A State Space Approach to the Design of Globally Optimal FIR Energy Compaction Filters

Jamal Tuqan, *Member, IEEE*, and P. P. Vaidyanathan, *Fellow, IEEE*

***Abstract*—We introduce a new approach for the least squared optimization of a weighted FIR filter of arbitrary order N under the constraint that its magnitude squared response be Nyquist(M). Although the new formulation is general enough to cover a wide variety of applications, the focus of the paper is on optimal energy compaction filters. The optimization of such filters has received considerable attention in the past due to the fact that they are the main building blocks in the design of principal component filter banks (PCFBs). The newly proposed method finds the optimum product filter $F_{opt}(z) = H_{opt}(z)H_{opt}(z^{-1})$ corresponding to the compaction filter $H_{opt}(z)$. By expressing $F(z)$ in the form $D(z) + D(z^{-1})$, we show that the compaction problem can be completely parameterized in terms of the state-space realization of the causal function $D(z)$. For a given input power spectrum, the resulting filter $F_{opt}(z)$ is guaranteed to be a *global* optimum solution due to the convexity of the new formulation. The new algorithm is universal in the sense that it works for any M, arbitrary filter length N, and any given input power spectrum. Furthermore, additional linear constraints such as wavelets regularity constraints can be incorporated into the design problem. Finally, obtaining $H_{opt}(z)$ from $F_{opt}(z)$ does not require an additional spectral factorization step. The minimum-phase spectral factor $H_{\min}(z)$ can be obtained automatically by relating the state space realization of $D_{opt}(z)$ to that of $H_{opt}(z)$.**

***Index Terms*—Discrete-time positive real lemma, energy compaction filters, Kalman–Yakubovich–Popov (KYP) lemma, linear matrix inequality (LMI), optimum orthonormal subband coder (SBC), principal components filter bank (PCFB), semi-definite programming (SDP).**

I. INTRODUCTION

CONSIDER the following optimization problem

$$\max_{H(e^{j\omega})} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 W(e^{j\omega}) \frac{d\omega}{2\pi} \tag{1}$$

subject to

$$\frac{1}{M}\sum_{k=0}^{M-1} \left| H\left(e^{j(\omega - 2\pi k/M)}\right)\right|^2 = |H(e^{j\omega})|^2|_{\downarrow M} = 1 \tag{2}$$

Manuscript received August 9, 1999; revised May 12, 2000. This work was supported in part by the National Science Foundation under Grant 0703755. The associate editor coordinating the review of this paper and approving it for publication was Prof. Hideaki Sakai.

J. Tuqan is with the Image and Video Communications Group, IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598 USA (e-mail: tuqan@us.ibm.com).

P. P. Vaidyanathan is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: ppvnath@sys.caltech.edu).

Publisher Item Identifier S 1053-587X(00)07688-1.

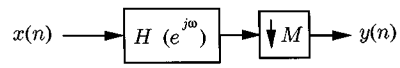


Fig. 1. Schematic of the FIR energy compaction problem.

where $H(e^{j\omega})$ is a real coefficient FIR filter of order N. The constraint (2) means that the magnitude squared response $|H(e^{j\omega})|^2$ is Nyquist(M) [1, pp. 151–152]].

The problem described above has received considerable attention in the past because of its wide occurrence in different disciplines depending on the choice of the frequency weight function $W(e^{j\omega})$. As an example, consider the problem of designing optimum FIR transmitter and receiver filters for data transmission over bandlimited channels [2]–[4]. Such filters are used in data modems realized predominantly in digital technology. The filters are designed so that maximum energy concentration is achieved in the transmission bandwidth of the channel, and zero intersymbol interference (ISI) is obtained when the filters operate in cascade. With a receiver filter $H_r(e^{j\omega})$ matched to the transmitter filter $H_t(e^{j\omega})$ and by choosing $W(e^{j\omega}) = rect(\omega/\omega_c)$, where ω_c is the cutoff frequency of the lowpass channel, the problem can be indeed expressed in the form described by (1) and (2) (see [4] for details). Other applications are in echo cancellation [5], the standard problem of designing FIR orthonormal filterbanks with good frequency selectivity [6], quantization of a class of nonbandlimited signals [7], optimization of wavelet basis [8], [9] and identification of time-varying systems [10], to name a few. Although the new method proposed in this paper is general enough to cover any of the previously mentioned applications, it is the design of FIR energy compaction filters that provides the main motivation of this work.

A. The FIR Energy Compaction Problem

Consider the scheme of Fig. 1, where $H(z)$ is a real coefficient FIR filter of order N. The input $x(n)$ is assumed to be a zero mean widesense stationary (WSS) random process with a power spectrum $S_{xx}(e^{j\omega})$. The output of the filter is decimated by M to produce $y(n)$. For a fixed pair (M, N), the FIR energy compaction problem is to maximize the output variance σ_y^2 subject to the Nyquist(M) constraint on $|H(e^{j\omega})|^2$. The optimal solution to the problem $H_{opt}(e^{j\omega})$ is termed an energy compaction filter. Since the decimator does not change the variance of the filter output, σ_y^2 is given by (1) with $W(e^{j\omega}) = S_{xx}(e^{j\omega})$.

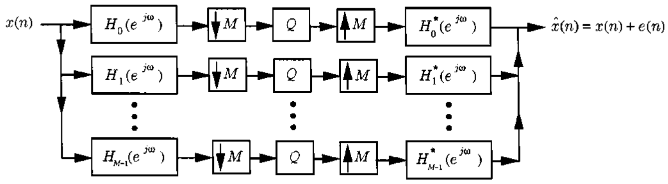


Fig. 2. M-channel FIR orthonormal filter bank with scalar quantizers.

A quantitative measure of performance (the compaction gain) is defined as follows:

$$G_{comp}(M, N) = \frac{\sigma_y^2}{\sigma_x^2} = \frac{\displaystyle\int_{-\pi}^{\pi} |H(e^{j\omega})|^2 S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi}}{\displaystyle\int_{-\pi}^{\pi} S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi}} \quad (3)$$

where σ_x^2 is the variance of $x(n)$. A compaction filter therefore maximizes the compaction gain. Note that by the Nyquist constraint, $G_{comp}(M, N) \leq M$. Note also that if the filter order is unconstrained, i.e., ideal filter solutions are permitted, an optimum filter has the following form [8], [11]–[13]: For all $\omega \in [0, 2\pi/M]$

$$H\left(e^{j(\omega + 2n\pi/M)}\right) = \begin{cases} 1, & \text{if } \omega \in \Omega \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where $\Omega = \{\omega \in [0, 2\pi/M]\colon W(e^{j(\omega+2n\pi/M)}) \geq W(e^{j(\omega+2m\pi/M)})\}$ for all $m \neq n$. For more details, see any of the above references.

B. Background and Motivation

FIR energy compaction filters, as defined above, play a key role in the statistical optimization of orthonormal filter banks when subband quantizers are present. To see this, consider the M-channel FIR orthonormal filter bank shown in Fig. 2, where the boxes labeled $\mathcal{Q}$ represent scalar uniform quantizers. Since the filter bank is orthonormal, the filters satisfy the following condition: $H_i(e^{j\omega})\ H_j^*(e^{j\omega})|_{\downarrow M} = \delta(i-j)$ [1], implying, in particular, that $|H_k(e^{j\omega})|^2$ is Nyquist(M) for each filter k (the superscript $*$ denotes complex conjugation). Given a fixed budget of b bits for the subband quantizers, the design of an optimum orthonormal subband coder (SBC) consists of simultaneously optimizing the analysis and synthesis filters as well as choosing a subband bit allocation strategy such that the average variance of the output error $e(n)$ is minimized. Under the *high bit rate* quantizer assumptions [14], and with the optimum bit allocation, the objective function is the well-known coding gain expression

$$\mathcal{G}_{\mathrm{SBC}}(M) = \sigma_x^2 \Bigg/ \left(\prod_{k=1}^{M} \sigma_{x_k}^2\right)^{1/M} \quad (5)$$

where $\sigma_{x_k}^2$ is the variance of the kth subband signal. Since σ_x^2 is fixed, the optimization of the analysis filters consists of minimizing the geometric mean of the subband variances under the orthonormality condition. For the unconstrained filter order case, Vaidyanathan derives a set of necessary and sufficient conditions for optimality of the filterbank and presents a procedure to obtain an optimum orthonormal SBC [12]. For the finite-order case, several design techniques have been proposed, but global optimality of the resulting filters is not guaranteed. One impetus for this is that the geometric mean is a concave function, making the above problem a difficult one to solve both theoretically and numerically. An alternative solution to the direct minimization of the geometric mean is the design of a so-called principal component filter bank (PCFB). PCFBs were first introduced in the context of optimal signal representation and are defined as follows [8], [15]: Consider Fig. 3, where $(M-P)$ channels are dropped in the synthesis part of an M-channel filter bank. A filterbank that minimizes the average mean square reconstruction error for *all* P is called a PCFB. For the ideal filter case, the solution to this problem was first derived by Unser [11] for $M = 2$ and then by Tsatsanis and Giannakis [8] for the general M-channel case. While Unser's formulation imposes *a priori* the orthonormality constraint on the filterbank, the work in [8] does not. Nevertheless, it turns out that the optimum PCFB is indeed orthonormal and therefore satisfies the so-called **majorization property**: An (orthonormal) PCFB produces a decreasing arrangement of the subband variances $\sigma_{x_1}^2 \geq \sigma_{x_2}^2 \cdots \geq \sigma_{x_P}^2$ such that, for all $1 \leq P < M$, $\sum_{k=1}^{P} \sigma_{x_k}^2$ is maximized. In particular, for $P = 1$, $\sigma_{x_1}^2$ must be maximized, that is, the objective function in (1) with $W(e^{j\omega}) = S_{xx}(e^{j\omega})$ must be maximized. Since we specifically consider the class of orthonormal filterbanks in this paper, the Nyquist constraint (2) is further imposed on the maximizing filter. Note that for $P = M$, $\sum_{k=1}^{M} \sigma_{x_k}^2 = M\sigma_x^2$ and is therefore fixed. The set of subband variances $\{\sigma_{x_k}^2\}$ generated by a PCFB is said to "majorize" any other possible set of subband variances $\{\sigma_{y_k}^2\}$. The connection between PCFB's and optimum orthonormal subband coders is established using a "majorization" theorem [16]. The result states that the majorization property of the subband variances of a PCFB implies, in particular, that $(\prod_{k=0}^{M-1} \sigma_{x_k}^2)^{1/M}$ is minimized. Note that the converse is, in general, not true. Therefore, instead of directly maximizing (5), we can, in principle, obtain an optimum orthonormal SBC by designing a PCFB. Unfortunately, the existence of a PCFB over the class of finite-order orthonormal filter banks is not, in general, guaranteed [17]. Nevertheless, if a PCFB exists, designing an optimal FIR energy compaction filter is a *necessary* first step in finding such a filterbank [18]. Finally, in a recent development, it is shown that a PCFB is optimal whenever the objective function to be minimized is a concave function of the subband variances produced by the orthonormal filter bank [19]. It follows that orthonormal PCFB's are also optimal for a variety of other signal processing applications, such as, for example, noise reduction.

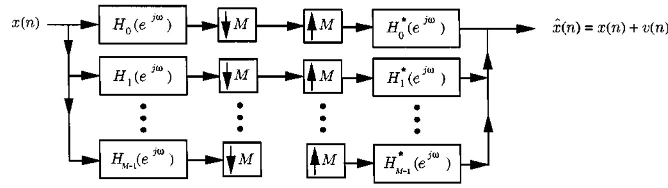


Fig. 3. M-channel FIR principal component filter bank. Here only the first two channels are retained in the synthesis part.

Although we have introduced the notion of a PCFB through its relation to the orthonormal subband coder problem, the topic itself is an active and equally important area of research. PCFB's usually play a fundamental role in multirate signal modeling [7], optimal signal representation [13], multitone digital communications and CDMA [4], and sampling applications [20]. We also note the work of Strintzis [21], who extends the above ideas to the class of multidimensional biorthogonal filterbanks and shows a close connection between the energy compaction problem and the problem of finding optimal analysis filters, given an arbitrary set of synthesis filters. We emphasize, however, that the energy compaction problem statement differs from the one proposed in this paper when the class of biorthogonal filter banks is considered.

C. Contribution and Organization of the Paper

The main contribution of this paper is the development of an efficient and numerically robust algorithm that finds the *global* optimum solution for the FIR energy compaction problem. The proposed method is universal in the sense that it works for any M, arbitrary filter length N, and the whole class of WSS random processes. The new method is expressed as a multiobjective semidefinite program that is convex and can be solved efficiently and with great accuracy using recently developed interior point methods [22]. The semidefinite program finds the optimum product filter $F_{opt}(z) = H_{opt}(z)H_{opt}(z^{-1})$ corresponding to the compaction filter $H_{opt}(z)$ and, in general, requires an additional spectral factorization step to obtain $H_{opt}(z)$. Spectral factorization is a procedure that is computationally expensive and numerically unstable. Nevertheless, we will show that if the minimum-phase spectral factor is desired, the spectral factorization step can be avoided. Finally, we believe that this paper has some tutorial value in the sense that it brings to the attention of signal processing researchers important and newly developed convex optimization techniques [22], particularly *semidefinite programming*. These optimization tools have been extensively used by the control community due to the natural occurrence of linear matrix inequalities (LMI's) (to be discussed shortly) in systems theory [23] but seem to have not been fully exploited in the field of signal processing.

The paper is organized as follows. In Section II, the difficulty in solving the general problem described by (1) and (2) is outlined by a brief overview of previous work. In Section III, by expressing the product filter $F(z)$ as $D(z) + D(z^{-1})$, we show that the FIR compaction problem is completely characterized by the state space realization of the *causal* function $D(z)$. The main advantage of this approach is that it fully exploits the rationality of the function to be optimized. The problem constraints can be now satisfied using a finite number of parameters, permitting the exact solution to be found. In Section IV, we study in detail the minimum-phase spectral factor and its properties. In particular, several theorems characterizing this special spectral factor are derived. The results of this section are important in order to avoid an additional spectral factorization step after obtaining $F_{opt}(z)$. Simplifications of some of the results of Sections III and IV for the particular FIR case under study are presented in Section V. In Section VI, we prove the convexity of the new formulation and, in Section VII, we show how regularity constraints can be formulated as linear matrix inequalities and equality constraints in terms of the state space realization of $D(z)$. Finally, in Section VIII, numerical examples are provided to illustrate the performance of the proposed algorithm. Part of this work has been presented in [24] and [25].

II. Summary of Previous Work

The general FIR optimization problem described in (1) and (2) has been considered by a number of authors. The different design approaches can be broadly classified into four main categories.

1) Optimizing the FIR Lattice Structure: It is well known that the class of two-channel FIR orthonormal filter bank is *completely* parameterized by a lattice structure [1, pp. 302–314]. One can therefore optimize the lattice coefficient, which is a set of N angles θ_k, $1 \leq k \leq N$, to obtain the compaction filter's impulse response $h(n)$. Since the Nyquist condition (2) is automatically enforced by the lattice structure, the problem is unconstrained, and unlike other approaches described below, no spectral factorization is required. The main drawback with this formulation is that it is highly nonlinear and cannot be expressed as a convex program. The quasi-Newton method used in [26] and the ring algorithm proposed in [27] both converge to a *local* maximum that depends on the starting point of the algorithm. Taubman and Zakhor [28] propose to use a multistart algorithm that generates several local optima over a subset of the parameter space.

2) Quadratically Constrained Quadratic Programming Method: The problem in this case is formulated in terms of the impulse response $h(n)$ of the filter $H(z)$ as follows:

$$\text{maximize} \quad h^T R_{xx} h \tag{6}$$

subject to the Nyquist(M) constraint, which is now expressed as

$$h^T P_l h = \delta(l) \quad \text{for } l = 0, 1, \cdots, \lfloor N/M \rfloor \tag{7}$$

where R_{xx} is a Toeplitz Hermitian matrix with first row equal to $[r(0)\, r(1) \cdots r(N)]$, and $r(n)$ is the autocorrelation sequence of $x(n)$, $h^T = [h(0)\, h(1) \cdots h(N)]$ are the filter coefficients, and P_l are matrices with $P_l(i, j) = 1$ for $i - j = Ml$ and zero otherwise. Note that $P_0 = I$, where I is the $N \times N$ identity matrix. Since the matrices P_l, $l \neq 0$ are singular, the above quadratically constrained quadratic optimization problem is nonconvex and is very hard to solve both theoretically and numerically due to the existence of local minima. Several authors have used the classical method of Lagrange multipliers, which leads to an iterative augmented Lagrangian algorithm (see, for example, [29] and [30] for $M = 2$ and [2] for arbitrary M).

3) Optimizing the Product Filter: Instead of directly optimizing the coefficients $h(n)$, the idea is to find the optimum product filter $F_{opt}(z) = H_{opt}(z)H_{opt}(z^{-1})$, and then obtain $H_{opt}(z)$ from $F_{opt}(z)$ by spectral factorization. This approach was first introduced by Vaidyanathan *et al.* [6] as part of the design of an M-channel FIR orthonormal filter bank. To obtain a

TABLE I
QUALITATIVE COMPARISON BETWEEN THE DIFFERENT FIR DESIGN METHODS

authors	Approach	M	objective function	problem type	solution	Nyquist constraint	Positivity constraint	spectral fact.
[29], [30]	QCQP	2	compaction filter	non linear non convex	local optimum	Yes	No	No
[2]	QCQP	arbitrary	ideal low pass filter	non linear non convex	local optimum	Yes	No	No
[26]-[28]	FIR Lattice	2	compaction filter	non linear non convex	local optimum	No	No	No
[6]	eigen filter [a]	arbitrary	ideal low pass filter	iterative [b]	suboptimum	Yes	No	Yes
[33]	window method [a]	arbitrary	compaction filter	iterative [b]	suboptimum	Yes	Yes	Yes
[31],[32],[45]	product filter	2	compaction filter	linear convex	global optimum [c]	Yes	Yes	Yes
[9]	analytical	2	ideal low pass filter	Non iterative	global optimum	Yes	No	Yes
[33]	analytical	2	compaction filter	non iterative	global optimum [d]	Yes	No	Yes
New method	state space	arbitrary	compaction filter	non linear convex	global optimum	Yes	Yes	No

[a]The eigen filter and window methods are special cases of the product filter approach where the product filter is assumed to be a cascade of two filters
[b]Since the product filter is assumed to be a cascade of two filters, the optimization procedure alternate between the two in an iterative manner
[c]Over only the defined discrete set of frequencies
[d]Only for a certain class of random processes

low pass subband filter with a sharp frequency response, the authors used the eigen filter method [6] which is not guaranteed to converge to the global optimum. Moulin *et al.* [31] considered the design of FIR energy compaction filters and observed that the problem reduces to a linear semi-infinite (SIP) program. The authors solve a "discretized" version of the SIP using standard linear programming methods. Other discretization methods can be found in [32], [33]. The main drawback with any discretization approach is that global optimality is not guaranteed. We emphasize again that in the product filter approach, a spectral factorization step is required to obtain $H_{opt}(z)$.

4) Analytical Methods: The goal in this case is to derive an analytical procedure to obtain $F_{opt}(z)$. The elegance of this approach lies in the fact that no iterative numerical optimization is involved. For $M=2$ and $W(e^{j\omega}) = rect(\omega/\omega_c)$ (ideal lowpass filter with cutoff frequency ω_c), Aas *et al.* [9] were able to identify the unit-circle zeros of $F_{opt}(z)$. Once these are known, the other zeros can be found using Gaussian quadrature theory. Kirac and Vaidyanathan [33] extend the results of [9] for $M=2$ and $W(e^{j\omega}) = S_{xx}(e^{j\omega})$, where $S_{xx}(e^{j\omega})$ is the power spectrum of $x(n)$. Unfortunately, the method works only for a certain class of WSS random processes. Note that in both cases, a spectral factorization step is still necessary at the end.

Table I provides a qualitative comparison between some previous work and the newly proposed method.

III. STATE-SPACE APPROACH

From (1) and (2), we can immediately observe that the optimum solution, if it exists, is only a function of $|H(e^{j\omega})|^2$. By denoting **the product filter** $H(z)H(z^{-1})$ as $F(z)$, the output variance σ_y^2 in (1) can be rewritten as

$$\sigma_y^2 = r(0) + 2\sum_{n=1}^{N} f(n)r(n) \tag{8}$$

and the constraint (2) becomes

$$f(Mn) = \delta(n) \tag{9}$$

$$F(e^{j\omega}) = 1 + 2\sum_{n=1}^{N} f(n)\cos(\omega n) \geq 0 \quad \forall\,\omega \tag{10}$$

where $r(i)$ denotes the ith autocorrelation coefficient of the input $x(n)$. The problem is now linear in the real optimization variables $f(n),\ n \geq 1$ at the expense of an additional constraint, namely, (10), which we will refer to as the positivity constraint. The positivity constraint has to be satisfied at each frequency ω and is therefore equivalent to an infinite number of inequality constraints. The above formulation has a finite number of variables and an infinite number of constraints, hence, the name semi-infinite programming (SIP). The semi-infinite program can be *approximated* by *sampling* or *discretizing* the continuous frequency axis. We choose a finite set of discrete frequencies $\{\omega_i,\, 0 \le i \le L\}$ that are often uniformly spaced and enforce the positivity constraint only at those frequencies. This approach was first suggested and analyzed in depth by Moulin *et al.* [31]. The authors solve the "discretized" version of the SIP using standard linear programming methods. Other discretization methods were proposed by Pesquet and Combettes [32], who use a projection onto convex sets (POCS) type of algorithm, and Kirac and Vaidyanathan [33], who use a fast algorithm called the window method. The main problem with the sampling approach is that we can no longer guarantee the positivity of $F_{opt}(z)$ between the discrete frequencies ω_i, *no matter how large L is*. This, in turn, can create an infeasible spectral factorization step. Indeed, the discretized version is an outer approximation of the original SIP problem; its feasible set includes the feasible set of the original SIP problem. There are, of course, several ways to get around this problem (see for example [1, pp. 219–220]), but the point is, no matter which method we choose, global optimality of the SIP described by (8)–(10) cannot be guaranteed. We show next, using the discrete-time KYP lemma, that the positivity constraint can be satisfied over all ω by adding $N(N+1)/2$ additional optimization variables to the N original variables $f(n)$.

A. Discrete-Time KYP Lemma

Since $F(z) = H(z)H(z^{-1})$, the product filter is a two sided symmetric sequence, and we can therefore write $F(z)$ as $D(z) + D(z^{-1})$, where $D(z)$ is a causal function, and $D(z^{-1})$ is an anticausal one. Clearly, $D(z)$ completely characterizes $F(z)$. It is therefore natural to wonder whether the positivity condition on $F(e^{j\omega})$ can be reformulated in terms of some other condition(s) on $D(e^{j\omega})$. The answer dates back to the work of Caratheodory and Schur [34]: $F(e^{j\omega}) \ge 0$ for all ω if, and only if, $D(z)$ is analytic in $|z| > 1$, and $D(z)$ is a discrete time positive real function. Moreover, Schur characterized all such functions in terms of the so-called Schur parameters (which are also known as the reflection coefficients). The results of Caratheodory and Schur, however, apply to functions that are not necessarily rational. Since $D(z)$ is *rational* and, furthermore *causal*, it has a state-space representation (A_d, B_d, C_d, D_d). The question then becomes: Can the positive real property, which is an *analytic* frequency domain constraint, be expressed in terms of *algebraic* conditions on the matrices (A_d, B_d, C_d, D_d)? The answer, for the continuous time case, is in the affirmative and is established by the famous KYP lemma. The discrete-time version was derived by Hitz and Anderson [35] and is also known as the **discrete time positive real lemma**. To state the lemma, we first start with the definition of discrete-time *rational* positive real functions.

Definition 1—Discrete-Time Positive Real Functions: A square transfer matrix (function) $D(z)$ whose elements are real rational functions analytic in $|z| > 1$ is discrete-time positive real if, and only if, it satisfies all the following conditions:

$$\text{poles of } D(z) \text{ on } |z| = 1 \text{ are simple} \tag{11}$$

and

$$D(e^{j\omega}) + D(e^{-j\omega}) \ge 0 \quad \forall\, \omega \text{ at which } D(e^{j\omega}) \text{ exists} \tag{12}$$

and furthermore, if $z_0 = e^{j\omega_0}$, ω_0 real is a pole of $D(z)$ and if K is the residue matrix of $D(z)$ at $z = z_0$, the matrix $S = e^{-j\omega_0}K$ is Hermitian positive semi-definite.

Assume now that $D(z)$ has the following state space realization:

$$\begin{aligned} x(n+1) &= A_d x(n) + B_d u(n) \\ y(n) &= C_d x(n) + D_d u(n) \end{aligned} \tag{13}$$

where

A_d $N \times N$;

B_d $N \times M$;

C_d $M \times N$;

D_d $M \times M$.

For our case, $M = 1$. Then, the following lemma can be established.

Fact 1—Discrete Time KYP Lemma [35]: Let $D(z)$ be a square transfer matrix (function) with real rational elements that is analytic in $|z| > 1$ with only simple poles on $|z| = 1$. Let (A_d, B_d, C_d, D_d) be a minimal realization of $D(z)$. Then, $D(z)$ is discrete time positive real if, and only if, there exist a real symmetric positive definite matrix P_d and real matrices W_d and L_d such that

$$P_d - A_d^T P_d A_d = L_d^T L_d \tag{14}$$

$$C_d^T - A_d^T P_d B_d = L_d^T W_d \tag{15}$$

$$D_d + D_d^T - B_d^T P_d B_d = W_d^T W_d. \tag{16}$$

The above equalities (14)–(16) can be rewritten as the following "linear" matrix inequality (LMI)

$$\begin{aligned} \mathcal{M}_d &= \begin{bmatrix} P_d - A_d^T P_d A_d & C_d^T - A_d^T P_d B_d \\ C_d - B_d^T P_d A_d & D_d + D_d^T - B_d^T P_d B_d \end{bmatrix} \\ &= \begin{bmatrix} L_d^T \\ W_d^T \end{bmatrix} [L_d \; W_d] \succeq 0 \end{aligned} \tag{17}$$

and therefore represent an equivalent condition for the positivity constraint to be satisfied. The symbols $\succ$ and $\succeq$ are *generalized inequalities*, which are defined as follows: $P \succeq 0$ if, and only if, P is positive semi-definite. Similarly, $P \succ 0$ if, and only if, P is positive definite. As usual with the product filter formulation, the major difficulty at this point is to deal simultaneously with the positivity and Nyquist constraints. It turns out that, in this case, the Nyquist constraint can be imposed as an equality constraint in a simple manner. To see this, assume that $D(z)$ is implemented in a direct-form structure with the following state-space representation:

$$\begin{aligned} A_d &= \begin{bmatrix} \mathbf{0} & I \\ 0 & \mathbf{0}^T \end{bmatrix}, \quad B_d = [0\; 0 \cdots\; 1]^T \\ C_d &= [f(N)\; \cdots\; f(1)], \quad D_d = \tfrac{1}{2} \end{aligned} \tag{18}$$

where $\mathbf{0}$ is the $(N-1) \times 1$ zero vector, and I is the $(N-1) \times (N-1)$ identity matrix. Clearly, this state-space realization is minimal since the number of delay elements is equal to the degree of $D(z)$. Then, the Nyquist constraint can be written as a linear equality constraint:

$$Q\, C_d^T = \mathbf{0} \tag{19}$$

where $\mathbf{0}$ is the $N \times 1$ zero vector, and Q is a diagonal matrix with diagonal elements $\in \{0, 1\}$ The positions of the unity elements are determined by M and N. For example, for $N = 5$ and $M = 2$, the diagonal elements are $\{0\,1\,0\,1\,0\}$. Summarizing, we can represent the positivity constraint as an LMI whose entries are affine functions of the variables (P_d and C_d), and the Nyquist constraint as an equality constraint on C_d. The compaction problem described by (8)–(10) can be rewritten as follows:

$$\max_{C_d} C_d R^T \tag{20}$$

where $R = [r(N) \cdots r(1)]$ and finds a symmetric positive definite matrix $P_d = P_d^T \succ 0$ such that

$$\begin{bmatrix} P_d - A_d^T P_d A_d & C_d^T - A_d^T P_d B_d \\ C_d - B_d^T P_d A_d & D_d + D_d^T - B_d^T P_d B_d \end{bmatrix} \succeq 0, \ QC_d^T = \mathbf{0}. \tag{21}$$

This new formulation is therefore a maximization problem in the variable vector C_d and a feasibility problem in the matrix P_d and can be solved using *semidefinite programming* (SDP). For more details on SDP, see the excellent survey paper by Vandenberghe and Boyd [36]. We would like to mention at this point that independent work in [37], which came to our attention after the submission of this paper, briefly uses the positive real lemma in a standard FIR filter design application. Nevertheless, the work in [37] does not address the multirate case nor the spectral factor analysis presented in this paper. More important, however, a discretization step is still necessary in [37] *even after* the use of the positive real lemma, which, in turn, sacrifices the global optimality of the resulting filter.

To summarize, the FIR energy compaction problem, which is expressed in terms of the coefficients of the filter $H(z)$, is a nonlinear *nonconvex* optimization problem. The product filter formulation is a semi-infinite, linear, and convex problem. The discretized version of the SIP is linear, convex, and can be solved using standard linear programming problem but is an *approximation* of the original problem. The state-space approach proposed in this paper is nonlinear, convex, and *semi-definite*. Using the rationality of $F(z)$, the infinite set of inequality constraints are replaced by a (finite-dimensional) positive semi-definite constraint (17) with the auxiliary variable P_d, permitting a globally optimal solution to be found. In principle, the problem as stated above can be solved. Specifically, we can write a SDP that returns a global optimum vector $C_{d_{opt}}$ and a feasible matrix P_d that will meet the constraints (21) and maximize the objective function (20). We can then spectrally factorize $F_{opt}(z)$ to obtain $H_{opt}(z)$ using any of the well-known algorithms (see, for example, [1, pp. 854–856]). It turns out, however, that this additional spectral factorization step can be completely avoided if the minimum-phase spectral factor is desired. Indeed, we show in the next section that the state-space representation of the *minimum-phase spectral factor*, $H_{\min}(z)$ can be expressed in terms of the matrices $(A_d, B_d, C_{d_{opt}}, D_d)$ and a particular P_d, namely, *the minimum element* $P_{d_{\min}}$ of the convex cone of positive definite matrices satisfying (21). Using this result, we then modify the objective function (20) in order for the program to return, along with a globally optimal vector $C_{d_{opt}}$, the specific matrix $P_{d_{\min}}$. Once $C_{d_{opt}}$ and $P_{d_{\min}}$ are found, $H_{\min}(z)$ is readily obtained, and the spectral factorization procedure is eliminated. It is important to keep in mind that although $H_{\min}(z)$ is unique (as we will show next), $F_{opt}(z)$ is **not** guaranteed to be so. The characterization of the optimal set of solutions of an SDP is an interesting and relevant issue but, due to space limitations, is outside the scope of this paper. Related material can be found at http://www.systems.caltech.edu/tuqan.

IV. Minimum-Phase Spectral Factor

We first derive an expression for a spectral factor.

Theorem 1: Assume that $D(z)$ satisfies the discrete time KYP lemma with a minimal realization (A_d, B_d, C_d, D_d). Then, a transfer function $H(z)$ in the form

$$H(z) = W_d + L_d(zI - A_d)^{-1} B_d \tag{22}$$

is a spectral factor of $F(z) \triangleq D(z) + D(z^{-1}) = H(z)H(z^{-1})$.

Proof: The proof is given in Appendix A. ■

The above theorem is the discrete-time counterpart of the continuous-time result found in [38, pp. 220–221]. The theorem indicates that if $D(z)$ satisfies the discrete-time KYP lemma, a spectral factor always exists and can be expressed in the form (22). It is important to note that in Theorem 1, the number of columns of W_d and the number of rows of L_d are unrestricted, where the dimensions of P_d and the other dimensions of L_d and W_d are automatically fixed. For example, in the single-input single-output (SISO) case, W_d can be a scalar or a row vector. The remainder of this section is dedicated to the study of the SISO minimum-phase spectral factor $H_{\min}(z)$. The motivation for such a study was given at the end of the last section. We first establish that the SISO minimum-phase spectral factor $H_{\min}(z)$ can be expressed in the form (22) with W_d being a scalar and L_d a row vector. We then present a characterization of $H_{\min}(z)$ in terms of the matrices $(A_d, B_d, C_{d_{opt}}, D_d)$ and the minimum element $P_{d_{\min}}$. The development of these results follows by applying the bilinear transformation $s = (z-1)/(z+1)$ to the continuous-time minimum-phase spectral factor and then by using some deep results proved for the continuous-time case by Willems [39] and Anderson [38], [40]. Unlike, however, the work in [38]–[40], the discussions and proofs presented here apply only to the scalar case, which is sufficient for the purpose of this paper. We now introduce some well-established facts.

Fact 2—KYP Lemma [38]: Let $D(s)$ be a square transfer matrix (function) with real rational elements that is analytic in Re$s > 0$ with only simple poles on Re$s = 0$. Let (A_c, B_c, C_c, D_c) be a minimal realization of $D(s)$. Then, $D(s)$ is positive real if, and only if, there exist a real symmetric positive definite matrix P_c and real matrices W_c and L_c such that

$$-A_c^T P_c - P_c A_c = L_c^T L_c \tag{23}$$

$$C_c^T - P_c B_c = L_c^T W_c \tag{24}$$

$$D_c + D_c^T = W_c^T W_c. \tag{25}$$

As in the discrete-time case, an equivalent condition for the above equalities is the following matrix inequality:

$$\mathcal{M}_c = \begin{bmatrix} -P_cA_c - A_c^TP_c & C_c^T - P_cB_c \\ C_c - B_c^TP_c & D_c + D_c^T \end{bmatrix} \succeq 0. \tag{26}$$

The definition of positive real functions for the continuous-time case can be found in [38, pp. 51–54].

Fact 3—Continuous-Time Minimum-Phase Spectral Factor [39], [40]: Let (A_c, B_c, C_c, D_c) be a minimal realization of a positive real transfer matrix $D(s)$. Then, the set of symmetric positive definite matrices $\{P_c = P_c^T \succ 0\}$ satisfying the LMI constraint (26) has a minimum element $P_{c_{\min}}$ (see the definition below). This minimum element is associated with a minimum-phase continuous-time spectral factor $H_{\min}(s)$, which is expressed as

$$H_{\min}(s) = W_c + L_c(sI - A_c)^{-1}B_c \tag{27}$$

where L_c and W_c satisfy equations (23)–(25) with $P_c = P_{c_{\min}}$.

Using the above facts and for the special SISO case, it immediately follows that the continuous-time minimum-phase spectral factor is unique, stable, and has no zeros in the right half plane (Re $s > 0$). Furthermore, if $H_{\min}(s) = W_c + L_c(sI - A_c)^{-1}B_c$, then W_c is a 1×1 scalar, and L_c is a $1 \times N$ row vector. Finally, all the eigenvalues of A_c have Re $\lambda_i < 0$. The following result can be then established.

Corollary 1: Assume that the continuous-time minimum-phase spectral factor $H_{\min}(s)$ is given in the form (27). Then, by applying the bilinear transformation $s = (z-1)/(z+1)$, $H_{\min}(s)$ maps to the *unique* discrete-time minimum-phase spectral factor $H_{\min}(z)$, which can be expressed in the form (22), with

$$\begin{aligned} A_d &= (I - A_c)^{-1}(I + A_c), \quad B_d = 2(I - A_c)^{-2}B_c, \\ L_d &= L_c, \quad W_d = W_c + L_c(I - A_c)^{-1}B_c. \end{aligned} \tag{28}$$

Furthermore, if (A_c, B_c, C_c, D_c) is a minimal realization, then (A_d, B_d, C_d, D_d) is also a minimal realization.

Proof: $H_{\min}(s)$ maps to $H_{\min}(z)$ is a consequence of the s-plane to z-plane mapping property of the bilinear transformation. The uniqueness of $H_{\min}(z)$ follows from the uniqueness of $H_{\min}(s)$. The proof of the other statements is given in Appendixes B and C, respectively. ■

Note that $(I - A_c)$ must be nonsingular. Otherwise, one of the eigenvalues of A_c is equal to one that contradicts the stability of $H_{\min}(s)$. Before stating the main theorem of this section, the following definitions are required.

Definition 2—Convex Cone: A set C is called a cone if for every $x \in C$ and scalar $\lambda \geq 0$, $\lambda x \in C$. A cone is convex if for $\lambda_1, \lambda_2 \geq 0$ and $x_1, x_2 \in C$, $\lambda_1 x_1 + \lambda_2 x_2 \in C$. The set of symmetric positive semi-definite matrices $\{P | P = P^T, P \succeq 0\}$ is a convex cone.

Definition 3—Partial Order: The convex cone of symmetric positive semi-definite matrices $K = \{P | P = P^T, P \succeq 0\}$ defines a partial order on the space of symmetric matrices in the following sense: $P_2 \succeq P_1$ if, and only if, $P_2 - P_1$ is positive semi-definite.

Definition 4—Minimum Element: We say that $P_{\min} \in S$ is a minimum element of S with respect to the generalized inequality $\preceq$ if for every $P \in S$ we have $P_{\min} \preceq P$. If a set has a minimum element, this element is *unique*.

Definition 5—Congruence: An $N \times N$ real matrix A is said to be congruent to B if there exits a nonsingular real matrix T such that $B = TAT^T$. The following property of congruence with respect to positive semi-definite matrices can be easily proved:

The partial order induced by the positive semi-definite cone is invariant under congruence, i.e.,

$$P_1 \preceq P_2 \Longrightarrow TP_1T^T \preceq TP_2T^T. \tag{29}$$

Assuming that T is *nonsingular*, a similar relation holds for the positive definite case with $\prec$ replacing $\preceq$. Note that by taking $P_1 = 0$, it follows that the cone of positive semi-definite matrices is invariant under a congruence transformation.

Theorem 2—Discrete-Time SISO Minimum-Phase Spectral Factor: Let $F(z) = D(z) + D(z^{-1})$ be a real rational function that is analytic in $|z| > 1$. Assume that $D(z)$ satisfies the discrete-time positive real lemma with a minimal realization (A_d, B_d, C_d, D_d). Then, the minimum-phase spectral factor $H_{\min}(z)$ can be expressed in the form $H_{\min}(z) = W_d + L_d(zI - A_d)^{-1}B_d$ with

$$W_d = (D_d + D_d^T - B_d^TP_{d_{\min}}B_d)^{1/2} \tag{30}$$

$$L_d = (D_d + D_d^T - B_d^TP_{d_{\min}}B_d)^{-1/2}(C_d - B_d^TP_{d_{\min}}A_d) \tag{31}$$

and $P_{d_{\min}}$ is the minimum element in the convex set of symmetric positive definite matrices satisfying (21).

Proof: The fact that the minimum-phase spectral factor has the form (22) has been established in Corollary 1. Equations (30) and (31) are obtained from (15) and (16) by recalling that for the SISO minimum-phase spectral factor, W_d is a scalar, and L_d is a column vector. The proof that the LMI and Nyquist constraints are satisfied with $P_d = P_{d_{\min}}$ for the case of $H_{\min}(z)$ is established through the following series of steps.

1) The Nyquist constraint (19) can be incorporated in the LMI by replacing C_d in (17) with $C_{d_{new}} = C_d(I - Q)$, where I is the $N \times N$ identity matrix. In the remainder of the proof, we will therefore only consider the LMI constraint, keeping the above substitution in mind.
2) The set $\{P_c = P_c^T \succ 0\}$ generating the cone of positive semi-definite matrices, $\{\mathcal{M}'_c | \mathcal{M}'_c = \mathcal{M}'^T_c, \mathcal{M}'_c \succeq 0\}$ defined by

$$\mathcal{M}'_c = \begin{bmatrix} -P_cA_c - A_c^TP_c & C_{c_{new}}^T - P_cB_c \\ C_{c_{new}} - B_c^TP_c & D_c + D_c^T \end{bmatrix} \succeq 0 \tag{32}$$

is the same set of symmetric positive definite matrices satisfying the following matrix inequality:

$$\begin{bmatrix} I & \mathbf{0} \\ B_d^T(I + A_d^T)^{-1} & 1 \end{bmatrix} \begin{bmatrix} -P_cA_c - A_c^TP_c & C_{c_{new}}^T - P_cB_c \\ C_{c_{new}} - B_c^TP_c & D_c + D_c^T \end{bmatrix} \cdot \begin{bmatrix} I & (I + A_d)^{-1}B_d \\ \mathbf{0}^T & 1 \end{bmatrix} \succeq 0 \tag{33}$$

where I is the $N \times N$ identity matrix, $\mathbf{0}$ is the $N \times 1$ zero vector, and $C_{c_{new}} = C_c(I - Q)$. To see this, observe that the left-hand side of (33) is congruent to $\mathcal{M}'_c$. Since the cone of symmetric positive semi-definite matrices is invariant under a congruence transformation and since the congruence transformation is independent of P_c, the claim follows automatically.

3) By multiplying the three matrices in (33) and performing the following substitutions:

$$A_c = (A_d + I)^{-1}(A_d - I), \quad B_c = 2(A_d + I)^{-2} B_d$$
$$C_{c_{new}} = C_{d_{new}}, \quad D_c = D_d - C_{d_{new}}(A_d + I)^{-1} B_d \qquad (34)$$

it can then be shown (see Appendix D) that these operations produce the LMI (17) with

$$P_d = 2(A_d^T + I)^{-1} P_c (A_d + I)^{-1}. \qquad (35)$$

Equation (35) describes another nonsingular congruence transformation applied this time on the set $\{P_c = P_c^T \succ 0\}$. The congruence transformation preserves the positive definiteness of the matrices as well as the partial order induced on the set.

4) Using Fact 3 (with $C_c \equiv C_{c_{new}}$) and steps 1) – 3) described above, we have therefore proven that the set $\{P_d = P_d^T \succ 0\}$ satisfying the constraints (21) has a minimum element $P_{d_{\min}}$ and that minimum element is given by $P_{d_{\min}} = 2(A_d^T + I)^{-1} P_{c_{\min}} (A_d + I)^{-1}$. It now remains to show that $P_{d_{\min}}$ is the solution associated with $H_{\min}(z)$. This can be done by starting with (23)–(25) with $P_c = P_{c_{\min}}$, applying the bilinear transformation on $H_{\min}(s)$, which produces the following relations:

$$A_c = (A_d + I)^{-1}(A_d - I), \quad B_c = 2(A_d + I)^{-2} B_d$$
$$L_c = L_d, \quad W_c = W_d - L_d (A_d + I)^{-1} B_d \qquad (36)$$

making the additional substitutions

$$C_{c_{new}} = C_{d_{new}}, \quad D_c = D_d - C_{d_{new}}(A_d + I)^{-1} B_d$$
$$P_{d_{\min}} = 2(A_d^T + I)^{-1} P_{c_{\min}} (A_d + I)^{-1}$$

and simplifying to obtain (14)–(16). The conclusion that these final equations are associated with $H_{\min}(z)$ follows from Corollary 1. The exact derivations of the above steps are algebraic in nature and very similar to the proofs found in Appendix D and are therefore omitted. ■

An alternative characterization of $P_{d_{\min}}$ is given by the following theorem.

Theorem 3: Assume that $D_d^T + D_d - B_d^T P_d B_d \neq 0$. Then, the minimum element $P_{d_{\min}}$ in the convex set of symmetric positive definite matrices satisfying the constraints (21) is the *unique* solution to the following *algebraic Riccati equations* (AREs):

$$\begin{aligned} P_d = & A_d^T P_d A_d + (C_{d_{new}}^T - A_d^T P_d B_d) \\ & \cdot (D_d + D_d^T - B_d^T P_d B_d)^{-1} (C_{d_{new}}^T - A_d^T P_d B_d)^T \end{aligned} \qquad (37)$$

$$\begin{aligned} P_d = & A_1^T P_d A_1 + A_1^T P_d B_d (R - B_d^T P_d B_d)^{-1} \\ & \cdot B_d^T P_d A_1 + C_{d_{new}}^T R^{-1} C_{d_{new}} \quad \text{where} \\ A_1 = & A_d - B_d R^{-1} C_{d_{new}}, \quad R = D_d + D_d^T \succ 0 \end{aligned} \qquad (38)$$

where $C_{d_{new}} = C_d (I - Q)$.

Proof: Equation (37) follows by substituting (31) in (14). Equation (38) is derived from (37), assuming that R is positive definite, and the proof can be found in Appendix E. ■

Corollary 2: $P_{d_{\min}}$ can be obtained from $P_{c_{\min}}$ using the congruence relation (35) and the fact $P_{c_{\min}}$ is the unique solution to the following equations:

$$\begin{aligned} & -A_c^T P_c - P_c A_c \\ & \quad = (C_c^T - P_c B_c)(D_c + D_c)^{-1} (C_c^T - P_c B_c)^T \end{aligned} \qquad (39)$$

$$\begin{aligned} & -A_2^T P_c - P_c A_2 \\ & \quad = P_c B_c R^{-1} B_c^T P_c + C_c^T R^{-1} C_c \quad \text{where} \\ & \qquad A_2 = A_c - B_c R^{-1} C_c, \quad R = D_c + D_c^T \end{aligned} \qquad (40)$$

and (A_c, B_c, C_c, D_c) are given by (34). The proof that $P_{c_{\min}}$ is the unique solution to (39) and (40) can be found in [38].

V. Some Simplifications for the SISO FIR Case

Assume that the positive real function $D(z)$ has the following minimal state-space realization:

$$A_d = \begin{bmatrix} \mathbf{0} & I \\ 0 & \mathbf{0}^T \end{bmatrix}, \quad B_d = [0\,0 \cdots 1]^T$$
$$C_d = [\, f(N) \cdots f(1)], \quad D_d = \tfrac{1}{2} \qquad (41)$$

where $\mathbf{0}$ is the $(N-1) \times 1$ zero vector. The minimum-phase spectral factor $H_{\min}(z)$ is then given by

$$\begin{aligned} & H_{\min}(z) \\ & \quad = \sqrt{1 - p_{d_{\min}}(N, N)} + \frac{(C_d - B_d^T P_{d_{\min}} A_d)}{\sqrt{1 - p_{d_{\min}}(N, N)}} \\ & \qquad \cdot (zI - A_d)^{-1} B_d \\ & \quad = \sqrt{1 - p_{d_{\min}}(N, N)} + \frac{(C_d - B_d^T P_{d_{\min}} A_d)}{\sqrt{1 - p_{d_{\min}}(N, N)}} \\ & \qquad \cdot \left[z^{-N} z^{-(N-1)} \cdots z^{-1} \right]^T \\ & \quad = \frac{1}{\sqrt{1 - p_{d_{\min}}(N, N)}} \{ 1 - p_{d_{\min}}(N, N) \\ & \qquad + (f(1) - p_{d_{\min}}(N, N-1)) z^{-1} \\ & \qquad + \cdots + (f(M) - p_{d_{\min}}(N, N-M)) z^{-M} \\ & \qquad + \cdots + f(N) z^{-N} \}. \end{aligned} \qquad (42)$$

The second equality follows by analogy with the transfer function of $D(z)$, and the third equality is obtained by direct substitution of (41). It is interesting to note that among all the elements of $P_{d_{\min}}$, only the last row affects the coefficients of $H_{\min}(z)$. Closed-form expressions for the continuous-time system A_c, B_c, D_c can be also derived and are given by

$$A_c = \begin{bmatrix} 1 & -1 & \cdots & (-1)^{N-1} \\ 0 & 1 & \cdots & (-1)^{N-2} \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & \cdots & 0 \\ 0 & -1 & \cdots & \vdots \\ \vdots & 0 & \ddots & 1 \\ 0 & 0 & \cdots & -1 \end{bmatrix} \qquad (43)$$

$$B_c = 2\begin{bmatrix} 1 & -1 & \cdots & (-1)^{N-1} \\ 0 & 1 & \cdots & (-1)^{N-2} \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}\begin{bmatrix} (-1)^{N-1} \\ (-1)^{N-2} \\ \vdots \\ 1 \end{bmatrix} = 2\begin{bmatrix} N(-1)^{N-1} \\ (N-1)(-1)^{N-2} \\ \vdots \\ 1 \end{bmatrix} \tag{44}$$

$$D_c = D(z)|_{z=-1} = 1/2 - f(1) + \cdots + (-1)^N f(N)$$
$$D_c + D_c^T = F(z)|_{z=-1} \tag{45}$$

The above follows by noticing that $A_d + I$ is an $N \times N$ Jordan block. It is easy to show that its inverse is equal to an upper triangular Toeplitz matrix with first row $[1\ -1\, 1 \cdots (-1)^{N-1}]$. The knowledge of the form of equations (43)–(45) is useful in order to avoid computing inverses during the optimization process (if the continuous-time characterization of Corollary 2 is to be used).

Corollary 3: For the special SISO FIR case under consideration, the minimum element $P_{d_{\min}}$ has the following form:

$$P_{d_{\min}} = \sum_{k=0}^{N-1} A_d^{T^k} L_d^T L_d A_d^k = \mathcal{O}_{L_d, A_d}^T \mathcal{O}_{L_d, A_d} \tag{46}$$

where $\mathcal{O}_{L_d, A_d}$ denotes the observability matrix of the realization $\{A_d, L_d^T\}$. The above result follows from the fact that $P_{d_{\min}}$ has to satisfy a discrete-time Lyapunov equation (14). The solution of a discrete time Lyapunov equation can be found in [1, pp. 684–685] and can be further simplified for this case using the fact that $A_d^N = 0$ to obtain (46). It also follows that $\mathrm{Tr}(P_{d_{\min}}) = \sum_{n=1}^N n l_d^2(n)$ and that (16) is the unit energy constraint enforced on the optimum filter.

Corollary 4: For the special SISO FIR case under consideration, the discrete-time positive real lemma is equivalent to the following condition: There exists an $(N+1) \times (N+1)$ matrix $P_1 = P_1^T \succeq 0$ such that

$$\mathrm{Tr}(M_k P_1 M_k^T) = \begin{cases} D_d + D_d^T, & \text{if } k = N+1 \\ C_d(k), & \text{otherwise} \end{cases} \tag{47}$$

with $M_k = [\begin{smallmatrix} \mathbf{0} & I_k \\ \boldsymbol{O} & \mathbf{0}^T \end{smallmatrix}]$, where

- $\mathbf{0}$ $1 \times (N+1-k)$ zero vector;
- I_k $k \times k$ identity matrix;
- $\boldsymbol{O}$ $(N+1-k) \times (N+1-k)$ zero matrix;
- $C_d(k)$ kth element of the vector C_d.

The above conditions mean, in particular, that the trace of $P_1 = 1$ and that the sum of the elements of each lower (or upper) diagonal of P_1 is equal to a coefficient $f(k)$. The result follows by substituting (41) into (17), which simplifies to the following form:

$$P_1 = \begin{bmatrix} P_d & \mathbf{0}^T \\ \mathbf{0} & 0 \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & P_d \end{bmatrix} + \begin{bmatrix} \boldsymbol{O} & C_d^T \\ C_d & D_d + D_d^T \end{bmatrix} \succeq 0 \tag{48}$$

where

- 0 scalar;
- $\mathbf{0}$ $1 \times N$ zero vector;
- $\boldsymbol{O}$ $N \times N$ zero matrix.

Since $D_d + D_d^T = 1$, $\mathrm{Tr}(P_1) = 1$. The other conditions are obtained by direct evaluation. The above corollary produces a more compact form for the positivity constraint and can be used to increase the *computational efficiency* of the SDP program [47]. Moreover, Corollary 4 can be used to generate new "theorems" and "forms" for the spectral factorization of polynomials [47]. The next example, while easily handled using elementary methods, serves to demonstrate the main points of the previous discussions.

Example—2 × 2 KLT: Assume that $N = 1$ and that $M = 2$. The state-space representation for $D(z)$ in this case is $A_d = 0$, $B_d = 1$, $C_d = f(1)$, and $D_d = 1/2$. Using (37) and this particular state-space realization, the optimization problem can be simplified and recast as follows. Maximize $f(1)R(1)$ subject to the equality constraint $\sqrt{P_{d_{\min}}(1-P_{d_{\min}})} = f(1)$, where $0 < P_{d_{\min}} < 1$. Note that by using (38) instead of (37) with $A_1 = -f(1)$ and $R = 1$, the same formulation is obtained. The problem can be reexpressed as an "unconstrained" problem in the variable $P_{d_{\min}}$, namely, maximize $\sqrt{P_{d_{\min}}(1-P_{d_{\min}})}R(1)$, where $0 < P_{d_{\min}} < 1$. Using the AM-GM inequality, the convex objective function is upper bounded by $R(1)/2$, which is independent of $f(1)$. The bound is achieved if, and only if, $1 - P_{d_{\min}} = P_{d_{\min}}$, i.e., $P_{d_{\min}} = 1/2$. From $\sqrt{P_{d_{\min}}(1-P_{d_{\min}})} = f(1)$, it then follows that $f(1) = 1/2$. Using (22), (30), and (31) with the above state-space representation, the minimum phase spectral factor has the form $H_{\min}(z) = (1-P_{d_{\min}})^{1/2} + f(1)(1-P_{d_{\min}})^{-1/2}z^{-1}$. By substituting $P_{d_{\min}} = f(1) = 1/2$, we get $H_{\min}(z) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}z^{-1}$, which corresponds to the first row of the 2×2 universal KLT. We also note that $H_{\min}(z)$ could have been obtained from Corollary 3 by using the two equations $1/2 = f(1) = l^2$ and $1 - l^2 = w^2$, where l and w are the filter coefficients of $H_{\min}(z)$. Neither the product filter nor the spectral factor coefficients depend on the value of $R(1)$. The compaction gain is, however, equal to $1 + |R(1)|/R(0)$. To check Corollary 2, note that $A_c = -1$, $B_c = 2$, $C_c = f(1)$, $D_c = 1/2 - f(1)$ and with $A_2 = -1/(1-2f(1))$, $R = 1 - 2f(1)$, (39) and (40) reduce to $\sqrt{2P_{c_{\min}}(1-2P_{c_{\min}})} = f(1)$, where $0 < P_{c_{\min}} < 1$. The problem can be put in the following form. Maximize $\sqrt{2P_{c_{\min}}(1-2P_{c_{\min}})}R(1)$ and solve in the same way as the discrete-time case. The final result is $C_c = f(1) = C_d = 1/2$ and $P_{c_{\min}} = 1/4$. We also note that $P_{d_{\min}} = 2(A_d^T + I)^{-1}P_c(A_d + I)^{-1} = 1/2$. Finally, the continuous-time spectral factor $H_{\min}(s)$ is equal to $\sqrt{2}/(s+1)$. It can be easily verified that this is the result we obtain by applying the bilinear transformation $z^{-1} = (1-s)/(1+s)$ to $H_{\min}(z)$.

Although the above example uses conditions (37) and (38) (which are nonlinear *nonconvex*) and/or their continuous-time equivalents to solve the maximization problem, it is actually the LMIs $\mathcal{M}_d$ in (17) and $\mathcal{M}_c$ in (26) that come into play when using a SDP to solve the general (N, M) case, as we discuss next.

VI. Spectral Factor Formulation

The minimum-phase spectral factor is determined by $(A_d, B_d, C_{d_{opt}}, D_d)$ and $P_{d_{\min}}$. Since A_d, B_d, and D_d are fixed by the choice (18), and since $C_{d_{opt}}$ is determined by the program, we can also include P_d in the objective function (20) to obtain $P_{d_{\min}}$. Minimizing P_d directly will produce a vector-valued objective function. To avoid this situation, we can instead minimize a scalar-valued function of P_d, and this can be established by the following observation.

Observation 1: Assume that $P_{d_{\min}}$ is the minimum element in the convex set of symmetric positive definite matrices satisfying the LMI constraint (17). Then, $P_d = P_{d_{\min}}$ if, and only if, $\mathrm{Tr}(WP_d)$ is minimum for any diagonal positive definite matrix W.

Proof: The necessary part is obvious because $P_1 \succeq P_2$ implies that $\mathrm{Tr}(WP_1) \geq \mathrm{Tr}(WP_2)$. For the sufficiency part, we proceed as follows: Assume there exists a matrix P_2 and a minimum element P_1 such that $\mathrm{Tr}(WP_2) = \mathrm{Tr}(WP_1) = \min \mathrm{Tr}(WP_d)$ over all P_d. Note that by the definition of the minimum element, $P_2 - P_1 \succeq 0$. We will show that P_2 must be equal to P_1. From $\mathrm{Tr}(WP_2) = \mathrm{Tr}(WP_1)$, it follows that $\mathrm{Tr}(W(P_2 - P_1)) = 0$. Since W is a diagonal matrix with positive elements, then the diagonal elements of the positive semi-definite matrix $P_2 - P_1$ are equal to zero. Using the fact that the principal minors of a positive semi-definite matrix must be non-negative, it follows that $P_2 - P_1$ must be identically zero. ∎

The optimization problem formulated at the end of Section III now reduces to the following final form:

$$\max_{C_d,\, P_d} C_d R^T - \mathrm{Tr}(WP_d) \tag{49}$$

where $R^T = [r(N) \cdots r(1)]^T$, and W is a diagonal positive definite weight matrix such that

$$\mathcal{M}_d = \begin{bmatrix} P_d - A_d^T P_d A_d & C_d^T - A_d^T P_d B_d \\ C_d - B_d^T P_d A_d & D_d + D_d^T - B_d^T P_d B_d \end{bmatrix} \succeq 0$$
$$QC_d^T = \mathbf{0} \tag{50}$$

and is therefore a maximization problem in the variable vector C_d and a minimization problem in the matrix P_d. The particular choice of the trace function $\mathrm{Tr}(\cdot)$ was intentional in order to use SDP. The weight matrix W is included in the objective function because, unlike in Section III, we now have two separate and *competing* objectives, namely, $C_d R^T$ and $\mathrm{Tr}(WP_d)$. The idea is to choose the weight so that optimality of C_d is never compromised, i.e., in order to prohibit $\mathrm{Tr}(WP_d)$ from becoming the dominant factor in (49). Finally, note that the continuous-time characterization can also be used. In particular, with the continuous-time state-space realization described in (43) and (44) and, with $C_c = C_d$, the optimization problem becomes

$$\max_{C_c,\, P_c} C_c R^T - \mathrm{Tr}(WP_c) \tag{51}$$

where R^T and W are defined as before such that

$$\mathcal{M}_c = \begin{bmatrix} -A_c^T P_c - P_c A_c & C_c^T - P_c B_c \\ C_c - B_c^T P_c & D_d + D_d^T - 1/2(C_c B_c + B_c^T C_c^T) \end{bmatrix} \succeq 0$$
$$QC_c^T = \mathbf{0}. \tag{52}$$

The matrix $P_{d_{\min}}$ is then obtained from $P_{c_{\min}}$ using (35).

Observation 2: The multiobjective optimization problems described, respectively, by (49) and (50) and by (51) and (52) are convex programs with respect to the variables C_d and P_d and C_c and P_c.

Proof: Since the two problems are identical in form, we only provide a proof for the discrete-time formulation (49) and (50). The objective function (49) is linear in both C_d and P_d and is therefore a convex function. The constraint set defined by (50) is a convex set with respect to the optimization variable C_d since for all C_{d_1} and $C_{d_2} \in R^N$ and for all $0 \leq \lambda \leq 1$

$$\mathcal{M}_d(\lambda C_{d_1} + (1-\lambda) C_{d_2}) = \lambda \mathcal{M}_d(C_{d_1}) + (1-\lambda)\mathcal{M}_d(C_{d_2}).$$

The same argument holds for P_d. The equality constraint is linear in C_d and is therefore convex. ∎

It follows that any local solution to these programs is also a global one.

Initialization and Strict Feasibility: The LMI control toolbox and the software package in [41] require a strictly feasible primal or dual problem (the so-called Slater conditions) to converge. Indeed, this is a sufficient condition for the duality gap to be zero [36]. For the design of compaction filters, we can use the following *strictly feasible point* as an initial solution to the primal problem:

$$C_d = \mathbf{0}, \quad P_d = \frac{1}{2N} \sum_{k=0}^{N-1} A_d^{T^k} A_d^k. \tag{53}$$

By definition, $P_d = P_d^T \succ 0$. With the above choice, the LMI in (50) is diagonal and positive definite.

VII. Regularity Constraints

The regularity property is important in wavelets applications such as image coding, numerical analysis, and computer graphics, to name a few. An orthonormal wavelet scaling function is obtained by cascading L subband filters $H(z^{Mi})$, where $H(z)$ is an FIR filter with a Nyquist(M) magnitude squared response. For certain applications, it is important that the product $\prod_{i=0}^{L-1} H(z^{Mi})$ converges to a "smooth" function. The degree of smoothness or regularity is characterized by the number of zeros that $H(z)$ has at the aliasing frequencies $\omega = 2\pi m/M$ for $1 \leq m < M$. For $M = 2$, this amounts to forcing L zeros at $z = -1$ ($\omega = \pi$). The first of these zeros ($r = 0$) is simply obtained from $F(-1) = 0$ [because $F(e^{j\omega}) \geq 0 \,\forall\, \omega$, $F(z)$ will automatically have a double zero at π]. The second zero ($r = 1$) is obtained by differentiating $F(e^{j\omega})$ twice with respect to ω, evaluating the result at π, and setting it to zero. Repeating this procedure, we derive the following equation:

$$D_d\delta(r) - C_d \begin{bmatrix} N^{2r} & 0 & \cdots & 0 \\ 0 & (N-1)^{2r} & \cdots & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} \cdot (A_d + I)^{-1}B_d = 0, \qquad 0 \le r \le L-1. \quad (54)$$

For general M, the regularity condition can be expressed as the following linear constraint on the filter product coefficients:

$$D_d\delta(r) - C_d \begin{bmatrix} (N)^{2r}\cos\left(\frac{2\pi Nm}{M}\right) \\ (N-1)^{2r}\cos\left(\frac{2\pi(N-1)m}{M}\right) \\ \vdots \\ \cos\left(\frac{2\pi m}{M}\right) \end{bmatrix} = 0$$
$$0 \le r \le L-1,\ 1 \le m < M. \quad (55)$$

In the remainder of this section, we will discuss only the case of $M = 2$. Most of the results can be easily extended for $M > 2$. We next show, using the continuous-time formulation (52), that the LMI in (50) becomes singular when adding (54) to the SDP. We then derive a new formulation for which a strict feasible solution (primal or dual) always exists for the case of a single zero at π ($r = 0$).

Theorem 4—SDP Formulation with the Regularity Constraint: Assume that (54) is satisfied for any $r \ge 0$. Then, the LMI in (50) is always singular. Nevertheless, the primal problem defined by (49) and (50) can be re-expressed as follows:

$$\max_{P_d} \operatorname{Tr}(R^T(A_d^T + I)P_d(A_d + I)^{-1}B_d - WP_d) \quad (56)$$

where $R^T = [r(N)\cdots r(1)]^T$, and W is a diagonal positive definite weight matrix subject to the following constraints:

$$P_d - A_d^T P_d A_d \succeq 0$$
$$D_d = B_d^T(A_d^T + I)^{-1}(A_d^T + I)P_d(A_d + I)^{-1}B_d$$
$$Q(A_d^T + I)P_d(A_d + I)^{-1}B_d = \mathbf{0}$$
$$B_d^T(A_d^T + I)^{-1} \begin{bmatrix} N^{2r} & 0 & \cdots & 0 \\ 0 & (N-1)^{2r} & \cdots & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} \cdot (A_d^T + I)P_d(A_d + I)^{-1}B_d = 0, \qquad 1 \le r \le L-1. \quad (57)$$

Proof: The key idea is to observe that for $r = 0$, the LMI (52) in the continuous-time formulation reduces to the following form:

$$D_c + D_c^T = 0, \quad C_c^T - P_cB_c = 0$$
$$-A_c^TP_c - P_cA_c \succeq 0. \quad (58)$$

The above follows from (24), (25), and (45). By applying (34), (58) can be rewritten as follows:

$$D_d = B_d^T(A_d^T + I)^{-1}P_dB_d,$$
$$C_d = B_d^T(A_d^T + I)^{-1}P_d(A_d + I), \quad P_d - A_d^TP_dA_d \succeq 0. \quad (59)$$

Substituting (59) in the LMI (50) and simplifying, we get (60), shown at the bottom of the page. The above matrix is singular *for all* P_d because the last column of $\mathcal{M}_d$ is a linear combination of the previous N ones. To see this, observe that $\mathcal{M}_d(N+1 : N+1) = \mathcal{M}_d(1 : N)(A_d + I)^{-1}B_d$, where the notation $\mathcal{M}_d(i : j)$ defines the $j - i + 1$ columns of $\mathcal{M}_d$, starting with column i. Some of the variables are therefore linearly dependent and have to be eliminated. This can be done by using (59), and the new formulation of Theorem 5 is immediately obtained. ■

Even in this more simplified form, the existence of a strictly feasible solution is still not guaranteed for $r > 0$, and an SDP software package (such as SDPT3 [42]) that does not require strict feasibility should be used. For $r = 0$, however, we propose the following initialization procedure.

i) Start with a vector $C_d = [f(N)\,0\,f(N-1)\,0\cdots f(1)]$ with $\sum_n f(n) = 1/2$.
ii) Assume that the matrix P_d is block diagonal with the first block being a scalar (a 1×1 block) and the remaining blocks of size 2×2. Since P_d must be positive definite, it follows that all its diagonal elements, as well as the determinants of the block matrices, are positive.
iii) Choose $P_d = P_d^T \succ 0$ such that $C_d = B_d^T(A_d^T + I)^{-1}P_d(A_d + I)$ and $P_d - A_d^TP_dA_d \succ 0$. It can be shown (with a good amount of matrix algebra) that we can always obtain a strictly feasible starting point by following the above process.

Compaction Gain Bounds in the Presence of Regularity Constraints: An upper bound for the compaction gain can be computed by running the SDP with $r = 0$. A strictly feasible point always exists, and the SDP converges to a global optimum. On the other hand, with all free zeros at π, a lower bound is obtained by solving the linear system of equations defined by (54) with $L = N(N+1)/2$. These bounds determine the range

$$\mathcal{M}_d = \begin{bmatrix} P_d - A_d^TP_dA_d & (P_d - A_d^TP_dA_d)(A_d + I)^{-1}B_d \\ B_d^T(I + A_d^T)^{-1}(P_d - A_d^TP_dA_d) & B_d^T(I + A_d^T)^{-1}(P_d - A_d^TP_dA_d)(A_d + I)^{-1}B_d \end{bmatrix} \succeq 0 \quad (60)$$

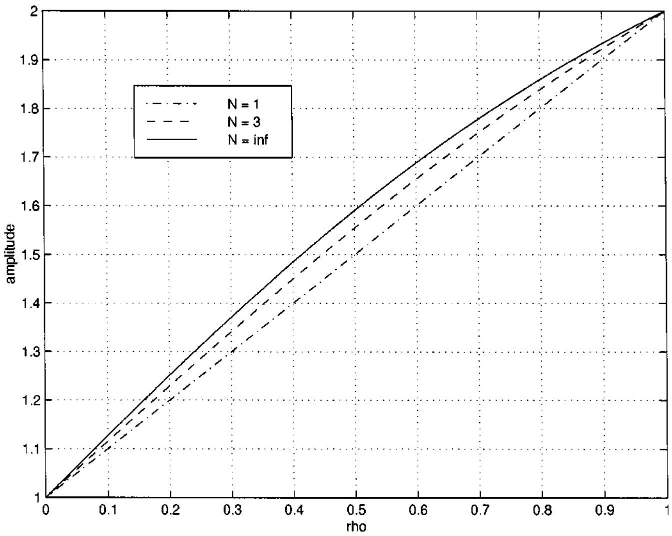


Fig. 4. Compaction gain curves for an AR(1) process for $N = 2$, 3 and ∞ with $M = 2$.

of all possible compaction gains as we increase the regularity degree.

VIII. Numerical Results

The results described here are obtained using the MATLAB LMI control toolbox. Due to space limitations, all the FIR energy compaction filter design programs and corresponding documentation can be found at http://www.systems.caltech.edu/tuqan. For all the following examples, $W = dI$, where $d = 10^{-6}$.

Example 1—AR(1) Process: Assume that the input $x(n)$ is a zero-mean AR(1) process with an autocorrelation sequence in the form $R_{xx}(k) = \rho^{|k|}$, where $0 < \rho < 1$. Let $M = 2$. The optimum compaction gain curves for $N = 2$ and 3 as a function of ρ are shown in Fig. 4. The curve for $N = 3$ coincides with the theoretical compaction gain formula $G_{comp}(2, 3) = 1 + (2\rho/\sqrt{3 + \rho^2})$ derived in [33]. The precise difference is actually on the order of 10^{-5}. The last curve denotes the compaction gain when $N = \infty$ (ideal lowpass filter case). A closed-form expression for the compaction gain can be obtained by evaluating the integral in (1) since the integrand is a Poisson kernel [43, p. 308]. The final result is $G_{comp}(2, \infty) = (4/\pi)\arctan((1 + \rho)/(1 - \rho))$. From Fig. 4, it is therefore very clear that for an AR(1) process, the margin of gain versus filter length is very small. Assume now that $\rho = 0.9$, $N = 3$, and $M = 2$. The theoretical optimum filter $F_{th}(z)$, which is obtained from [33], is the same as the SDP one (the difference in the numerical accuracy of the coefficients is in the order of 10^{-8}) and is given by $F_{opt}(z) = -0.067\,233(z^{-3} + z^3) + 0.566\,774(z^{-1} + z^1) + 1$. The compaction gain in both cases is equal to 1.922. The minimum-phase filter in this case is $H_{\min}(z) = 0.4939 + 0.827\,9z^{-1} + 0.2282z^{-2} - 0.1361z^{-3}$. Note that $H_{\min}(z)\,H_{\min}(z^{-1}) = F_{opt}(z) = F_{th}(z)$, indicating a numerically accurate spectral factor. The positivity of $F_{opt}(z)$ is demonstrated by the double roots in the Z-plane plot of Fig. 5(a). The compaction gain of $H_{\min}(z)H_{\min}(z^{-1})$ remains equal to 1.922, and the positivity property of $H_{\min}(z)H_{\min}(z^{-1})$ is not lost, as we can clearly see in Fig. 5(b).

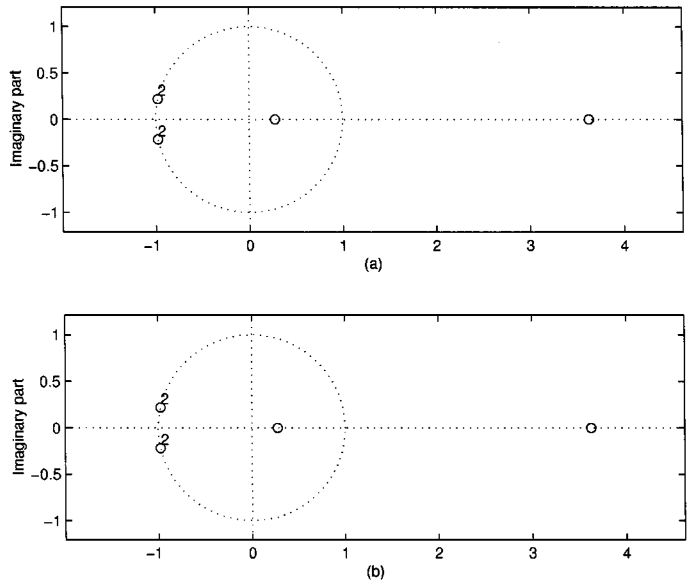


Fig. 5. Double roots on the unit circle, indicating the positivity of the product filter $F(z)$ (a) as the output of the program (b) as a result of convolving $h_{\min}(n)$ with its flipped version.

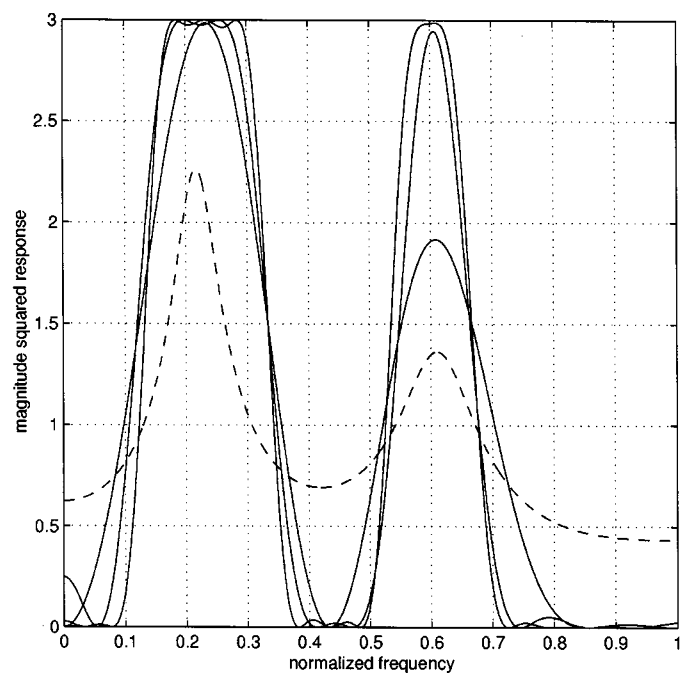


Fig. 6. Magnitude squared responses of the optimum compaction filters corresponding to the multiband AR(5) process (dashed curve) of order $N = 7$, 17, and 27 with $M = 2$.

Example 2—Multiband AR(5) Process: Assume that the input $x(n)$ is a zero-mean multiband AR(5) process (dashed

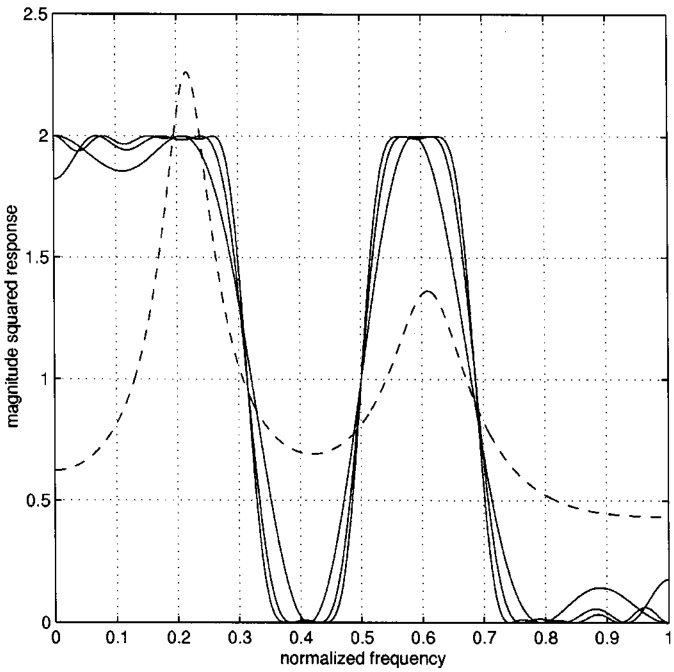


Fig. 7. Magnitude squared responses of the optimum compaction filters corresponding to the multiband AR(5) process (dashed curve) of order $N = 7$, 17, and 27 with $M = 3$.

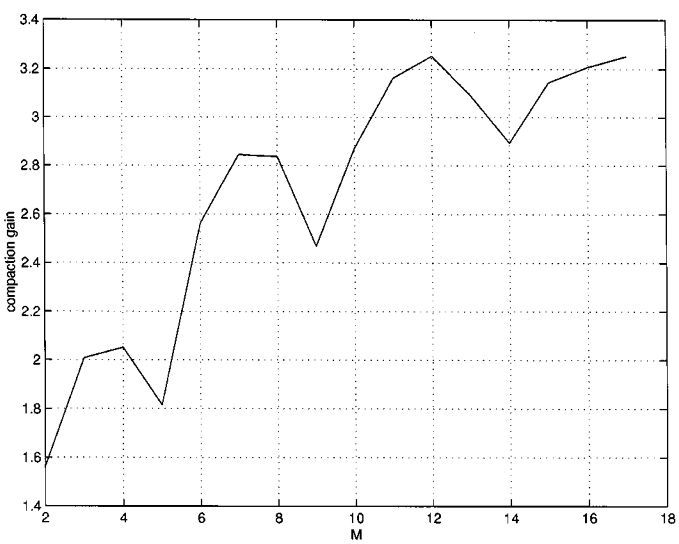


Fig. 8. Nonmonotone behavior of the compaction gain as a function of the number of channels M with a filter of fixed order $N = 17$.

curve in Figs. 6 and 7). The magnitude squared responses of the resulting optimum compaction filters are shown in Fig. 6 for $N = 7$, 17, and 27 and $M = 2$. The corresponding compaction gains are 1.5243, 1.5633, and 1.5748. Similarly, the magnitude squared responses of the resulting optimum compaction filters with N =7, 17, 27, and $M = 3$ are shown in Fig. 7. In this case, the compaction gains are 1.866, 2.007, and 2.045. With a fixed filter order $N = 17$, a plot of G_{comp} as a function of M is shown in Fig. 8, indicating a nonmonotonic behavior. When $M \geq N + 1$, the Nyquist constraint reduces to a unit energy constraint, and the optimum compaction filter is the eigenvector corresponding to the largest eigen value of the Toeplitz symmetric autocorrelation matrix with first row $[R(0)R(1)\cdots R(17)]$. The maximum eigenvalue of the 17 $\times$ 17 Toeplitz symmetric autocorrelation matrix is therefore an upper bound on the compaction gain as M increases from 2 to 17. The overall incremental behavior of the compaction gain should be intuitively acceptable because as M increases with N fixed, the constraints on the filter coefficients become less stringent. In fact, we can easily prove that $G_{comp}((k+1)M, N) \geq G_{comp}(kM, N)$.

Example 3—Regularity versus Compaction Gain: Consider again the AR(5) process of Example 2 with $M = 2$, $N = 3$. The compaction gain is 1.387 with no zeros at π and 1.374 with one or two zeros at π. The compaction filter converges to the same solution $f(1) = 9/16$, $f(3) = -1/16$ in the latter two cases. For $N = 5$, the compaction gain is 1.453 with no regularity constraints and drops to 1.384 when forcing a single zero at π. However, it (almost) remains constant, even when forcing all three zeros at π. In general, we have found that the compaction gain can drop substantially when forcing a zero at π but then usually remains constant as we increase the smoothness degree.

IX. Concluding Remarks

Using state-space theory, we have proposed a new approach for designing globally optimal FIR energy compaction filters. The design of such filters is important because they are the basic building blocks of an M-channel FIR orthonormal PCFB. In particular, for the two-channel case ($M = 2$), the optimum compaction filter determines the optimum orthonormal filterbank. Therefore, by using any of the proposed formulations in this paper, *the optimal two-channel FIR PCFB* is found [24]. The issue regarding which spectral factor to choose is, however, unclear. Different spectral factors exhibit different subband decorrelation properties. For the M-channel case, some progress regarding the design of the filterbank has been reported in [18], but the problem, in its full generality, remains open at the moment of this writing.

The tradeoff between any global optimum algorithm and a suboptimal one is typically complex. In general, SDP's implemented using interior point methods are more computationally expensive than (for example) linear programs. One way to see this is to note that a linear program is a special case of an SDP, where the matrices defining the LMI are diagonal. The added structure produces more efficient algorithms. Nevertheless, SDP's come in all sorts of different forms and implementations. We have already displayed several formulations of the same problem in a single paper, and we certainly have not tried all of them in our simulations (see, however, additional documentation at http://www.systems.caltech.edu/tuqan/). Since the main goal was to validate our new approach, we felt that finding more efficient semidefinite programs, although important, plays a secondary role with respect to the other results of the paper. Indeed, changing the form of an SDP to another is usually not that difficult. Moreover, there is a whole community of numerical analysis researchers looking for faster implementations of SDPs, and substantial speedups are expected in the future.

APPENDIX A
PROOF OF THEOREM 1

Substituting (14)–(16) in the expression of $F(z)$, we can write the following sequence of equations:

$$\begin{aligned} F(z) &= D(z)+D(z^{-1}) \\ &= D_d+D_d^T+C_d(zI-A_d)^{-1}B_d \\ &\quad +B_d^T(z^{-1}I-A_d^T)^{-1}C_d^T \\ &= W_d^TW_d+B_d^TP_dB_d+(W_d^TL_d^T+B_d^TP_dA_d) \\ &\quad \cdot (zI-A_d)^{-1}B_d+B_d^T \\ &\quad \cdot (z^{-1}I-A_d^T)^{-1}(L_dW_d+A_d^TP_dB_d) \\ &= W_d^TW_d+B_d^TP_dB_d+W_d^TL_d^T(zI-A_d)^{-1}B_d \\ &\quad +B_d^T(z^{-1}I-A_d^T)^{-1}L_dW_d+B_d^TP_dA_d(zI-A_d)^{-1} \\ &\quad \cdot B_d+B_d^T(z^{-1}I-A_d^T)^{-1}A_d^TP_dB_d \\ &= W_d^TW_d+W_d^TL_d^T(zI-A_d)^{-1}B_d \\ &\quad +B_d^T(z^{-1}I-A_d^T)^{-1}L_dW_d+B_d^T \\ &\quad \cdot [P_d+P_dA_d(zI-A_d)^{-1}+(z^{-1}I-A_d^T)^{-1}A_d^TP_d]B_d \\ &= W_d^TW_d+W_d^TL_d^T(zI-A_d)^{-1}B_d \\ &\quad +B_d^T(z^{-1}I-A_d^T)^{-1}L_dW_d+B_d^T(z^{-1}I-A_d^T)^{-1} \\ &\quad \cdot [(z^{-1}I-A_d^T)P_d(zI-A_d)+(z^{-1}I-A_d^T)P_dA_d \\ &\quad +A_d^TP_d(zI-A_d)](zI-A_d)^{-1}B_d. \end{aligned}$$

Now, note that

$$\begin{aligned} &[(z^{-1}I-A_d^T)P_d(zI-A_d)+(z^{-1}I-A_d^T)P_dA_d \\ &\quad +A_d^TP_d(zI-A_d)] \\ &\quad = P_d-A_d^TP_dA_d. \end{aligned} \tag{61}$$

This result is obtained by simply multiplying out the matrices on the left in (61). By using (61) and (14)

$$\begin{aligned} F(z) &= D(z)+D(z^{-1}) \\ &= W_d^TW_d+W_d^TL_d^T(zI-A_d)^{-1}B_d \\ &\quad +B_d^T(z^{-1}I-A_d^T)^{-1}L_dW_d+B_d^T(z^{-1}I-A_d^T)^{-1} \\ &\quad \cdot [P_d-A_d^TP_dA_d](zI-A_d)^{-1}B_d \\ &= W_d^TW_d+W_d^TL_d^T(zI-A_d)^{-1}B_d \\ &\quad +B_d^T(z^{-1}I-A_d^T)^{-1}L_dW_d+B_d^T(z^{-1}I-A_d^T)^{-1} \\ &\quad \cdot L_dL_d^T(zI-A_d)^{-1}B_d \\ &= (W_d^T+B_d^T(z^{-1}I-A_d^T)^{-1}L_d) \\ &\quad \cdot (W_d+L_d^T(zI-A_d)^{-1}B_d). \end{aligned}$$

APPENDIX B
PROOF OF THE DISCRETE TIME MINIMUM PHASE SPECTRAL FACTOR FORM

Substituting $s=(1-z^{-1})/(1+z^{-1})$ into $H_{\min}(s)=W_c+L_c(sI-A_c)^{-1}B_c$ and simplifying, we obtain

$$\begin{aligned} &H_{\min}(z) \\ &\quad = W_c+L_c(1+z^{-1})(I-z^{-1}(I-A_c)^{-1}(I+A_c))^{-1} \\ &\quad\quad \cdot (I-A_c)^{-1}B_c. \end{aligned} \tag{62}$$

Now, the expression $H_{\min}(z)=W_d+L_dz^{-1}\,(I-z^{-1}A_d)^{-1}B_d$ implies the power series

$$H_{\min}(z)=W_d+\sum_{n=1}^{\infty}L_d(z^{-1}A_d)^nB_d. \tag{63}$$

Equating the constant terms in the above expressions, we obtain $W_d=W_c+L_c(I-A_c)^{-1}B_c$. The coefficient of z^{-1} in (62) can be simplified to $2L_c(I-A_c)^2B_c$ so that $L_dB_d=2L_c(I-A_c)^2B_c$. Finally, since the nth term in the power series (62) has the form $S[(I-A_c)^{-1}(I+A_c)]^nQ$ for constant S and Q, the choice $A_d=(I-A_c)^{-1}(I+A_c), L_d=L_c, B_d=2(I-A_c)^2B_c$ yields a realization of $H_{\min}(z)$.

APPENDIX C
PROOF OF MINIMALITY

A state-space realization is minimal if and only if it is jointly observable and controllable. Assuming the minimality of the triple $\{A_c,\, B_c,\, C_c\}$, we use the (PBH) test [44, pp. 135–136] to prove the minimality of the triple $\{A_d,\, B_d,\, C_d\}$ given by (28). In particular, since $(A_c,\, B_c)$ is controllable, then there does not exist a row vector $q\neq 0$ such that $qB_c=0$ and $qA_c=\mu q$. Now, assume that $(A_d,\, B_d)$ is not controllable. Then, there exists a row vector $x\neq 0$ such that $xB_d=x(I-A_c)^{-2}B_c=0$ and $x(I-A_c)^{-1}(I+A_c)=\lambda x$. Let $y=x(I-A_c)^{-2}$. Then, $yB_c=0$ and $y(I-A_c)(I+A_c)=\lambda y(I-A_c)^2$. By observing that the matrices $(I-A_c)(I+A_c)$ commute, the last expression therefore simplifies to $y(I+A_c)=\lambda y(I-A_c)$. This in turn implies that $yA=((\lambda-1)/(\lambda+1))y$. If $\lambda=-1$, then $y=x=0$, which is a contradiction. If $\lambda\neq -1$, then the assumption that $(A_d,\, B_d)$ is not controllable implies that $(A_c,\, B_c)$ is also not controllable, which is again a contradiction. The observability of $(A_d,\, C_d)$ can be established in a similar way.

APPENDIX D
SIMPLIFYING (33)

It is not difficult to see that by multiplying the matrices in (33), we get the LMI as shown at the bottom of the next page, where

$$\begin{aligned} X &= (C_c-B_c^TP_c)(I+A_d)^{-1}B_d+B_d^T(I+A_d^T)^{-1} \\ &\quad \cdot (C_c^T-P_cB_c)+B_d^T(I+A_d^T)^{-1} \\ &\quad \cdot (P_d-A_d^TP_dA_d)(I+A_d)^{-1}B_d. \end{aligned}$$

Making the substitutions (34) and (35), the first term in the above matrix becomes

$$\begin{aligned}
&-P_cA_c - A_c^TP_c\\
&\quad= P_c(A_d+I)^{-1}(I-A_d) + (I-A_d^T)(A_d^T+I)^{-1}P_c\\
&\quad= (A_d^T+I)(A_d^T+I)^{-1}P_c(A_d+I)^{-1}(I-A_d)\\
&\qquad+(I-A_d^T)(A_d^T+I)^{-1}P_c(I+A_d)^{-1}(I+A_d)\\
&\quad= \tfrac{1}{2}\{(A_d^T+I)P_d(I-A_d)+(I-A_d^T)P_d(I+A_d^T)\}\\
&\quad= P_d - A_d^TP_dA_d.
\end{aligned}$$

Similarly, the second term simplifies as follows:

$$\begin{aligned}
&C_c^T - P_cB_c - (P_cA_c + A_c^TP_c)(I+A_d)^{-1}B_d\\
&\quad= C_d^T - 2P_c(I+A_d)^{-2}B_d\\
&\qquad-(P_d - A_d^TP_dA_d)(I+A_d)^{-1}B_d\\
&\quad= C_d^T - (I+A_d^T)P_d(I+A_d)^{-1}B_d\\
&\qquad-(P_d - A_d^TP_dA_d)(I+A_d)^{-1}B_d\\
&\quad= C_d^T - \{(I+A_d^T)P_d - P_d + A_d^TP_dA_d\}(I+A_d)^{-1}B_d\\
&\quad= C_d^T - \{A_d^TP_d + A_d^TP_dA_d\}(I+A_d)^{-1}B_d\\
&\quad= C_d^T - A_d^TP_dB_d.
\end{aligned}$$

The third term is simply the transpose of the second term. Finally, the fourth term reduces to

$$\begin{aligned}
&D_c + D_c^T + X\\
&\quad= B_d^T(I+A_d^T)^{-1}(P_d - A_d^TP_dA_d)(I+A_d)^{-1}B_d\\
&\qquad+C_d(I+A_d)^{-1}B_d - 2B_d^T(I+A_d^T)^{-2}P_c(I+A_d)^{-1}B_d\\
&\qquad+B_d^T(I+A_d^T)^{-1}C_d^T 2B_d^T(I+A_d^T)^{-1}P_c(I+A_d)^{-2}B_d\\
&\qquad+D_d + D_d^T - C_d(I+A_d)^{-1}B_d - B_d^T(I+A_d^T)^{-1}C_d^T\\
&\quad= D_d + D_d^T + B_d^T(I+A_d^T)^{-1}\{P_d - A_d^TP_dA_d\\
&\qquad-P_d(I+A_d) - (I+A_d^T)P_d\}(I+A_d)^{-1}B_d\\
&\quad= D_d + D_d^T - B_d^T(I+A_d^T)^{-1}\{(I+A_d^T)P_d(I+A_d)\}\\
&\qquad\cdot(I+A_d)^{-1}B_d\\
&\quad= D_d + D_d^T - B_d^TP_dB_d.
\end{aligned}$$

APPENDIX E
PROOF OF CONDITION (38)

Since P_d and R are positive definite, the following identity can be established:

$$\begin{aligned}
&(R - B_d^TP_dB_d)^{-1}\\
&\quad= R^{-1} + R^{-1}B_d^T(P_d^{-1} - B_dR^{-1}B_d^T)^{-1}B_dR^{-1}\\
&\quad= R^{-1} + R^{-1}B_d^TP_dB_dR^{-1} + R^{-1}B_d^TP_dB_d\\
&\qquad\cdot(R - B_d^TP_dB_d)^{-1}B_d^TP_dB_dR^{-1}.
\end{aligned} \tag{64}$$

The identity is obtained by applying twice the matrix inversion lemma [44, p. 656]. Starting with (38)

$$\begin{aligned}
P_d &= (A_d - B_dR^{-1}C_d)^TP_d(A_d - B_dR^{-1}C_d)\\
&\quad+C_d^TR^{-1}C_d + (A_d - B_dR^{-1}C_d)^TP_dB_d\\
&\quad\cdot(R - B_d^TP_dB_d)^{-1}\\
&\quad\cdot B_d^TP_d(A_d - B_dR^{-1}C_d)\\
&= C_d^TR^{-1}C_d + C_d^TR^{-1}B_d^TP_dB_dR^{-1}C_d\\
&\quad+C_d^TR^{-1}B_d^TP_dB_d(R - B_d^TP_dB_d)^{-1}\\
&\quad\cdot B_d^TP_dB_dR^{-1}C_d + \text{other terms}\\
&= C_d^T(R - B_d^TP_dB_d)^{-1}C_d + \text{other terms}
\end{aligned}$$

where the last equation follows from (64). Substituting now the other terms, we get

$$\begin{aligned}
P_d &= A_d^TP_dA_d + C_d^T(R - B_d^TP_dB_d)^{-1}C_d\\
&\quad+A_d^TP_dB_d(R - B_d^TP_dB_d)^{-1}B_d^TP_dA_d\\
&\quad-A_d^TP_dB_dR^{-1}C_d - A_d^TP_dB_d\\
&\quad\cdot(R - B_d^TP_dB_d)^{-1}B_d^TP_dB_dR^{-1}C_d\\
&\quad-C_d^TR^{-1}B_d^TP_dA_d - C_d^TR^{-1}B_d^TP_dB_d\\
&\quad\cdot(R - B_d^TP_dB_d)^{-1}B_d^TP_dA_d\\
&= A_d^TP_dA_d + C_d^T(R - B_d^TP_dB_d)^{-1}C_d\\
&\quad+A_d^TP_dB_d(R - B_d^TP_dB_d)^{-1}B_d^TP_dA_d\\
&\quad-A_d^TP_dB_d\{R^{-1} + (R - B_d^TP_dB_d)^{-1}B_d^TP_dB_dR^{-1}\}C_d\\
&\quad-C_d^T\{R^{-1} + R^{-1}B_d^TP_dB_d(R - B_d^TP_dB_d)^{-1}\}B_d^TP_dA_d.
\end{aligned} \tag{65}$$

We only simplify one of the cross terms [last two lines of (65)] since they are the transpose of each other. Recalling that $R \succ 0$ (by assumption), we can then write

$$\begin{aligned}
&C_d^T\{R^{-1} + R^{-1}B_d^TP_dB_d(R - B_d^TP_dB_d)^{-1}\}B_d^TP_dA_d\\
&\quad= C_d^TR^{-1/2}\{I + R^{-1/2}B_d^TP_dB_d(R - B_d^TP_dB_d)^{-1}R^{1/2}\}\\
&\qquad\cdot R^{-1/2}B_d^TP_dA_d\\
&\quad= C_d^TR^{-1/2}\{I + R^{-1/2}B_d^TP_dB_dR^{-1/2}\\
&\qquad\qquad\cdot(I - R^{-1/2}B_d^TP_dB_dR^{-1/2})^{-1}\}\\
&\qquad\cdot R^{-1/2}B_d^TP_dA_d\\
&\quad= C_d^TR^{-1/2}\{(I - R^{-1/2}B_d^TP_dB_dR^{-1/2})\\
&\qquad\qquad+R^{-1/2}B_d^TP_dB_dR^{-1/2}\}\\
&\qquad\cdot(I - R^{-1/2}B_d^TP_dB_dR^{-1/2})^{-1}R^{-1/2}B_d^TP_dA_d\\
&\quad= C_d^TR^{-1/2}(I - R^{-1/2}B_d^TP_dB_dR^{-1/2})^{-1}\\
&\qquad\cdot R^{-1/2}B_d^TP_dA_d\\
&\quad= C_d^T(R - B_d^TP_dB_d)^{-1}B_d^TP_dA_d.
\end{aligned} \tag{66}$$

By substituting (66) and its transpose into (65), (37) is easily obtained, and the result follows.

$$\begin{bmatrix} -P_cA_c - A_c^TP_c & C_c^T - P_cB_c - (P_cA_c + A_c^TP_c)(I+A_d)^{-1}B_d\\ C_c - B_c^TP_c - B_d^T(I+A_d^T)^{-1}(P_cA_c + A_c^TP_c) & X + D_c + D_c^T \end{bmatrix} \succeq 0$$

ACKNOWLEDGMENT

The authors would like to thank Prof. S. Boyd (Stanford University) and Prof. L. Vandenberghe (UCLA) for generously making available their SDP software which was used to validate earlier versions of our results. They would also like to thank Prof. M. G. Strintzis for sending a preprint of [21].

REFERENCES

[1] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[2] P. R. Chevillat and G. Ungerboeck, "Optimum FIR transmitter and receiver filters for data transmission over band-limited channels," *IEEE Trans. Commun.*, vol. COMM-30, pp. 1909–1915, Aug. 1982.

[3] J. Tuqan and P. P. Vaidyanathan, "Globally optimal FIR filters with applications in source and channel coding," in *Proc. Int. Symp. Inform. Theory*, Cambridge, MA, 1998, p. 475.

[4] J. Tuqan, "Global optimization of orthogonal FIR transmitter and receiver filters for data transmission over noisy channels," in *Proc. 33rd Asilomar Conf. Signals, Syst., Comput.*, vol. 1, 1999, pp. 202–206.

[5] Q. Jin, Z.-Q. Luo, and K. M. Wong, "Optimum filter banks for signal decomposition and its applications in adaptive echo cancellation," *IEEE Trans. Signal Processing*, vol. 44, pp. 1669–1689, July 1996.

[6] P. P. Vaidyanathan, T. Q. Nguyen, Z. Doganata, and T. Saramaki, "Improved technique for design of perfect reconstruction FIR QMF banks with lossless polyphase matrices," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1042–1056, 1989.

[7] J. Tuqan and P. P. Vaidyanathan, "Oversampling PCM techniques and optimum noise shapers for quantizing a class of nonbandlimited signals," *IEEE Trans. Signal Processing*, vol. 47, pp. 389–407, Feb. 1999.

[8] M. K. Tsatsanis and G. B. Giannakis, "Principal component filter banks for optimal multiresolution analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1766–1777, Aug. 1995.

[9] K. C. Aas, K. A. Duell, and C. T. Mullis, "Synthesis of extremal wavelet-generating filters using Gaussian quadrature," *IEEE Trans. Signal Processing*, vol. 43, pp. 1045–1057, May 1995.

[10] M. K. Tsatsanis and G. B. Giannakis, "Time-varying system identification and model validation using wavelets," *IEEE Trans. Signal Processing*, vol. 41, pp. 3512–3523, Dec. 1993.

[11] M. Unser, "On the optimality of ideal filters for pyramid and wavelet signal approximation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3591–3596, Dec. 1993.

[12] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders," *IEEE Trans. Signal Processing*, vol. 46, pp. 1528–1543, June 1998.

[13] M. G. Strintzis, "Optimal biorthogonal wavelet bases for signal decomposition," *IEEE Trans. Signal Processing*, vol. 44, pp. 1406–1416, June 1996.

[14] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[15] M. Unser, "An extension of the Karhunen–Loeve transform for wavelets and perfect reconstruction filter banks," *SPIE Math. Imag.*, vol. 2034, pp. 45–56, 1993.

[16] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*. New York: Academic, 1979.

[17] A. Kirac and P. P. Vaidyanathan, "On existence of FIR principal component filter banks," *Proc. ICASSP*, vol. 3, pp. 1329–1332, May 1998.

[18] P. Moulin and M. K. Mihcak, "Theory and design of signal adapted FIR paraunitary filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 920–929, Apr. 1998.

[19] S. Akkarakaran and P. P. Vaidyanathan, "On optimization of filter banks with denoising applications," in *Proc. ISCASS*, 1999.

[20] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, pp. 2915–2925, Nov. 1994.

[21] M. G. Strintzis, "Optimal pyramidal and subband decompositions for hierarchical coding of noisy and quantized images," *IEEE Trans. Image Processing*, vol. 7, pp. 155–166, Feb. 1998.

[22] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994, vol. 13.

[23] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994, vol. 15.

[24] J. Tuqan and P. P. Vaidyanathan, "Globally optimal two-channel FIR orthonormal filter banks adapted to the input signal statistics," *Proc. ICASSP*, vol. 3, pp. 1353–1356, May 1998.

[25] ——, "The role of the discrete-time Kalman–Yakubovich–Popov (KYP) lemma in designing statistically optimum FIR orthonormal filter banks," in *Proc. ISCAS*, vol. 5, June 1998, pp. 122–125.

[26] P. P. Vaidyanathan and P. Q. Hoang, "Lattice structures for optimal design and robust implementation of two channel perfect reconstruction QMF banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 81–94, Jan. 1988.

[27] P. Delsarte, B. Macq, and D. T. M. Slock, "Signal-adapted multiresolution transform for image coding," *IEEE Trans. Inform. Theory*, vol. 38, pp. 897–904, Mar. 1992.

[28] D. Taubman and A. Zakhor, "A multi-start algorithm for signal adaptive subband systems," *Proc. ICASSP*, vol. 3, pp. 213–216, 1992.

[29] H. Caglar, Y. Liu, and A. N. Akansu, "Statistically optimized PR-QMF design," *SPIE, Visual Commun. Image Process.*, vol. 1605, pp. 86–94, 1991.

[30] L. Vandendorpe, "CQF filter banks matched to signal statistics," *Signal Process.*, vol. 29, pp. 237–249, 1992.

[31] P. Moulin, M. Anitescu, K. Kortanek, and F. A. Potra, "The role of linear semi-infinite programming in signal dapted QMF bank design," *IEEE Trans. Signal Processing*, vol. 45, pp. 2160–2174, Sept. 1997.

[32] J.-C. Pesquet and P. L. Combettes, "Wavelet synthesis by alternating projections," *IEEE Trans. Signal Processing*, vol. 44, pp. 728–732, Mar. 1996.

[33] A. Kirac and P. P. Vaidyanathan, "Theory and design of optimum FIR compaction filters," *IEEE Trans. Signal Processing*, vol. 46, pp. 903–919, Apr. 1998.

[34] I. Gigberg, *Methods in Operator Theory and Signal Processing*, I. Schur, Ed. Boston, MA: Birkhauser-Verlag, 1986.

[35] B. E. L. Hitz and B. D. O. Anderson, "Discrete positive-real functions and their application to system stability," *Proc. Inst. Elect. Eng.*, pp. 153–155, Jan. 1969.

[36] L. Vandenberghe and S. Boyd, "Semidefinite programming," *Siam Rev.*, vol. 38, pp. 49–95, Mar. 1996.

[37] S.-P. Wu, S. Boyd, and L. Vandenberghe, "FIR filter design via semidefinite programming and spectral factorization," in *Proc. IEEE Conf. Decision Contr.*, 1996.

[38] B. D. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis: A Modern System Theory Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1973.

[39] J. C. Willems, "Least squares stationary optimal control and the algebraic Ricatti equation," *IEEE Trans. Automat. Contr.*, vol. 7, pp. 621–634, Dec. 1971.

[40] B. D. Anderson, "Algebraic properties of minimal degree spectral factors," *Automatica*, pp. 491–500, June 1973.

[41] L. Vandenberghe and S. Boyd, "SP: Software for semi definite programming, beta version,", Oct. 1994.

[42] K. C. Toh, M. J. Todd, and R. H. Tutuncu, "SDPT3—A Matlab software package for semi definite programming," version 2.1, Sept. 1999.

[43] R. V. Churchill, *Complex Variables and Applications*. New York: McGraw-Hill, 1990.

[44] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[45] D. Xuan and R. Bamberger, "Complete FIR principal component filter banks," in *Proc. ISCAS*, 1996, pp. 417–420.

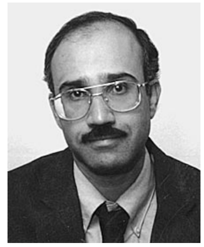
[46] W.-S. Lu and A. Antoniou, "Optimized orthogonal and biorthogonal wavelets using linear parametrization of half band filters," *Proc. ISCAS*, vol. 5, pp. 102–105, June 1998.

[47] J. Tuqan, "On the design of optimal orthogonal finite order transmitter and receiver filters over noisy channels," *Proc. 34th Asilomar Conf. Signals, Syst., Comput.*, Nov. 2000.

Jamal Tuqan (S'91–M'98) received the B.Sc. degree (with honors) in electrical engineering from Cairo University, Cairo, Egypt in 1989, the M.S.E.E. degree from the Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree in electrical engineering from the California Institute of Technology (Caltech), Pasadena, in 1997.

From September 1989 to July 1990, he worked at the IBM Research Center, Cairo, writing software for Arabic speech compression based on linear predictive coding techniques. From July 1994 to December 1997, he was a Research and Teaching Assistant at Caltech. From January until March 1998, he has been a Technical Staff Member with the Digital Signal Processing Group at the same university. During the spring quarter 1998, he was appointed as a Lecturer at Caltech to teach a class on linear estimation theory and adaptive filtering. He is currently with the Image and Video Communications Department at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY. His main research interests are in the general area of digital and statistical signal processing, signal processing applications in emerging communication technologies, data compression, and applied mathematics. The focus of his Ph.D. dissertation research was on the optimization of multirate systems and filterbanks according to the input signal statistics for compression and communication applications.

Dr. Tuqan is a member of the Technical Program Committee for the 2000 IEEE International Conference on Multimedia and Exposition.



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, all from the University of Calcutta, in 1974, 1977, and 1979, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara (UCSB), in 1982.

He was a Post-Doctoral Fellow at UCSB from September 1982 to March 1983. In March 1983, he joined the Electrical Engineering Department, California Institute of Technology (Caltech), Pasadena, as an Assistant Professor, and since 1993, he has been 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 the Technical Program Chairman for the 1992 IEEE International Symposium on Circuits and Systems. He was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS ffrom 1985 to 1987 and is currently an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS and a Consulting Editor for the journal *Applied and Computational Harmonic Analysis*. He has been a Guest Editor in 1998 for special issues of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II on the topics of filter banks, wavelets, and subband coders. He has authored a number of papers in IEEE journals and is the author of the book *Multirate Systems and Filter Banks* (Englewood Cliffs, NJ: Prentice-Hall, 1993). He has written several chapters for various signal processing handbooks. He was a recipient of the Award for Excellence in Teaching at Caltech for the years 1983–1984, 1992–1993 and 1993–1994. 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 was recipient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommuncations Engineers, India, for his joint paper in the *IETE Journal*. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE TRANSACTIONS ON SIGNAL PROCESSING for which the first author (T. Nguyen) received the Young Outstanding Author award in 1993. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by Hewlett Packard Co., for his contributions to engineering education, especially the book *Multirate Systems and Filter Banks*. He has given several plenary talks, including at the Eusipco'98, Asimolar'88, and SPCOM'95 conferences on signal processing. He has been chosen a Distinguished Lecturer for the IEEE Signal Processing Society for 1996–1997. In 1999, he received the IEEE CAS Society's Golden Jubilee Medal.