

Effects of Multirate Systems on the Statistical Properties of Random Signals

Vinay P. Sathe, *Member, IEEE*, and P. P. Vaidyanathan, *Fellow, IEEE*

Abstract—In multirate digital signal processing, we often encounter time-varying linear systems such as decimators, interpolators, and modulators. In many applications, these building blocks are interconnected with linear filters to form more complicated systems. It is often necessary to understand the way in which the statistical behavior of a signal changes as it passes through such systems. While some issues in this context have an obvious answer, the analysis becomes more involved with complicated interconnections. For example, consider this question: if we pass a cyclostationary signal with period K through a fractional sampling rate-changing device (implemented with an interpolator, a *nonideal low-pass filter* and a decimator), what can we say about the statistical properties of the output? How does the behavior change if the filter is replaced by an ideal low-pass filter? In this paper, we answer questions of this nature. As an application, we consider a new adaptive filtering structure, which is well suited for the identification of band-limited channels. This structure exploits the band-limited nature of the channel, and embeds the adaptive filter into a multirate system. The advantages are that the adaptive filter has a smaller length, and the adaptation as well as the filtering are performed at a lower rate. Using the theory developed in this paper, we show that a matrix adaptive filter (dimension determined by the decimator and interpolator) gives better performance in terms of lower error energy at convergence than a traditional adaptive filter. Even though matrix adaptive filters are, in general, computationally more expensive, they offer a performance bound that can be used as a yardstick to judge more practical “scalar multirate adaptation” schemes.

I. INTRODUCTION

MULTIRATE digital filtering is used in a variety of applications such as subband coding, voice privacy systems [1], [2], transmultiplexers, and adaptive filtering [3], [4], to name a few. In multirate digital signal processing, we encounter time-varying linear systems such as decimators, interpolators, and modulators [2]. In many applications, these building blocks are interconnected with linear filters to form more complicated systems. Consider, for example, some of the simple interconnections shown in Fig. 1. The M -fold decimator is shown in Fig. 1(a). Fig. 1(b) shows the L -fold interpolator. Typically,

Manuscript received October 21, 1990; revised December 10, 1991. This work was supported by the National Science Foundation under Grants MIP 8919196 and MIP 8604456 and by matching funds by Tektronix, Inc.

V. P. Sathe was with the Department of Electrical Engineering, California Institute of Technology. He is now with Video and Electronic Systems Laboratories, Tektronix, Inc., Beaverton, OR 97077.

P. P. Vaidyanathan is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

IEEE Log Number 9203332.

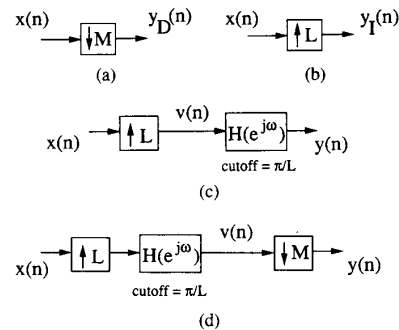


Fig. 1. Some typical interconnections analyzed in the paper.

a low-pass filter is used after an interpolator, to suppress the images created by interpolation. This is shown in Fig. 1(c). The operation of fractional decimation (or sampling-rate conversion) is shown in Fig. 1(d).

It is often necessary to understand the way in which the statistical behavior of a signal changes as it passes through such systems. While some issues in this context have an obvious answer, the analysis becomes more involved with complicated interconnections. For example, it is easy to see that the decimated version $x(nM)$ of a wide-sense-stationary (WSS) signal $x(n)$ remains WSS. But the following question is more complicated. Consider Fig. 1(d), which represents a fractional sampling rate changing device. If the input $x(n)$ is cyclo-wide-sense stationary [6] with period K , then what can we say about the stationary (or otherwise) of the output? Is the answer to this question dependent on whether the low-pass filter is ideal or not, and if so, how?

In this paper, we answer questions of this nature, starting from a small set of elementary observations. When we make transition from single-rate to multirate systems, the assumption that the output signals (in response to WSS input) are WSS is not valid even in theoretical study. We shall find it more natural to assume that the signals are cyclo-wide-sense stationary (CWSS). For example, the output of an L -fold interpolation filter is CWSS with period L , and becomes WSS only if the filter has certain ideal band-limiting properties (as we shall specify in Theorem 4.1). Occurrence of CWSS signals in signal processing and communications applications has been recently discussed in [15]. In Section IV, we shall study the effects of multirate filters with reference to CWSS signals.

As an application, we consider an adaptive filtering structure of identification of band-limited channels. This scheme is shown in Fig. 16. As we shall see in Section V, this structure exploits the band-limited nature of the channel and embeds the adaptive filter into a multirate system. (In the past, the band-limited nature, as well as the spectral energy distribution have been exploited in adaptive filtering, e.g., [3], [4], [10], [11]. However, the fractional multirate adaptive structure for this purpose, shown in Fig. 16, is new). The advantages are that the adaptive filter has a smaller length and the adaptation, as well as the filtering take place at a speed lower than the input data rate, resulting in improved computational efficiency. In the theoretical analysis of this system, due to its multirate nature we cannot assume that the input to the adaptive filter is WSS (even if the primary input $x(n)$ to the channel is WSS). Using the theory developed in this paper, we show that the input to the adaptive filter is CWSS, and that a matrix adaptive filter gives better performance than a traditional scalar filter. The fact that a matrix adaptive filter is computationally more expensive clearly places in evidence the tradeoff involved when we switch from single-rate to multirate systems. The matrix adaptive filter offers a theoretical performance bound that cannot be exceeded by a scalar filter of comparable complexity. Finally, it is shown that if the nonadaptive filters in this system are close to ideal (in the sense of having a sharp cutoff and good stopband attenuation), then the matrix adaptive filter can be replaced by a scalar adaptive filter without a significant loss of performance.

A. Outline of the Paper

The paper is organized as follows. In Section II, we include definitions of various statistical and multirate concepts used in the paper. The effects of the basic multirate building blocks (such as decimators, interpolators, and modulators) are investigated in Section III. In Section IV, we derive similar results for some useful interconnections of the basic building blocks. The multirate adaptive filtering scheme for identification of band-limited channels is discussed in detail in Section V. Simulation results are included. Throughout the paper, all WSS and CWSS processes are assumed to be zero mean.

B. Notations used in the Paper

The superscript T stands for matrix (or vector) transposition whereas the superscript dagger (\dagger) stands for transposition followed by complex conjugation. Boldface italic letters indicate matrices and vectors. The superscript asterisk ($*$) stands for complex conjugation. The tilde accent on a function $F(z)$ is defined such that, on the unit circle, $\tilde{F}(z) = F^{\dagger}(z)$. Thus, for arbitrary z , $\tilde{F}(z) = F^T_*(z^{-1})$, where the subscript asterisk denotes complex conjugation of coefficients of the function. Following the standard signal processing convention, the operators z and z^{-1} are used in flow graphs to represent advance and delay operations, respectively. Whenever it is convenient, we

use the notation $((x))_Q$ to denote $(x \text{ modulo } Q)$. Here, Q is frequently an integer, but on occasion we will need to use $Q = 2\pi$.

II. PRELIMINARIES

We first review some basic concepts and definitions from multirate signal processing and system theory.

A. M -Fold Decimator

A decimator is a device that takes an input sequence $x(n)$ and produces the output sequence

$$y_D(n) = x(nM). \quad (2.1)$$

This means that only those samples of $x(n)$ that occur at sample locations equal to integer multiples of M are retained. In the transform domain, the Fourier transforms are related as

$$Y_D(e^{j\omega}) = \frac{1}{M} \sum_{k=0}^{M-1} X(e^{j\omega/M} W_M^k) \quad (2.2)$$

where $W_M = e^{-2j\pi/M}$. Thus, in general, decimation causes aliasing.

B. L -Fold Interpolator

The interpolator takes an input sequence $x(n)$ and produces an output sequence

$$y_I(n) = \begin{cases} x(n/L), & \text{if } n \text{ is an integer multiple of } L \\ 0, & \text{otherwise.} \end{cases} \quad (2.3)$$

In the frequency domain, we can write

$$Y_I(e^{j\omega}) = X(e^{j\omega L}). \quad (2.4)$$

This is the well-known imaging effect [1]; we now have L "squeezed" copies of the spectrum $X(e^{j\omega})$ in the region $0 \leq \omega < 2\pi$.

C. Blocking a Signal

Given a scalar signal $x(n)$, we define its M -fold blocked version $\mathbf{x}(n)$ by

$$\mathbf{x}(n) = [x(nM)x(nM-1) \cdots x(nM-M+1)]^T. \quad (2.5)$$

Using decimators and delays, the blocking mechanism can be represented as in Fig. 2(a). The signal $x(n)$ is called the *unblocked version* of the vector process $\mathbf{x}(n)$. The unblocking operation can be represented in terms of multirate building blocks as in Fig. 2(b). We will see below that the components of the blocked version are precisely the "polyphase components" of $x(n)$.

D. Wide-Sense-Stationary (WSS) Process

A vector stochastic process $\mathbf{x}(n)$ is said to be a wide-sense-stationary process if 1) $E[\mathbf{x}(n)] = E[\mathbf{x}(n+k)]$ for all integers n and k ; and 2) the autocorrelation function

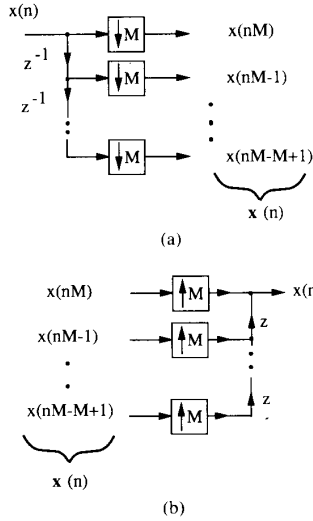


Fig. 2. (a) M -fold blocking of a signal. (b) Unblocking of an $M \times 1$ vector signal.

depends only on the time difference between the two samples, i.e.,

$$E[x(n)x^*(n-k)] = R_{xx}(k), \quad \forall n, \forall k. \quad (2.6)$$

The mean value $E[x(n)]$ will usually not enter our discussion because it is normally assumed to be zero.

E. Jointly WSS Processes

Two processes $v(n)$ and $w(n)$ are said to be *jointly WSS* if the process $u(n) = [v^T(n) w^T(n)]^T$ is WSS.

F. Power Spectral Density

The power spectral density $S_{xx}(z)$ of a WSS process $x(n)$ is defined as the z transform of its autocorrelation matrix defined in (2.6), i.e.,

$$S_{xx}(z) = \sum_{k=-\infty}^{\infty} R_{xx}(k) z^{-k}. \quad (2.7)$$

Thus, each entry of this matrix is the z transform of the corresponding entry of $R_{xx}(k)$.

G. Cyclo-WSS Process

1) *Definition 1*: A stochastic process $x(n)$ is said to be *cyclo-WSS* with period L (abbreviated $(CWSS)_L$), if the L -fold blocked version $x(n)$ is WSS.

2) *Definition 2*: Let $R_{xx}(n, k) = E[x(n)x^*(n-k)]$ denote the autocorrelation function of a process $x(n)$. The process is said to be $(CWSS)_L$ if

$$E[x(n)] = E[x(n+kL)], \quad \forall n, \forall k \quad (2.8a)$$

$$R_{xx}(n, k) = R_{xx}(n+L, k), \quad \forall n, \forall k. \quad (2.8b)$$

(Again, mean values such as (2.8a) will not enter our discussions, as they are normally assumed to be zero.) A proof of the equivalence of these definitions is given in

Appendix I. Notice in particular that a $(CWSS)_L$ process is also $(CWSS)_{mL}$ for any positive integer m .

H. Linear Periodically Time-Varying (LPTV) System

A system is said to be LPTV with period L (denoted as $(LPTV)_L$) if the output $y(n)$ in response to input $x(n)$ can be written as

$$y(n) = \sum_{-\infty}^{\infty} h(n, k) x(n-k) \quad (2.9)$$

where

$$h(n, k) = h(n+L, k), \quad \forall n, \forall k. \quad (2.10)$$

An $(LPTV)_L$ system is also an $(LPTV)_{mL}$ system for any positive integer m . An implementation of an $(LPTV)_L$ system is shown in Fig. 3. The output at time n is the output of one of the L filters, as governed by the value of $(n \text{ modulo } L)$.

I. Pseudocirculant Matrices

An $M \times M$ matrix $A(e^{j\omega})$ is said to be *pseudocirculant* if the entries $a_{i,l}(e^{j\omega})$ ($i = 0, \dots, M-1, l = 0, \dots, M-1$) satisfy the following relation:

$$a_{i,l}(e^{j\omega}) = \begin{cases} a_{0,l-i}(e^{j\omega}), & 0 \leq i \leq l \\ e^{-j\omega} a_{0,l-i+M}(e^{j\omega}), & l < i \leq M-1. \end{cases} \quad (2.11)$$

In words, a pseudocirculant matrix is a circulant matrix with elements under the diagonal multiplied by $e^{-j\omega}$. Here is an example of a 3×3 pseudocirculant matrix

$$A(e^{j\omega}) = \begin{pmatrix} a_0(e^{j\omega}) & a_1(e^{j\omega}) & a_2(e^{j\omega}) \\ e^{-j\omega} a_2(e^{j\omega}) & a_0(e^{j\omega}) & a_1(e^{j\omega}) \\ e^{-j\omega} a_1(e^{j\omega}) & e^{-j\omega} a_2(e^{j\omega}) & a_0(e^{j\omega}) \end{pmatrix}. \quad (2.12)$$

We will be using the following properties of pseudocirculants, which can be verified from [7]:

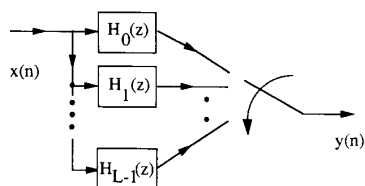
- 1) If $A(e^{j\omega})$ is pseudocirculant, then so is $A^\dagger(e^{j\omega})$. If the inverse $[A(e^{j\omega})]^{-1}$ exists, it is also pseudocirculant.
- 2) The product of pseudocirculant matrices is also pseudocirculant.

The above definitions and properties hold true in the z -domain if we use the substitution $z = e^{j\omega}$, and replace $A^\dagger(e^{j\omega})$ with $\tilde{A}(z)$ (assuming, of course, that the z transforms exist).

J. Polyphase Decomposition

Let $X(z)$ be the z transform of a signal $x(n)$. The polyphase decomposition with respect to M is

$$X(z) = z^{-(M-1)} R_0(z^M) + z^{-(M-2)} R_1(z^M) + \dots + R_{M-1}(z^M). \quad (2.13)$$

Fig. 3. Implementation of an $(LPTV)_l$ system.

Each function $R_i(z)$, $0 \leq i \leq M-1$ is called a polyphase component of $X(z)$. In the time domain, the k th polyphase component is obtained as

$$r_k(n) = x(nM + M - 1 - k). \quad (2.14)$$

From Fig. 2(a) it is easy to see that the components of the blocked version are also the polyphase components of the input signal.

We now tie together seemingly unrelated concepts such as pseudocirculant matrices and wide-sense stationarity by proving some interesting relations. These results also bring out the importance of pseudocirculants in the analysis involving WSS signals.

Fact 2.1: Let $x(n)$ be an $N \times 1$ vector WSS process, input to an $M \times N$ transfer matrix $H(z)$ as shown in Fig. 4. Then the power-spectral density of the $M \times 1$ vector WSS process $y(n)$ is given by

$$S_{yy}(z) = H(z) S_{xx}(z) H^H(z). \quad (2.15)$$

(Reminder: The WSS nature of the vector $x(n)$ is equivalent to the property that its unblocked version $x(n)$ will be $(CWSS)_N$.)

Proof: This follows using the convolution expression and the definition of $S_{yy}(z)$. ■

Fact 2.2. Relation between WSS and Pseudocirculant Properties: Let $x(n)$ be the M -fold blocked version of a zero-mean stochastic process $x(n)$. Then the following statements are true:

a) If $x(n)$ is WSS, then the power-spectral density $S_{xx}(z)$ of $x(n)$ is pseudocirculant.

b) If, for some M , the blocked version $x(n)$ is WSS and $S_{xx}(z)$ pseudocirculant, then $x(n)$ is a WSS process.

Proof:

a) Let $x(n)$ be WSS. Then, the (i, l) th element of $S_{xx}(z)$ can be written as

$$[S_{xx}(z)]_{i,l} = \sum_{k=-\infty}^{\infty} E[x(nM - i) \cdot x^*(nM - kM - l)] z^{-k}. \quad (2.16)$$

Using the WSS property, we can now rewrite this. For $0 \leq i \leq l$, we can write (2.16) as

$$[S_{xx}(z)]_{i,l} = \sum_{k=-\infty}^{\infty} E[x(nM) x^*(nM - kM - (l - i))] z^{-k} = [S_{xx}(z)]_{0,l-i}. \quad (2.17)$$

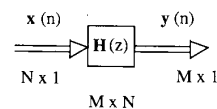


Fig. 4. A multiinput-multioutput system.

For $l < i \leq M-1$, $i-l$ is positive, but $i-l-M$ is negative, hence we can write (2.16) as

$$\begin{aligned} [S_{xx}(z)]_{i,l} &= \sum_{k=-\infty}^{\infty} E[x(nM) x^*(nM - kM + M \\ &\quad - (l - i + M))] z^{-k} \\ &= \sum_{k=-\infty}^{\infty} E[x(nM) x^*(nM - (k - 1)M \\ &\quad - (l - i + M))] z^{-k} \\ &= [z^{-1} S_{xx}(z)]_{0,l-i+M}. \end{aligned} \quad (2.18)$$

From the definition of a pseudocirculant, we conclude that $S_{xx}(z)$ is pseudocirculant.¹

b) If $x(n)$ is a WSS process, then we can indeed write a valid autocorrelation matrix as in (2.6). The entries of this matrix are

$$\begin{aligned} [R_{xx}(k)]_{i,l} &= E[x(nM - i) x^*(nM - kM - l)] \\ &= E[x(-i) x^*(-kM - l)]. \end{aligned} \quad (2.19)$$

The pseudocirculant property of $S_{xx}(z)$ implies

$$[R_{xx}(k)]_{i,l} = \begin{cases} [R_{xx}(k)]_{0,l-i}, & l \geq i \geq 0 \\ [R_{xx}(k-1)]_{0,l-i+M}, & l < i \leq M-1. \end{cases} \quad (2.20)$$

Hence, we can write

$$\begin{aligned} E[x(-i) x^*(-kM - l)] &= E[x(0) x^*(-kM + i - l)], \\ &0 \leq i, l \leq M-1. \end{aligned} \quad (2.21)$$

Now consider $E[x(n) x^*(m)]$. We can write the time indices as $n = n_0M - i$ and $m = m_0M - l$, where $0 \leq i, l \leq M-1$. So

$$\begin{aligned} E[x(n) x^*(m)] &= E[x(n_0M - i) x^*(m_0M - l)] \\ &= E[x(-i) x^*((m_0 - n_0)M - l)] \\ &= E[x(0) x^*((m_0 - n_0)M + i - l)] \\ &= E[x(0) x^*(m - n)]. \end{aligned} \quad (2.22)$$

Hence, $x(n)$ is indeed a WSS process. ■

Remarks: We prove in Appendix II that the $(0, m)$ th entry of $S_{xx}(z)$ is the $(M-1-m)$ th polyphase component of $S_{xx}(z)$. This implies that we can write down $S_{xx}(z)$ from the 0th row of $S_{xx}(z)$.

¹Notice the difference between $S_{xx}(z)$ and $S_{xx}(z)$ in our discussions: The former corresponds to a vector process and the latter to a scalar process.

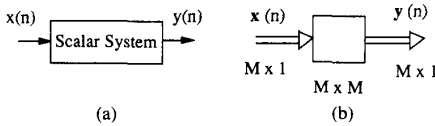


Fig. 5. (a) A scalar system. (b) Corresponding blocked version.

Fact 2.3. Relation between Blocked Versions and Pseudocirculants: Consider the scalar system of Fig. 5(a), with input $x(n)$ and output $y(n)$. With $x(n)$ and $y(n)$ denoting the blocked versions of $x(n)$ and $y(n)$, the M -folded blocked system is shown in Fig 5(b). Clearly, for an arbitrary scalar system, the blocked version is not a linear and time invariant (LTI) system. However, we can make the following statements: 1) the blocked version is LTI [i.e., it can be described by a transfer matrix, say $H(z)$] if and only if the scalar system is (LPTV) $_M$; and 2) the blocked version is LTI with a pseudocirculant transfer matrix $H(z)$, if and only if the scalar system is LTI. The proofs can be found in [7].

Now assume that the system in Fig. 5(a) is LTI and that $x(n)$ is WSS. Then the blocked version $x(n)$ has a pseudocirculant power spectrum $S_{xx}(z)$. Since the blocked version $H(z)$ is also pseudocirculant, the product (2.15) is pseudocirculant. This is consistent with the fact that the unblocked version $y(n)$ is WSS.

Finally, we mention a result that gives the solution of an optimal filtering problem involving WSS signals.

Fact 2.4. Matrix Version of Wiener Filtering: Consider the system shown in Fig. 6. If the signals $w(n)$ and $v(n)$ are jointly WSS, then the best filter $A(z)$ in terms of minimizing error variance $E[e^+(n)e(n)]$ is given by

$$A(z) = \bar{S}_{wv}(z) S_{ww}^{-1}(z). \quad (2.23)$$

The matrix $S_{wv}(z)$ is the cross-power spectrum, i.e., the z transform of cross correlation $R_{wv}(k) = E[w(n)v^+(n-k)]$. This solution is called the Wiener solution to the problem mentioned above.

Proof: This is a simple extension of the scalar version, which can be found, for example, in [8, p. 265]. ■

Remarks: It is assumed here that the determinant of $S_{ww}(z)$ is not identically zero for all z (i.e., that $S_{ww}(z)$ has full normal rank). The stability and realizability of the solution $A(z)$ are in general not guaranteed, and one often tries to replace $A(e^{j\omega})$ with a stable practical approximation.

We know that if a matrix filter $A(z)$ is pseudocirculant, then, in fact, there exists a scalar transfer function corresponding to the unblocked input-output signal description (Fact 2.3). On the other hand, if the blocked (matrix) transfer function is not pseudocirculant, then the corresponding unblocked transfer function is an LPTV system. So for an optimal matrix-filtering problem, if the solution given by (2.23) is pseudocirculant, then it is in fact a scalar LTI system. Otherwise, the optimal solution is an LPTV system. We will use this fact later when we discuss the adaptive filtering application.

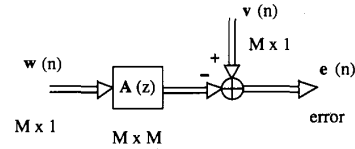


Fig. 6. Optimal filtering setup.

III. BASIC RESULTS

Using the concepts defined in the previous section, we now derive some useful results pertaining to multirate building blocks. First, we prove two results about modulation of a WSS signal by a deterministic signal.

Fact 3.1: Let $x(n)$ be an $M \times 1$ vector WSS process, modulated by a vector function $f(n)$ as shown in Fig. 7. More precisely, the components of $y(n)$, $f(n)$, and $x(n)$ are related as $y_i(n) = f_i(n)x_i(n)$. Then $y(n)$ is WSS if and only if $f(n)$ is of the form

$$f(n) = d e^{j\theta n}, \quad d \text{ (possibly a complex) constant, } \theta \text{ real.} \quad (3.1)$$

This equation says that the time dependence is identical for all the components of $f(n)$.

Proof: The “if” part can be easily verified by direct substitution. We prove the “only if” part. Let us represent all the vector quantities in term of their individual components as

$$\begin{aligned} x(n) &= \begin{pmatrix} x_0(n) \\ \vdots \\ x_{M-1}(n) \end{pmatrix}, & f(n) &= \begin{pmatrix} f_0(n) \\ \vdots \\ f_{M-1}(n) \end{pmatrix}, \\ y(n) &= \begin{pmatrix} y_0(n) \\ \vdots \\ y_{M-1}(n) \end{pmatrix}. \end{aligned} \quad (3.2)$$

We can write the input-output relation as

$$y(n) = \Lambda(n)x(n) \quad (3.3)$$

where $\Lambda(n)$ is an $M \times M$ diagonal matrix with entries $(f_0(n) \cdots f_{M-1}(n))$ on the diagonal. The autocorrelation function for $y(n)$ can thus be written as

$$\begin{aligned} E[y(n)y^+(n-k)] \\ = \Lambda(n)E[x(n)x^+(n-k)]\Lambda^+(n-k). \end{aligned} \quad (3.4)$$

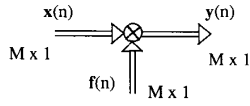
This gives the autocorrelation matrix as

$$R_{yy}(n, k) = \Lambda(n)R_{xx}(k)\Lambda^+(n-k) \quad (3.5)$$

whose (i, l) th entry is therefore

$$[R_{yy}(n, k)]_{i,l} = f_i(n)[R_{xx}(k)]_{i,l}f_l^*(n-k). \quad (3.6)$$

For $y(n)$ to be WSS, we want all the above entries to be free from the time index n . Consider the diagonal elements first. The i th diagonal element will not be a function of n , if and only if $f_i(n)f_i^*(n-k)$ is independent of

Fig. 7. Modulation of signal $x(n)$.

n , for all k . Consider $k = 0$. Then this implies, in particular, that $|f_i(n)|^2$ is independent of n . Hence, $f_i(n)$ must have the form

$$f_i(n) = c_i e^{j\alpha_i(n)}, \quad \alpha_i(n) \text{ real.} \quad (3.7)$$

The (i, i) th element of (3.6) thus becomes

$$\begin{aligned} [\mathbf{R}_{xx}(k)]_{i,i} f_i(n) f_i^*(n-k) \\ = [\mathbf{R}_{xx}(k)]_{i,i} |c_i|^2 e^{j(\alpha_i(n) - \alpha_i(n-k))}. \end{aligned} \quad (3.8)$$

Using the fact that this has to be independent of n , and using a particular value of k ($k = 1$), we get a recursion of the type

$$\alpha_i(n) = \alpha_i(0) + n\beta_i, \quad \beta_i \text{ constant.} \quad (3.9)$$

Hence, we can rewrite (3.7) as

$$f_i(n) = d_i e^{j\theta_i n}, \quad \theta_i \text{ real, } d_i \text{ constant.} \quad (3.10)$$

Now, if we use the expression in (3.10) for the (i, l) th element, we get

$$f_i(n) f_l^*(n-k) = d_i d_l^* e^{j(\theta_i n - \theta_l(n-k))}. \quad (3.11)$$

For this to be independent of n , we should have $((\theta_i - \theta_l))_{2\pi} = 0$. Summarizing, $f_i(n) = d_i e^{j\theta_i n}$, so that (3.1) follows.

Remarks: This result implies that translating the power spectrum of a scalar WSS process by *unequal* amounts generates processes that are WSS themselves, but are not jointly WSS.

Fact 3.2: Let $x(n)$ be a $(\text{CWSS})_M$ signal. Then the signal $y(n) = f(n)x(n)$ is $(\text{CWSS})_M$ if and only if each polyphase component of the modulating function $f(n)$ with respect to M has the form $\alpha_i e^{j\theta_i n}$ (α_i possibly complex, and θ_i real).

Proof: If $x(n)$ is $(\text{CWSS})_M$, then its blocked version of length M will be a vector WSS process. Similarly, if we block the modulating function $f(n)$, this problem reduces to the setup mentioned in Fact 3.1. Since the $(\text{CWSS})_M$ property of $y(n)$ is equivalent to the WSS property of its blocked version $y(n)$, the result follows. ■

Remarks: These results imply, in particular, that if we modulate a WSS signal by a cosine wave ($y(n) = x(n) \cos(\omega_0 n)$), then the output $y(n)$ is not WSS even if $x(n)$ is WSS (unless $\omega_0 = 0, \pi$).

We now turn attention to the remaining multirate building blocks.

Fact 3.3: Let $y(n) = x(nM)$, where $x(n)$ is $(\text{CWSS})_L$. Then $y(n)$ is CWSS with period K , where $K = L/\text{gcd}(L, M)$.

Remarks: Notice in particular, that $y(n)$ is also

$(\text{CWSS})_L$. The following special cases are worth mentioning.

- 1) Let $L = M$. Then $K = 1$, so $y(n)$ is WSS.
- 2) Let $L = 1$ (i.e., $x(n)$ is WSS). Then $K = 1$, and $y(n)$ is WSS.
- 3) Let L and M be relatively prime. Then $\text{gcd}(L, M) = 1$ and $K = L$ regardless of M .

Proof of Fact 3.3: Using the input-output relation of a decimator, we can write the autocorrelation for $y(n)$ as

$$\begin{aligned} E\{y(n) y^*(n - n_0)\} \\ = E\{x(nM) x^*(nM - n_0 M)\}. \end{aligned} \quad (3.12)$$

If K is the period of cyclo-wide-sense stationarity, then

$$\begin{aligned} E\{y(n + K) y^*(n + K - n_0)\} \\ = E\{y(n) y^*(n - n_0)\}. \end{aligned} \quad (3.13)$$

Thus, from (3.12) we get

$$\begin{aligned} E\{x(nM) x^*(nM - n_0 M)\} \\ = E\{x(nM + KM) x^*(nM + KM - n_0 M)\}. \end{aligned} \quad (3.14)$$

Since $x(n)$ is $(\text{CWSS})_L$, (3.14) is satisfied if $KM = lL$ for some integer l . The smallest K is such that all prime factors of L are accounted for by the left-hand side. Since M has $\text{gcd}(M, L)$ as the largest factor common with L , we get $K = L/\text{gcd}(M, L)$. ■

Fact 3.4: Passing a $(\text{CWSS})_L$ signal $x(n)$ through an $(\text{LPTV})_L$ system gives a signal $y(n)$, which is $(\text{CWSS})_L$.

Proof: To prove the result, consider a corresponding blocked system obtained by applying the L -fold blocked input $x(n)$ to the L -fold blocked version of the $(\text{LPTV})_L$ system above. We know that the blocked version of the $(\text{LPTV})_L$ system is an LTI system and $x(n)$ is WSS by definition of cyclo-wide-sense stationarity. Hence, the output of the blocked system will be a WSS signal, and the corresponding unblocked version $y(n)$ will be $(\text{CWSS})_L$. The result thus follows. ■

Fact 3.5: If we pass a WSS signal $x(n)$ through an L -fold interpolator, the output $y(n)$ is $(\text{CWSS})_L$. This setup is shown in Fig. 1(b).

Proof: Let us block the output $y(n)$ into the vector process $\mathbf{y}(n) = [y(nL) \cdots y(nL - L + 1)]^T$. Since $y(n)$ is the interpolated version of $x(n)$, we get $\mathbf{y}(n) = (x(n) \ 0 \cdots 0)^T$. Hence

$$E\{\mathbf{y}(n) \mathbf{y}^\dagger(n - k)\} = \begin{pmatrix} R_{xx}(k) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}. \quad (3.15)$$

The right-hand side of independent of n . This means that $\mathbf{y}(n)$ is a vector WSS process. From the definition of a CWSS process, it follows that $y(n)$ is a $(\text{CWSS})_L$ process. ■

So far we have seen the effects of basic multirate building blocks on stationary random inputs. In multirate-filtering applications, these building blocks are interconnected to form more complex systems. So it is of interest to study similar properties for some standard interconnections of these building blocks.

IV. INTERCONNECTIONS OF THE MULTIRATE BUILDING BLOCKS

We first consider the operation of the L -fold interpolation filter. We prove that, in general, the output of this multirate interconnection is not WSS even if the input is.

Fact 4.1: Consider the L -fold interpolation filter shown in Fig. 1(c). If $x(n)$ is WSS, then $y(n)$ is $(\text{CWSS})_L$.

Proof: From Fact 3.5, we know that $v(n)$ is $(\text{CWSS})_L$. Hence, from Fact 3.4, we can conclude that $y(n)$ is $(\text{CWSS})_L$.

Another important multirate operation is fractional decimation or sampling-rate conversion. This is used in a variety of multirate applications. We prove that this operation in general does not produce a WSS output for a WSS input.

Fact 4.2: For the multirate filter shown in Fig. 1(d), if input $x(n)$ is WSS, then the output $y(n)$ is $(\text{CWSS})_K$ where $K = L/\text{gcd}(L, M)$.

Proof: From Fact 4.1, we know that $v(n)$ is $(\text{CWSS})_L$. Hence, the result follows from Fact 3.3. ■

A. Necessary and Sufficient Condition for Wide-Sense Stationarity of $y(n)$

We now prove an important result. From Fact 4.1, we know that the output of an L -fold interpolation filter in response to a WSS input is $(\text{CWSS})_L$. We now find the necessary and sufficient conditions on the interpolation filter, such that this $(\text{CWSS})_L$ process actually becomes WSS.

Let us split $H(e^{j\omega})$ into its polyphase components as follows:

$$H(e^{j\omega}) = e^{-j\omega(L-1)} R_0(e^{j\omega L}) + \dots + e^{-j\omega} R_{L-2}(e^{j\omega L}) + R_{L-1}(e^{j\omega L}). \quad (4.1)$$

We have explained, in Definition 2.10, how to obtain the above representation. Using this, we can redraw Fig. 1(c) as Fig. 8. Then using the multirate identity of Fig. 9, we can simplify the implementation as shown in Fig. 10. Consider the signal $t(n) \triangleq y(n + L - 1)$. Fig. 11 shows the generation of $t(n)$ from $x(n)$. It is trivially true that $y(n)$ is WSS if and only if $t(n)$ is WSS. From Fig. 11 and Fact 2.2, we know that $t(n)$ is WSS if and only if the blocked version $\mathbf{t}(n) \triangleq [t_0(n) \dots t_{L-1}(n)]^T$ is WSS and $S_{\mathbf{t}}(z)$ is pseudocirculant. Let us define the matrix transfer function $\mathbf{G}(e^{j\omega}) \triangleq [R_0(e^{j\omega}) \dots R_{L-1}(e^{j\omega})]^T$. We now derive the necessary and sufficient conditions for the wide-sense stationarity of $t(n)$ in the following steps.

Step 1: The signal $t(n)$ is WSS if and only if $\mathbf{G}(e^{j\omega})\mathbf{G}^\dagger(e^{j\omega})$ is pseudocirculant.

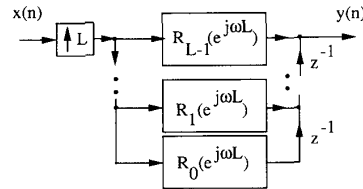


Fig. 8. Redrawing Fig. 1(c) using polyphase components.

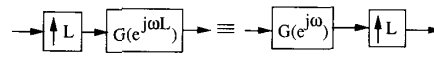


Fig. 9. A well-known multirate identity.

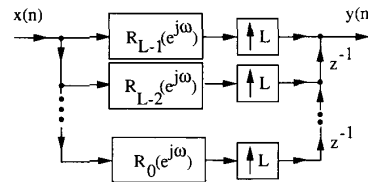


Fig. 10. Simplification of the implementation in Fig. 8 using the multirate identity.

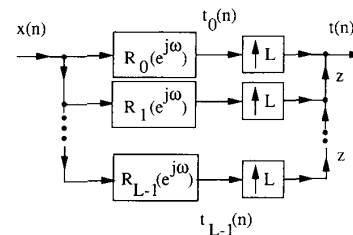


Fig. 11. The generation of $t(n)$ from $x(n)$.

Step 2: $\mathbf{G}(e^{j\omega})\mathbf{G}^\dagger(e^{j\omega})$ is pseudocirculant if and only if the polyphase components can be represented as

$$R_k(e^{j\omega}) = R_{L-1}(e^{j\omega}) e^{j(L-1-k)((\omega/L) - (2\pi P(\omega)/L))}, \quad \forall \omega$$

where $P(\omega)$ is an integer valued function satisfying the property that

$$((P(\omega + 2\pi) - P(\omega) - 1))_L = 0, \quad \forall \omega.$$

In other words, $P(\omega + 2\pi) - P(\omega) - 1$ is an integer multiple of L .

Step 3: The above representation of $R_k(e^{j\omega})$ is possible if and only if the filter $H(e^{j\omega})$ has the following property: no aliasing occurs if we perform L -fold decimation of the impulse response $h(n)$. This is equivalent to the following condition (Appendix III): the frequency regions where $H(e^{j\omega})$ is nonzero do not overlap, if the frequency region $0 \leq \omega < 2\pi$ is reduced modulo $2\pi/L$.

Remark: For the interpolation scheme of Fig. 1(c), the condition for wide-sense stationarity of the output of the statistical case is the same as the condition for image-free interpolation for the deterministic case. (See Appendix III for an explanation of the term *image-free interpolation*.)

Proof of Step 1: Since the vector $\mathbf{t}(n)$ is the output

of a linear system to which $x(n)$ is the input, it is WSS. Using (2.15), we can write

$$\begin{aligned} S_H(e^{j\omega}) &= \mathbf{G}(e^{j\omega}) S_{xx}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega}) \\ &= S_{xx}(e^{j\omega}) \mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega}) \end{aligned} \quad (4.2)$$

since $S_{xx}(e^{j\omega})$ is scalar. Thus, $t(n)$ [and hence $y(n)$] is WSS if and only if $\mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega})$ is pseudocirculant.

Proof of Step 2:

The "Necessary" Part: We first assume that $\mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega})$ is pseudocirculant, and prove that the polyphase components are related as stated above. Using the decimation relation (2.2), we can write

$$\begin{aligned} R_k(e^{j\omega}) &= \frac{1}{L} \sum_{m=0}^{L-1} e^{j(\omega - 2\pi m)/L(L-1-k)} H(e^{j(\omega - 2\pi m)/L}), \\ k &= 0, \dots, L-1. \end{aligned} \quad (4.3)$$

Let us represent the polyphase components in their magnitude-phase form as

$$\begin{aligned} R_k(e^{j\omega}) &= |R_k(e^{j\omega})| e^{j\phi_k(\omega)}, \\ \forall \omega, \quad k &= 0, \dots, L-1. \end{aligned} \quad (4.4)$$

The (i, l) th entry of $\mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega})$ is given by

$$\begin{aligned} [\mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega})]_{i,l} &= |R_i(e^{j\omega})| |R_l(e^{j\omega})| e^{j(\phi_i(\omega) - \phi_l(\omega))}. \end{aligned} \quad (4.5)$$

If $\mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega})$ is pseudocirculant, all the elements on the main diagonal have to be equal. This gives

$$\begin{aligned} |R_k(e^{j\omega})| &= |R_{L-1}(e^{j\omega})|, \\ \forall \omega, \quad k &= 0, \dots, L-1. \end{aligned} \quad (4.6)$$

Using the fact that a pseudocirculant is, in particular, Toeplitz, we can say that the $(k, k+1)$ th elements have to be equal for $k = 0, \dots, L-2$. This gives the following condition on the phases:

$$\begin{aligned} (\phi_k(\omega) - \phi_{k+1}(\omega))_{2\pi} &= \phi(\omega), \\ \forall \omega, \quad k &= 0, \dots, L-2. \end{aligned} \quad (4.7)$$

Using the property of pseudocirculants that the $(1, 0)$ th entry is obtained by multiplying the $(0, L-1)$ th entry by $e^{-j\omega}$, we get

$$\begin{aligned} (\phi_0(\omega) - \phi_{L-1}(\omega))_{2\pi} &= (\omega + \phi_1(\omega) - \phi_0(\omega))_{2\pi}, \quad \forall \omega. \end{aligned} \quad (4.8)$$

Using the recursion (4.7) we get

$$\begin{aligned} (\phi_k(\omega))_{2\pi} &= (\phi_{L-1}(\omega) + (L-1-k)\phi(\omega))_{2\pi}, \\ \forall \omega, \quad k &= 0, \dots, L-2. \end{aligned} \quad (4.9)$$

Setting $k = 0$ in (4.9) and subtracting from (4.8), we get

$$(L\phi(\omega) - \omega)_{2\pi} = 0, \quad \forall \omega. \quad (4.10)$$

This can be written as an exact functional equality as

$$L\phi(\omega) = \omega - 2\pi P(\omega), \quad \forall \omega \quad (4.11)$$

for an appropriate integer-valued function $P(\omega)$. Substituting (4.11) in (4.9) we get the following relation among the polyphase components:

$$\begin{aligned} R_k(e^{j\omega}) &= R_{L-1}(e^{j\omega}) e^{j(L-1-k)((\omega/L) - (2\pi P(\omega)/L))}, \\ \forall \omega, \quad k &= 0, \dots, L-1. \end{aligned} \quad (4.12)$$

Let $\Phi(\omega) = ((\omega/L) - (2\pi P(\omega)/L))$. Since this is a phase function, the quantity $(\Phi(\omega))_{2\pi}$ (i.e., the value modulo 2π) is periodic with period 2π . This gives the following condition on the integer-valued function $P(\omega)$

$$P(\omega + 2\pi) = P(\omega) + 1 \quad \text{modulo } L, \quad \forall \omega. \quad (4.13)$$

Once again, the modulo notation means that $P(\omega + 2\pi) - P(\omega) - 1$ is an integer multiple of L . This proves the "necessary" part.

An example of $P(\omega)$ is shown in Fig. 12 for $L = 4$. It can be seen that for this example, $P(\omega)$ takes three different integer values in the interval $0 \leq \omega < 2\pi$. This pattern repeats as in (4.13).

The "Sufficient" Part: We have to prove that (4.12) implies that $\mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega})$ is pseudocirculant. If we assume that (4.12) holds for the polyphase components, then the (i, l) th entry is given by

$$\begin{aligned} [\mathbf{G}(e^{j\omega}) \mathbf{G}^\dagger(e^{j\omega})]_{i,l} &= |R_{L-1}(e^{j\omega})|^2 e^{j(l-i)((\omega - 2\pi P(\omega))/L)}. \end{aligned} \quad (4.14)$$

It is easy to see that this satisfies the pseudocirculant conditions (2.11).

Proof of Step 3:

The "Necessary" Part: We first prove that (4.12) implies the spectral properties mentioned in Step 3. Substituting the relation (4.12) in (4.1), we get

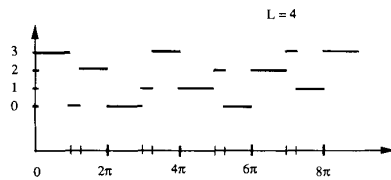
$$\begin{aligned} H(e^{j\omega}) &= \sum_{k=0}^{L-1} R_k(e^{j\omega L}) e^{-j\omega(L-1-k)} \\ &= R_{L-1}(e^{j\omega L}) e^{-(j2\pi(L-1)P(\omega L)/L)} \sum_{k=0}^{L-1} W_L^{-kP(\omega L)}. \end{aligned} \quad (4.15)$$

Consider the sum

$$\sum_{k=0}^{L-1} W_L^{-kP(\omega L)} \quad (4.16)$$

in (4.15). This sum will be nonzero for a frequency ω if and only if $P(\omega L) = 0 \pmod L$. Let us study the behavior of $P(\omega L)$ in the interval $0 \leq \omega < 2\pi$. Consider a set of L equispaced frequencies

$$\hat{\omega}_l = \hat{\omega}_0 + \frac{2\pi l}{L}, \quad 0 \leq l \leq L-1.$$


 Fig. 12. An example of the function $P(\omega)$.

In view of (4.13), the integer-valued function $P(\omega)$ will take on the value (0 modulo L) for one and only one of the above frequencies $\hat{\omega}_l$. (This statement is true for any initial choice $\hat{\omega}_0$). The transfer function $H(e^{j\omega})$ can therefore be nonzero only at one such frequency. This means that if for two distinct frequencies ω_1 and ω_2 ($\omega_1 \neq \omega_2$ modulo 2π), $H(e^{j\omega_1})$ and $H(e^{j\omega_2})$ are nonzero, then $\omega_1 - \omega_2 \neq (2\pi k/L)$ unless k is an integer multiple of L .

Fig. 13 indicates an example of $H(e^{j\omega})$, corresponding to the function $P(\omega)$ shown in Fig. 12. The filter has nonzero frequency response only in the interval where $P(\omega L)$ has the value (0 modulo L).

The "Sufficient" Part: We assume that $H(e^{j\omega})$ has the above mentioned frequency characteristics and prove that the relation (4.3) holds.

Consider (4.3), which shows how each polyphase component $R_k(e^{j\omega})$ is obtained from $H(e^{j\omega})$. Each polyphase component is obtained by the weighted addition of L stretched and shifted copies of $H(e^{j\omega})$. The stretching is by a factor of L and the shifting is by the amount $2\pi m$, $m = 0, \dots, L-1$. For the particular frequency characteristics of $H(e^{j\omega})$, the stretched and shifted versions of $H(e^{j\omega})$ will not overlap. Hence, for each frequency ω , there will only be one nonzero term in the summation (4.3). We can represent the index of that term as a function $m(\omega)$ of the frequency ω . Furthermore, $m(\omega)$ is independent of k . Hence, (4.3) can be rewritten as

$$R_k(e^{j\omega}) = \frac{1}{L} e^{j(\omega - 2\pi m(\omega))/L(L-1-k)} H(e^{j(\omega - 2\pi m(\omega))/L}),$$

$$k = 0, \dots, L-1. \quad (4.17)$$

This shows that the polyphase components satisfy the relation (4.12) with $P(\omega) = m(\omega)$. This completes the proof.

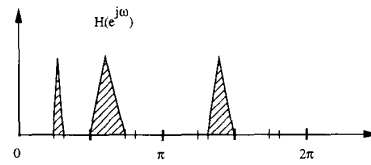
We can summarize the results proved above in the following theorem.

Theorem 4.1: The output of Fig. 1(c) is WSS for a WSS input $x(n)$ if and only if $H(e^{j\omega})$ is such that the L -fold decimation of its impulse response does not create aliasing. This condition is equivalent to the statement that $H(e^{j\omega})$ is an ideal image-free (or image-suppressing) interpolation filter, a term explained in Appendix III.

Remarks:

1) Assuming that $H(e^{j\omega})$ satisfies the conditions of the above theorem, the power spectrum of the WSS process $y(n)$ is given by (Appendix IV)

$$S_{yy}(e^{j\omega}) = \frac{1}{L} S_{xx}(e^{j\omega L}) |H(e^{j\omega})|^2. \quad (4.18)$$


 Fig. 13. A corresponding possible $H(e^{j\omega})$.

2) Suppose $y_h(n)$ and $y_g(n)$ are two signals generated from the WSS signal $x(n)$, by use of two interpolation filters $H(e^{j\omega})$ and $G(e^{j\omega})$. If $H(e^{j\omega})$ and $G(e^{j\omega})$ satisfy the conditions of Theorem 4.1, then $y_h(n)$ and $y_g(n)$ are WSS. Suppose, however, that the filter $H(e^{j\omega}) + G(e^{j\omega})$ does not satisfy the frequency occupancy conditions of Theorem 4.1. Then, according to the theorem, the sum $y(n) \triangleq y_h(n) + y_g(n)$ is not WSS. This is consistent with the fact that the sum of two WSS processes is not necessarily WSS (unless the two processes are *jointly* WSS).

Summarizing, we have shown that for a scalar WSS process $x(n)$ and corresponding power spectrum $S_{xx}(e^{j\omega})$, we can perform the following operations on $S_{xx}(e^{j\omega})$ without changing the WSS property at the output:

- 1) filtering (or "distortion" of the power spectrum);
- 2) translation by any amount (time-domain modulation);
- 3) stretching (decimation in time domain);
- 4) compression, and piecewise translation as explained in Theorem 4.1.

Since ideally band-limited filters cannot be implemented in practice, Theorem 4.1 implies that if we use interpolators in a multirate filtering scheme, the subsequent signals will not be WSS (their appropriately blocked versions will be WSS). This fact has some interesting implications. One of these is illustrated next.

V. AN APPLICATION: ADAPTIVE IDENTIFICATION OF BAND-LIMITED CHANNELS USING A MULTIRATE ADAPTIVE FILTER

Consider the channel identification scheme of Fig. 14. The channel is assumed to be a linear time invariant system with transfer function $C(z)$. The adaptive filter is represented by coefficients $a_{i,n}$, ($i = 0, \dots, L_a - 1$), and L_a is the filter length. The second subscript indicates the time instant of adaptation. In the adaptive identification procedure, the coefficients are updated to minimize the appropriate measure of the error $e(n)$. If the adaptation identification procedure converges to some steady state values a_i^l , then the frequency response of the corresponding filter $|A'(e^{j\omega})|$ resembles the channel frequency response.

Suppose that the channel is band limited to frequency σ (i.e., σ -BL). In that case, $|A'(e^{j\omega})|$ will also be close to being σ -BL. We can use this information to modify the adaptive filter as follows: split the adaptive filter as a cascade of a fixed σ -BL filter $H(z)$ and an adaptive part (Fig. 15). One advantage of this is that the adaptive part would

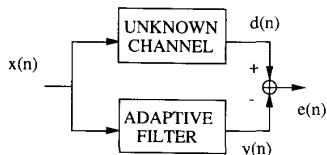


Fig. 14. Adaptive identification of an unknown channel.

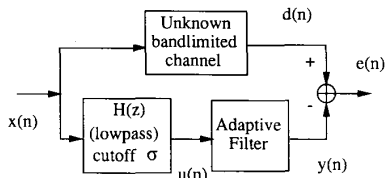


Fig. 15. Splitting the adaptive filter into fixed and adaptive parts.

now typically have fewer coefficients as it is required to match the passband shape of $|C(e^{j\omega})|$ only.

We can further note that the input $u(n)$ to the adaptive filter is a σ -BL signal. This suggests that there is some redundancy in $u(n)$ due to oversampling. We can thus decimate the signal using a fractional decimator before feeding to the adaptive filter. If L and M are two (relatively prime) integers such that $\sigma \leq \pi L/M$, then we can decimate the signal by the ratio M/L . The filter $H(z)$ can now be combined with the fractional decimation circuit as shown in Fig. 16. The advantage of the decimation is that the adaptive filter now operates at a lower rate. (The band-limited property has been exploited in the past in adaptive filtering literature. See, for example, [10], and [11]. Also, subband adaptation [3], [4] is a more sophisticated, though conceptually different, way to exploit band-limitedness and more generally spectral nonuniformity.) To match the data rates from both the branches in Fig. 16, we perform fractional decimation at the output of the band-limited channel also. Qualitatively, we can say that the adaptive filter is now required to match the M/L -fold stretched response $C(e^{j\omega L/M})$ in the region $0 \leq \omega \leq \pi$.

Summarizing, the multirate adaptive filtering structure of Fig. 16 has the following advantages:

- 1) the adaptation takes place at a lower rate;
- 2) the adaptive filter has a smaller length because it has to match only the passband of the channel's frequency response.

A. Computational Complexity of the Proposed Method

Suppose the unknown channel impulse response is estimated to have length L_c . For running the LMS algorithm [9] on the adaptation scheme shown in Fig. 14, we have to perform $2L_c + 1$ multiplications per input sample (MPU) and $2L_c + 1$ additions per input sample (APU).

For the multirate method, we have to calculate computational complexity for 1) the adaptation procedure; and 2) the implementation of the nonadaptive filters. Let L_f be

the length of the interpolation filter $H_c(z)$. Using the efficient polyphase technique mentioned in [12], we require L_f/M MPU and $(L_f - 1)/M$ APU. Similarly, if L_h is the length of $H_a(z)$, we need L_h/M MPU and $(L_h - 1)/M$ APU for implementing $H_a(z)$. One way to choose the length of the adaptive filter is to use the length constraint

$$\begin{aligned} L_h - 1 + M \times (L_a - 1) \\ = L_f - 1 + L \times (L_c - 1). \end{aligned} \quad (5.1)$$

This constraint implies that the highest powers of z^{-1} in the transfer functions in the upper and lower branches of Fig. 16 are equal. However, in our experience, the adaptive filter can have a smaller length without degrading performance significantly. Using (5.1), and using the fact that the adaptive filter is implemented at a rate slower than the input data rate, we get the computational complexity for the algorithm as

$$\begin{aligned} \text{MPU} &= 2 \left(\frac{L}{M} \right)^2 L_c + \frac{2L}{M^2} (L_f - L_h) \\ &+ \frac{2L}{M} \left(2 - \frac{L}{M} \right) \end{aligned} \quad (5.2a)$$

$$\begin{aligned} \text{APU} &= 2 \left(\frac{L}{M} \right)^2 L_c + \frac{2L}{M^2} (L_f - L_h) \\ &+ \frac{2L}{M} \left(2 - \frac{L}{M} \right) - \frac{2}{M} \end{aligned} \quad (5.2b)$$

This method thus offers savings in computations by a factor $(M/L)^2$ for a sufficiently long length L_c of the channel impulse response.

B. Need for a Block Adaptive Filter

Even though the above scheme has its own advantages, it also brings with it some disadvantages. To explain this, let us analyze the scheme of Fig. 16 to get the optimal set of coefficients a_i' under the assumption that the input $x(n)$ is WSS. From Fact 4.1, we know that the signal $w(k)$ input to the adaptive filter is CWSS with period $K = L/\text{gcd}(L, M) = L$ (the signal would be WSS if and only if $H_a(e^{j\omega})$ satisfies the conditions of Theorem 4.1). Due to this, we cannot use traditional Wiener filter theory to derive the optimal set of coefficients a_i' . The L -fold blocked version $w(n)$ is a vector WSS process. We can thus pose a matrix-Wiener filtering problem (Fig. 17) to solve for the $L \times L$ optimal coefficient matrix a_n'' . The $L \times 1$ output $y(n)$ of the block filter is compared with the blocked signal $v(n)$ to obtain error $e(n)$. The error $e(n)$ will be WSS and we can obtain the Wiener solution $A(z)$, which minimizes error energy $E[e^T(n)e(n)]$. We summarize the discussion above by proving the following result.

Fact 5.1: Consider the channel identification problem of Fig. 16, where L and M relatively prime, and $H_a(e^{j\omega})$ and $H_c(e^{j\omega})$ are low pass with same passband region. The Wiener (optimal) solution is, in general, an $(\text{LPTV})_L$ sys-

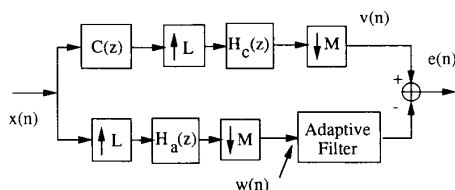


Fig. 16. Multirate adaptive filtering scheme for identification of band-limited channel.

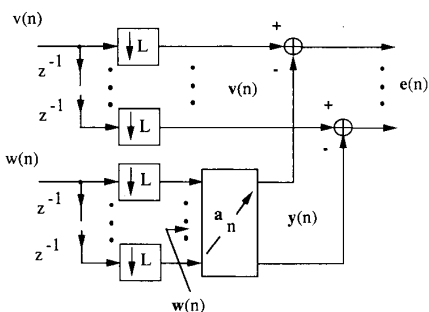


Fig. 17. Optimal filtering problem for the scheme in Fig. 16.

tem and is a scalar LTI if and only if both the filters $H_a(e^{j\omega})$ and $H_c(e^{j\omega})$ are such that the decimation of their impulse responses do not result in aliasing (i.e., they are image-free interpolation filters; see Appendix III).

Remark: Since we consider only low-pass filters, we can say that the Wiener solution is scalar LTI if and only if the filters are *ideally band limited* to $\pi L/M$.

Proof of Fact 5.1: From Theorem 4.1, for nonideal low-pass filters $H_c(e^{j\omega})$ or $H_a(e^{j\omega})$, the input $w(n)$ and the desired signal $v(n)$ will be $(\text{CWSS})_L$. Both the signals are WSS if and only if the filters are ideal low-pass filters. Since the signals are $(\text{CWSS})_L$ in general, the blocked versions $v(n)$ and $w(n)$ are vector WSS.

The vector processes $w(n)$ and $v(n)$ are jointly WSS, as justified in Appendix V. So we can apply the vector version of Wiener-filtering theory; from (2.23) we know that the Wiener solution is given by

$$A(e^{j\omega}) = S_{wv}^\dagger(e^{j\omega}) S_{ww}^{-1}(e^{j\omega}) \quad (5.3)$$

(assuming the full normal-rank condition mentioned after Fact 2.4). The matrices $S_{wv}(e^{j\omega})$ and $S_{ww}^{-1}(e^{j\omega})$ in general are not pseudocirculant, so $A(e^{j\omega})$ in general need not be pseudocirculant. From Fact 2.3, it can be seen that the optimal filter cannot be written as an LTI system. If the optimal solution has a matrix form as in (5.3), then the output at time n is the output of one of L filters [each filter being a row of $A(e^{j\omega})$] depending on the value (k modulo L). Thus, the optimal filter is an $(\text{LPTV})_L$ system.

However, in the case when both the low-pass filters are ideal, both the signals $w(n)$ and $v(n)$ are WSS. In this case, the matrices $S_{wv}(e^{j\omega})$ and $S_{ww}(e^{j\omega})$ are pseudocirculant (as shown in Appendix V). As a result, the inverses and transposed conjugates (as well as their products) are pseudocirculant. Thus, the optimal filter $A(e^{j\omega})$ will be a pseudocirculant, i.e., its unblocked version is an LTI system for this case. The transfer function of the optimal fil-

ter in this case would be $A(e^{j\omega}) = \sum_{k=0}^{L-1} e^{-j\omega k} A_k(e^{j\omega L})$, where $A_k(e^{j\omega})$ are entries of the 0th row of $A(e^{j\omega L})$.

For a real-time setup, the above result leads us to the following conclusions.

1) To be able to converge to the optimal solution, we should use an $L \times L$ adaptive filter a_n . This would result in a better performance compared to a scalar adaptive filter. However, a matrix filter is inherently more complicated than a scalar filter and might offset the advantages offered by the multirate approach.

2) As the stopband attenuation of filters $H_a(z)$ and $H_c(z)$ increases, the performance of the scalar filter will approach that of the matrix filter. As a result, the use of a scalar adaptive filter would result in little loss in performance.

Clearly there is a tradeoff involved in designing $H_a(z)$ and $H_c(z)$. We shall not further discuss the optimal choice in this tradeoff (which appears to require a careful study) but proceed to demonstrate the above ideas with an example. It should be noted that the passband ripple size for $|H_a(e^{j\omega})|$ is not very crucial, as the adaptive filter will perform passband equalization. This fact can be exploited to reduce cost of the prefilter $H_a(z)$. However, this passband ripple cannot be excessively large, because the adaptive filter itself would then fail to represent the channel faithfully. The ideal responses of $H_a(z)$ and $H_c(z)$, which we should approximate, are shown in Fig. 18.

C. A Simulation Example

We now demonstrate the above results by simulation. We have used the LMS algorithm for updating the adaptive filter. The channel $C(z)$ was simulated using an FIR filter of length 77. The channel was designed to be band-limited to frequency $3\pi/4$. We thus chose $L = 3$, $M = 4$ for the simulation. The filters $H_a(z)$ and $H_c(z)$ were designed to be linear phase FIR filters of length 47 and 31, respectively, designed by the Parks-McClellan program [13]. Two different cases were studied: 1) scalar adaptive filter; and 2) matrix adaptive filter. For the scalar adaptive filtering case, the adaptive filter was chosen to have length 30, since this length gave satisfactory performance [although this length does not satisfy (5.1)]. For the matrix adaptive filter, the adaptive filter was a 3×3 matrix with each entry being a length 10 filter. The reason for this choice was that if the low-pass filters were ideal, blocking a length 30 filter (chosen for the scalar adaptive case) would have resulted in a block matrix, each of whose entries would have 10 coefficients. The blocked signals $w(n)$ and $v(n)$ were used in the adaptive updating. The adaptation procedure was run for various step sizes μ for both the methods. The results presented here are for the case where $\mu = 0.1$ for the scalar and $\mu = 0.3$ for the matrix adaptive filter. These step-size values were chosen because the algorithms gave minimum error energy at convergence (over a wide choice of μ). The error energy at convergence of these two cases was compared for differ-

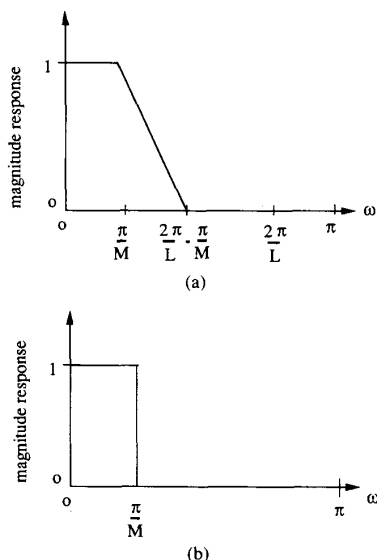


Fig. 18. Ideal magnitude response for (a) $H_c(z)$; and (b) $H_a(z)$.

ent attenuations of the low-pass filter $H_a(z)$. The filter $H_c(z)$ had 51-dB stopband attenuation and was not changed throughout the simulations. Table I gives the results of the simulation. A typical plot of the error energy versus iteration number is shown in Fig. 19, for both methods. The normalized error energy at the n th iteration is computed as

$$\frac{\sum_{k=0}^n e^2(k) \lambda^{n-k}}{\sum_{k=0}^n v^2(k) \lambda^{n-k}} \quad (5.4)$$

where the value $\lambda = 0.99$ was used for the simulations. The quantity in the numerator represents the exponential window estimate of the error energy at time n . This is normalized by the exponential window estimate of the desired signal energy in the denominator.

The results show that the matrix adaptive filter performs better than the scalar adaptive filter in terms of minimizing the error energy at convergence. This agrees with Fact 5.1, because the scalar adaptive filter converges to an LTI system that is suboptimal if $H_a(z)$ is not ideally band limited. On the other hand, the matrix filter converges to an (LPTV)_L filter. As the stopband attenuation of $H_a(z)$ increases, the signal $w(n)$ gets closer and closer to being WSS. So the relative degradation in the performance of the scalar adaptive filter reduces, as seen from Table I. This shows that designing $H_a(z)$ to have "good" attenuation reduces the degradation in the performance.

D. Special Case when $\sigma = \pi/M$

Now we consider the special case where the channel bandwidth is $\sigma = \pi/M$. This means that we now decimate the appropriate signals by the integer factor M . The adap-

TABLE I
A COMPARISON OF THE PERFORMANCES OF THE SCALAR AND MATRIX ADAPTIVE FILTERS

Stopband attenuation of $H_a(z)$ (dB)	Error energy at convergence (dB)	
	Scalar	Matrix
41	-15.4	-14.1
29	-13.1	-13.4
21	-9.5	-11.4
18	-7.4	-9.0

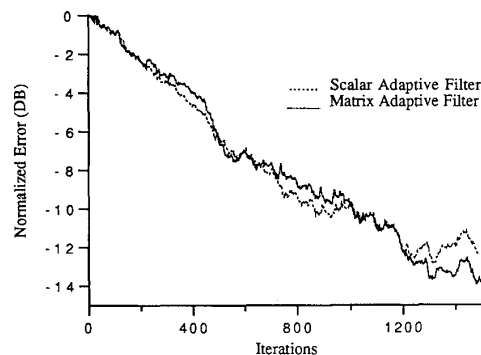


Fig. 19. Error energy versus iterations curves for scalar and matrix adaptive filtering cases.

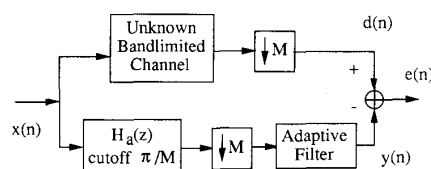


Fig. 20. Adaptive filtering scheme for a π/M band-limited channel.

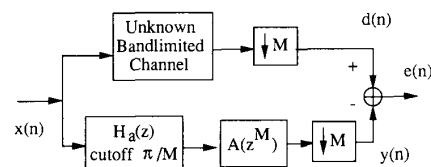


Fig. 21. Redrawing the adaptation scheme of Fig. 20 after convergence.

tive filtering scheme now has the form shown in Fig. 20. The adaptive filter has approximately M times fewer coefficients, and operates at an M times lower rate. So we achieve computational savings by a factor of M^2 . At convergence, the system is equivalent to Fig. 21. The cascade transfer function $H_a(z)A(z^M)$ now approximates the channel $C(z)$. This scheme can therefore be considered an extension, to the adaptive regime, of the interpolated FIR (IFIR) approach for efficient design and implementation of narrow-band FIR filters [14]. Once again, the main purpose of $H_a(z)$ is to provide satisfactory out-of-band attenuation. Its passband ripple size is not crucial.

VI. CONCLUSIONS

In this paper, we have addressed the question of the effects of multirate systems on a few statistical properties of random inputs. Starting from the basic multirate building blocks, viz., decimators, interpolators and modulators, we have derived results for more complex interconnections. We saw that when we start analyzing multirate systems, the assumption that signals are wide sense stationary is not valid even in theoretical study. It is more natural to assume that signals are CWSS. For example, the output $y(n)$ of an L -fold interpolation filter (in response to a WSS input) is $(\text{CWSS})_L$. We showed that $y(n)$ reduces to a WSS process if and only if the filter coefficients are such that the L -fold decimation of these coefficients results in no aliasing.

We have illustrated an application of the theoretical analysis to a multirate adaptive filtering scheme for identification of band-limited channels. This scheme exploits the fact that the channel is band limited and embeds the adaptive filter in a multirate configuration. We saw that this scheme is computationally efficient because the adaptive filter has smaller length, and is implemented at a lower speed. Using the results derived in the paper, we proved that the optimal filter for this scheme is a matrix filter. A matrix adaptive filter is computationally more expensive. However, the performance of a scalar adaptive filter (which is computationally less expensive) tends to be close to the matrix case if the fixed low-pass filters in the multirate scheme are designed to have good stopband attenuation and narrow transition bandwidth.

APPENDIX I

We prove the equivalence of the two definitions of cyclostationarity.

1) *Definition 1* \Rightarrow *Definition 2*: From Definition 1, the (i, l) th entry of the autocorrelation matrix is independent of n . Hence, the quantity

$$E[x(n_0L - i)x^*(n_0L - k_0L - l)] \quad (\text{A1.1})$$

is independent of n_0 , for all k_0 , for $0 \leq i, l \leq L - 1$. Now let $n = n_0L - i$ and $k = k_0L + l$. Then

$$\begin{aligned} E[x(n)x^*(n - k)] &= E[x(n_0L - i) \\ &\quad \cdot x^*(n_0L - k_0L - i - l)] \\ &= E[x((n_0 + 1)L - i)x^*((n_0 + 1) \\ &\quad \cdot L - k_0L - i - l)] \quad (\text{A1.2}) \end{aligned}$$

(Since the expression is independent of n_0)

$$= E[x(n + L)x^*(n + L - k)] \quad (\text{A1.3})$$

so that $R_{xx}(n, k) = R_{xx}(n + L, k)$.

2) *Definition 2* \Rightarrow *Definition 1*: The (i, l) th entry of the

autocorrelation matrix in Definition 1 is given by

$$\begin{aligned} E[x(nL - i)x^*(nL - kL - l)] \\ = E[x(nL + L - i)x^*(nL + L - kL - l)] \\ \text{(by Definition 2)} \quad (\text{A1.4}) \end{aligned}$$

$$= E[x((n + 1)L - i)x^*((n + 1)L - kL - l)]. \quad (\text{A1.5})$$

Clearly, the (i, l) th entry is independent of n . ■

APPENDIX II

Reproducing (2.17) for the $(0, m)$ th entry of $S_{xx}(z)$ we get

$$[S_{xx}]_{0,m} = \sum_{k=-\infty}^{\infty} E[x(nM)x^*(nM - kM - m)]z^{-k}. \quad (\text{A2.1})$$

If $R_{xx}(k)$ is the autocorrelation function of $x(n)$, then (A2.1) can be written as

$$[S_{xx}]_{0,m} = \sum_{k=-\infty}^{\infty} R_{xx}(kM + m)z^{-k}. \quad (\text{A2.2})$$

From (2.14) we can clearly see that this is the z -transform of the $(M - 1 - m)$ th polyphase component. Thus, the $(0, m)$ th entry of $S_{xx}(z)$ is the $(M - 1 - m)$ th polyphase component of $S_{xx}(z)$. ■

APPENDIX III

We first prove that if the signal $h(n)$ has the spectral characteristics mentioned in Theorem 4.1, then L -fold decimation of the signal does not result in aliasing. In the frequency domain, the decimation process creates L copies of the original function stretched by the factor L , shifted by $2\pi m$, $m = 0, \dots, L - 1$ and added. Clearly, no aliasing results for the spectral conditions of Theorem 4.1.

Conversely, if no aliasing results after L -fold decimation, then the L -fold stretched and $2\pi m$ -shifted versions do not overlap. This in turn implies that the frequency response has a total spectral occupancy of at most $2\pi/L$, and that its frequency shifted versions (in integer multiples of $2\pi/L$) do not overlap. This is precisely the condition given earlier in Theorem 4.1. ■

Remarks: A filter $H(e^{j\omega})$ with the frequency characteristics as above has some special properties in the deterministic multirate filtering case. The filter can be used before the L -fold decimator as shown in Fig. 22(a) to ensure that no aliasing takes place after decimation. Similarly, if this filter is used following an L -fold interpolator [Fig. 22(b)], it suppresses unwanted images. So we call it an image-free interpolation filter or an ideal image suppressor.

General Meaning of "Image-Free Interpolator": In a traditional image suppressing interpolation filter ([1], [2]), we suppress $L - 1$ images and retain one of the L images.

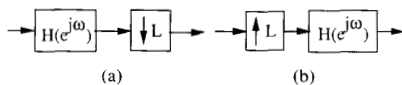


Fig. 22. Deterministic cases. (a) Alias-free decimation. (b) Image-free interpolation.

However, the filter $H(e^{j\omega})$ satisfying the conditions of Theorem 4.1 can be more general. The Fourier transform of its output will in general consist of different pieces from different images such that these pieces can be put together by frequency translation, in order to obtain an L -fold compressed version of the input power spectrum. Such a filter $H(e^{j\omega})$ is called a generalized image-free interpolation filter (or image suppressor).

APPENDIX IV

With $H(e^{j\omega})$ as in Theorem 4.1, we now derive an expression for the output power-spectral density.

We assume that $y(n)$ [and hence $t(n)$] is WSS. We derive the expression for $S_{tt}(e^{j\omega}) [= S_{yy}(e^{j\omega})]$. In Appendix II, we have proved a result that implies that the entries of the 0th row of the matrix $S_{tt}(e^{j\omega})$ are the polyphase components of $S_{tt}(e^{j\omega})$. Using this and the expression (4.5) we can write

$$\begin{aligned} S_{tt}(e^{j\omega}) &= \sum_{l=0}^{L-1} R_0(e^{j\omega L}) R_l^*(e^{j\omega L}) S_{xx}(e^{j\omega L}) e^{-j\omega l} \\ &= R_0(e^{j\omega L}) S_{xx}(e^{j\omega L}) \left(\sum_{l=0}^{L-1} R_l(e^{j\omega L}) e^{j\omega l} \right)^* \\ &= R_0(e^{j\omega L}) S_{xx}(e^{j\omega L}) e^{-j\omega(L-1)} \\ &\quad \cdot \left(\sum_{l=0}^{L-1} R_l(e^{j\omega L}) e^{-j\omega(L-1-l)} \right)^* \\ &= R_0(e^{j\omega L}) S_{xx}(e^{j\omega L}) e^{-j\omega(L-1)} H^*(e^{j\omega}). \end{aligned} \quad (\text{A4.1})$$

Since

$$R_0(e^{j\omega}) = R_{L-1}(e^{j\omega}) e^{j(L-1)\omega/L} W_L^{P(\omega)(L-1)} \quad (\text{A4.2})$$

we can rewrite (A4.1) as

$$S_{tt}(e^{j\omega}) = S_{xx}(e^{j\omega L}) H^*(e^{j\omega}) R_{L-1}(e^{j\omega L}) W_L^{P(\omega L)(L-1)}. \quad (\text{A4.3})$$

Using (4.15) we can write

$$S_{tt}(e^{j\omega}) = S_{xx}(e^{j\omega}) |R_{L-1}(e^{j\omega L})|^2 \sum_{l=0}^{L-1} W_L^{-lP(\omega L)}. \quad (\text{A4.4})$$

Using the fact that the sum (4.16) can only take two values (0 and L), we can write this down as

$$S_{tt}(e^{j\omega}) = S_{yy}(e^{j\omega}) = \frac{1}{L} S_{xx}(e^{j\omega L}) |H(e^{j\omega})|^2. \quad (\text{A4.5})$$

APPENDIX V

Many of the definitions and properties pertaining to $(\text{CWSS})_L$ signals extend easily to joint statistics of two random processes. We now summarize these.

Consider two scalar random processes $x(n)$ and $y(n)$. We say that they are jointly $(\text{CWSS})_L$ if the L -fold blocked versions $x(n)$ and $y(n)$ are jointly WSS, i.e., $x(n)$ and $y(n)$ are WSS and

$$E[x(n)y^\dagger(n-k)] \quad \text{is independent of } n. \quad (\text{A5.1})$$

Properties:

1) Notice, in particular that the joint CWSS property means that each process is individually $(\text{CWSS})_L$. Defining

$$R_{xy}(n, k) = E[x(n)y^*(n-k)] \quad (\text{A5.2a})$$

it can be shown that (A5.1) is equivalent to the property

$$R_{xy}(n+L, k) = R_{xy}(n, k), \quad \forall n, k. \quad (\text{A5.2b})$$

(This is proved by simple modifications of the proof given in Appendix I).

2) Let $x(n)$ and $y(n)$ be jointly $(\text{CWSS})_L$. Then the decimated versions of $x(Mn)$ and $y(Mn)$ are individually $(\text{CWSS})_L$ (Fact 3.3). In fact $x(Mn)$ and $y(Mn)$ are jointly $(\text{CWSS})_L$, since it can be shown that they satisfy a result similar to (A5.2b).

3) Let the L -fold blocked versions $x(n)$ and $y(n)$ be jointly WSS. Let $S_{xy}(e^{j\omega})$ be their cross-power-spectral density. If the scalar processes $x(n)$ and $y(n)$ are jointly WSS, then $S_{xy}(e^{j\omega})$ is a pseudocirculant. Conversely, if $x(n)$ and $y(n)$ are WSS and $S_{xy}(e^{j\omega})$ is pseudocirculant, then $x(n)$ and $y(n)$ are jointly WSS. The proof is identical to that of Fact 2.2.

4) Now consider Fig. 23 where we generate the vector processes $v(n)$ and $w(n)$ by passing a scalar process $x(n)$ through two LTI system with $L \times 1$ transfer matrices $b(e^{j\omega})$ and $a(e^{j\omega})$. Let $x(n)$ be WSS. Then

$$S_{wv}(e^{j\omega}) = S_{xx}(e^{j\omega}) a(e^{j\omega}) b^\dagger(e^{j\omega})$$

where $S_{xx}(e^{j\omega})$ is the (scalar) power spectral density of $x(n)$. So $S_{wv}(e^{j\omega})$ is pseudocirculant if and only if the $L \times L$ matrix $a(e^{j\omega}) b^\dagger(e^{j\omega})$ is pseudocirculant.

5) Let $H_a(e^{j\omega})$ and $H_b(e^{j\omega})$ be filters with polyphase components $a_i(e^{j\omega})$ and $b_i(e^{j\omega})$, $0 \leq i \leq L-1$. Define

$$a(e^{j\omega}) = [a_0(e^{j\omega}) \cdots a_{L-1}(e^{j\omega})]^T,$$

$$b(e^{j\omega}) = [b_0(e^{j\omega}) \cdots b_{L-1}(e^{j\omega})]^T. \quad (\text{A5.3})$$

Let $H_a(e^{j\omega})$ and $H_b(e^{j\omega})$ be ideal filters satisfying the conditions of Theorem 4.1, with same $P(\omega)$ (i.e., same pass-band regions). Then $a(e^{j\omega}) b^\dagger(e^{j\omega})$ is pseudocirculant. The reason for this is sketched as follows: The k, l element of $a(e^{j\omega}) b^\dagger(e^{j\omega})$ is given by $a_k(e^{j\omega}) b_l^*(e^{j\omega})$. Since $H_a(e^{j\omega})$ and $H_b(e^{j\omega})$ satisfy the conditions of Theorem 4.1, the polyphase components $a_i(e^{j\omega})$ and $b_i(e^{j\omega})$ have forms similar to (4.12). By using this in the product $a_k(e^{j\omega}) b_l^*(e^{j\omega})$ we can show that $a(e^{j\omega}) b^\dagger(e^{j\omega})$ is pseudocirculant indeed. ■

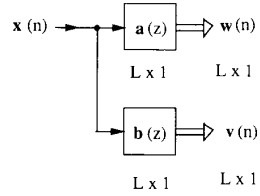


Fig. 23. Two vector processes generated from a scalar process.

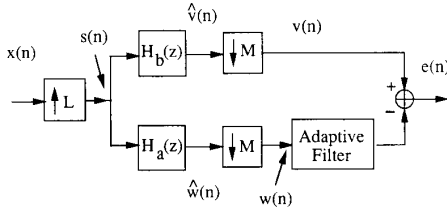


Fig. 24. A redrawing of Fig. 16.

Applications: The above results on joint statistics have been applied in our discussion of the matrix Wiener filter of Fig. 17, which in turn was derived from Fig. 16 by blocking the scalar signals $w(n)$ and $v(n)$.

Application 1: First note that Fig. 16 can be redrawn as in Fig. 24 where $H_b(e^{j\omega}) = C(e^{j\omega L})H_c(e^{j\omega})$. With $H_b(e^{j\omega})$ and $H_a(e^{j\omega})$ denoting the blocked versions of $H_b(e^{j\omega})$ and $H_a(e^{j\omega})$, we can redraw the filtering part as in Fig. 25. Notice that $\hat{v}(n)$ and $\hat{w}(n)$ are the blocked versions of the inputs to the M -fold decimators. If $x(n)$ is WSS then $s(n)$ is $(CWSS)_L$, i.e., $s(n)$ is WSS. Now the vector random process $\begin{pmatrix} \hat{v}(n) \\ \hat{w}(n) \end{pmatrix}$ is the output of the LTI system $\begin{pmatrix} H_b(e^{j\omega}) \\ H_a(e^{j\omega}) \end{pmatrix}$, and is therefore WSS. So the vectors

$\hat{v}(n)$ and $\hat{w}(n)$ are jointly WSS, i.e., the unblocked versions $v(n)$ and $w(n)$ are jointly $(CWSS)_L$. In view of Property 2 above, the decimated versions $v(n)$ and $w(n)$ are jointly $(CWSS)_L$, i.e., their blocked versions $\hat{v}(n)$ and $\hat{w}(n)$ are jointly WSS. We used this joint WSS property in writing down the Wiener solution (5.3).

Application 2: We claimed that when the filters $H_b(e^{j\omega})$ and $H_a(e^{j\omega})$ are ideal low pass, the matrices $S_{ww}(e^{j\omega})$ and $S_{vv}(e^{j\omega})$ are pseudocirculant (assuming of course that $x(n)$ in Fig. 16 is WSS). This is proved as follows (refer to Fig. 24): the ideal nature of the filters ensures that $\hat{w}(n)$, $\hat{v}(n)$, $w(n)$, and $v(n)$ are individually WSS. In particular, this means that the power-spectral density $S_{ww}(e^{j\omega})$ of the blocked version $\hat{w}(n)$ of $w(n)$ is pseudocirculant [Fact 2.2(a)].

We now prove the subtler fact that $S_{vv}(e^{j\omega})$ is pseudocirculant. The generation of $\hat{v}(n)$ and $\hat{w}(n)$ can be represented as in Fig. 26 where $b(e^{j\omega})$ and $a(e^{j\omega})$ are the polyphase vectors of $H_b(e^{j\omega})$ and $H_a(e^{j\omega})$, respectively. Since $H_b(e^{j\omega})$ and $H_a(e^{j\omega})$ have the same passband region, the matrix $b(e^{j\omega})a^\dagger(e^{j\omega})$ is pseudocirculant (Property 5). So the joint power-spectral density of the outputs

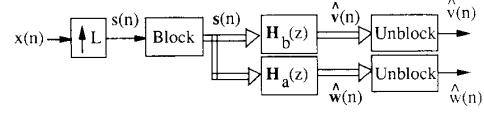


Fig. 25. A schematic for the generation of the signals $\hat{v}(n)$ and $\hat{w}(n)$ in Fig. 24.

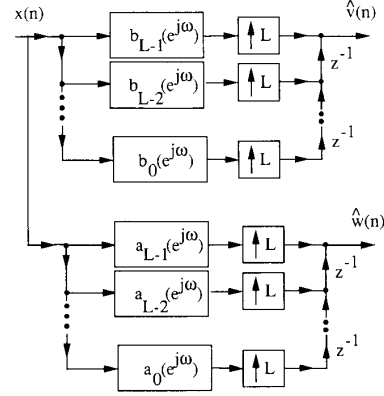


Fig. 26. Polyphase implementation of $H_b(z)$ and $H_a(z)$.

of $b(e^{j\omega})$ and $a(e^{j\omega})$ is pseudocirculant. This means that $\hat{v}(n)$ and $\hat{w}(n)$ are jointly WSS (Property 3). So the decimated versions $v(n)$ and $w(n)$ are jointly WSS as well. By invoking Property 3 again, we finally conclude that $S_{vv}(e^{j\omega})$ is pseudocirculant indeed!

REFERENCES

- [1] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [2] P. P. Vaidyanathan, "Multirate-digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proc. IEEE*, vol. 78, no. 1, Jan. 1990.
- [3] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1988, pp. 1572-1575.
- [4] W. Kellerman, "Analysis and design of multirate systems for cancellation of acoustic echoes," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1988, pp. 2570-2573.
- [5] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1965.
- [6] W. B. Davenport, Jr., and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*. New York: McGraw Hill, 1958.
- [7] P. P. Vaidyanathan and S. K. Mitra, "Polyphase networks, block digital filtering, LPTV systems, and alias free QMF banks: A unified approach based on pseudocirculants," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 381-391, Mar. 1988.
- [8] P. Z. Peebles, *Probability, random variables, and random signal principles*. New York: McGraw-Hill, 1987.
- [9] B. Widrow and S. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, N.J.: 1985.
- [10] J. C. Lee and C. K. Un, "Block realization of multirate adaptive digital filters," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 34, Feb. 1986.
- [11] G. Ungerboeck, "Fractional tap-spacing equalizer and consequences for clock recovery in data modems," *IEEE Trans. Commun.*, vol. COM-24, Aug. 1976.
- [12] C. Hsiao, "Polyphase filter matrix for rational rate conversions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Dallas, TX, Apr. 1987, pp. 2173-2176.
- [13] J. H. McClellan, T. W. Parks, and L. R. Rabiner, "A computer program for designing optimum FIR linear phase digital filters," *IEEE Trans. Audio Electroacoust.*, vol. 21, pp. 506-526, Dec. 1973.

- [14] Y. Neuvo, C.-Y. Dong, and S. K. Mitra, "Interpolated finite impulse response filters," *IEEE Trans. Acoust., Speech, Signal Processing*, pp. 563-570, June 1984.
- [15] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Processing Mag.*, Apr. 1991.



Vinay P. Sathe (S'85-M'90) received the B.Tech. degree from the Indian Institute of Technology, Bombay, in 1986, and the M.S. and the Ph.D. degrees from the California Institute of Technology in 1987 and 1991, respectively, all in electrical engineering.

From 1986 to 1987, he was the recipient of the Caltech Graduate Fellowship. He is currently with the Video and Electronic Systems Labs (VESL) at Tektronix Inc., Beaverton, OR. His current areas of work include video signal processing, multirate adaptive filtering, and digital communication systems.



P. P. Vaidyanathan (S'80-M'83-SM'88-F'91) was born in Calcutta, India, on October 16, 1954. He received the B.Sc. (Hons.) degree in physics and the B.Tech. and M.Tech. degrees in radiophysics and electronics, from the University of Calcutta, India, in 1974, 1977, and 1979, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 1982.

From 1982 to 1983 he was a postdoctoral fellow at the University of California, Santa Barbara. In 1983 he joined the Electrical Engineering Department of the California Institute of Technology as an Assistant Professor, and since 1988 has been an Associate Professor of electrical engineering there. His main research interests are in digital signal processing, multirate systems, wavelet transforms, and adaptive filtering.

Dr. Vaidyanathan served as Vice-Chairman of the Technical Program committee for the 1983 IEEE International Symposium on Circuits and Systems, and as an Associate Editor for the *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS* for the period of 1985-1987. He also served as the Technical Program Chairman for the 1992 IEEE International Symposium on Circuits and Systems. He was a recipient of the Award for Excellence in Teaching at the California Institute of Technology for the year of 1983-1984. He also received the NSF's Presidential Young Investigator award in 1986. In 1989 he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990 he received the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommunications Engineers, India, for his joint paper in the *IETE Journal*.