possible to choose the filters $\{F_k(z)\}$ and $\{H_k(z)\}$ or equivalently $\mathbf{R}(z)$ and $\mathbf{E}(z)$ such that there is perfect symbol recovery ($\widehat{s}_k(n) = s_k(n)$) in absence of channel noise. Instead of perfect symbol recovery, it is also possible to design the receiver filters for minimum mean square error [2], [5] or for minimum bit error rate.

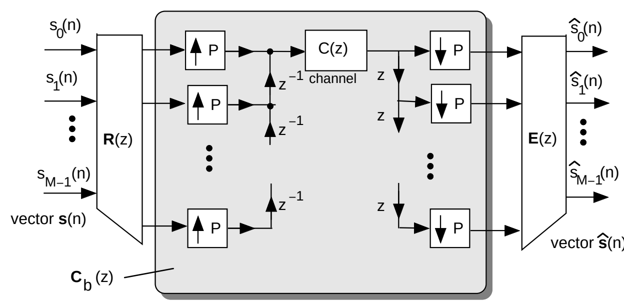

Fig. 2. Polyphase representation of Fig. 1.

If the channel transfer function is unknown, it has to be estimated before equalization. Such estimation can be done either with the help of training signals or by blind identification methods. It is well known that with non redundant systems ($P = M$) blind identification is not possible unless we use fourth order moments such as the Kurtosis of the data [6]. However blind identification is indeed possible without the use of fourth order moments, if we use *redundant* transmultiplexers or filter bank precoders [5].

In this paper we discuss a number of issues pertaining to blind identification of channels. The basics of blind identification are first discussed and a method called Vandermonde method is presented which is based on elementary linear system principles. We then make some remarks about precoders with paraunitary antipodal preprocessors and argue that paraunitary preprocessors usually destroy signal richness which is a necessary feature in blind identification systems. The reader not familiar with recent literature can get a quick tour from a recent tutorial on filter bank precoders and blind identification systems [12].

2. BLIND IDENTIFICATION

First a few words on blind identification based on filter bank precoders. Assume that the channel is FIR with

$$C(z) = \sum_{n=0}^{L} c(n) z^{-n},$$

and that $P \geq M + L$. Assume further that the receiver filters have order $\leq P-1$, and that the transmitting filters have order $\leq M-1$. In particular therefore $\mathbf{R}(z)$ has the form

$$\mathbf{R}(z) = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} \tag{1}$$

Figure 3 shows the path from the transmitted symbols to the channel output $y(n)$. For convenience we consider the blocked version $\mathbf{y}(n)$ as indicated. With the vector $\mathbf{s}(n)$ as in the figure, it can be shown that [5], [2], [12]

$$\mathbf{y}(n) = \mathbf{A}\mathbf{R}_1\mathbf{s}(n)$$

where $\mathbf{A}$ is a fully banded Toeplitz matrix given by

$$\mathbf{A} = \begin{bmatrix} c(0) & 0 & \dots & 0 \\ c(1) & c(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c(L) & & & \\ 0 & c(L) & & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & c(L) \end{bmatrix} \tag{2}$$

Notice that $\mathbf{A}$ is $P \times M$.

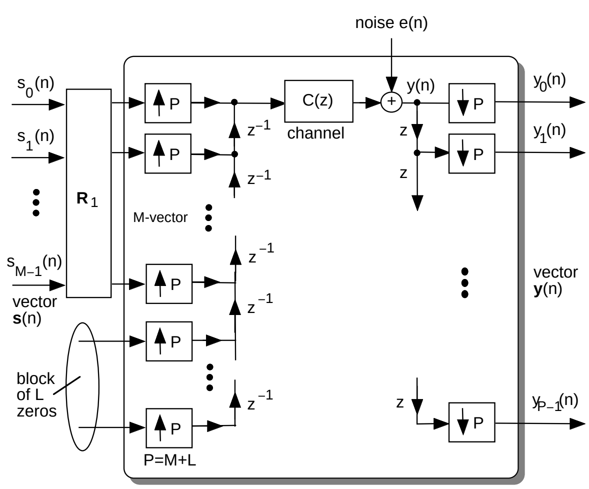

Fig. 3. The zero padding system with precoder $\mathbf{R}_1$.

Assume the channel $c(n)$ is unknown. Under certain conditions it can be argued that the observation of $\mathbf{y}(n)$ can be used to identify the channel $c(n)$ upto a scale-factor ambiguity [5], [12]. This is called *blind identification* because the input stream $\mathbf{s}(n)$ is unknown, unlike in training-based channel identification. Imagine we observe the output vector $\mathbf{y}(n)$ for a certain duration, say $0 \leq n \leq J-1$, and write the equation

$$\underbrace{\begin{bmatrix} \mathbf{y}(0) & \mathbf{y}(1) & \dots & \mathbf{y}(J-1) \end{bmatrix}}_{\mathbf{Y} \text{ matrix; size } P\times J}$$

$$= \underbrace{\mathbf{A}}_{P\times M} \underbrace{\mathbf{R}_1}_{M\times M} \underbrace{\begin{bmatrix} \mathbf{s}(0) & \mathbf{s}(1) & \dots & \mathbf{s}(J-1) \end{bmatrix}}_{\mathbf{S} \text{ matrix; size } M\times J} \tag{3}$$

In order for blind identification to succeed it is necessary [5,12] for the input to be *rich,* that is, there exists a J such that $\mathbf{S}$ has full rank M. Since $\mathbf{A}$ and $\mathbf{R}_1$ have rank M, the product on the right hand side of Eq. (3) has rank M. So the $P \times J$ data matrix $\mathbf{Y}$ has rank M, and there are $P-M$ linearly independent vectors orthogonal to all the columns in $\mathbf{Y}$, that is,

$$\mathbf{V}^\dagger \mathbf{Y} = \mathbf{0} \tag{4}$$

for some $(P-M)\times P$ matrix $\mathbf{V}$ with rank $P-M$. Starting from this idea it is possible to identify the leftmost column of (2) upto a scale-factor ambiguity. For the case where $P = M + L$ this is especially simple, and is explained in great detail in [5]. The method can also be modified to work under the practical situation where there is channel noise [5], [12]. Variations of the method which work in the frequency domain have been reported in [13]. The advantage of such variations is that they are more readily applicable to OFDM systems based on cyclic-prefix which uses frequency-domain equalizers.

3. THE VANDERMONDE METHOD

We now argue that the principle of blind identification follows from elementary system theory. The discussion also places in evidence a general method. Equation (4) is true for any $P \geq M + L$. Since $\mathbf{Y} = \mathbf{A}\mathbf{R}_1\mathbf{S}$ where $\mathbf{S}$ has full rank M, this implies

$$\mathbf{V}^\dagger \mathbf{A}\mathbf{R}_1 = \mathbf{0} \tag{5}$$

Observe now that we have the identity

$$\begin{bmatrix} 1 & z^{-1} & z^{-2}\dots & z^{-(P-1)} \end{bmatrix} \underbrace{\begin{bmatrix} c(0) & 0 & \dots & 0 \\ c(1) & c(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c(L) & & & \\ 0 & c(L) & & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & c(L) \end{bmatrix}}_{\mathbf{A}}$$

$$= C(z) \begin{bmatrix} 1 & z^{-1} & z^{-2}\dots & z^{-(P-L-1)} \end{bmatrix} \tag{6}$$

for any z. This follows from the fact that $\mathbf{A}$ is a fully banded Toeplitz matrix. Eq. (6) is a direct consequence of the fact that exponentials are eigenfunctions of LTI systems. If z_k is a zero of $C(z)$ it then follows that

$$\begin{bmatrix} 1 & z_k^{-1} & z_k^{-2}\dots & z_k^{-(P-1)} \end{bmatrix} \mathbf{A} = \mathbf{0}$$

If $C(z)$ has L *distinct zeros* z_k it follows therefore that there are L vectors of the form

$$\mathbf{z}_k^\dagger = \begin{bmatrix} 1 & z_k^{-1} & z_k^{-2}\dots & z_k^{-(P-1)} \end{bmatrix} \tag{7}$$

which annihilate $\mathbf{A}$ from the left. These therefore annihilate $\mathbf{Y}$ as well. Vectors of the form (7) will be referred to as *Vandermonde vectors.* Assume for a moment that there are no other Vandermonde vectors annihilating $\mathbf{Y}$. Then, simply by identifying the L Vandermonde vectors annihilating the data $\mathbf{Y}$ we can find all the L zeros of $C(z)$. This identifies $C(z)$ upto a scale factor.

We now explain how the numbers z_k can be identified in the first place. Notice that a Vandermonde vector (7) annihilates a column vector with elements $a(0), a(1), \ldots$ if and only if the polynomial (or FIR filter) $\sum_{n=0}^{P-1} a(n)z^{-n}$ has a zero at z_k. Thus, once we have the data matrix $\mathbf{Y}$ we simply identify the $P-1$ zeros of each of the J polynomials defined by the J columns of $\mathbf{Y}$. From this we can identify all the zeros which are common to all the J polynomials. There are at least L such common zeros, namely the L distinct zeros of the channel, and these identify the channel as explained above.

It only remains to address the situation where there are more than L common zeros among the polynomials defined by the columns of $\mathbf{Y}$. For example let z_L be such a zero. This means that the Vandermonde vector

$$\mathbf{z}_L^\dagger = \begin{bmatrix} 1 & z_L^{-1} & z_L^{-2} \ldots & z_L^{-(P-1)} \end{bmatrix}$$

annihilates $\mathbf{Y}$ though $C(z_L) \neq 0$. Observe that

$$\mathbf{z}_L^\dagger \mathbf{Y} = \mathbf{z}_L^\dagger \mathbf{A} \mathbf{R}_1 \mathbf{S} = C(z_L) \widehat{\mathbf{z}}_L^\dagger \mathbf{R}_1 \mathbf{S}$$

where the second equality follows from (6). Here $\widehat{\mathbf{z}}_L^\dagger$ is similar to $\mathbf{z}_L^\dagger$ but only has the first M entries. Thus the condition $\mathbf{z}_L^\dagger \mathbf{Y} = \mathbf{0}$ implies that $\widehat{\mathbf{z}}_L^\dagger \mathbf{R}_1 \mathbf{S} = \mathbf{0}$ unless $C(z_L) = 0$. Since $\mathbf{S}$ has full rank M, the preceding therefore implies that $\widehat{\mathbf{z}}_L^\dagger \mathbf{R}_1 = \mathbf{0}$. Thus

$$\begin{bmatrix} 1 & z_L^{-1} & z_L^{-2} \ldots & z_L^{-(M-1)} \end{bmatrix} \mathbf{R}_1 = \mathbf{0}$$

That is, z_L is a *common zero of all the M polynomials* defined by the M columns of the precoder $\mathbf{R}_1$. So the most typical situation is like this: we identify a certain number of common zeros of the columns of $\mathbf{Y}$. Of these, there are some which are zeros of all columns of $\mathbf{R}_1$. Rejecting these, there would still be L zeros and these are the zeros of $C(z)$. From these we can identify $c(n)$ upto scale.

This procedure fails if one or both of the following things happen: (a) the L zeros of the channel are not distinct. or (b) there is a zero of the channel which is also a zero of *all* the columns of $\mathbf{R}_1$. Since these are both unlikely, failure of the method has low probability.

The preceding method for blind identification will be referred to as the *Vandermonde annihilation method*. The main advantage of this method is that it works for any $P \geq M + L$ unlike the method in [5] which requires $P = M+L$. The situation $P > M+L$ arises for example in null resistance precoders [9]. Before concluding this section, we would like to draw attention to another important point. Even though the Vandermonde annihilation method works for the general case of $P > L+M$, it is not trivial to extend the method to the case of noisy channels. Furthermore it involves computation of common zeros of polynomials, and is subject to numerical error for large sized matrices. The method is therefore not commonly used.

4. PRECONDITIONERS

The performance of filter bank precoders with zero padding or cyclic prefixing can often be improved by using a preconditioning matrix $\mathbf{G}(z)$ in the transmitter and its inverse $\mathbf{G}^{-1}(z)$ in the receiver as shown in Fig. 4.

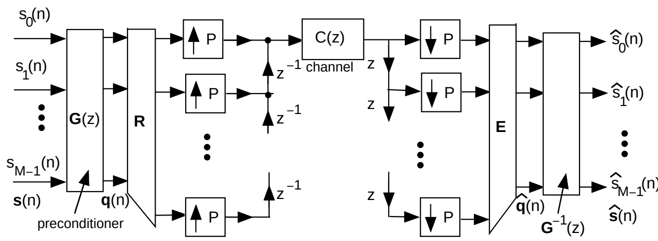

Fig. 4. The precoder system with a preconditioner matrix $\mathbf{G}(z)$.

In absence of noise the preconditioner and its inverse do not change the behavior of the system from $\{s_k(n)\}$ to $\{\widehat{s}_k(n)\}$ at all. It has been shown in [3] that a class of preconditioners called antipodal paraunitary matrices (APU) are especially useful when there is channel noise. Such matrices are paraunitary (i.e., $\mathbf{G}^\dagger(e^{j\omega})\mathbf{G}(e^{j\omega}) = \mathbf{I}$) and furthermore the entries of the coefficient matrices $\mathbf{g}(n)$ are restricted to be $\pm c$, where c is fixed for all coefficients. For example in the FIR case

$$\mathbf{G}(z) = \sum_{n=0}^{N} \mathbf{g}(n) z^{-n}$$

we have $c = 1/\sqrt{M(N+1)}$. The consequence of using APU matrices is explained next. Assume perfect symbol recovery in absence of noise. The error due to the channel in absence of the preconditioner is $\mathbf{e}(n) = \widehat{\mathbf{q}}(n) - \mathbf{q}(n)$ (see Fig. 4). Let $e_k(n)$ denote the kth component of this vector for $0 \leq k \leq M-1$. In general the mean square error

$$\beta_k(n) = E|e_k(n)|^2$$

depends both on the block-time n and the index k (which we refer to as the frequency-bin index because in cyclic prefix systems such as the DMT system this is the interpretation). If we use an FIR APU matrix of order N then the recovery error is $\mathbf{r}(n) = \widehat{\mathbf{s}}(n) - \mathbf{s}(n)$. Let $r_k(n)$ denote the kth component of this vector for $0 \leq k \leq M-1$. Then its mean square error is given by [3]

$$E|r_k(n)|^2 = \frac{1}{M(N+1)} \sum_{\ell=0}^{N} \sum_{k=0}^{M-1} \beta_k(n+\ell) \qquad (8)$$

The fact that $\mathbf{G}(z)$ is APU has been used to derive this expression [3]; note that the exact details of the APU coefficients $\mathbf{g}(n)$ do not enter this expression. The beautiful implication of Eq. (8) is that the mean square error has been *averaged* out. There is averaging across frequency bins (i.e., with respect to k) and across block-time index (i.e., with respect to ℓ). In particular the mean square error is identical for all k. This averaging does not necessarily result in improved performace (e.g., bit error rate) but it has been observed that for time-varying channels (especially fast-varying ones) this is the case [3].

The use of antipodal paraunitary (APU) matrices has been extended to a generalization called distributed APU matrices [7]. In these systems the coefficients of the paraunitary matrix need not occupy successive powers of z^{-1} and have the form $\sum_{n=0}^{N} \mathbf{g}(n)z^{-k_n}$ where $k_0 < k_1 < k_2 \ldots$ It can be shown that preconditioners based on DAPU systems have similar averaging properties except that we now have more flexibility to choose the time domain samples that are averaged. This has been shown to be more useful for time varying channels [7].

The main point of interest for us here is that there exist certain useful filter bank precoder schemes where the memoryless precoder (such as a zero padding precoder) is preceded by a LTI system $\mathbf{G}(z)$ with memory. If we have devised a scheme for blind identification of the channel $C(z)$ in the absence of the preconditioner $\mathbf{G}(z)$ (such as in Sec. 2), will the scheme continue to work when there is a preconditioner?

This is indeed the case as the following argument shows: In Fig. 4 the symbol stream $\widehat{\mathbf{q}}(n)$ is available at the receiver. This can be used instead of $\widehat{\mathbf{s}}(n)$ for blind identification of the channel because we can regard $\mathbf{q}(n)$ as the transmitter symbol stream instead of $\mathbf{s}(n)$. This makes perfect sense because $\mathbf{G}(z)$ is an invertible LTI system; for every sequence $\mathbf{q}(n)$ there is a unique sequence $\mathbf{s}(n)$ and vice versa.

Deeper thought shows that there is a slight technical difficulty here. Recall that the schemes described in Secs. 2, 3 have been possible under the assumption that the transmitted symbol stream $\mathbf{s}(n)$ is rich (see remarks after Eq. (3)). Similarly if we have to identify the channel from $\widehat{\mathbf{q}}(n)$ then $\mathbf{q}(n)$ has to be rich. The question now is this: if we pass a rich signal $\mathbf{s}(n)$ through an LTI system $\mathbf{G}(z)$, will the output $\mathbf{q}(n)$ remain rich? Intuition tells us that this should be so because $\mathbf{G}(z)$ is invertible. Surprisingly however this is not the case as explained next.

5. STAYING RICH

In this section we consider $M \times M$ transfer matrices of the form

$$\mathbf{G}(z) = \sum_{n=0}^{N} \mathbf{g}(n)z^{-n}$$

Denote the input and output signals by $\mathbf{x}(n)$ and $\mathbf{y}(n)$ respectively. Assume $\mathbf{x}(n)$ rich, that is, the matrix

$$[\,\mathbf{x}(0) \quad \mathbf{x}(1) \quad \ldots \quad \mathbf{x}(K_x)\,]$$

has rank M for sufficiently large K_x. The output $\mathbf{y}(n)$ is rich if there exists an integer K_y such that

$$[\,\mathbf{y}(0) \quad \mathbf{y}(1) \quad \ldots \quad \mathbf{y}(K_y)\,]$$

has rank M. We say that $\mathbf{G}(z)$ is **richness-preserving**, abbreviated **RP**, if richness of $\mathbf{x}(n)$ always implies richness of $\mathbf{y}(n)$. If $\mathbf{G}(z)$ is **memoryless**, that is, $\mathbf{G}(z) = \mathbf{g}(0)$ then

$$[\,\mathbf{y}(0) \;\; \mathbf{y}(1) \;\; \ldots \;\; \mathbf{y}(K_y)\,] = \mathbf{g}(0)\,[\,\mathbf{x}(0) \;\; \mathbf{x}(1) \;\; \ldots \;\; \mathbf{x}(K_y)\,]$$

So we see that for a rich input, $\mathbf{y}(n)$ is rich if and only if $\mathbf{g}(0)$ is nonsingular. If $\mathbf{G}(z)$ has memory then richness preservation is nontrivial. In fact examples which do not preserve richness are readily constructed. Thus consider the very simple system

$$\mathbf{G}(z) = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \tag{9}$$

Let the input sequence be defined as

$$\mathbf{x}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}(1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

with $\mathbf{x}(n) = \mathbf{0}$ otherwise. Then it is readily verified that $\mathbf{y}(n) = \mathbf{0}$ unless $n = 1$. This shows that the output is not rich though the input is. Notice that the system in Eq. (9) is a simple first order **paraunitary** matrix [10]. Thus it is invertible; in fact $\mathbf{G}(e^{j\omega})$ is unitary for all ω, and yet it does not preserve richness! Next consider the example

$$\mathbf{G}(z) = \mathbf{I} + z^{-1} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \tag{10}$$

Even though $\mathbf{g}(0)$ is nonsingular this is not sufficient for preserving richness. For example define the input as

$$\mathbf{x}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}(1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with $\mathbf{x}(n) = \mathbf{0}$ otherwise. Then $\mathbf{y}(n) = \mathbf{0}$ for all $n \neq 0$, so the output is not rich. In this example $\mathbf{G}(z)$ is a unimodular matrix, that is, det $\mathbf{G}(z) = 1$. This implies in particular that the inverse exists and is a polynomial matrix as well. We see that even *unimodular matrices may not preserve richness.*

5.1. Degree one paraunitary systems

It is well known[2] that a degree-one FIR paraunitary matrix can be written in the form $\mathbf{G}(z) = \mathbf{V}(z)\mathbf{U}$ where $\mathbf{U}$ is a constant unitary matrix and the matrix $\mathbf{V}(z)$ has the form

$$\mathbf{V}(z) = \mathbf{I} - \mathbf{v}\mathbf{v}^{\dagger} + z^{-1}\mathbf{v}\mathbf{v}^{\dagger} \tag{11}$$

where $\mathbf{v}$ is a unit-norm vector. Can a degree-one building block preserve richness under some conditions? Since $\mathbf{U}$ does not affect richness we only have to check if $\mathbf{V}(z)$ is richness-preserving. Thus consider the input singal with

$$\mathbf{x}(0) = \mathbf{v}, \;\; \mathbf{x}(1) = \mathbf{v}_1, \;\; \ldots \;\; , \mathbf{x}(M-1) = \mathbf{v}_{M-1}$$

where $\mathbf{v}_k$ are unit-norm vectors orthongonal to each other and to $\mathbf{v}$. Assume $\mathbf{x}(n) = \mathbf{0}$ otherwise. This is clearly a rich input. Observe that

$$[\mathbf{I} - \mathbf{v}\mathbf{v}^{\dagger} + z^{-1}\mathbf{v}\mathbf{v}^{\dagger}]\mathbf{v} = z^{-1}\mathbf{v}$$

whereas

$$[\mathbf{I} - \mathbf{v}\mathbf{v}^{\dagger} + z^{-1}\mathbf{v}\mathbf{v}^{\dagger}]\mathbf{v}_k = \mathbf{v}_k$$

[2] Degree should not be confused with order. By definition a degree-d system can be implemented with d delays. The system $\mathbf{G}(z) = z^{-1}\mathbf{I}_M$ has order one but degree M because it requires M delays for its implementation.

Using these it follows that the output of $\mathbf{V}(z)$ in response to such an input $\mathbf{x}(n)$ is such that

$$\mathbf{y}(0) = \mathbf{0}, \quad \text{and} \quad \mathbf{y}(n) = \mathbf{0}, n \geq M$$

so that the matrix $[\,\mathbf{y}(0) \;\; \mathbf{y}(1) \;\; \ldots]$ has rank at most $M-1$ though the input is rich. *So degree one paraunitary systems are necessarily non richness-preserving!*

5.2. Cascaded systems

It is easy to construct examples of systems which turn nonrich inputs into rich outputs. For example consider Eq. (9) which was shown to be non RP. This same system can sometimes turn a nonrich input into rich output. For example suppose $\mathbf{x}(n)$ is defined as

$$\mathbf{x}(n) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \delta(n)$$

which is not rich. Then the output has the samples

$$\mathbf{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{y}(1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

showing that $\mathbf{y}(n)$ is rich. For this reason we cannot argue that a cascade of systems destroys richness if one member of the cascade destroys richness. A simple example is this: suppose we cascade the unimodular system Eq. (10) with its inverse (which is also causal unimodular). Then the result is the identity matrix which obviously preserves richness (though one of the elements in the cascade is not RP).

One consequence of this observation is this: even though an arbitrary FIR paraunitary matrix is a cascade [10] of degree one building blocks of the form (11), the fact that (11) is not RP is not sufficient to prove that all FIR paraunitary matrices are non-RP. However it can (unfortunately) be proved by other means that nonconstant FIR paraunitary matrices can never be richness-preserving [8].

6. CONCLUDING REMARKS

More complete necessary and suffiicient conditions for richness preservation are currently being investigated. It can be shown in particular [8] that if $\mathbf{g}(0)$ is nonsingular then the necessary and sufficient is that $\mathbf{G}(z)$ be of the form

$$\mathbf{G}(z) = (a_0 + a_1 z^{-1} + \ldots + a_N z^{-N})\mathbf{g}(0)$$

with $a_0 \neq 0$. Next, if the transfer matrix $\mathbf{G}(z)$ is a constant, that is $\mathbf{G}(z) = \mathbf{A}$, we know that it preserves richness if and only if $\mathbf{A}$ is nonsingular. Now imagine that the entries of $\mathbf{A}$ are chosen at random (say with a Gaussian distribution). Then the probability of any column being an exact linear combination of the other columns is zero, so the matrix is nonsingular with probability one. In this sense, memoryless systems are almost always richness-preserving.

Consider again $\mathbf{G}(z) = \sum_{n=0}^{N} \mathbf{g}(n) z^{-n}$ with $N > 0$, so the system has memory. Assume that the entries of the matrices $\mathbf{g}(n)$ are chosen at random. Then $\mathbf{g}(0)$ is nonsingular with probability one. But the probability that the other matrices simultaneously have the form $\mathbf{g}(1) = a_1\mathbf{g}(0), \mathbf{g}(2) = a_2\mathbf{g}(0), \ldots$ is clearly zero. This shows that a system with memory is almost always non richness-preserving. This does not however mean that it will destroy the richness of every rich input; it only assures the existence of a rich input for which the output is non rich. Indeed it might turn out that if the inputs are also chosen at random using, say, a Gaussian distribution or a discrete signal constellation, then the output would stay rich for almost all rich inputs. In that sense the preconditioning matrices in Fig. 4 may not affect blind identifiability at all! Finally, there are other possible definitions of richness, and will be explored elsewhere.

REFERENCES

[1] Horn, R. A., and Johnson, C. R *Matrix analysis*, Cambridge University Press, 1985.

[2] Lin, Y.-P., and Phoong, S.-M., "Perfect discrete multitone modulation with optimal transceivers," IEEE Trans. SP, vol. 48, pp. 1702–1712, June. 2000.

[3] Phoong, S-M., Chang, K.-Y., and Lin, Y.-P, "Antipodal paraunitary precoding for OFDM application", Proc. IEEE Intl. Symp. on Circuits and Systems, Vancouver, Canada, May 2004.

[4] Proakis, J. G. *Digital communications,* McGraw Hill 1995.

[5] Scaglione, A., Giannakis, G. B., and Barbarossa, S. "Redundant filter bank precoders and equalizers Part II: Synchronization and direct equalization", IEEE Trans. Signal Processing, vol. 47, pp. 2007-2022, July 1999.

[6] Shalvi, O., and Weinstein, E. "New criteria for blind deconvolution of nonminimum phase systems (channels)", IEEE Trans. Info. Th., vol. 36, pp. 312-321, Mar. 1990.

[7] Su. B, and Vaidyanathan, P. P., "Distributed antipodal paraunitary precoders for OFDM systems", submitted, IEEE ICASSP, March 2005.

[8] Su. B, and Vaidyanathan, P. P., "Necessary and sufficient conditions for LTI systems to preserve signal richness", submitted, IEEE ISCAS, May 2005.

[9] Liang, J., and Tran, T. D., "A partial DFT based OFDM system resistant to channel nulls", Proc. IEEE Asil. Conf. Signals, Systems, and Computers, Nov. 2002.

[10] Vaidyanathan, P. P. *Multirate systems and filter banks,* Prentice Hall, Inc., 1993.

[11] Vaidyanathan, P. P., "Filter banks in digital communications," IEEE CAS Mag., vol. 1, pp. 4–25, 2001.

[12] Vaidyanathan, P. P., and Vrcelj, Bojan, "Transmultiplexers as precoders in modern digital communication: a review," IEEE Int. Symp. Circuits and Systems, Vancouver, Canada, May 2004.

[13] Vaidyanathan, P. P., and Vrcelj, B. "A frequency domain approach for blind identification with filter bank precoders," IEEE Int. Symp. Circuits and Systems, Vancouver, Canada, May 2004.

[14] Xia, X-G. "Modulated coding for intersymbol interference channels," Marcel Dekker, Inc., 2001.

[15] X.-G. Xia, W. Su, and H. Liu, "Filterbank precoders for blind equalization: polynomial ambiguity resistant precoders (PARP)," IEEE Trans. CAS-I, Feb. 2001.