

Theory and Design of Optimum FIR Compaction Filters

Ahmet Kırac, *Student Member, IEEE*, and P. P. Vaidyanathan, *Fellow, IEEE*

Abstract—The problem of optimum FIR energy compaction filter design for a given number of channels M and a filter order N is considered. The special cases where $N < M$ and $N = \infty$ have analytical solutions that involve eigenvector decomposition of the autocorrelation matrix and the power spectrum matrix, respectively. In this paper, we deal with the more difficult case of $M < N < \infty$. For the two-channel case and for a restricted but important class of random processes, we give an *analytical solution* for the compaction filter that is characterized by its zeros on the unit circle. This also corresponds to the optimal two-channel FIR filter bank that maximizes the coding gain under the traditional quantization noise assumptions. With a minor extension, this can also be used to generate optimal wavelets. For the arbitrary M -channel case, we provide a very efficient suboptimal design method called the *window method*. The method involves two stages that are associated with the above two special cases. As the order increases, the suboptimality becomes negligible, and the filter converges to the ideal optimal solution. We compare the window method with a recently introduced technique based on linear programming.

Index Terms—Energy compaction, Nyquist filters, orthonormal filter banks, subband coding, wavelets.

I. INTRODUCTION

THE DESIGN OF optimum energy compaction filters has been of interest in the recent past because of their known connection to optimal subband coding (SBC) and principal component filter banks (PCFB's) [1]–[9]. When there is no order constraint on the filters, the optimization of uniform orthonormal filter banks for given second-order statistics has been solved [6], [7], [9]. The minimum mean-squared error solution is such that each filter is an optimal compaction filter corresponding to a power spectral density (psd) that is derived from the input psd. In particular, the filter corresponding to the largest subband variance has to be an optimal compaction filter for the input psd itself.

Consider an M -channel uniform orthonormal (or paraunitary) filter bank shown in Fig. 1. In terms of the filters, we can express the orthonormality as $H_i(e^{j\omega})H_j^*(e^{j\omega})|_{\perp M} = \delta(i-j)$ [10]. This, in particular, implies that each filter satisfies the Nyquist(M) property (see Section I-B for notations and

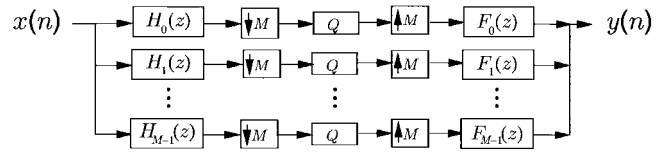


Fig. 1. M -channel orthonormal filter bank. $F_i(e^{j\omega}) = H_i^*(e^{j\omega})$, and $|H_i(e^{j\omega})|^2$ is Nyquist(M).

terminology)

$$|H_i(e^{j\omega})|^2|_{\perp M} = 1. \quad (1)$$

Using the mean-squared error (mse) as the criterion, with high-bit rate assumptions on the quantization noise sources, and with optimal bit allocation, one can write the coding gain as [11]

$$G_{\text{coding}} = \frac{\sigma_x^2}{\left(\prod_{i=0}^{M-1} \sigma_{x_i}^2\right)^{1/M}} \quad (2)$$

where $\sigma_{x_i}^2$ is the variance at the output of $H_i(z)$, and $\sigma_x^2 = (1/M) \sum_{i=0}^{M-1} \sigma_{x_i}^2$ by the orthonormality.

The optimum orthonormal filter bank that maximizes (2) is well known for the case where filter orders are constrained to be less than M . This is the famous Karhunen–Loeve transform coder (KLT), and it diagonalizes the $M \times M$ autocorrelation matrix of the input. The solution for the case where the filter orders are unconstrained (ideal SBC) has recently been established. The polyphase matrix [10] of the solution diagonalizes the *psd matrix* in the frequency domain. This, in particular, implies the diagonalization of the autocorrelation matrix (which was both necessary and sufficient condition for the transform coding case). Diagonalization of the psd matrix at each frequency, however, is not sufficient for the unconstrained filter bank to be optimum [9]. There should be an additional ordering of the eigenvalues of the psd matrix at each frequency (spectral majorization) [9]. At a frequency ω , these eigenvalues are $\{S_{xx}(e^{j(\omega+k(2\pi/M))})\}$, $k = 0, \dots, M-1$, where $S_{xx}(e^{j\omega})$ is the input psd.

For a uniform filter bank to be optimum, it is only required that the product of the subband variances $\prod_{i=0}^{M-1} \sigma_{x_i}^2$ be a minimum [see (2)]. It turns out that both in the transform coder case and in the ideal SBC case, the optimal solutions achieve a fascinating *majorization* property described as follows. Let us order the subband variances such that

$$\sigma_{x_0}^2 \geq \sigma_{x_1}^2 \geq \dots \geq \sigma_{x_{M-1}}^2. \quad (3)$$

Manuscript received February 15, 1997; revised November 30, 1997. This work was supported in part by the Office of Naval Research under Grant N00014-93-1-0231 and Tektronix, Inc. The associate editor coordinating the review of this paper and approving it for publication was Prof. Mark J. T. Smith.

The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA.

Publisher Item Identifier S 1053-587X(98)02694-4.

Among all orthogonal transform coders, the KLT has the property that the partial sum $\sum_{i=0}^K \sigma_{x_i}^2$ is maximized for each K . The same property holds for orthonormal subband coders with no order constraints. That is, for each K , $\sum_{i=0}^K \sigma_{x_i}^2$ is the largest for the optimal one. In particular, when $K = 0$, this says that $\sigma_{x_0}^2$ should be maximized by the choice of $H_0(e^{j\omega})$. That is, $H_0(e^{j\omega})$ should be an *optimum compaction filter*. A filter bank with this property was first constructed by Tsatsanis and Giannakis [6] as a solution to a different problem: Assume we keep only K of the M filters in Fig. 1, and without quantizing the subbands, we try to reconstruct the original input. What is the best filter bank that minimizes the mse of the reconstruction for each K ? The solution is named to be a PCFB. PCFB's have also been extended for the multidimensional case [12]–[15].

A. Motivation for Compaction Filter Design

One can prove directly that a PCFB maximizes the coding gain of uniform orthonormal filter banks. For this, let $\hat{\sigma}_{x_0}^2 \geq \hat{\sigma}_{x_1}^2 \geq \dots \geq \hat{\sigma}_{x_{M-1}}^2$ be set of subband variances corresponding to another orthonormal filter bank. Assume that

$$\sum_{i=0}^K \sigma_{x_i}^2 \geq \sum_{i=0}^K \hat{\sigma}_{x_i}^2, \quad K = 0, 1, \dots, M-1. \quad (4)$$

By orthonormality, we have equality for $K = M-1$ [10]. Then, by a well-known result from linear algebra (see [16, p. 199]), we have

$$\prod_{i=0}^{M-1} \sigma_{x_i}^2 \leq \prod_{i=0}^{M-1} \hat{\sigma}_{x_i}^2. \quad (5)$$

Hence, the product of variances is minimized by a PCFB. This was also independently shown in [15]. If we think of the collection of the set of subband variances obtainable by a certain class of orthonormal filter banks, then the PCFB has the set of variances that majorizes every other set in the collection. Hence, it has the minimum product $\prod_{i=0}^{M-1} \sigma_{x_i}^2$ and, therefore, the maximum coding gain.

When some of the subband variances turn out to be smaller than a certain threshold, the corresponding channels should be dropped. In this case, the coding gain expression (2) is not applicable, and the total error is the sum of the quantization error and the error due to dropping. This is the case when high bit-rate assumption on the quantization noise sources is not satisfied. Recently, it has been shown that the PCFB's are optimal for subband coding for all bit rates and for all bit allocation strategies [17]. Furthermore, the optimality of PCFB's carries over to the nonuniform case as well [18]. Hence, designing an energy compaction filter is a right step in any subband or wavelet-based coding scheme.

If the class in the definition of a PCFB is the class of orthonormal block transforms (filter orders less than M) or the orthonormal filter banks with unconstrained filters, the existence of a PCFB is assured by its very construction. In the intermediate case (i.e., finite-order filter banks), unfortunately, the existence of such a filter bank is not always guaranteed. Refer to [19] for examples where no FIR PCFB of a given

degree exists for a particular input psd. However, if it exists, then designing an optimum FIR compaction filter $H_0(e^{j\omega})$ is the first step of finding such a filter bank. In that case, Moulin *et al.* [20] uses a result due to Vaidyanathan [21] to optimally complete the filter bank. This is based on the fact that if one filter $H_0(e^{j\omega})$ in an FIR orthonormal filter bank of a given degree is known, then the number of freedoms available for the design of the remaining filters is limited. This remaining freedom can, in fact, be captured with a simple constant unitary matrix \mathbf{U} . Essentially, the last $M-1$ rows of \mathbf{U} are free and should be chosen to maximize the coding gain. The optimum \mathbf{U} is the KLT corresponding to its input vector, which is determined by the first filter $H_0(z)$ and the original input $x(n)$. As we will see, the optimality of a compaction filter depends only on its magnitude-squared frequency response. Hence, for an optimum magnitude-squared frequency response, one has the choice of selecting a particular spectral factor. It turns out that this choice affects the coding gain [19], and one has to choose the best spectral factor.

For the two-channel case, the existence of a PCFB is assured even if the filters are order constrained. To see this, note that a two-channel PCFB maximizes only $\sigma_{x_0}^2$, corresponding to $K = 0$. By orthonormality, the sum $\sigma_{x_0}^2 + \sigma_{x_1}^2$ is constant. Once one order-constrained filter that maximizes $\sigma_{x_0}^2$ is found, all that remains is to find another filter such that the two filters form an orthonormal filter bank. It is very well known that the second filter is determined from the first filter by simple flipping and sign changes [see (30)]. Hence, in the two-channel case, designing an optimal FIR compaction filter is the same as designing an optimal FIR orthonormal filter bank for subband coding. By a recent result in [17], the optimality of this filter bank is independent of the bit rates involved. In the high-bit rate case, the coding gain expression becomes $G_{\text{coding}} = \sigma_x^2 / \sqrt{\sigma_{x_0}^2 \sigma_{x_1}^2} = \sigma_x^2 / \sqrt{\sigma_{x_0}^2 (2\sigma_x^2 - \sigma_{x_0}^2)}$. In this case, we can write $G_{\text{coding}} = 1 / \sqrt{G_{\text{comp}}(2 - G_{\text{comp}})}$, where $G_{\text{comp}} = \sigma_{x_0}^2 / \sigma_x^2$ is the compaction gain defined more precisely in Section II.

In this paper, we focus on the design of an optimum FIR compaction filter when the order is such that $M < N < \infty$. As we discussed, for $M = 2$, this is equivalent to the design of optimum orthonormal filter banks for subband coding and, with trivial extensions, to the design of optimal wavelet generating filters. For arbitrary M , the design in [20] can be used to obtain a good orthonormal filter bank using the compaction filter. The usefulness of signal-adapted designs in image coding with the mse as the criterion is demonstrated in [4], [12], and [22]. We are currently working on image compression algorithms to see the extent of improvement by our optimal filter designs.

Other Applications of Compaction Filters: In view of principal component analysis, in addition to subband coding and data compression, other immediate applications of compaction filters are signal modeling and model reduction, low-resolution data representation (multimedia databases), and classification. Two other interesting applications of compaction filters are adaptive echo cancellation [23] and time-varying system identification [24]. Consider the design of zero intersymbol-interference (ISI) transmitter and receiver filters for data

transmission over bandlimited communication channels. Let $H(e^{j\omega})$ and $F(e^{j\omega})$ be the transmitter and receiver filters, respectively. To maximize SNR in the presence of additive white noise, *matched filters* should be used, that is $F(e^{j\omega}) = H^*(e^{j\omega})$. With this, the zero ISI property becomes $|H(e^{j\omega})|^2|_{\downarrow M} = 1$, which is nothing but a Nyquist(M) property. Such optimally designed filters are used, for example, in voiceband data modem applications [25].

B. Notations and Terminology

- 1) The notation $\tilde{X}(z)$ denotes the z transform of $x^*(-n)$, where $*$ stands for complex conjugation. If $x(n)$ is real, then $\tilde{X}(z) = X(z^{-1})$. Notice that $\tilde{X}(z) = X^*(1/z^*)$, and the FT of $x^*(-n)$ is $X^*(e^{j\omega})$.
- 2) The symbols $\downarrow M$ and $\uparrow M$ denote M -fold decimation and expansion, as defined in [10]. The notation $X(z)|_{\downarrow M}$ denotes the z transform of the decimated sequence $x(Mn)$.
- 3) *Nyquist(M) Property*: A sequence $x(n)$ is said to be Nyquist(M) if $x(Mn) = \delta(n)$ or, equivalently, $X(z)|_{\downarrow M} = 1$. This can be rewritten in the form [10] $\sum_{k=0}^{M-1} X(zW^k) = M$, where $W = e^{-j2\pi/M}$.
- 4) The notation $x_L(n)$ stands for a periodic sequence with periodicity L . If there is a reference to a finite sequence $x(n)$ as well, then it is to be understood that $x_L(n)$ is the periodical expansion of $x(n)$, i.e., $x_L(n) = \sum_{i=-\infty}^{\infty} x(n+Li)$. The Fourier series coefficients (FSC) of $x_L(n)$ is denoted by $X_L(k)$. For L a multiple of M , a periodic sequence $x_L(n)$ is said to be Nyquist(M) if $x_L(Mn) = \delta_K(n) \triangleq \sum_{i=-\infty}^{\infty} \delta(n+Ki)$, where $K = L/M$.
- 5) *Positive Definite Sequences*: Let a sequence $\{x(n), n = 0, \dots, N\}$ be given, and let \mathbf{P} be the Hermitian Toeplitz matrix whose first row is $[x(0) \ x(1) \ \dots \ x(N)]$. The sequence $\{x(n)\}$ is called positive definite if \mathbf{P} is positive definite. Let $[a(0) \ a(1) \ \dots \ a(N)]^T$ denote an eigenvector corresponding to the maximum eigenvalue of \mathbf{P} . Then, the filter $A(z) = \sum_{n=0}^N a(n)z^{-n}$ is called a *maximal eigenfilter* of \mathbf{P} . The definitions for negative definite sequences and minimal eigenfilters are analogous.

C. New Results and Outline of the Paper

In Section II, we formulate the optimum FIR energy compaction problem and present a brief review of existing work. The remaining sections contain new results. In Section III, we give an extension of the technique in [26] for the analytical solution of the FIR energy compaction problem in the two-channel case. This is equivalent to the problem of the optimal two-channel orthonormal filter bank that maximizes the coding gain and with a trivial extension (constraining some zeros at $\omega = \pi$) to the optimum wavelet generating filter problem. The method involves Levinson recursion and two spectral factorizations of half the filter order. We will see that the *analytical method* is related to the well-known line-spectral theory in signal processing [27].

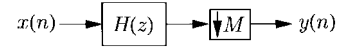


Fig. 2. M -channel compaction filter. $|H(e^{j\omega})|^2$ is Nyquist(M).

We develop a new technique called the *window method* for the design of FIR compaction filters for the M -channel case (Section IV). The window method has the advantage that no optimization tools or iterative numerical techniques are necessary. The solution is generated in a finite number of elementary steps, the crucial step being a simple comparison operation on a *finite frequency grid*. In Section V, we briefly review the linear programming (LP) method and mention some of its drawbacks. Comparison of the window and LP methods is done in Section IV-B. *Matlab programs can be found at our webpage [28] for the algorithms described in this paper.*

The three techniques (the analytical method, the window method, and the LP method) are complementary rather than in competition with each other. For the two-channel case, the analytical method should be the choice whenever it is successful. If it is not or if $M > 2$, for high filter orders, the window method should be preferred. If the filter orders are low, then linear programming should also be considered, although sometimes, the window method performs as well as LP, even for low filter orders (see Example 12).

II. THE FIR ENERGY COMPACTION PROBLEM

An FIR filter $H(z)$ of order N will be called a valid *compaction filter* for the pair (M, N) if $|H(e^{j\omega})|^2$ is Nyquist(M), that is, $|H(e^{j\omega})|^2|_{\downarrow M} = 1$. Let $G(e^{j\omega}) = |H(e^{j\omega})|^2$. We will call $G(z)$ the *product filter* corresponding to $H(z)$. Conversely, $G(z)$ is the product filter of a valid compaction filter for the pair (M, N) if it is of *symmetric order* N , that is $G(z) = \sum_{n=-N}^N g(n)z^{-n}$, and satisfies the following two conditions:

$$\begin{aligned} g(Mn) &= \delta(n) \text{ (Nyquist(M) condition)} \quad \text{and} \\ G(e^{j\omega}) &\geq 0 \text{ (non-negativity)}. \end{aligned} \quad (6)$$

Now, consider Fig. 2, where $H(z)$ is applied to a zero-mean WSS input $x(n)$ with psd $S_{xx}(e^{j\omega})$, and the output is decimated by M . The optimum FIR compaction problem is to find a valid compaction filter $H(z)$ for the pair (M, N) such that the variance σ_y^2 of $y(n)$ is maximized. Since decimation of a WSS process does not alter its variance, we have

$$\begin{aligned} \sigma_y^2 &= \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi} \\ &= \int_{-\pi}^{\pi} G(e^{j\omega}) S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi}. \end{aligned} \quad (7)$$

We define the *compaction gain* as

$$G_{\text{comp}}(M, N) = \frac{\sigma_y^2}{\sigma_x^2} = \frac{\int_{-\pi}^{\pi} G(e^{j\omega}) S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi}}{\int_{-\pi}^{\pi} S_{xx}(e^{j\omega}) \frac{d\omega}{2\pi}}. \quad (8)$$

The aim therefore is to maximize the compaction gain under the constraints (6).

Two Extreme Special Cases: Let us consider the two special cases:

- a) the case where $N < M$;
- b) the ideal case $N = \infty$.

In the first case, the condition $g(Mn) = \delta(n)$ is the same as $g(0) = 1$. This is equivalent to saying that $H(e^{j\omega})$ has unit energy. Let \mathbf{h} be $(N+1) \times 1$ vector formed by $h(n)$, and let \mathbf{R}_{xx} be the $(N+1) \times (N+1)$ autocorrelation matrix of $x(n)$. Then, the problem is to maximize $\mathbf{h}^\dagger \mathbf{R}_{xx} \mathbf{h}$ subject to the condition $\mathbf{h}^\dagger \mathbf{h} = 1$. By Rayleigh's principle [16], the optimum \mathbf{h} is the *maximal eigenvector* of \mathbf{R}_{xx} . In other words, $H(z)$ is the maximal eigenfilter of \mathbf{R}_{xx} . Let the maximum eigenvalue of \mathbf{R}_{xx} be denoted by $\lambda_{\max}\{r(n)\}_0^N$, where $r(n)$ is the input autocorrelation sequence. Then, the optimum compaction gain is $\lambda_{\max}\{r(n)\}_0^N / \sigma_x^2$. The second case has the following solution [6], [7], [9]: If we write $H(z)$ in polyphase form [10], $H(z) = \sum_{k=0}^{M-1} z^{-k} E_k(z^M)$ and if $\mathbf{S}_{xx}(e^{j\omega})$ denotes the $M \times M$ psd matrix of $x(n)$, then for each ω , $\mathbf{e}(e^{j\omega}) = [E_0(e^{j\omega}) \cdots E_{M-1}(e^{j\omega})]^\dagger$ is the *maximal eigenvector* of $\mathbf{S}_{xx}(e^{j\omega})$. Equivalently, for each $\omega \in [0, (2\pi/M))$, let $\mathbf{S}_{xx}(e^{j(\omega+i_0(2\pi/M))})$ be the maximum of the set $\{\mathbf{S}_{xx}(e^{j(\omega+i(2\pi/M))}), i = 0, 1, \dots, M-1\}$. Then, $H(e^{j(\omega+i_0(2\pi/M))}) = \sqrt{M}$, and $H(e^{j(\omega+i(2\pi/M))}) = 0$ for $i \neq i_0$. Note that the eigenvalues of $\mathbf{S}_{xx}(e^{j\omega})$ are $\{\mathbf{S}_{xx}(e^{j(\omega+i(2\pi/M))}), i = 0, 1, \dots, M-1\}$. Let $G_{\text{ideal}}(M)$ denote the corresponding compaction gain. We can write $G_{\text{ideal}}(M) = M \int_{\Omega} \mathbf{S}_{xx}(e^{j\omega})(d\omega/2\pi) / \sigma_x^2$, where Ω is the passband of $H(e^{j\omega})$.

In the N th-order FIR compaction problem, we do not have the flexibility of assigning values to $H(e^{j\omega})$ independently for each ω . This is because $H(e^{j\omega})$ is determined by its $N+1$ frequency samples. For $N > M$, the problem is not an eigenfilter problem either, as the condition $g(Mn) = \delta(n)$ implies more than the simple unit-energy condition. In Section IV, we will introduce a suboptimal method called the window method. Interestingly enough, the method involves two stages that can be associated with the above special cases. Although the method is suboptimal, it produces compaction gains that are very close to the optimum ones, especially for high filter orders.

Upper Bounds on the Compaction Gain: We have the following bounds for the compaction gain.

$$G_{\text{opt}}(M, N) \leq \lambda_{\max}\{r(n)\}_0^N, \\ G_{\text{opt}}(M, N) \leq G_{\text{ideal}}(M), \quad \text{and} \quad G_{\text{opt}}(M, N) \leq M. \quad (9)$$

For the first inequality, let k be an integer such that $kM > N$. From the first special case above, we have $G_{\text{opt}}(kM, N) = \lambda_{\max}\{r(n)\}_0^N$. Since Nyquist(M) property implies Nyquist(kM) property, $G_{\text{opt}}(M, N) \leq G_{\text{opt}}(kM, N)$. The second inequality follows because $G_{\text{opt}}(M, N) \leq G_{\text{opt}}(M, N+1)$. For the last inequality, first observe that $G(e^{j\omega}) \leq M$. Hence, $\sigma_y^2 = \int_{-\pi}^{\pi} G(e^{j\omega}) \mathbf{S}_{xx}(e^{j\omega})(d\omega/2\pi) \leq M \int_{-\pi}^{\pi} \mathbf{S}_{xx}(e^{j\omega})(d\omega/2\pi) = M\sigma_x^2$. The equality holds if and only if $G(e^{j\omega}) = M$ for all ω for which $\mathbf{S}_{xx}(e^{j\omega}) \neq 0$. If $\mathbf{S}_{xx}(e^{j\omega})$ is not line spectral, this requires $G(e^{j\omega})$ to

be identically zero for some region of frequency that is impossible since the order is assumed to be finite. Hence, for a process that is not line spectral, the last inequality is strict. That is, $G_{\text{opt}}(M, N) < M$. For $M = 2$, we will derive another upper bound for $G_{\text{opt}}(2, N)$ in Section III-B [see (23)] for a class of random processes. Whenever the analytical method of Section III succeeds, this bound is in fact achieved.

A. Previous Work

Here is a brief review of the existing methods for FIR compaction filter design.

- 1) *Lattice Parametrization:* Two-channel real-coefficient orthonormal filter banks can be completely parametrized by a lattice structure [10]. Each stage in the lattice has an angle parameter θ_k . The objective function, however, is highly nonlinear function of these angles. Delsarte *et al.* [4] propose an iterative algorithm called the ring algorithm to optimize the lattice stage by stage. Taubman and Zakhor [22] propose an algorithm aimed at finding a globally optimum solution for small filter orders. They extend the results to two-dimensional (2-D) nonseparable filters as well.
- 2) *Quadratically Constrained Optimization:* One can formulate the problem in terms of the compaction filter impulse response $h(n)$: Maximize $\mathbf{h}^\dagger \mathbf{R}_{xx} \mathbf{h}$ subject to $\mathbf{h}^\dagger \mathbf{A}^i \mathbf{h} = \delta(i)$, $i = 0, \dots, K$, where \mathbf{A} is an appropriately chosen singular matrix, and $\mathbf{A}^0 = \mathbf{I}$. Here, \mathbf{R}_{xx} is the input autocorrelation matrix, and \mathbf{h} is the vector of filter coefficients. The authors in [29] and [30] use Lagrangian techniques to solve the problem for the two-channel case and for small filter orders. Chevillat and Ungerboeck [25] provide an iterative projected gradient algorithm for the M -channel case and for moderate filter orders.
- 3) *Eigenfilter Method:* In [21], the authors design one filter of an M -channel orthonormal filter bank using the so-called eigenfilter method. The objective in their design is to have a good frequency response. However, one can modify the technique to incorporate the input statistics. This can be done by using the psd $\mathbf{S}_{xx}(e^{j\omega})$ as a weighting function in the optimization. The paper also discusses how to design a good orthonormal filter bank using the remaining degrees of freedom. In [20], Moulin *et al.* show how to use this idea for the statistical optimization of orthonormal filter banks.
- 4) *Linear Programming:* The objective is a linear function of the impulse response $g(n)$ of $G(e^{j\omega}) = |H(e^{j\omega})|^2$. The Nyquist(M) property can be trivially achieved. However, we need to impose $G(e^{j\omega}) \geq 0$ for all ω . This can be written as a linear inequality for each ω in terms of $g(n)$. Hence, the problem is a linear programming (LP) problem with infinitely many inequality constraints, hence, we have the name semi-infinite programming (SIP). Although we used LP independently to design compaction filters at the early stages of this project, it was first proposed and examined in depth by Moulin *et al.* [20], [31], [32].

5) *Analytical Methods*: Aas *et al.* [26] worked on a closely related problem for the two-channel case. They have constructed a Nyquist(2) real filter $H(e^{j\omega})$ that maximizes the baseband energy $\int_{-\pi/2}^{\pi/2} |H(e^{j\omega})|^2 (d\omega/2\pi)$. Based on the fundamentals of Gaussian quadrature, the authors were able to obtain an analytical method to identify the unit-circle zeros of $H(z)$ that uniquely determine it. In our paper, this method will be referred to as *analytical method*. In Section III, we present extensions of the analytical method. Although the original method primarily addresses conventional half-band filter design, we will show how to adapt the idea for the case of FIR compaction filter design for a given input psd. Interestingly enough, we shall show that the analytical method is related to the well-known line-spectral theory in the signal processing community [27]. An analytical expression for the compaction gain for $N = 3$ is presented in [33]. See also [34] for $N = 3$.

The major disadvantage of the first three methods is that they are iterative and that there is a possibility of reaching a locally optimum solution. Nonlinearity of the objective is very severe in the first technique. A milestone in the design approaches is the formulation of the problem in terms of the product filter. This is done in the last two methods above. In this paper, we also design the product filters. A spectral factorization step is necessary to find the compaction filter coefficients in contrast with the first three methods. In a newly developed technique, Tuqan and Vaidyanathan [35] use state space theory to cast the problem into a semi-definite programming problem. The formulation is such that the spectral factorization is automatically achieved within the algorithm. In [36], we consider the design of FIR compaction filters in multiple stages. This is efficient both in terms of design and implementation complexity. Some of the results of this paper have been presented at recent conferences [37], [38].

III. ANALYTICAL METHOD FOR $M = 2$

In this section, we consider the special case of two channels ($M = 2$) and assume that the input $x(n)$ is real so that the compaction filter coefficients $h(n)$ can be assumed to be real. For this two-channel case, we will show that the optimal product filter $G(e^{j\omega})$ can sometimes be obtained using an analytical method instead of going through a numerical optimization procedure. We will also present a number of examples that demonstrate the usefulness of the method. Examples where the analytical method can be shown to fail are also presented. As in [26], one can modify the algorithms of this section to constrain the filters to have specified number of zeros at $\omega = \pi$ to generate optimal wavelets.

The analytical method is motivated by the fact that under some conditions to be explained, the objective function (7) can be conveniently expressed as a summation over a finite number of frequencies determined by the psd $S_{xx}(e^{j\omega})$. The summation involves the samples of a modified polyphase component of $G(e^{j\omega})$. This will allow us to optimize the modified polyphase component and, hence $G(z)$, essentially *by inspection*. Using these observations, we come up with

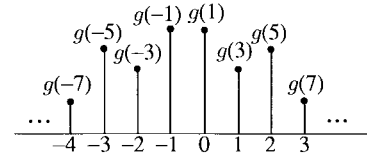


Fig. 3. Coefficients of the polyphase component $E_1(z)$ of the product filter $G(z)$. Because of the symmetry $g(n) = g(-n)$, we have $E_1(-1) = 0$.

an algorithm that determines the unit-circle zeros of the compaction filter. Using the Nyquist(2) condition, this in turn determines the filter itself.

The inspiration for our work in this section comes from the recent contribution by Aas *et al.* [26], where the Gaussian quadrature technique is cleverly used to address the problem of maximizing the baseband energy of halfband filters. Our work in this section differs in a number of respects. First, we do not use Gaussian quadrature but take advantage of an elegant representation for positive definite sequences that results from the theory of line-spectral processes. Second, we take into account the knowledge of the input psd in the optimization process. We give the analytical solutions for some practically important classes of random processes.

Let us represent the product filter $G(z) = \sum_{n=-N}^N g(n)z^{-n}$ in the traditional polyphase form [10] for $M = 2$: $G(z) = E_0(z^2) + z^{-1}E_1(z^2)$. By the Nyquist(2) property, we have $E_0(z) = 1$. For the real coefficient case, we have $g(n) = g(-n)$, and it follows that the coefficients of the FIR filter $E_1(z)$ have the symmetry demonstrated in Fig. 3. This implies, in particular, that $E_1(z) = 0$ for $z = -1$. By factoring the zero at $z = -1$, we can write $E_1(z) = (1+z)G_1(z)$, where $G_1(z)$ has symmetric real coefficients. Hence, we can write

$$G(z) = 1 + (z + z^{-1})G_1(z^2), \quad \text{i.e.,} \\ G(e^{j\omega}) = 1 + 2\cos\omega G_1(e^{j2\omega}). \quad (10)$$

Since Nyquist condition and nonnegativity of $G(e^{j\omega})$ together imply $0 \leq G(e^{j\omega}) \leq 2$, the modified polyphase component $G_1(e^{j\omega})$ is bounded as

$$-\frac{1}{2\cos(\omega/2)} \leq G_1(e^{j\omega}) \leq \frac{1}{2\cos(\omega/2)}, \quad -\pi < \omega < \pi. \quad (11)$$

Notice that $G(z)$ and $G_1(z)$ can be determined from each other uniquely. We shall express the output variance σ_y^2 in terms of $G_1(e^{j\omega})$ so that we can see how to optimize the coefficients of $G_1(z)$. For this, write the input psd in the traditional polyphase form as $S_{xx}(z) = S_0(z^2) + z^{-1}S_1(z^2)$. Then, σ_y^2 can be simplified into the form $\sigma_y^2 = r(0) + \int_{-\pi}^{\pi} G_1(e^{j\omega})\Psi_x(e^{j\omega})(d\omega/2\pi)$, where $\Psi_x(z) = (1+z^{-1})S_1(z)$, or equivalently

$$\Psi_x(e^{j\omega}) = \cos(\omega/2)(S_{xx}(e^{j\omega/2}) - S_{xx}(e^{j(\pi-\omega/2)})). \quad (12)$$

Using Parseval's relation, the objective can be written as

$$\sigma_y^2 = r(0) + \sum_{n=-(N-1)/2}^{(N-1)/2} g_1(n)\psi_x(n) \quad (13)$$

where $\psi_x(n)$ is the inverse transform of $\Psi_x(z)$, which is produced below explicitly for convenience.

$$\begin{aligned}\psi_x(0) &= 2r(1), \\ \psi_x(1) &= r(1) + r(3), \dots \\ \psi_x\left(\frac{N-1}{2}\right) &= r(N-2) + r(N)\end{aligned}\quad (14)$$

and $\psi_x(n) = \psi_x(-n)$. Our aim is to maximize the second term in (13) for fixed $\psi_x(n)$ (i.e., fixed input) by choosing $g_1(n)$ under the constraint (11) and the usual filter-order constraint. Under the assumption that the input-dependent sequence $\psi_x(n)$ is *positive or negative definite* (see Section I-B for definition), we will show how this can be done analytically. The significance of this assumption on $\psi_x(n)$ is explained in Section III-D. We will need the representation theorem of the next section for positive definite sequences.

A. Representation of Positive Definite Sequences

Theorem 1: Given a positive definite sequence of $m+1$ complex numbers $\{\phi(n), n = 0, \dots, m\}$, there exists a representation of the form

$$\phi(n) = \sum_{k=0}^m \alpha_k e^{j\omega_k n}, \quad n = 0, \dots, m \quad (15)$$

where $\alpha_k > 0, k = 0, \dots, m$, and ω_k 's are all distinct.

Comments: Note that this is different from the Caratheodory representation theorem, which is the basis for the Pisarenko method [39] for identifying sinusoidal signals under noise: Given $\{\phi_n, n = 1, \dots, m\}$, there exists a representation of the form $\phi_n = \sum_{k=1}^m \alpha_n e^{j\omega_k n}, n = 1, \dots, m$, where α_n 's are nonnegative. The frequencies ω_n 's are the angles of the unit magnitude roots of the minimal eigenpolynomial of a matrix \mathbf{Q} . The matrix \mathbf{Q} is $(m+1) \times (m+1)$ Hermitian Toeplitz with the first row $[\phi_0 \ \phi_1 \ \dots \ \phi_m]$, where $\phi_0 = \sum_{k=1}^m \alpha_k$ is the positive number that makes the matrix singular. Here, the number of distinct frequencies depends on the multiplicity of the minimum eigenvalue of the so-obtained matrix. If the multiplicity is 1, there are m distinct frequencies. If we start with a positive definite sequence $\phi_n, n = 0, \dots, m$, then Caratheodory representation takes the form

$$\begin{aligned}\phi_n &= \left(\phi_0 - \sum_{k=1}^m \alpha_k \right) \delta(n) \\ &+ \sum_{k=1}^m \alpha_k e^{j\omega_k n}, \quad n = 0, \dots, m.\end{aligned}\quad (16)$$

This is obviously not the same as (15) and is not suitable for our purposes. Although Theorem 1 turns out to be well known in the literature [40], we include our proof below for two reasons.

- 1) It is elegant and uses the theory of line-spectral processes.
- 2) It reveals us the algorithmic steps of the analytical method.

Proof of Theorem 1: Let \mathbf{P} be the $(m+1) \times (m+1)$ Hermitian Toeplitz matrix whose first row is $\Phi^T = [\phi(0) \ \phi(1) \ \dots \ \phi(m)]$. Consider the extension of \mathbf{P} into a singular $(m+2) \times (m+2)$ Hermitian Toeplitz matrix $\hat{\mathbf{P}}$ such that its $(m+1) \times (m+1)$ principal submatrix is \mathbf{P} . This extension is merely augmenting an extra element $\phi(m+1)$ to the end of Φ and forming the corresponding Hermitian Toeplitz matrix. The number $\phi(m+1)$ is chosen to make $\hat{\mathbf{P}}$ singular. This can always be done because of the following reason. For the positive definite matrix \mathbf{P} , one can run the well-known Levinson recursion procedure [27] to obtain the optimal m th-order predictor polynomial $A_m(z)$. If one now considers the following continuation of the recursion $P_c(z) = A_m(z) + cz^{-(m+1)}\hat{A}_m(z)$ with $|c| = 1$, then this corresponds to the singular predictor polynomial of a random process with singular autocorrelation matrix $\hat{\mathbf{P}}$. The result now follows from a well-established fact [27], which states that a WSS process is line spectral with exactly $m+1$ lines if and only if its $(m+1) \times (m+1)$ autocorrelation matrix is nonsingular and $(m+2) \times (m+2)$ autocorrelation matrix is singular. Applying this result to a process with autocorrelation matrix $\hat{\mathbf{P}}$, we get (15). ■

Remarks: It is clear that $P_c(z)$ defined in the above proof is also the minimal eigenfilter of $\hat{\mathbf{P}}$. The zeros of $P_c(z)$ are all on the unit circle and distinct. Let $\{e^{j\omega_k}, k = 0, \dots, m\}$ be these zeros. The distinct frequencies $\{\omega_k, k = 0, \dots, m\}$ are referred to as the line-spectral frequencies, and α_k is the power at the frequency ω_k . The representation (15) is not unique because of the nonuniqueness of the unit-magnitude constant c in the proof.

Real Case: For real $x(n)$, the predictor polynomial $A_m(z)$ and the constant c are real. Hence, we have two cases: $c = \pm 1$. The case $c = 1$ leads to a symmetric polynomial $P_1(z)$, whereas the case $c = -1$ leads to an antisymmetric polynomial $P_{-1}(z)$. It is a well-known fact that the distinct unit-circle zeros of these two polynomials are interleaved. For simplicity, assume that m is odd. Then, $P_{-1}(z)$ has both of the zeros $z = 1$ and $z = -1$, and $P_1(z)$ has none of them. Using $P_1(z)$, we have the following representation for a real positive definite sequence $\phi(n)$:

$$\phi(n) = \sum_{k=0}^{(m-1)/2} \beta_k \cos \omega_k n, \quad n = 0, \dots, m \quad (17)$$

where $\beta_k > 0, k = 0, \dots, (m-1)/2$, and ω_k 's are all distinct, and $0 < \omega_k < \pi, k = 0, \dots, (m-1)/2$.

B. Derivation of the Analytical Method

Assume for simplicity that $(N-1)/2$ is odd, and assume that $\{\psi_x(n), n = 0, \dots, (N-1)/2\}$ is positive definite. Applying the real form of the representation, we have

$$\psi_x(n) = \sum_{k=0}^{(N-3)/4} \beta_k \cos \omega_k n, \quad n = 0, \dots, \frac{N-1}{2}. \quad (18)$$

The objective (13) can therefore be written as

$$\begin{aligned}\sigma_y^2 &= r(0) + \sum_{k=0}^{(N-3)/4} \beta_k \sum_{n=-(N-1)/2}^{(N-1)/2} g_1(n) \cos \omega_k n \\ &= r(0) + \sum_{k=0}^{(N-3)/4} \beta_k G_1(e^{j\omega_k}).\end{aligned}\quad (19)$$

From (11), the output variance (19) is maximized if

$$G_1(e^{j\omega_k}) = \frac{1}{2 \cos(\omega_k/2)}, \quad k = 0, \dots, (N-3)/4. \quad (20)$$

This implies $G(e^{j\omega_k/2}) = 2$, and by Nyquist(2) property

$$G(e^{j(\pi-\omega_k/2)}) = 0, \quad k = 0, \dots, (N-3)/4. \quad (21)$$

Notice that these zeros are all located in the region $(\pi/2, \pi)$. Since $0 \leq G(e^{j\omega}) \leq 2$, the derivatives of $G(e^{j\omega})$ should vanish at the above frequencies. Hence, we should have $G'(e^{j\omega_k/2}) = 0, k = 0, \dots, (N-3)/4$. In view of (10), this in turn implies

$$G'_1(e^{j\omega_k}) = \frac{\sin(\omega_k/2)}{4 \cos^2(\omega_k/2)}, \quad k = 0, \dots, (N-3)/4. \quad (22)$$

From the two sets of constraints (20) and (22), $G_1(z)$ is determined uniquely. To see this, note that $G_1(e^{j\omega}) = g_1(0) + 2 \sum_{n=1}^{(N-1)/2} g_1(n) \cos \omega n$ is a polynomial in $x = \cos \omega$ of degree $(N-1)/2$. Since ω_k 's are all distinct and $0 < \omega_k < \pi$, the constraints (20) and (22) translate into a similar set of constraints for $G_1(x)$ and $G'_1(x)$ and by simple Hermite interpolation [41, p. 28], $G_1(x)$ is determined uniquely. The corresponding solution $G(e^{j\omega})$ is necessarily nonnegative in the frequency region $[\pi/2, \pi]$ (Appendix A). If it is nonnegative in the region $[0, \pi/2]$ as well, then it is the optimum compaction filter with the corresponding compaction gain

$$G_{\text{opt}}(2, N) = 1 + \frac{\sum_{k=0}^{(N-3)/4} \frac{\beta_k}{2 \cos(\omega_k/2)}}{r(0)}. \quad (23)$$

If, however, $G(e^{j\omega})$ turns out to be negative at some frequencies in $[0, \pi/2]$, then it is not a valid solution, and the above RHS is only an upper bound for $G_{\text{opt}}(2, N)$. Assume that $G(e^{j\omega})$ obtained by the method is indeed non-negative. Then, it is *the unique solution*. To see this, assume there is another optimal product filter $K(z)$. Assume that $K_1(z)$ is its modified polyphase component. Then, there exists a frequency ω_k among the line-spectral frequencies such that $K_1(e^{j\omega_k}) < 1/2 \cos(\omega_k/2)$. Hence, the summation (19) for $K_1(e^{j\omega})$ is necessarily less than that for $G_1(e^{j\omega})$, resulting in contradiction. Notice finally that $H(z)$, which is an arbitrary spectral factor of the unique solution $G(z)$, is not unique.

C. Completion of the Optimal $G(z)$

Consider the following factorization of $G(z)$.

$$G(z) = \hat{G}_0(z) \hat{G}_1(z) \quad (24)$$

where $\hat{G}_0(z)$ contains the unit-circle zeros determined by the above procedure. From (21), we have

$$\hat{G}_0(z) = \prod_{k=0}^{(N-3)/4} (z + 2 \cos(\omega_k/2) + z^{-1})^2. \quad (25)$$

Using the Nyquist(2) property, it is possible to determine $\hat{G}_1(z)$ and, hence, $G(z)$. For this, let $\hat{g}_0(n)$ and $\hat{g}_1(n)$ be the impulse responses of $\hat{G}_0(z)$ and $\hat{G}_1(z)$, respectively. The product (24) in z domain is equivalent to the convolution in time domain. Using the convolution matrix and taking into account the symmetries, we get

$$\mathbf{g} = \mathbf{A} \hat{\mathbf{g}}_1 \quad (26)$$

where the vectors $\mathbf{g}, \hat{\mathbf{g}}_1$ have the components $g_n = g(2n), \hat{g}_{1n} = g_1(n), n = 0, \dots, (N-1)/2$, and \mathbf{A} is obtained from the impulse response $g_0(n)$. From the Nyquist(2) property, it is clear that $\mathbf{g} = [1 \ 0 \ \dots \ 0]^T$. Hence, $\hat{\mathbf{g}}_1$ is the first column of the matrix \mathbf{A}^{-1} . To see that \mathbf{A} is invertible, it suffices to show that a unique solution to $\hat{\mathbf{g}}_1$ exists for a given $g_0(n)$. For this, write the Nyquist(2) condition for $G(z)$.

$$\hat{G}_0(z) \hat{G}_1(z) + \hat{G}_0(-z) \hat{G}_1(-z) = 2. \quad (27)$$

The zeros of $\hat{G}_0(z)$ lie on the left half of the unit circle. Hence, the zeros of $\hat{G}_0(-z)$ lie on the right half of the unit circle. This implies that $\hat{G}_0(z)$ and $\hat{G}_0(-z)$ are coprime. It is now easy to show that a unique solution to $\hat{G}_1(z)$ of symmetric degree less than or equal to $(N-1)/2$ exists [26]. Actually, this is an efficient way of determining $\hat{G}_1(z)$ (see [26] for details).

Efficient Determination of $\hat{G}_0(z)$: We will show that we can obtain $\hat{G}_0(z)$ from the singular predictor polynomial $P_1(z)$ without having to find its roots. For this, let us write $P_1(z)$ explicitly:

$$\begin{aligned}P_1(z) &= dz^{-(N+1)/2} \prod_{k=0}^{(N-3)/4} (z - e^{j\omega_k})(z - e^{-j\omega_k}) \\ &= dz^{-(N+1)/4} \prod_{k=0}^{(N-3)/4} (z - 2 \cos \omega_k + z^{-1}).\end{aligned}\quad (28)$$

Now, consider the upsampled polynomial $P_1(z^2)$. This can be written in the form $P_1(z^2) = \pm P_0(z) P_0(-z)$, where $P_0(z)$ is a polynomial in z^{-1} of order $(N+1)/2$ with all its zeros in the left half plane. To be explicit

$$P_0(z) = z^{-(N+1)/4} \prod_{k=0}^{(N-3)/4} (z + 2 \cos(\omega_k/2) + z^{-1}). \quad (29)$$

Hence, from (25), it follows that $\hat{G}_0(z) = z^{(N+1)/2} P_0^2(z)$. Therefore, given the singular predictor polynomial $P_1(z)$, one can apply a continuous-time spectral factorization algorithm [42] to $P_1(z^2)$ to obtain $P_0(z)$ and, therefore, $\hat{G}_0(z)$. Since $G(z)$ can be determined from $\hat{G}_0(z)$, we observe that there is no need to find the roots of $P_1(z)$.

Spectral Factorization: To find the compaction filter $H(z)$, we need to spectrally factorize $G(z)$. It is clear that we can write $H(z)$ as

$$H(z) = H_0(z)H_1(z)$$

where $H_0(z)$ and $H_1(z)$ are the spectral factors of $\hat{G}_0(z)$ and $\hat{G}_1(z)$, respectively. We can deduce $H_0(z)$ immediately: $H_0(z) = P_0(z)$. Hence, all we need to do is to determine $H_1(z)$, which is of order $(N-1)/2$. This can be done by a discrete-time spectral factorization of $\hat{G}_1(z)$ [43]. Although the phase of the compaction filter is immaterial for the compaction gain, it is important in the design of an optimal orthonormal filter bank for subband coding [19]. For some applications like image coding, the linear-phase property might be important. Although it is not possible to have linear-phase compaction filter in the two-channel case [10], one can achieve close-to-linear-phase response by a careful grouping of the roots of $\hat{G}_1(z)$.

The case where $(N-1)/2$ is even can be treated in a very similar manner. In this case, we use the singular polynomial $P_{-1}(z)$ corresponding to $c = -1$, and one of the line-spectral frequencies is 0, that is, $z = 1$ is a root of $P_{-1}(z)$. The resulting product filter $G(e^{j\omega})$ continues to be non-negative in $[\pi/2, \pi]$. We skip the details and give the summary of the algorithm for both cases.

Summary of the Analytical Method: Given the autocorrelation sequence $r(n)$, $n = 0, \dots, N$, where N is odd, let $m = (N-1)/2$. First, obtain the sequence $\psi_x(n)$, $n = 0, \dots, m$ using the relations (14). If this sequence is positive definite, then go to Step 1.

Step 1: Calculate $A_m(z)$, which is the optimum predictor polynomial of order m corresponding to the sequence $\psi_x(n)$ and obtain $P_c(z)$ from $P_c(z) = A_m(z) + cz^{-(m+1)}A_m(z^{-1})$, where $c = 1$ if m is odd, and $c = -1$ otherwise.

Step 2: Obtain the spectral factor, $P_0(z)$ of $P_c(z^2)$ using a continuous time spectral factorization algorithm, and determine $\hat{G}_0(z) = z^{(m+1)}P_0^2(z)$.

Step 3: Calculate $\hat{G}_1(z)$ using (26) or (27), and find its spectral factor $H_1(z)$. The optimum compaction filter is $H(z) = P_0(z)H_1(z)$.

See our webpage [28] for a matlab program that implements the algorithm.

Decorrelation in Optimal Subband Coding: Let us form a two-channel orthonormal filter bank by letting the first filter be the optimal FIR compaction filter $H(z)$ designed above and by having the second filter as [10]

$$F(z) = z^{-N}\tilde{H}(-z). \quad (30)$$

Let $S_{x_0x_1}(z)$ be the cross spectral density of the subband signals after decimation. Then, we have

$$\begin{aligned} S_{x_0x_1}(z) &= [S_{xx}(z)H(z)\tilde{F}(z)]_{\downarrow 2} \\ &= [z^N S_{xx}(z)H(z)H(-z)]_{\downarrow 2} \\ &= [z^N S_{xx}(z)P_c(z^2)H_1(z)H_1(-z)]_{\downarrow 2} \\ &= [z^N S_{xx}(z)\hat{G}_1(z)]_{\downarrow 2} P_c(z). \end{aligned} \quad (31)$$

Hence, we have

$$S_{x_0x_1}(e^{j\omega_k}) = 0, \quad k = 0, \dots, m \quad (32)$$

TABLE I

OPTIMUM COMPACTION FILTER COEFFICIENTS $h(n)$ AND THE CORRESPONDING COMPACTION GAINS FOR AR(1) PROCESS WITH $\rho = 0.1, 0.5$, AND 0.9 . THE FILTER ORDER IS $N = 3$, AND THE NUMBER OF CHANNELS IS $M = 2$

$\rho = 0.1$

n	analytical method	window method	linear programming
0	0.5494144350	0.6940928372	0.5839818982
1	0.7789293967	0.7136056607	0.7658293099
2	0.2470689810	0.0680132766	0.2140666953
3	-0.1742690225	-0.0661535225	-0.1632362113
xcompaction gain	1.1153	1.1078	1.1151

$\rho = 0.5$

n	analytical method	window method	linear programming
0	0.5308991349	0.6817974052	0.5693221037
1	0.7963487023	0.7258587819	0.7821354608
2	0.2411149862	0.0663296736	0.2047520302
3	-0.1607433241	-0.0623033026	-0.1490404954
compaction gain	1.5547	1.5283	1.5537

$\rho = 0.9$

n	analytical method	window method	linear programming
0	0.4938994371	0.6550553981	0.5605331011
1	0.8279263239	0.7510864372	0.8017336546
2	0.2281902949	0.0620169861	0.1699982390
3	-0.1361269173	-0.0540877314	-0.1188544843
compaction gain	1.9222	1.9118	1.9207

where ω_k 's are the line-spectral frequencies (see Remarks after Theorem 1). This is the form of decorrelation that takes place in optimal subband coding with FIR filters.

Case where $\psi_x(n)$ Is Negative Definite: From our developments for the positive definite case, and using the sequence $-\psi_x(n)$, it can be proven that the optimum compaction filter is $H(z) = \hat{H}(-z)$, where $\hat{H}(z)$ is the optimum compaction filter for the positive definite sequence $\hat{\psi}_x(n) = -\psi_x(n)$. However, it is easier to see this directly by looking at the objective in time domain: $\sigma_y^2 = r(0) + 2\sum_{n=1}^N g(n)r(n)$. First, note that $\hat{\psi}_x(n)$ corresponds to the autocorrelation sequence $\hat{r}(n) = -r(n)$, $n \neq 0$. Let $g(n)$ and $\hat{g}(n)$ be the product filter coefficients for $H(z)$ and $\hat{H}(z)$, respectively. The objective is then to maximize $\sum_{n=1}^N -g(n)\hat{r}(n)$. This has the solution $-g(n) = \hat{g}(n)$, $n \neq 0$. Hence, we have $G(z) = \hat{G}(-z)$, and therefore, $H(z) = \hat{H}(-z)$.

Example 1: AR(1) Process: Let the input process be AR(1) with the autocorrelation sequence $r(n) = \rho^n$, $0 < \rho < 1$. This is also called Markov-1 process and is a good model for many of the practical signals including images and speech signals [44]. Let the compaction filter order be $N = 3$. Then, $m = 1$, which is odd. We have $\psi_x(0) = 2\rho$ and $\psi_x(1) = \rho(1+\rho^2)$. The Hermitian Toeplitz matrix corresponding to $\{\psi_x(n), n = 0, 1\}$ is $\mathbf{P} = \rho \begin{bmatrix} 2 & 1+\rho^2 \\ 1+\rho^2 & 2 \end{bmatrix}$, which is positive definite. Hence, we can apply the analytical method:

Step 1: Running the Levinson recursion, we have $A_1(z) = 1 - (1 + \rho^2/2)z^{-1}$ and using $c = 1$ (m is odd), we have $P_1(z) = 1 - (1 + \rho^2)z^{-1} + z^{-2}$.

Step 2: By straightforward calculation $P_0(z) = 1 + \sqrt{3 + \rho^2}z^{-1} + z^{-2}$, and $\hat{G}_0(z) = (z + \sqrt{3 + \rho^2} + z^{-1})^2$.

Step 3: Using the Nyquist(2) constraint, we find $\hat{G}_1(z) = -(1/2(3 + \rho^2)^{3/2})(z - 2\sqrt{3 + \rho^2} + z^{-1})$. It is readily verified that $\hat{G}_1(e^{j\omega}) \geq 0$, $\forall \omega$ for all values of ρ . The spectral factor of $\hat{G}_1(z)$ turns out to be $(1/\sqrt{2}(3 + \rho^2)^{3/4})(a + bz^{-1})$, where $a = \sqrt{\sqrt{3 + \rho^2} + \sqrt{2 + \rho^2}}$, and $b = -\sqrt{\sqrt{3 + \rho^2} - \sqrt{2 + \rho^2}}$.

TABLE II
COMPACTION FILTER COEFFICIENTS AND CORRESPONDING GAINS FOR MA(1) PROCESSES FOR $M = 2$

n	N = 3	N = 9	N = 15	N = 21
0	0.5502267080	0.3472380509	0.2619442448	0.2135948251
1	0.7781380728	0.7212669193	0.6444985282	0.5837095513
2	0.2473212614	0.5313628729	0.6197371546	0.6522949264
3	-0.1748825411	-0.0301418144	0.1178983343	0.2237822545
4		-0.2357012104	-0.2498547909	-0.2185003959
5		0.0008621669	-0.1127984531	-0.1849242462
6		0.1250275869	0.1367316336	0.1009864921
7		-0.0141611881	0.0800586128	0.1357184335
8		-0.0608205190	-0.0879123348	-0.0574793351
9		0.0292806975	-0.0512638394	-0.0996883819
10			0.0616834351	0.0386787054
11			0.0272577935	0.0732876341
12			-0.0441106561	-0.0298852666
13			-0.0065141401	-0.0529392251
14			0.0275486991	0.0255464898
15			-0.0111966480	0.0362662187
16				-0.0230508869
17				-0.0216340483
18				0.0206081274
19				0.0077883397
20				-0.0156868986
21				0.0057402527
$\rho = 0.1$	1.1155	1.1244	1.1260	1.1266
compaction gains $\rho = 0.3$	1.3464	1.3732	1.3781	1.3798
$\rho = 0.5$	1.5774	1.6220	1.6301	1.6330

Step 4: The optimum compaction filter is

$$\begin{aligned}
 H(z) &= P_0(z)H_1(z) \\
 &= \frac{1}{\sqrt{2}(3 + \rho^2)^{3/4}}(a + (b + a\sqrt{3 + \rho^2})z^{-1} \\
 &\quad + (a + b\sqrt{3 + \rho^2})z^{-2} + bz^{-3}). \quad (33)
 \end{aligned}$$

The product filter is $G(z) = (1/2(3 + \rho^2)^{3/2})(-z^3 + 3(2 + \rho^2)z + 2(3 + \rho^2)^{3/2} + 3(2 + \rho^2)z^{-1} - z^{-3})$. If $-1 < \rho < 0$, then the optimum compaction filter is $H(-z)$. The optimum compaction gain for both cases is

$$G_{\text{opt}}(2, 3) = 1 + \frac{2|\rho|}{\sqrt{3 + \rho^2}}. \quad (34)$$

See Table I for the numerical values of the filter coefficients and the compaction gains for various values of ρ . We have found that the analytical method is successful for any filter order N for AR(1) processes.

Example 2—MA(1) Process: Let $N = 3, r(0) = 1, r(1) = \rho > 0$, and $r(n) = 0, n > 1$. The sequence $\psi_x(n)$ is therefore $\psi_x(0) = 2\rho, \psi_x(1) = \rho, \mathbf{P} = \rho \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, which is positive definite. Hence, applying the algorithm, we find $A_1(z) = 1 - \frac{1}{2}z^{-1}, c = 1, P_1(z) = 1 - z^{-1} + z^{-2}, P_0(z) = 1 + 2\cos(\pi/6)z^{-1} + z^{-2} = 1 + \sqrt{3}z^{-1} + z^{-2}, \hat{G}_0(z) = (z + \sqrt{3} + z^{-1})^2, \hat{G}_1(z) = -(\sqrt{3}/18)(z - 2\sqrt{3} + z^{-1}), H_1(z) = 3^{-3/4}2^{-1/2}(\sqrt{\sqrt{3} + \sqrt{2}} - \sqrt{\sqrt{3} - \sqrt{2}}z^{-1})$, and the compaction filter $H(z)$ is

$$\begin{aligned}
 &3^{-3/4}2^{-1/2}(\sqrt{\sqrt{3} + \sqrt{2}} + (\sqrt{3 + \sqrt{6}} - \sqrt{\sqrt{3} - \sqrt{2}})z^{-1} \\
 &\quad + (\sqrt{\sqrt{3} + \sqrt{2}} - \sqrt{3 - \sqrt{6}})z^{-2} - \sqrt{\sqrt{3} - \sqrt{2}}z^{-3}).
 \end{aligned}$$

The product filter is $G(z) = -(\sqrt{3}/18)z^3 + (\sqrt{3}/3)z + 1 + (\sqrt{3}/3)z^{-1} - (\sqrt{3}/18)z^{-3}$. If $\rho < 0$, then the optimal filter is $H(-z)$. The optimum compaction gain for both cases is $G_{\text{opt}}(2, 3) = 1 + (2/\sqrt{3})|\rho|$.

Example 3—MA(1) Process, Arbitrary Order N : Following the steps of the algorithm, we have

$$P_c(z) = 1 - z^{-1} + z^{-2} - \dots + (-1)^{(N+1)/2}z^{-(N+1)/2}.$$

If $(N - 1)/2$ is odd, then the zeros of $P_1(z)$ are $e^{\pm j\omega_k}, \omega_k = (2k - 1)2\pi/(N + 3), k = 1, \dots, (N + 1)/4$. Therefore, the roots of $P_0(z)$ and, hence, the unit-circle zeros of the optimum compaction filter $H(z)$ are

$$e^{\pm j\Omega_k}, \quad \Omega_k = \pi - (2k - 1)\frac{\pi}{N + 3}, \quad k = 1, \dots, \frac{N + 1}{4}.$$

Similarly, if $(N - 1)/2$ is even, the unit-circle zeros of the optimum compaction filter $H(z)$ are $\pi, e^{\pm j\Omega_k}, \Omega_k = \pi - 2k\pi/(N + 3), k = 1, \dots, (N - 1)/4$. The rest of the procedure involves spectral factorization, and it is not easy to see what $H_1(z)$ will be in closed form. However, we note that the algorithm successfully finds the optimum compaction filter for any order N . Table II shows the compaction filters and the corresponding compaction gains for various filter orders. The optimum compaction filter for $\rho < 0$ is $H(-z)$. Note that the filters do not depend on the value of ρ but only on the sign. The optimum compaction gains, on the other hand, depend on ρ .

Example 4—KLT: If $N = 1$, then the algorithm yields $H(z) = (1/\sqrt{2})(1 + z^{-1})$ if $r(1) > 0$ and $H(-z) = (1/\sqrt{2})(1 - z^{-1})$ if $r(1) < 0$. Notice that these correspond to the two-channel transform coder, which is known to be

fixed. The corresponding compaction gain is $G_{\text{opt}}(2, 1) = 1 + |r(1)|/r(0)$. It is also true that the above filters and the corresponding compaction gains are optimal for any psd and for any number of channels M . Hence

$$G_{\text{opt}}(M, 1) = 1 + \frac{|r(1)|}{r(0)}, \quad M \geq 2.$$

If $r(m)$ is maximum of all $r(n)$ where n is not a multiple of M , then one can achieve the compaction gain of $1 + |r(m)|/r(0)$ by using the filter $(1/\sqrt{2})(1 + z^{-m})$ if $r(m) > 0$ and the filter $(1/\sqrt{2})(1 - z^{-m})$ if $r(m) < 0$.

Case Where $\psi_x(n)$ Is Semidefinite: Assume that $\psi_x(n)$ is positive semidefinite. Then, there exists an integer $P < (N - 1)/2$ such that $\{\psi_x(n), n = 0, 1, \dots, P\}$ is positive definite, and $\{\psi_x(n), n = 0, 1, \dots, P+1\}$ is only positive semidefinite. Then, we can replace $(N - 1)/2$ in the above arguments with P and write the objective (19) in terms of $P + 1$ corresponding line-spectral frequencies. This enables us to determine a product filter of symmetric order $2P + 1 < N$. If this resulting filter is nonnegative, then we have found the unique minimum symmetric order product filter that is optimum among the filters of symmetric order less than or equal to N . The case where $\psi_x(n)$ is negative semidefinite is similar, and the details are omitted.

Example 5—Case Where $\psi_x(n)$ Is Positive Semidefinite: Let $N = 3$, $r(0) = 1$ and $r(1) = r(3) = \rho > 0$. Then, $\psi_x(0) = \psi_x(1) = 2\rho$. The associated Toeplitz matrix is $2\rho \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, which is positive semidefinite and singular. The number P is 0 in this case, and the objective (19) is $1 + 2\rho G_1(e^{j0})$. By letting $G_1(e^{j0}) = \frac{1}{2}$, the product filter $G(z)$ of symmetric order 1 can easily be seen to be $\frac{1}{2}z + 1 + \frac{1}{2}z^{-1}$, and it is readily verified that $G(e^{j\omega}) \geq 0$. In fact, this is the KLT solution with the compaction filter $H(z) = (1/\sqrt{2})(1 + z^{-1})$. The corresponding optimum compaction gain is $1 + \rho$. No third-order solution can achieve better gain than this.

D. Characterization of Processes for Which the Analytical Method Is Applicable for All N

For the analytical method to be applicable for all N , the sequence $\psi_x(n)$ has to be positive or negative definite for all N . The sequence $\psi_x(n)$ is positive definite for all N if and only if $\Psi_x(e^{j\omega})$ is not a line spectrum, and $\Psi_x(e^{j\omega}) \geq 0$. Using (12), this is true if and only if $S_{xx}(e^{j\omega})$ is not a line spectrum, and

$$S_{xx}(e^{j\omega}) \geq S_{xx}(e^{j(\pi-\omega)}), \quad \omega \in [0, \pi/2]. \quad (35)$$

We will say that the process is “lowpass” if its psd satisfies the above condition. A nonincreasing psd is an example of this. However, a psd may not be nonincreasing but may still be lowpass. In the ideal case, the optimum compaction filter for that type of process is the ideal halfband lowpass filter [4], [6], [8]. For the case where $\psi_x(n)$ is negative definite for all N , the preceding is replaced with $S_{xx}(e^{j\omega}) \leq S_{xx}(e^{j(\pi-\omega)})$, $\omega \in [0, \pi/2]$. This type of process will be called “highpass” since the ideal halfband highpass filter is

optimum for such a process. Notice that for the algorithm to be applicable for a particular N , it is only necessary that $\psi_x(n), n = 0, \dots, (N - 1)/2$ is positive or negative definite. For a small order N , this corresponds to a much broader class than that of lowpass and highpass processes.

Cases Where the Algorithm Fails: Assume that the process is such that the sequence $\{\psi_x(n), n = 0, \dots, (N - 1)/2\}$ is positive definite, and therefore, the algorithm is applicable for the filter order N . Assume, however, that one of the line-spectral frequencies ω_k is close to π . The algorithm will require $e^{j(\pi-\omega_k/2)}$ to be a zero of $G(z)$. Hence, $G(e^{j\omega})$ will have a zero close to $\pi/2$. However, this may be impossible if the order N is low. To see this, note that $G(e^{j\pi/2}) = 1$ from the Nyquist(2) property, and therefore, requiring $G(e^{j\omega})$ to have a zero close to the frequency $\pi/2$ is the same as requiring a narrow transition band for $G(e^{j\omega})$, which is impossible if the order is not sufficiently high. One can however, increase the filter order to overcome the problem.

Example 6: Let $N = 3$ and $r(n) = \cos \omega_1 n$, $\omega_1 \in [0, \pi/2)$. Hence, $\psi_x(0) = 2 \cos \omega_1$, $\psi_x(1) = \cos 3\omega_1 + \cos \omega_1$, and $\psi_x(n)$ is positive definite. Using the algorithm, we find $\hat{G}_0(z) = (z + 2 \cos \omega_1 + z^{-1})^2$ from which it follows that $\hat{G}_1(z) = -(1/16 \cos^3 \omega_1)(z - 4 \cos \omega_1 + z^{-1})$. This has single unit-circle zeros if $\omega_1 \in (\pi/3, \pi/2)$, and therefore, $\hat{G}_1(e^{j\omega})$ is not non-negative. Hence, the algorithm fails if the impulse is within the $\pi/6$ neighborhood of $\pi/2$. We have designed optimum compaction filters for the above autocorrelation sequence using LP for various values of ω_1 . We have observed that the optimum compaction filters agree with the above analytical solution if $\omega_1 \in (0, \pi/3]$. For the complementary case of $\omega_1 \in (\pi/3, \pi/2)$, where the analytical method fails for $N = 3$, LP yields the solution $G(z) = -\frac{1}{2}z^3 + 1 - \frac{1}{2}z^{-3}$, regardless of the exact value of ω_1 . The factors $\hat{G}_0(z)$ and $\hat{G}_1(z)$ of $G(z)$ are $\hat{G}_0(z) = (z + 2 \cos(\pi/3) + z^{-1})^2$, and $\hat{G}_1(z) = -(1/16 \cos^3(\pi/3))(z - 4 \cos(\pi/3) + z^{-1})$. This is the same as the previous solution, except that ω_1 in the previous solution is replaced with a constant value equal to $\pi/3$.

As another example, let us fix $\omega_1 = 2\pi/5 > \pi/3$, and find the optimal FIR compaction filter of order 5. The corresponding product filter is $G(z) = \frac{1}{2}z^5 + 1 + \frac{1}{2}z^{-5}$, and the compaction gain is $G_{\text{opt}}(2, 5) = 2$, which is the largest possible gain for $M = 2$. Since the process is line spectral, this is not surprising. *The important point here is that while the algorithm is not successful for the filter order 3, it is successful for a higher order 5.*

Example 7—Case Where the Process Is Multiband: Finally, we will consider an example in which the input is neither lowpass nor highpass, but rather, it is of multiband nature. Let $r(0) = 1$, $r(1) = \frac{1}{10}$, $r(2) = 0$ and $r(3) = -\frac{1}{4}$. The sequence $\psi_x(n)$ is positive definite for $N = 3$ so that the algorithm is applicable. There is more than one way to extrapolate this sequence and find the corresponding psd. For example, one can consider MA(3), AR(3), or line spectra(4). In all three cases, we have verified that the psd is neither lowpass nor highpass. Rather, it is of multiband nature. Applying the algorithm steps, we have $\hat{G}_0(z) = (z + (1/\sqrt{2}) + z^{-1})^2$,

from which it follows that $\hat{G}_1(z) = -\sqrt{2}(z - \sqrt{2} + z^{-1})$. This has single unit-circle zeros. Hence, $\hat{G}_1(e^{j\omega})$ is not non-negative, and therefore, $G(e^{j\omega}) = \hat{G}_0(e^{j\omega})\hat{G}_1(e^{j\omega})$ is not non-negative either. The algorithm halts because $\hat{G}_1(e^{j\omega})$ cannot be spectrally factorized.

IV. WINDOW METHOD

In this section, we will describe a new method to design FIR compaction filters. The method is applicable for arbitrary filter order N , arbitrary number of channels M , and for any given psd (including complex and multiband spectra). The technique is quite simple while the resulting compaction gains are very close to the optimum ones especially for high filter orders.

A common practice in filter design is to approximate ideal filter responses by windowing their impulse responses. Consider the ideal compaction filter design. For each $\omega \in [0, 2\pi/M]$, let $S_{xx}(e^{j(\omega+i_0(2\pi/M))})$ be the maximum of the set $\{S_{xx}(e^{j(\omega+i(2\pi/M))}), i = 0, \dots, M-1\}$. Then, $H_i(e^{j(\omega+i_0(2\pi/M))}) = \sqrt{M}$, and $H_i(e^{j(\omega+i(2\pi/M))}) = 0$ for $i \neq i_0$. Let $h_i(n)$ be the impulse response of $H_i(e^{j\omega})$, and consider $h(n) = w(n)h_i(n)$ for a given finite length window $w(n)$. Let $H(e^{j\omega})$ be the FT of $h(n)$. Then, $G(e^{j\omega}) = |H(e^{j\omega})|^2$ is no longer Nyquist(M). Instead of windowing $h_i(n)$, let us try to window the coefficients of the product filter: $g(n) = w(n)g_i(n)$. Here, $g_i(n)$ is the impulse response of $G_i(e^{j\omega}) = |H_i(e^{j\omega})|^2$. Then, $G(e^{j\omega})$ is Nyquist(M), but it may no longer be non-negative. The non-negativity can also be assured by constraining the FT of $w(n)$ to be nonnegative. A compaction filter can then be successfully obtained by spectrally factorizing $G(e^{j\omega})$. This can be considered to be the approximation of the ideal compaction filter response.

In this section, we extend this idea to design compaction filters that perform better than the above *ad hoc* windowing of ideal compaction filters. We will replace $g_i(n)$ with a periodic sequence $f_L(n)$, which will be determined by applying the ideal design algorithm at L uniform DFT frequencies. If $L = \infty$, then we have $f_L(n) = g_i(n)$, and the above *ad hoc* method results as a special case. It turns out that the experimentally optimum value of L for the best compaction gain is $L = M\lceil 2N/M \rceil$ (see Section IV-B).

A. Derivation of the Window Method

To formalize the above ideas, let us write the product filter coefficients $g(n)$ in the form

$$g(n) = w(n)f_L(n) \quad (36)$$

where $w(n)$ has the same length as $g(n)$, namely, $2N+1$, and $f_L(n)$ is a periodic sequence with period $L = KM \geq 2N$ for some K (see Fig. 4). Let $W(e^{j\omega})$ be the FT of $w(n)$ and $F_L(k)$ be the Fourier series coefficients (FSC) of $f_L(n)$, that is, $F_L(k) = \sum_{n=0}^{L-1} f_L(n)W_L^{kn}$, $W_L = e^{-j2\pi/L}$. The first period of $F_L(k)$ is just the DFT of the first period of $f_L(n)$. We make the following observation.

Lemma 1: Consider (36). If

- 1) $w(0) = 1$;
- 2) $W(e^{j\omega}) \geq 0$;

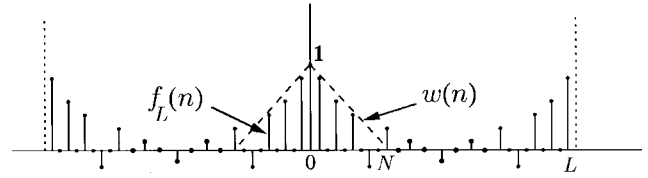


Fig. 4. Decomposition of $g(n)$ as $w(n)f_L(n)$, where $W(e^{j\omega}) \geq 0$ and $F_L(k) \geq 0$.

- 3) $F_L(k) \geq 0$;
- 4) $f_L(n)$ is Nyquist(M);

then $G(z)$ is the product filter of a valid compaction filter. That is, $g(Mn) = \delta(n)$, and $G(e^{j\omega}) \geq 0$.

Proof: It is readily verified that $G(e^{j\omega}) = (1/L) \sum_{k=0}^{L-1} F_L(k)W(e^{j(\omega-(2\pi/L)k)})$. Since $F_L(k) \geq 0$ and $W(e^{j\omega}) \geq 0$, it follows that $G(e^{j\omega}) \geq 0$. If $f_L(n)$ is Nyquist(M), then so is $g(n)$ because $g(Mn) = w(Mn)f_L(Mn) = \delta(n)$. ■

Assume the conditions of the lemma hold so that $G(z)$ is the product filter of a valid compaction filter. If $w(n)$ and L is fixed, what is the best $f_L(n)$ that maximizes the compaction gain? To answer the question, first note the following lemma.

Lemma 2: A periodic sequence $f_L(n)$ with period $L = KM$ is Nyquist(M), that is, $f_L(Mn) = \delta_K(n)$ if and only if its FSC $F_L(k)$ satisfy the following.

$$\sum_{i=0}^{M-1} F_L(k+iK) = M, \quad k=0, \dots, K-1. \quad (37)$$

Proof: Let us find the FSC $Y_K(k)$ of the decimated sequence $y_K(n) = f_L(Mn)$: $Y_K(k) = \sum_{n=0}^{K-1} f_L(Mn)W_K^{kn}$. This can be written as

$$Y_K(k) = \frac{1}{M} \sum_{i=0}^{M-1} \sum_{j=0}^{K-1} F_L(j+iK) \frac{1}{K} \sum_{n=0}^{K-1} W_K^{(k-j)n}.$$

Using $(1/K) \sum_{n=0}^{K-1} W_K^{mn} = \delta_K(m)$, we have $Y_K(k) = (1/M) \sum_{i=0}^{M-1} F_L(k+iK)$. The FSC of $\delta_K(n)$ are all 1. Hence, $f_L(Mn) = \delta_K(n)$ if and only if $(1/M) \sum_{i=0}^{M-1} F_L(k+iK) = 1, \forall k=0, \dots, K-1$. ■

To obtain the best $f_L(n)$, let $\hat{r}(n) = w^*(n)r(n)$, and let $\hat{S}_L(k)$ be the FSC of its periodic expansion $\hat{r}_L(n)$. For simplicity, assume that $L > 2N$. The objective (7) becomes $\sigma_y^2 = \sum_{n=0}^{L-1} f_L(n)\hat{r}_L^*(n) = (1/L) \sum_{k=0}^{L-1} F_L(k)\hat{S}_L(k)$. Both $F_L(k)$ and $\hat{S}_L(k)$ are real. Now, to incorporate the Nyquist(M) constraint, we write the preceding as

$$\sigma_y^2 = \frac{1}{L} \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} F_L(k+iK)\hat{S}_L(k+iK). \quad (38)$$

For a fixed k , let $\hat{S}_L(k+i_0K)$ be the maximum of the set $\{\hat{S}_L(k+iK), i=0, \dots, M-1\}$. Then, by (37), and noting that $F_L(k) \geq 0$, the objective (38) is maximized if we assign

$$F_L(k+i_0K) = M, \quad \text{and} \\ F_L(k+i_lK) = 0, \quad l=1, \dots, M-1. \quad (39)$$

Repeating the process for each $k=0, \dots, K-1$, the FSC of the best $f_L(n)$ is determined. The procedure is illustrated

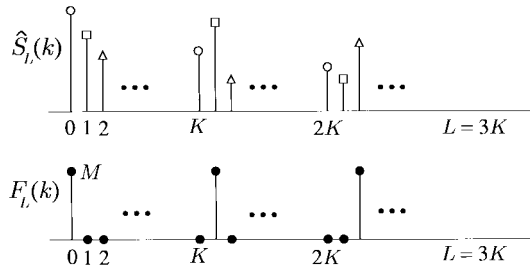


Fig. 5. Procedure to find $F_L(k)$: $\hat{S}_L(0)$ is maximum among $\{\hat{S}_L(iK)\}$; hence, $F_L(0) = M$, $F_L(lK) = 0$, $l \neq 0$. $\hat{S}_L(1+K)$ is maximum among $\{\hat{S}_L(1+iK)\}$; hence, $F_L(1+K) = M$, $F_L(1+lK) = 0$, $l \neq 1$, and so on.

in Fig. 5. The sequence $f_L(n)$ is just the inverse DFT of $F_L(k)$: $f_L(n) = (1/L) \sum_{k=0}^{L-1} F_L(k) W_L^{-nk}$.

Summary of the Window Algorithm: Assume a window $w(n)$ of the same length as $g(n)$ with non-negative FT has been chosen. Let $L = KM > 2N$. Then the algorithm steps are as follows.

Step 1: Calculate $\hat{S}_L(k)$, the L -point DFT of the conjugate-symmetric sequence $\hat{r}(n) = w^*(n)r(n)$ (which is the same as the FSC of the periodical expansion $\hat{r}(n)$ of $\hat{r}(n)$).

Step 2: For each $k = 0, \dots, K-1$, determine the index i_0 for which $\hat{S}_L(k + i_0K)$ is maximum, and assign $F_L(k + i_0K) = M$ and $F_L(k + i_lK) = 0$, $l = 1, \dots, M-1$.

Step 3: Calculate $f_L(n)$ by the inverse DFT. We need only to determine $f_L(n)$ for $n = 1, \dots, N$.

Step 4: Form the product filter $g(n) = w(n)f_L(n)$, and spectrally factorize it to find $H(z)$.

Real Case: If the input is real, the above algorithm can be modified to produce real-coefficient compaction filters. Consider the set $\{\hat{S}_L(k + iK), i = 0, \dots, M-1\}$ for each $k = 0, \dots, K-1$. Since $\hat{S}_L(k) = \hat{S}_L(L-k)$ if the process is real, this set is equivalent to $\{\hat{S}_L(L-k-iK) = \hat{S}_L(K-k+K(M-1-i)), i = 0, \dots, M-1\}$. Hence, in the comparison, we need to consider only $k = 0, \dots, P$, where $P = K/2$ if K is even, and $P = (K-1)/2$ if it is odd. Let $\hat{S}_L(k + i_0K)$ be the maximum of this set for each $k = 0, \dots, P$. We need to be careful in the assignments. The symmetric frequencies may end up in the same set, and we cannot assign different values to them. There are two cases to consider: **i)** The index $L-k-i_0K$ is among the set $\{k+iK, i = 0, \dots, M-1\}$; **ii)** it is not. The first case happens if and only if $2k \bmod K = 0$. This happens if $k = 0$ or $k = K/2$. We assign $F_L(k + i_0K) = F_L(L-k-i_0K) = M/2$ if $k + i_0K \neq L/2$, and $F_L(k + i_0K) = M$ if $k + i_0K = L/2$. In the second case, we assign $F_L(k + i_0K) = F_L(L-k-i_0K) = M$ if $k + i_0K \neq 0$ and $F_L(k + i_0K) = M$ if $k + i_0K = 0$. In either case, we set the remaining values in the set $\{F_L(k + iM)\}$ to zeros. This will maximize the objective (38). $f_L(n)$ calculated by the inverse DFT is the best sequence, and it is real.

Summary of the Window Algorithm for the Real Case: Assume a real symmetric window $w(n)$ of order N , with non-negative FT is given. Let $L = KM > 2N$, as before. Let P be as explained above. Then, Step 2 of the previous algorithm should be replaced by the following two steps.

Step 2.1: For each $k = 0, \dots, P$, determine the index i_0 for which $\hat{S}_L(k + i_0K)$ is maximum.

Step 2.2: If $k + i_0K = 0$ or $k + i_0K = L/2$, then set $F_L(k + i_0K) = M$; else, if $k = 0$ or $k = K/2$, then set $F_L(k + i_0K) = F_L(L-k-i_0K) = M/2$; else, set $F_L(k + i_0K) = F_L(L-k-i_0K) = M$. Set the remaining to zeros.

Optimization of the Window: The algorithm produces very good compaction gains, especially when the filter order is high, as we shall demonstrate shortly. However, one can get better compaction gains by optimizing the window $w(n)$. Consider the representation (36) again, and let $w(n)$ and $f_L(n)$ satisfy the conditions of Lemma 1. If we fix $f_L(n)$, what is the best window $w(n)$? The objective (7) can be written as

$$\sigma_y^2 = \int_{-\pi}^{\pi} \hat{S}_{xx}(e^{j\omega}) W(e^{j\omega}) \frac{d\omega}{2\pi} \quad (40)$$

where $W(e^{j\omega})$ is the FT of $w(n)$, and $\hat{S}_{xx}(e^{j\omega})$ is the FT of $f^*(n)r(n)$, where $f(n)$ is one period of $f_L(n)$ centered at $n = 0$. Let $W(e^{j\omega}) = |A(e^{j\omega})|^2$, where $A(z) = \sum_{n=0}^N a(n)z^{-n}$ is the spectral factor of $W(z)$. The only constraint on $A(e^{j\omega})$ is that it has to have unit energy in view of $w(0) = \int_{-\pi}^{\pi} |A(e^{j\omega})|^2 (d\omega/2\pi) = 1$. Let \mathbf{P} be the $(N+1) \times (N+1)$ Hermitian Toeplitz matrix corresponding to the sequence $\{f_L^*(n)r(n)\}_0^N$. Then, by Rayleigh's principle [16], (40) is maximized if $A(z)$ is the maximal eigenfilter of \mathbf{P} . The corresponding compaction gain is $\lambda_{\max}\{f_L^*(n)r(n)\}_0^N / \sigma_x^2$.

Corollary—A Lower Bound on the Compaction Gain: Let $f_L(n)$ be any Nyquist(M) sequence with non-negative FSC. Assume $L > N$. Then

$$G_{\text{opt}}(M, N) \geq \lambda_{\max}\{f_L^*(n)r(n)\}_0^N. \quad (41)$$

To see this, note that $g(n) = w(n)f_L(n)$ achieves that bound by choosing $w(n)$ as the optimum window for the sequence $f_L(n)$. If we replace $f_L(n)$ by a positive definite Nyquist(M) sequence $f(n)$ of order N , the inequality continues to be valid because $w(n)f(n)$ is still a product filter of a valid compaction filter. To see this, note that the sequence $f(n)$ can be extended to an infinite sequence (e.g., using autoregressive extrapolation) such that its FT is non-negative. Hence, the product $w(n)f(n)$ has non-negative FT. The Nyquist(M) property of the product follows from that of $f(n)$.

We have described how to optimize $w(n)$ given $f_L(n)$, and vice versa. It is reasonable to expect that one can iterate and obtain better compaction gains at each stage. We have observed that this is not the case. We started with a triangular window and found that $f_L(n)$ did not change after the reoptimization of the window. Notice that the use of an initial window is not necessary if one is willing to optimize the window after finding $f_L(n)$. However, in most of the design examples we considered, using an initial window with non-negative FT (in particular, the triangular window) and then reoptimizing the window resulted in better compaction gains. A matlab program that implements the window method can be found at our webpage [28]. Here is a simple example to illustrate how the window method works.

Example 8—MA(1) Process: Let $N = 5$, $M = 4$, $r(0) = 1$, $r(1) = \rho$, and $r(n) = 0$, $n > 1$. Assume the process is real

so that $r(-n) = r(n)$. Let the window be triangular, i.e.,

$$w(n) = \begin{cases} 1 - \frac{|n|}{6}, & n = 0, \pm 1, \dots, \pm 5 \\ 0, & \text{elsewhere.} \end{cases} \quad (42)$$

The FSC $\hat{S}_L(k)$ of $\hat{r}_L(n)$ in Step 1 are

$$\begin{aligned} \hat{S}_L(k) &= \hat{S}(e^{j\omega})|_{\omega=(2\pi/L)k} \\ &= 1 + \frac{5}{3}\rho \cos \frac{2\pi}{L}k, \quad k = 0, \dots, L-1. \end{aligned} \quad (43)$$

Now, assume that $L = 12 > 10$ so that $K = 3$ and $P = 1$. Therefore, we have the following sets to consider in Step 2.

$$\begin{aligned} &\{\hat{S}_L(0), \hat{S}_L(3), \hat{S}_L(6), \hat{S}_L(9)\} \\ &\{\hat{S}_L(1), \hat{S}_L(4), \hat{S}_L(7), \hat{S}_L(10)\} \end{aligned} \quad (44)$$

which are evaluated below, respectively.

$$\begin{aligned} &\left\{1 + \frac{5}{3}\rho, 1, 1 - \frac{5}{3}\rho, 1\right\}, \\ &\left\{1 + \frac{5\sqrt{3}}{6}\rho, 1 - \frac{5}{6}\rho, 1 - \frac{5\sqrt{3}}{6}\rho, 1 + \frac{5}{6}\rho\right\}. \end{aligned} \quad (45)$$

First, assume $\rho > 0$. The maximum of the first set is $\hat{S}_L(0)$, and the maximum of the second set is $\hat{S}_L(1)$. Hence, applying Step 3 of the algorithm, we have

$$\{F_L(k), k = 0, \dots, L-1\} = \{4, 4, 0, 0, 0, 0, 0, 0, 0, 0, 4\}. \quad (46)$$

Taking the inverse DFT of $F_L(k)$, we calculate in Step 4

$$\{f_L(n), n = 0, \dots, N\} = \left\{1, \frac{1+\sqrt{3}}{3}, \frac{2}{3}, \frac{1}{3}, 0, \frac{1-\sqrt{3}}{3}\right\}. \quad (47)$$

Hence, the product filter $g(n) = w(n)f_L(n)$ has been found, and

$$\begin{aligned} G(z) &= \frac{1-\sqrt{3}}{18}z^5 + \frac{1}{6}z^3 + \frac{4}{9}z^2 + \frac{5(1+\sqrt{3})}{18}z + 1 \\ &\quad + \frac{5(1+\sqrt{3})}{18}z^{-1} + \frac{4}{9}z^{-2} + \frac{1}{6}z^{-3} + \frac{1-\sqrt{3}}{18}z^{-5}. \end{aligned} \quad (48)$$

Next, consider the case $\rho < 0$. Referring to (45), $\hat{S}_L(6)$ in the first set and $\hat{S}_L(7)$ in the second set is maximum. Hence, $\{\hat{F}_L(k), k = 0, \dots, L-1\} = \{0, 0, 0, 0, 0, 4, 4, 4, 0, 0, 0\}$, which is equal to $F_L(k-6)$, where $F_L(k)$ is the previous solution. Hence, $\hat{f}_L(n) = (-1)^n f_L(n)$, and therefore, $\hat{G}(z) = G(-z)$. By spectrally factorizing the product filter, an optimum compaction filter is obtained. The compaction gain is $1 + (5(1+\sqrt{3})/9)|\rho| \simeq 1 + 1.5178|\rho|$.

Compaction gain versus L , $M = 2$

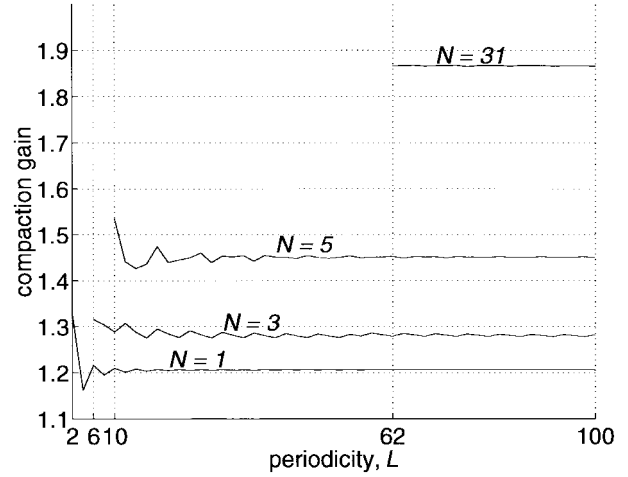


Fig. 6. Compaction gain versus periodicity L .

Let us find the improvement we can get by optimizing the window when we fix $f_L(n)$. Since $\sigma_x^2 = r(0) = 1$, the compaction gain is the maximum eigenvalue of the 6×6 symmetric Toeplitz matrix with the first row $[1 \ f_L(1) \ \rho \ 0 \ 0 \ 0]$, which is $1 + 1.8019f_L(1)|\rho|$. Using $f_L(1)$ given in (47), the improved compaction gain is $1 + 1.6410|\rho|$. With this optimum window fixed, one can verify that $f_L(n)$ in (47) is still the optimum sequence.

B. Choice of the Periodicity L

The window method will produce compaction filters as long as L is a multiple of M and is greater than N . This choice of L will ensure that $f_L(n)$ is Nyquist(M). The smallest such period is $L = M\lceil N/M \rceil$, and the largest is $L = \infty$. The choice $L = M\lceil N/M \rceil$ leads to an additional symmetry in $f_L(n)$, and according to our experience, the corresponding compaction gains are not good. If we use $L = \infty$, then we get the ideal solution for $f_L(n)$: $f_L(n) = g_i(n)$. The corresponding compaction filter obtained after windowing is not optimal either. If L is chosen to be the smallest multiple of M such that $L \geq 2N$, then we obtain very good compaction gains. This choice can be compactly written as

$$L = M\lceil 2N/M \rceil.$$

If $M = 2$, then this choice reduces to $L = 2N$. In Example 8, we increased L from 12 to 16 and found that the compaction gain decreased. When we used the ideal filter for $f_L(n)$, which corresponds to $L = \infty$, the compaction gain was better than that of the case $L = 16$ but worse than that of the case $L = 12$.

Example 9—Dependence on L : We have designed compaction filters using the window method for an AR(5) process, whose psd is shown in Fig. 8. We have chosen this psd because it is multiband, and the capture of the signal energy can be illustrated clearly. The number of channels is $M = 2$. We considered the filter orders $N = 1, 3, 5$, and $N = 31$. For each order N , we increased L from $2N$ to 100 in steps of 2. The resulting compaction gains are plotted in Fig. 6. From the plot, we see that the best compaction gain is for $L = 2N$.

V. LINEAR PROGRAMMING METHOD

The use of linear programming (LP) method in compaction filter design was proposed by Moulin *et al.* [20], [31], [32]. We briefly review the method and propose some improvements. Assume that the input process $x(n)$ is real. The output variance is $\sigma_y^2 = r(0) + 2 \sum_{n=1}^N g(n)r(n)$. Let \mathbf{g}_d and \mathbf{r}_d be the vectors formed by deleting every M th coefficients of $g(n)$ and $r(n)$ for $n = 1, \dots, N$. Then, the objective can be written as $\sigma_y^2 = r(0) + 2\mathbf{r}_d^T \mathbf{g}_d$. This incorporates the Nyquist(M) condition but not the nonnegativity constraint in (6). Let $\mathbf{c}_d(\omega) \triangleq [\cos(\omega) \cos(2\omega) \dots \cos((M-1)\omega) \cos((M+1)\omega) \dots \cos(N\omega)]^T$. Then, $G(e^{j\omega}) = 1 + 2\mathbf{c}_d^T(\omega) \mathbf{g}_d$. Hence, the problem is equivalent to

$$\begin{aligned} & \text{maximize } \mathbf{r}_d^T \mathbf{g}_d \\ & \text{subject to } \mathbf{c}_d^T(\omega) \mathbf{g}_d \geq -0.5, \forall \omega \in [0, \pi]. \end{aligned} \quad (49)$$

This type of problem is typically classified as semi-infinite linear programming (SIP) [32] because there are infinitely many inequality constraints on finitely many variables. By discretizing the frequency, one reduces this to a well known standard LP problem.

Drawbacks of the Technique: No matter how dense the frequency grid is, LP guarantees the nonnegativity of $G(e^{j\omega})$ only on this grid. Hence, one has to modify the solution to have $G(e^{j\omega}) \geq 0, \forall \omega$. One can numerically determine the unit-circle zeros of $G(e^{j\omega})$ and merge the pairs of zeros that are close to each other. Yet another way is to “lift” $G(e^{j\omega})$ by increasing $g(0)$ relative to other coefficients. Since $g(0)$ has to be 1, in effect, we scale $g(n)$ for $n \neq 0$ by a constant $c < 1$. This can also be considered as windowing with $w(n) = c, n \neq 0$, and $w(0) = 1$. In the next section, we propose another windowing technique to modify $G(e^{j\omega})$. The advantage of this is to avoid having to locate any zeros or the minimum of $G(e^{j\omega})$. The non-negativity of $G(e^{j\omega})$ is guaranteed by that of $W(e^{j\omega})$ as in Section IV-A. If the filter order N and the number of discrete frequencies L are small, using an optimum window performs better than the other techniques. In principal, as $L \rightarrow \infty$, the LP solution approaches the optimal solution. However, as stated in [32], there will be numerical problems if L is too high. Another drawback of LP is that the complexity is prohibitively high for high filter orders. We should note here that the window method that we proposed in Section IV does not have this problem. The window method is very fast, even with very high filter orders, and the resulting filters are very close to the optimal ones.

A. Windowing of the Linear Programming Solution

Let L uniform frequencies $\{\omega_k = (2\pi/L)k, k = 0, \dots, L-1\}$ be used in LP, and let $g_L(n)$ be the periodical expansion of the resulting product filter. Assume that $L > 2N$. Linear programming assures that $G(e^{j\omega})$ is non-negative at the frequencies $\{\omega_k\}$. Hence, the FSC $G_L(k)$ of $g_L(n)$ are non-negative. Now, consider the product

$$g(n) = w(n)g_L(n) \quad (50)$$

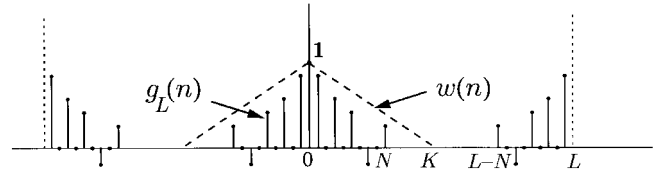


Fig. 7. Windowing of the linear programming solution.

where $w(n)$ is a symmetric window of order $K < L - N$ (length $2K + 1$) with non-negative FT (see Fig. 7); then, from Section IV-A, we conclude that $G(e^{j\omega}) \geq 0, \forall \omega$. The Nyquist(M) property of $g(n)$ is assured by that of $g_L(n)$. In contrast to the window method, here, we can have $K > N$. This is because the LP solution already has the desired order. For maximum compaction gain, the symmetric order of $w(n)$ is chosen to be maximum, namely, $K = L - N - 1$. Note that when $L = 2N$, we have $g_L(N) = 2g(N)$. One can use a fixed window like a triangular window, as depicted in the figure, and get a satisfactory compaction gain. However, one can always optimize the window as in Section IV-A. If L is very large, optimization should be avoided as the performance loss becomes negligible. The loss can be quantified as follows: Assuming $\sigma_x^2 = 1$, when a fixed window is used, the compaction gain is $G_w = 1 + 2\mathbf{g}^T \hat{\mathbf{r}}$, where \mathbf{g} and $\hat{\mathbf{r}}$ are the vectors formed by the sequences $g_L(n)$ and $w(n)r(n), n = 0, \dots, N$. If, for example, a triangular window of symmetric order $K = L - N - 1$ is used, we have $w(n) = 1 - n/(L - N), n = 0, \dots, N$. When the optimum window is used, the compaction gain is $G_o = \lambda_{\max}\{g_L(n)r(n)\}_0^N$. Hence, the loss is $G_o - G_w = \lambda_{\max}\{g_L(n)r(n)\}_0^N - 2\mathbf{g}^T \hat{\mathbf{r}} - 1$. As $L \rightarrow \infty, w(n)r(n) \rightarrow r(n)$, and $g_L(n) \rightarrow g_{\text{opt}}(n)$. Hence, $G_w \rightarrow G_{\text{opt}}$. Since $G_w \leq G_o \leq G_{\text{opt}}$, we see that $G_o \rightarrow G_{\text{opt}}$ as well. Hence, $G_o - G_w \rightarrow 0$ as $L \rightarrow \infty$.

Example 10: Let the input psd be as in Fig. 8, and let $N = 65$ and $M = 2$. In the same figure, we plot the magnitude square $|H(e^{j\omega})|^2$ of the compaction filter $H(z)$ designed by LP. The number of frequencies used in the design process was $L = 512$. We have used triangular window of symmetric order $K = L - N - 1 = 446$ and found that the resulting compaction gain is $G_w = 1.8698$. If we optimize the window, the compaction gain becomes $G_o = 1.8744$. Hence, the loss is $G_o - G_w = 0.0046$. One can verify that the compaction gain of the ideal ($L = \infty$) filter is $G_{\text{ideal}} = 1.8754$.

B. Connection Between the Linear Programming and Window Methods

In both the LP and window methods, we use windows to assure the nonnegativity of $G(e^{j\omega})$. Consider the equations (50) and (36). When L is a multiple of M , a periodic sequence $g_L(n)$ in the linear programming method and a periodic sequence $f_L(n)$ in the window method are found such that they are Nyquist(M) and their FSC are all non-negative. For $L > 2N$, the two problems are not the same because $g_L(n)$ is order constrained, whereas $f_L(n)$ is not. If, however, $L = 2N$, then the two problems are exactly the same. If windowing is done in the same way in both methods, then we see that the resulting compaction gains should be the same.

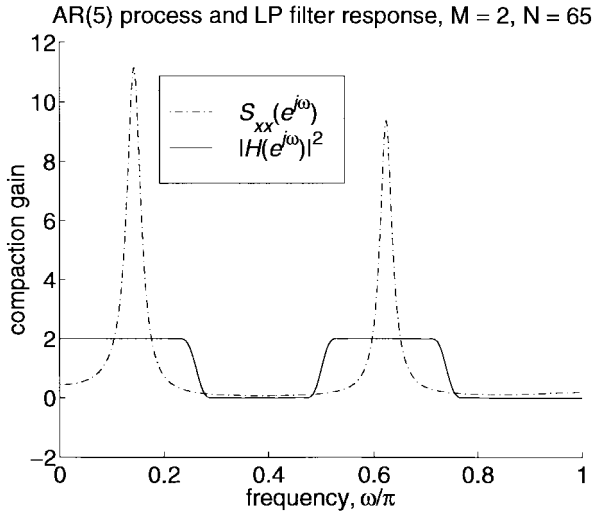


Fig. 8. Power spectral density of an AR(5) process and the magnitude square of an optimal compaction filter for $N = 65$ and $M = 2$, designed by LP. The parameter L is 512 and a triangular window is used.

Hence, one can view the window method as an efficient and noniterative technique to solve an LP problem when $L = 2N$. If L is increased, we saw that the window method does not necessarily yield better gains, whereas this is the case for the LP method, provided the window order is increased as well. However, optimization of the window in LP becomes costly as the order increases. If one uses a fixed triangular window (with highest possible order) in LP, and if the windows are optimized in the window method, then the window method is very close and sometimes superior to LP, as we demonstrate in the following example.

Example 11: Comparison of Linear Programming and Window Methods: Let the input psd be as in Fig. 8. In Fig. 9(a), the compaction gains of both the LP and the window method versus the filter order are plotted for $M = 2$. The number of frequencies used in LP is $L = 512$, whereas the periodicity used in the window method is $L = 2N$. The windows used in LP are triangular windows with symmetric order $L - N - 1$. In the window method, the autocorrelation sequence is first windowed by a triangular window of symmetric order N to find $f_L(n)$, and then, the window is reoptimized. From the figure, we observe that if the order is high, one has slightly better compaction gains using the window method. This implies that if one optimizes the window, there is no need to use large number of frequencies in LP. More importantly, there is no need to use LP for high filter orders. However, it should be emphasized that if the windows are optimized in LP, one can get slightly better compaction gains than the window method. In Fig. 9(b), we show the plots of the compaction gains of the two methods for various values of M for a fixed filter order of 65. We observe that the window method performs very close to LP, especially for low values of M . We show the upper bounds on compaction gains in both plots. The upper bound in the first plot is achieved by an ideal compaction filter, and that in the second plot is achieved by a maximal eigenfilter, as discussed in Section II.

Example 12: Let the input be AR(1) as in Example 1. For $N = 3$ and $M = 2$, we have designed compaction filters

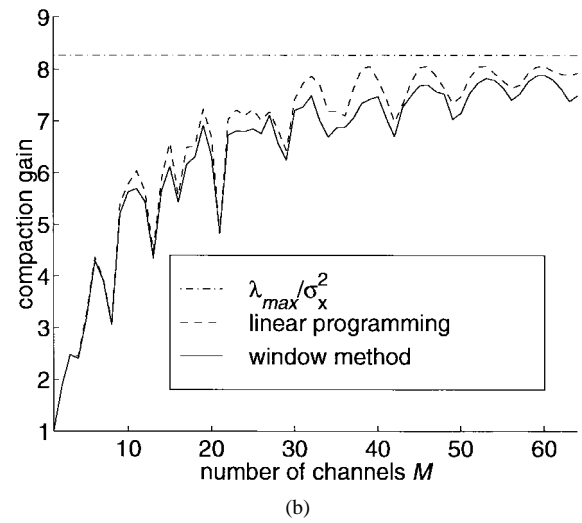
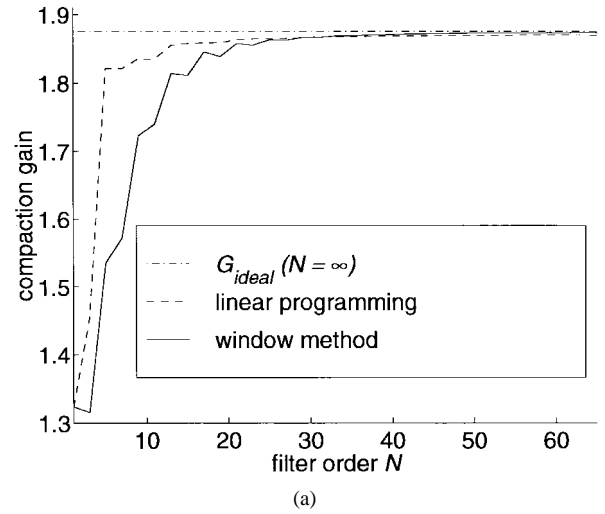


Fig. 9. Comparison of the window and linear programming methods. The input power spectrum is as shown in Fig. 8. (a) Compaction gain versus N for $M = 2$. (b) Compaction gain versus M for $N = 65$.

using the window and LP methods. We present in Table I the resulting filter coefficients and the corresponding compaction gains for $\rho = 0.1, 0.5$, and 0.9 . The analytically optimum coefficients (33) and the corresponding compaction gains (34) are also presented in the same table. We see in this case that the compaction gain of the window method is not too far from the optimal one and slightly worse than that of LP, even for such a small order. The discrepancy between the window and LP compaction gains is maximum when $\rho = 0.5$.

VI. CONCLUDING REMARKS

We have presented two new techniques for the design of optimum FIR compaction filters. First, we have proposed an analytical method in the two-channel case. The technique is applicable for a rather restricted but practically important class of signals. The method involves Levinson recursion and two spectral factorizations of half the filter order. As examples, we have produced analytical expressions for the compaction filter coefficients for AR(1) and MA(1) processes. Next, we have proposed a method called the *window method*.

It is applicable for any given spectra and for any given number of channels. It is very efficient since it is noniterative and involves only comparison of some DFT coefficients and windowing. We have given its relation to the LP method. As the filter order becomes higher, the computational complexity of the LP method grows rapidly. The window method, on the other hand, is very fast, even when the filter orders are very high. Furthermore, the suboptimality of the window method diminishes as the filter order increases.

Future work will incorporate these methods in the design of optimal FIR orthonormal uniform and nonuniform subband coders. In the two-channel case, the optimum compaction filter already determines the optimum filter bank. Hence, the algorithms in this paper can readily be used in applications like wavelet-based image coding. In particular, it would be interesting to investigate the performance of our filters in zero-tree coding and wavelet-package coding. For such applications, we expect that the analytical method of Section III will be quite useful. In the M -channel case, we mentioned one method [20] that efficiently finds the rest of the filter bank if the first filter is given. In speech and audio coding applications, M -channel uniform filter banks are commonly used, and the filters have high orders. We expect that the window method of Section IV will be very useful for such applications. Needless to say, there are many other important applications of compaction filters, some of which are mentioned in the last paragraph of the introduction. Hence, our design algorithms can directly be used in such applications as well. All the algorithms described in this paper can be found at our webpage [28].

APPENDIX PROOF OF NON-NEGATIVITY

We will show that $G(e^{j\omega})$ obtained by the procedure in Section III is necessarily nonnegative in the region $[\pi/2, \pi]$. The Nyquist(2) property of $G(e^{j\omega})$ implies $G'(e^{j\omega}) = G'(e^{j(\pi-\omega)})$. We therefore have $G'(e^{j\omega_k/2}) = G'(e^{j(\pi-\omega_k/2)}) = 0, k = 0, \dots, (N-3)/4$. Now, by the mean value theorem in calculus, we also have $G'(e^{j\hat{\omega}_k/2}) = G'(e^{j(\pi-\hat{\omega}_k/2)}) = 0, k = 0, \dots, (N-7)/4$ for some $\hat{\omega}_k \in (\omega_k, \omega_{k+1})$. Notice that since ω_k 's are all distinct and lie in the open region $(0, \pi)$, all of the above zeros are distinct. The total number of such zeros is therefore $N-1$. Since $G(e^{j\omega})$ is a cosine polynomial of order N , $G'(e^{j\omega})$ is a sine polynomial of order N , and therefore, it can be written in the form $G'(e^{j\omega}) = \sin \omega T(\cos \omega)$, where $T(x)$ is a polynomial of order $N-1$. Excluding the zeros at 0 and π , the total number of zeros $G'(e^{j\omega})$ can have in $[0, \pi]$ is $N-1$. Hence, $G'(e^{j\omega})$ cannot have any other zero on the unit circle. If $G(z)$ has a zero at $\pi - \omega_k/2$ with multiplicity greater than 2, then $G'(e^{j\omega})$ has at least double zero at that frequency, implying that the total number of its zeros is more than $N-1$, which is a contradiction. If $G(z)$ has a single zero in the region $(\pi/2, \pi)$, which is different from all ω_k 's, then by applying the mean value theorem once more, $G'(e^{j\omega})$ has to have another zero, which is again a contradiction. Hence, we have proved that $G'(e^{j\omega})$ has double zeros at $\pi - \omega_k/2, k = 0, \dots, (N-3)/4$ and that it does not have any other unit circle zeros in $[\pi/2, \pi]$.

This, in particular, implies $G(e^{j\omega}) \geq 0$ for $\omega \in [\pi/2, \pi]$. The proof for the case of even $(N-1)/2$ is similar; the details are omitted. ■

REFERENCES

- [1] K. C. Aas and C. T. Mullis, "Minimum mean-squared error transform coding and subband coding," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1179–1192, July 1996.
- [2] S. O. Aase and T. A. Ramstad, "On the optimality of nonunitary filter banks in subband coders," *IEEE Trans. Image Processing*, vol. 4, pp. 1585–1591, Dec. 1995.
- [3] A. N. Akansu and Y. Liu, "On signal decomposition techniques," *Opt. Eng.*, vol. 30, pp. 912–920, July 1991.
- [4] 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.
- [5] J. Katto and Y. Yasuda, "Performance evaluation of subband coding and optimization of its filter coefficients," in *Proc. SPIE, Visual Commun. Image Process.: Visual Commun.*, 1991, vol. 1605, pp. 95–106.
- [6] 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.
- [7] M. Unser, "An extension of the Karhunen-Loeve transform for wavelets and perfect reconstruction filterbanks," in *Proc. SPIE Math. Imaging*, 1993, vol. 2034, pp. 45–56.
- [8] ———, "On the optimality of ideal filters for pyramid and wavelet signal approximation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3591–3596, Dec. 1993.
- [9] P. P. Vaidyanathan, "Theory of optimal orthonormal filter banks," in *Proc. IEEE ICASSP*, May 1996, Atlanta, GA, vol. 3, pp. 1487–1490.
- [10] ———, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [11] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [12] A. Tirakis, A. Delopoulos, and S. Kollias, "Two-dimensional filter bank design for optimal reconstruction using limited subband information," *IEEE Trans. Image Processing*, vol. 4, pp. 1160–1165, Aug. 1995.
- [13] B. Xuan and R. H. Bamberger, "2D factorizable FIR principal component filter banks," in *Proc. IEEE ICASSP*, May 1996, Atlanta, GA, pp. 1538–1541.
- [14] ———, "Complete FIR principal component filter banks," in *Proc. IEEE ISCAS*, May 1996, Atlanta, GA, pp. 417–420.
- [15] ———, "Multi-dimensional paraunitary principal component filter banks," in *Proc. IEEE ICASSP*, May 1995, Detroit, MI, pp. 1488–1491.
- [16] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [17] A. Kirac and P. P. Vaidyanathan, "Optimality of orthonormal transforms for subband coding," in *Proc. IEEE DSP Workshop*, Snowbird, UT, 1998.
- [18] ———, "Optimal nonuniform orthonormal filter banks for subband coding and signal representation," submitted for publication.
- [19] ———, "On existence of FIR principal component filter banks," in *Proc. IEEE ICASSP*, Seattle, WA, 1998.
- [20] P. Moulin, M. Anitescu, K. Kortanek, and F. Potra, "Design of signal-adapted FIR paraunitary filter banks," in *Proc. IEEE ICASSP*, May 1996, Atlanta, GA, vol. 3, pp. 1519–1522.
- [21] 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.
- [22] D. Taubman and A. Zakhori, "A multi-start algorithm for signal adaptive subband systems," in *Proc. IEEE ICASSP*, Mar. 1992, San Francisco, CA, vol. 3, pp. 213–216.
- [23] Q. Jin, T. Luo, and K. M. Wong, "Optimum filter banks for signal decomposition and its application in adaptive echo cancellation," *IEEE Trans. Signal Processing*, vol. 44, pp. 1669–1680, July 1996.
- [24] 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.
- [25] 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.
- [26] 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.

- [27] R. A. Roberts and C. T. Mullis, *Digital Signal Processing*. Reading, MA, Addison-Wesley, 1987.
- [28] <http://www.systems.caltech.edu/kirac/>
- [29] H. Caglar, Y. Liu, and A. N. Akansu, "Statistically optimized PR-QMF design," in *Proc. SPIE, Visual Commun. Image Process.: Visual Commun.*, 1991, vol. 1605, pp. 86–94.
- [30] L. Vandendorpe, "CQF filter banks matched to signal statistics," *Signal Process.*, vol. 29, pp. 237–249, 1992.
- [31] P. Moulin, "A new look at signal-adapted QMF bank design," in *Proc. IEEE ICASSP*, May 1995, Detroit, MI, vol. 5, pp. 1312–1315.
- [32] P. Moulin, M. Anitescu, K. Kortanek, and F. Potra, "The role of linear semi-infinite programming in signal-adapted QMF bank design," *IEEE Trans. Signal Processing*, vol. 45, pp. 2160–2174, Sept. 1997.
- [33] B. E. Usevitch and M. T. Orchard, "Smooth wavelets, transform coding, and markov-1 processes," *IEEE Trans. Signal Processing*, vol. 43, pp. 2561–2569, Nov. 1995.
- [34] V. Silva and L. de Sa, "Analytical optimization of CQF filter banks," *IEEE Trans. Signal Processing*, vol. 44, pp. 1564–1568, June 1996.
- [35] J. Tuqan and P. P. Vaidyanathan, "A state space approach for the design of optimum FIR energy compaction filters," in preparation.
- [36] A. Kirac and P. P. Vaidyanathan, "Efficient design methods of optimal FIR compaction filters for M-channel FIR subband coders," in *Proc. 30th Asilomar Conf. Signals, Syst., Comput.*, 1996.
- [37] ———, "Analytical method for 2-channel optimum FIR orthonormal filter banks," in *Proc. IEEE ISCAS*, June 1997, Hong Kong.
- [38] ———, "FIR compaction filters: New design methods and properties," in *Proc. IEEE ICASSP*, May 1997, Munich, Germany, pp. 2229–2232.
- [39] V. F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophys. J. Roy. Astron. Soc.*, vol. 33, pp. 347–366, 1973.
- [40] N. I. Ahiezer and M. Krein, *Some Questions in the Theory of Moments*. Translations of Mathematical Monographs, Vol. 2. Providence, RI: Amer. Math. Soc., 1962.
- [41] P. J. Davis, *Interpolation and Approximation*. New York: Dover, 1975.
- [42] F. L. Bauer, "A direct iterative process for the Hurwitz-decomposition of a polynomial," *Arch. Elekt. Uebertr.*, vol. 9, pp. 285–290, June 1955; *IRE Trans. Circuit Theory*, vol. CT-2, pp. 370–371, Dec. 1955.
- [43] T. Q. Nguyen, "A novel spectral factorization method and its application in the design of filter banks and wavelets," in *Proc. IEEE-SP Int. Symp. Time-Freq. Time-Scale Anal.*, Oct. 1992, pp. 303–306.
- [44] A. K. Jain, *Fundamentals of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.



Ahmet Kirac (S'94) was born in Afsin, Turkey, on October 10, 1971. He received the B.S. degree in 1993 in electrical engineering from Bilkent University, Ankara, Turkey, and the M.S. degree in electrical engineering from the California Institute of Technology, Pasadena, in 1994. He is currently pursuing the Ph.D. degree at the California Institute of Technology.

His research interests are in digital signal processing with applications to coding and communications. Optimization of filter banks and wavelets

for data compression and signal representation is his main focus.



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. He received 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 Department of Electrical Engineering, California Institute of Technology, 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 from 1985 to 1987, is currently an associate editor for the IEEE SIGNAL PROCESSING LETTERS, and is a consulting editor for the journal *Applied and Computational Harmonic Analysis*. He is a guest editor for this Special Issue of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the Special Issue of 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 the California Institute of Technology 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 Awards for his paper on multirate perfect-reconstruction filter banks. In 1990, he was the recipient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommunications 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 been chosen a distinguished lecturer for the IEEE Signal Processing Society for 1996–1997.