ITERATIVE EIGENFILTER METHOD FOR DESIGNING OPTIMUM OVERDECIMATED ORTHONORMAL FIR COMPACTION FILTER BANKS

Andre Tkacenko and P. P. Vaidyanathan

Dept. of Electrical Engineering
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: andre@systems.caltech.edu, ppvnath@systems.caltech.edu

ABSTRACT

Much attention has recently been given to the construction of signal-adapted orthonormal filter banks designed to optimize a particular objective function such as coding gain or a multiresolution criterion. For certain classes of filter banks, the optimum solution is a principal component filter bank (PCFB), which is *simultaneously* optimal for several objectives including the ones mentioned above. However, for the class of finite impulse response (FIR) filter banks, a PCFB in general does not exist. For this case, numerical techniques must be employed to find an optimum filter bank for a particular objective. In this paper, we present an iterative method for designing an overdecimated FIR filter bank optimized for energy compaction. The proposed algorithm is an eigenfilter method that is low in computational complexity. Simulation results show the merit of the proposed method, in that as the filter order increases, the filters designed behave more and more like those of the infinite order PCFB.[1]

1. INTRODUCTION

The problem of the design of optimal signal-adapted multirate filter banks has been of interest to the signal processing community on account of its applications in signal representation and data compression [5, 1, 8]. Such filter banks are typically chosen to optimize a particular objective, such as coding gain or a multiresolution criterion, adapted to the input signal statistics. If no constraints are imposed on the orders of the analysis/synthesis filters, then the optimal filter bank for a number of objectives is an infinite order principal component filter bank (PCFB) [5, 1]. Such a filter bank is *simultaneously* optimal for a wide variety of objectives, including those mentioned above. The corresponding optimal analysis/synthesis filters in this case turn out to be a series of *compaction filters* [8]. For many practical cases of inputs, these filters have an ideal bandpass response [8] and as such are unrealizable.

If we restrict the analysis/synthesis filters to be finite impulse response (FIR) filters, then designing an optimal signal-adapted filter bank becomes far more difficult, since a PCFB in general does not exist [1]. As a result, the filters must be designed to optimize a *specific* objective. In [10], the authors focused on maximizing the coding gain, whereas in [4], the authors considered optimizing a multiresolution criterion. For both methods, suboptimal numerical techniques were employed. Despite this, in [4], it was shown that for the multiresolution criterion under consideration, the design problem could be greatly simplified since the entire filter bank could be constructed by designing a single FIR compaction filter, followed by an appropriate Karhunen-Loève transform (KLT).

[1]Work supported in part by the ONR grant N00014-99-1-1002, USA.

As such, much attention in the signal processing community has focused on the design of FIR compaction filters [4, 3, 6]. In [4], the authors proposed a linear semi-infinite programming (SIP) method for the design of such filters. Though their method is globally optimal for the problem considered, the required orthonormality condition is only satisfied on a discrete set of frequencies. A window method for designing FIR compaction filters was proposed in [3]. Despite its low computational complexity and good performance, it is inherently suboptimal. In [6], a method for the design of *globally* optimal FIR compaction filters was presented. This method, though globally optimal, becomes very computationally intensive as the filter order increases, and only applies for the specific case where the input is wide sense stationary (WSS).

In this paper, we present an iterative method to solve the FIR compaction problem for *overdecimated* filter banks, described below. The overdecimated filter bank problem we focus on here is a generalization of the classical compaction filter problem in that it can be applied to design an arbitrary number of subbands (instead of only one for the classical problem). Here, the iterative method comes about by linearizing the quadratic orthonormality constraints. Having done this, the problem can be easily solved using the *eigenfilter* approach [7], which is well known for its low computational complexity and numerical robustness (i.e. no inversion of ill conditioned matrices is required). Though the method is not guaranteed to be globally optimal, simulation results provided show its merits. In particular, it can be seen that as the order increases, the designed filters behave more and more like the ideal compaction filters of the infinite order PCFB, consistent with intuition.

1.1. Overdecimated Filter Bank Signal Model

Here, we focus on the overdecimated uniform filter bank shown in Fig. 1(a). By overdecimated, we mean that the number of channels L satisfies $L < M$, i.e. the number of subbands is strictly less than the decimation ratio [7]. In such a system, alias cancellation and perfect reconstruction are in general impossible. If we consider the following polyphase decompositions [7] of the analysis filters $H_k(z)$ and synthesis filters $F_k(z)$ for $0 \leq k \leq L-1$,

$$H_k(z) = \sum_{\ell=0}^{M-1} z^{\ell} H_{k,\ell}(z^M) \quad \text{(Type II)}$$

$$F_k(z) = \sum_{\ell=0}^{M-1} z^{-\ell} F_{k,\ell}(z^M) \quad \text{(Type I)}$$

then the system of Fig. 1(a) can be redrawn as in Fig. 1(b), where,

$$[\mathbf{H}(z)]_{\ell,m} = H_{\ell,m}(z)\,,\ [\mathbf{F}(z)]_{m,\ell} = F_{\ell,m}(z)$$

for $0 \leq \ell \leq L-1$ and $0 \leq m \leq M-1$. Note that here, the vector signals $\mathbf{x}(n)$ and $\mathbf{y}(n)$ denote, respectively, the M-fold blocked versions [7] of the filter bank input $x(n)$ and output $y(n)$.

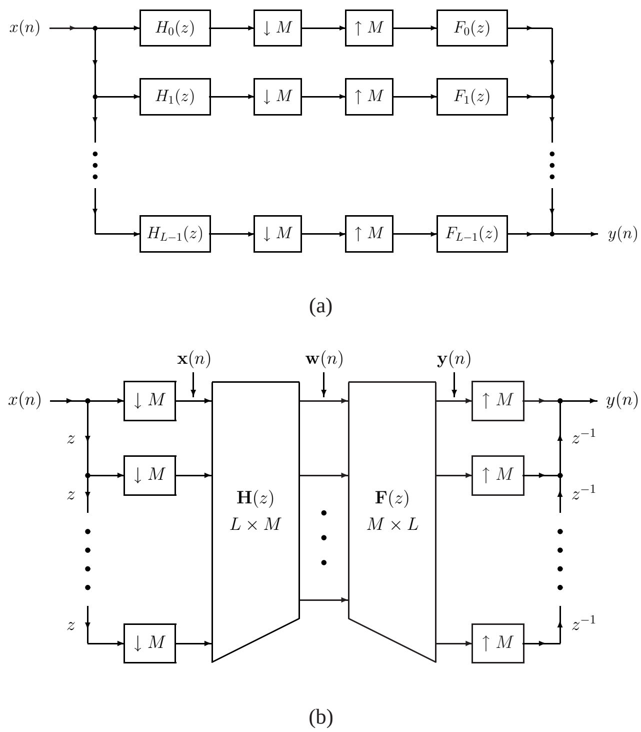

Fig. 1. (a) Uniform overdecimated filter bank ($L < M$), (b) polyphase representation.

2. ORTHONORMAL FIR ENERGY COMPACTION PROBLEM

2.1. Derivation Of The Energy Compaction Problem

From here on in, we will assume that the input $x(n)$ to the overdecimated filter bank of Fig. 1 is cyclo wide sense stationary with period M (CWSS(M)) [7], which is tantamount to saying that its M-fold blocked version $\mathbf{x}(n)$ from Fig. 1(b) is WSS. Let us denote the autocorrelation sequence and power spectral density (psd) of $\mathbf{x}(n)$ by $\mathbf{R}_{\mathbf{xx}}(k)$ and $\mathbf{S}_{\mathbf{xx}}(z)$, respectively. In addition to this stationarity assumption on $x(n)$, we will also assume that the filter bank is *orthonormal*. This means that the matrices $\mathbf{H}(z)$ and $\mathbf{F}(z)$ from Fig. 1(b) satisfy [7],

$$\mathbf{H}(z) = \widetilde{\mathbf{F}}(z)\,,\ \widetilde{\mathbf{F}}(z)\mathbf{F}(z) = \mathbf{I}_L \tag{1}$$

(For the tilde notation [7] recall that $\widetilde{\mathbf{A}}(z) \triangleq \mathbf{A}^{\dagger}\left(1/z^{*}\right)$ for any $\mathbf{A}(z)$.) In other words, the matrix $\mathbf{F}(z)$ is *paraunitary*. With the above assumptions on the input and filter bank, it can easily be shown that minimizing the error of the output is equivalent to *compacting* the energy of the signal $\mathbf{w}(n)$.

Suppose that we wish to choose $\mathbf{H}(z)$ and $\mathbf{F}(z)$ subject to the orthonormality constraint of (1) to minimize the expected mean squared error between $\mathbf{x}(n)$ and $\mathbf{y}(n)$, defined as follows.

$$\xi \triangleq E\left[||\mathbf{x}(n) - \mathbf{y}(n)||^2\right] \tag{2}$$

If we define the blocked filter error $\mathbf{e}(n)$ as $\mathbf{e}(n) \triangleq \mathbf{x}(n) - \mathbf{y}(n)$ and denote the psd of $\mathbf{e}(n)$ by $\mathbf{S}_{\mathbf{ee}}(z)$, then from (2), we have,

$$\xi = \text{Tr}\left[E\left[\mathbf{e}(n)\mathbf{e}^{\dagger}(n)\right]\right] = \frac{1}{2\pi}\int_0^{2\pi} \text{Tr}\left[\mathbf{S}_{\mathbf{ee}}(e^{j\omega})\right]\,d\omega \tag{3}$$

From Fig. 1(b) and [7], it can be shown that we have,

$$\begin{aligned}\mathbf{S}_{\mathbf{ee}}(z) \;=\; & \mathbf{S}_{\mathbf{xx}}(z) - \mathbf{F}(z)\mathbf{H}(z)\mathbf{S}_{\mathbf{xx}}(z) - \mathbf{S}_{\mathbf{xx}}(z)\widetilde{\mathbf{H}}(z)\widetilde{\mathbf{F}}(z) \\ & + \mathbf{F}(z)\mathbf{H}(z)\mathbf{S}_{\mathbf{xx}}(z)\widetilde{\mathbf{H}}(z)\widetilde{\mathbf{F}}(z)\end{aligned} \tag{4}$$

Imposing the orthonormality constraint of (1) in (4) yields,

$$\text{Tr}\left[\mathbf{S}_{\mathbf{ee}}(z)\right] = \text{Tr}\left[\mathbf{S}_{\mathbf{xx}}(z)\right] - \text{Tr}\left[\widetilde{\mathbf{F}}(z)\mathbf{S}_{\mathbf{xx}}(z)\mathbf{F}(z)\right]$$

Substituting this into (3) leads to the following.

$$\begin{aligned}\xi \;=\; & \frac{1}{2\pi}\int_0^{2\pi} \text{Tr}\left[\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\right]\,d\omega \\ & \underbrace{-\,\frac{1}{2\pi}\int_0^{2\pi} \text{Tr}\left[\mathbf{F}^{\dagger}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{F}(e^{j\omega})\right]\,d\omega}_{\sigma^2_{\mathbf{w}}} \qquad (5) \\ =\; & \text{Tr}\left[\mathbf{R}_{\mathbf{xx}}(0)\right] - \sigma^2_{\mathbf{w}} \qquad (6)\end{aligned}$$

Hence, from (6), with the orthonormality constraint of (1) in effect, minimizing ξ from (2) is equivalent to maximizing $\sigma^2_{\mathbf{w}}$. But $\sigma^2_{\mathbf{w}}$ is just the energy of the subband vector process $\mathbf{w}(n)$ from Fig. 1(b), i.e. $\sigma^2_{\mathbf{w}} = \text{Tr}\left[\mathbf{R}_{\mathbf{ww}}(0)\right]$, where $\mathbf{R}_{\mathbf{ww}}(k)$ denotes the autocorrelation of $\mathbf{w}(n)$. Thus, minimizing the mean squared error of the overdecimated filter bank is equivalent to maximizing or *compacting* the energy of the subband process $\mathbf{w}(n)$. It can be shown that if no length constraints are made on the matrix $\mathbf{F}(z)$ from Fig. 1(b), then an optimal set of synthesis filters $F_k(z)$ for $0 \le k \le L-1$ from Fig. 1(a) which maximize $\sigma^2_{\mathbf{w}}$ from (6) are the first L ideal compaction filters appearing in the infinite order PCFB for $\mathbf{S}_{\mathbf{xx}}(z)$ [5, 8].

2.2. Imposing The FIR Constraint On The Matrix $\mathbf{F}(z)$

Suppose now that in addition to the orthonormality constraint of (1), the matrix $\mathbf{F}(z)$ is causal and FIR of length N. In other words, suppose that we have the following,

$$\mathbf{F}(z) = \sum_{n=0}^{N-1} \mathbf{f}(n)z^{-n} \tag{7}$$

where $\mathbf{f}(n)$ is the $M \times L$ impulse response of $\mathbf{F}(z)$. Define the $MN \times L$ impulse response matrix $\widehat{\mathbf{f}}$ and $M \times MN$ block delay matrix $\mathbf{d}(z)$ as follows.

$$\begin{aligned}\widehat{\mathbf{f}} \;&\triangleq\; \begin{bmatrix} \mathbf{f}^T(0) & \mathbf{f}^T(1) & \cdots & \mathbf{f}^T(N-1) \end{bmatrix}^T \\ \mathbf{d}(z) \;&\triangleq\; \begin{bmatrix} \mathbf{I}_M & z^{-1}\mathbf{I}_M & \cdots & z^{-(N-1)}\mathbf{I}_M \end{bmatrix}\end{aligned}$$

From (7), we clearly have $\mathbf{F}(z) = \mathbf{d}(z)\widehat{\mathbf{f}}$ and $\widetilde{\mathbf{F}}(z) = \widehat{\mathbf{f}}^{\dagger}\widetilde{\mathbf{d}}(z)$. Substituting this into (5) yields the following.

$$\sigma^2_{\mathbf{w}} = \text{Tr}\left[\,\widehat{\mathbf{f}}^{\dagger}\underbrace{\left(\frac{1}{2\pi}\int_0^{2\pi}\mathbf{d}^{\dagger}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{d}(e^{j\omega})\,d\omega\right)}_{\widehat{\mathbf{R}}}\widehat{\mathbf{f}}\,\right] \tag{8}$$

Here, the $MN \times MN$ matrix $\widehat{\mathbf{R}}$ is positive semidefinite and can be expressed in terms of the autocorrelation of $\mathbf{x}(n)$ as follows.

$$\widehat{\mathbf{R}} = \begin{bmatrix} \mathbf{R}_{\mathbf{xx}}(0) & \mathbf{R}_{\mathbf{xx}}(-1) & \cdots & \mathbf{R}_{\mathbf{xx}}(-(N-1)) \\ \mathbf{R}_{\mathbf{xx}}(1) & \mathbf{R}_{\mathbf{xx}}(0) & \cdots & \mathbf{R}_{\mathbf{xx}}(-(N-2)) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbf{xx}}(N-1) & \mathbf{R}_{\mathbf{xx}}(N-2) & \cdots & \mathbf{R}_{\mathbf{xx}}(0) \end{bmatrix} \tag{9}$$

From (9), note that $\widehat{\mathbf{R}}$ is the N-fold block autocorrelation matrix corresponding to $\mathbf{x}(n)$ and that $\widehat{\mathbf{R}}$ is a *block Toeplitz* matrix [2]. In the special case where the scalar input signal $x(n)$ is WSS with autocorrelation $R_{xx}(k)$, then we have,

$$[\mathbf{R}_{\mathbf{xx}}(k)]_{\ell,m} = R_{xx}(Mk + \ell - m)\,,\ 0 \leq \ell, m \leq M-1$$

and so $\widehat{\mathbf{R}}$ in this case is actually Toeplitz.

To analyze the orthonormality condition of (1) with the FIR constraint on $\mathbf{F}(z)$ in effect, define $\mathbf{G}(z) \triangleq \widetilde{\mathbf{F}}(z)\mathbf{F}(z)$. Then, from (1), in the time domain, we require,

$$\mathbf{g}(n) = \mathbf{f}^{\dagger}(-n) * \mathbf{f}(n) = \sum_{m} \mathbf{f}^{\dagger}(m)\mathbf{f}(m+n) = \mathbf{I}_L \delta(n) \tag{10}$$

where $\mathbf{g}(n)$ is the impulse response of $\mathbf{G}(z)$. Assuming $\mathbf{F}(z)$ to be causal and FIR as in (7), then $\mathbf{g}(n)$ can only be nonzero for $-(N-1) \leq n \leq (N-1)$. As $\mathbf{g}^{\dagger}(-n) = \mathbf{g}(n)$, the orthonormality conditions of (10) only need to be satisfied for $0 \leq n \leq N-1$. For $n = 0$, (10) can be expressed in terms of the matrix $\widehat{\mathbf{f}}$ as,

$$\widehat{\mathbf{f}}^{\dagger}\widehat{\mathbf{f}} = \mathbf{I}_L \tag{11}$$

For $1 \leq n \leq N-1$, (10) can be expressed in terms of $\widehat{\mathbf{f}}$ as follows.

$$\underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{f}^{\dagger}(0) & \mathbf{f}^{\dagger}(1) & \cdots & \mathbf{f}^{\dagger}(N-2) \\ \mathbf{0} & \mathbf{0} & \mathbf{f}^{\dagger}(0) & \cdots & \mathbf{f}^{\dagger}(N-3) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}^{\dagger}(0) \end{bmatrix}}_{\mathbf{C}} \underbrace{\begin{bmatrix} \mathbf{f}(0) \\ \mathbf{f}(1) \\ \vdots \\ \mathbf{f}(N-1) \end{bmatrix}}_{\widehat{\mathbf{f}}} = \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\mathbf{0}} \tag{12}$$

It should be noted that $\mathbf{C}$ from (12) is a function of the impulse response coefficients $\mathbf{f}(n)$. As such, the constraint in (12) is an *implicit quadratic* constraint. Combining (8), (11), and (12), the energy compaction problem in the presence of the FIR constraint on $\mathbf{F}(z)$ can be expressed as follows.

$$\begin{gathered} \text{Maximize } \sigma^2_{\mathbf{w}} = \text{Tr}\left[\widehat{\mathbf{f}}^{\dagger}\widehat{\mathbf{R}}\widehat{\mathbf{f}}\right] \\ \text{subject to } \widehat{\mathbf{f}}^{\dagger}\widehat{\mathbf{f}} = \mathbf{I}_L \text{ and } \mathbf{C}\widehat{\mathbf{f}} = \mathbf{0}_{L(N-1)\times L} \end{gathered} \tag{13}$$

with $\widehat{\mathbf{R}}$ and $\mathbf{C}$ as in (9) and (12), respectively.

In general, the optimization problem of (13) is *nonlinear* and *nonconvex* in terms of the elements of the matrix $\widehat{\mathbf{f}}$. What makes the problem difficult to solve is the implicit quadratic constraint $\mathbf{C}\widehat{\mathbf{f}} = \mathbf{0}$ from (12). Using the iterative approach for solving the optimization problem of (13) to be discussed in the next section, it is possible to turn this implicit quadratic constraint into an *explicit linear* constraint. Once this constraint becomes linear, the optimization at each iteration can be solved exactly using the *eigenfilter* technique [7], which is low in complexity and numerically stable. Before showing this, we will first formally present the iterative algorithm for solving the optimization problem of (13).

3. PROPOSED ITERATIVE ALGORITHM FOR SOLVING THE FIR COMPACTION PROBLEM

In what follows, let $\mathbf{f}_k(n)$ denote the impulse response $\mathbf{f}(n)$ at the k-th iteration. Also, define the $MN \times L$ matrix $\widehat{\mathbf{f}}_k$ and $L(N-1) \times MN$ matrix $\mathbf{C}_k$ as follows.

$$\widehat{\mathbf{f}}_k \triangleq \begin{bmatrix} \mathbf{f}_k^T(0) & \mathbf{f}_k^T(1) & \cdots & \mathbf{f}_k^T(N-1) \end{bmatrix}^T \tag{14}$$

$$\mathbf{C}_k \triangleq \begin{bmatrix} \mathbf{0} & \mathbf{f}_k^{\dagger}(0) & \mathbf{f}_k^{\dagger}(1) & \cdots & \mathbf{f}_k^{\dagger}(N-2) \\ \mathbf{0} & \mathbf{0} & \mathbf{f}_k^{\dagger}(0) & \cdots & \mathbf{f}_k^{\dagger}(N-3) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{f}_k^{\dagger}(0) \end{bmatrix} \tag{15}$$

Then, the proposed iterative algorithm is as follows.

Initialization:

Choose any $\mathbf{f}_0(n)$ which satisfies the orthonormality constraints of (10). This can be easily done using the complete characterization of FIR paraunitary systems in terms of degree one Householder-like building blocks [9, 7]. Compute $\widehat{\mathbf{f}}_0$ from $\mathbf{f}_0(n)$.

Iteration: For $k \geq 1$, do the following.

1. Compute the constraint matrix $\mathbf{C}_{k-1}$.
2. Solve the linearized optimization problem,

$$\begin{gathered} \text{Maximize } \sigma^2_{\mathbf{w},k} = \text{Tr}\left[\widehat{\mathbf{f}}_k^{\dagger}\widehat{\mathbf{R}}\widehat{\mathbf{f}}_k\right] \\ \text{subject to } \widehat{\mathbf{f}}_k^{\dagger}\widehat{\mathbf{f}}_k = \mathbf{I}_L \text{ and } \mathbf{C}_{k-1}\widehat{\mathbf{f}}_k = \mathbf{0}_{L(N-1)\times L} \end{gathered} \tag{16}$$

3. To measure the convergence of the iteration to an orthonormal solution, calculate the *orthonormality error matrix* at the k-th iteration defined by,

$$\boldsymbol{\epsilon}_k \triangleq \mathbf{C}_k\widehat{\mathbf{f}}_k \tag{17}$$

As we need $\boldsymbol{\epsilon}_k = \mathbf{0}$ in theory (from (12), (14), (15), and (17)), terminate the iteration when we have,

$$||\boldsymbol{\epsilon}_k||_F < \delta_T \tag{18}$$

where $||\boldsymbol{\epsilon}_k||_F$ denotes the *Frobenius norm* of $\boldsymbol{\epsilon}_k$ [2] and δ_T is a some small threshold value.

Before proceeding, it should be noted that there is no guarantee that the iterative algorithm will converge to an orthonormal solution, although in simulations it often does so as shown below. At present, there is no known method as to what should be done if the iteration fails to converge to an orthonormal solution. Furthermore, even if there is convergence, there is no guarantee that the resulting solution is globally optimal.

Despite this, often times in simulations such as those presented below, the algorithm performs well in terms of approaching the behavior of the ideal compaction filters of the infinite order PCFB. Also, the algorithm can be used for relatively large orders N, as the linearized optimization problem of (16) can be solved using the eigenfilter approach [7]. We now proceed to show how to solve the linearized optimization problem of (16).

4. SOLUTION TO THE ITERATIVE OPTIMIZATION PROBLEM

Consider the linear constraint $\mathbf{C}_{k-1}\widehat{\mathbf{f}}_k = \mathbf{0}$ from (16). This constraint holds iff the columns of $\widehat{\mathbf{f}}_k$ lie in the *null space* of $\mathbf{C}_{k-1}$ [2]. Let $\mathbf{U}_{k-1}$ denote a unitary matrix whose columns span the null space of $\mathbf{C}_{k-1}$. If ρ denotes the dimension of the null space of $\mathbf{C}_{k-1}$, then $\mathbf{U}_{k-1}$ is $MN \times \rho$. As the columns of $\widehat{\mathbf{f}}_k$ must lie in the null space of $\mathbf{C}_{k-1}$, $\widehat{\mathbf{f}}_k$ must be of the form $\widehat{\mathbf{f}}_k = \mathbf{U}_{k-1}\mathbf{a}$ for some arbitrary $\rho \times L$ matrix $\mathbf{a}$. Hence, we have,

$$\mathbf{C}_{k-1}\widehat{\mathbf{f}}_k = \mathbf{0} \Longleftrightarrow \widehat{\mathbf{f}}_k = \mathbf{U}_{k-1}\mathbf{a} \qquad (19)$$

Given that $\mathbf{C}_{k-1}$ is $L(N-1) \times MN$, we can easily argue that the dimension of its null space ρ satisfies $\rho \geq L$. Hence, the linear constraint $\mathbf{C}_{k-1}\widehat{\mathbf{f}}_k = \mathbf{0}$ transforms the problem of finding $\widehat{\mathbf{f}}_k$ into that of finding the $\rho \times L$ matrix $\mathbf{a}$. The quantity $\mathbf{a}$ is arbitrary but must be such that the unitary constraint $\widehat{\mathbf{f}}_k^\dagger \widehat{\mathbf{f}}_k = \mathbf{I}_L$ from (16) is satisfied. Clearly, from (19), we have,

$$\widehat{\mathbf{f}}_k^\dagger \widehat{\mathbf{f}}_k = \mathbf{I}_L \Longleftrightarrow \mathbf{a}^\dagger \mathbf{a} = \mathbf{I}_L$$

upon exploiting the unitarity of $\mathbf{U}_{k-1}$. As can be seen, the constraints of (16) transform the problem of finding $\widehat{\mathbf{f}}_k$ into that of finding $\mathbf{a}$ where $\mathbf{a}$ is allowed to by any $\rho \times L$ unitary matrix. Hence, the optimization problem of (16) can be recast as follows.

$$\text{Maximize } \sigma^2_{\mathbf{w},k} = \text{Tr}\left[\mathbf{a}^\dagger \overline{\mathbf{R}}_{k-1}\mathbf{a}\right] \text{ where } \overline{\mathbf{R}}_{k-1} \triangleq \mathbf{U}^\dagger_{k-1}\widehat{\mathbf{R}}\mathbf{U}_{k-1}$$
$$\text{subject to the constraint } \mathbf{a}^\dagger \mathbf{a} = \mathbf{I}_L \qquad (20)$$

The solution to this problem follows from a generalization of Rayleigh's principle [2, p. 191] and is as follows. Suppose that $\overline{\mathbf{R}}_{k-1}$ has the following unitary diagonalization.

$$\overline{\mathbf{R}}_{k-1} = \mathbf{V}_{k-1}\mathbf{\Lambda}_{k-1}\mathbf{V}^\dagger_{k-1}$$

where $\mathbf{V}_{k-1}$ is a $\rho \times \rho$ matrix of eigenvectors of $\overline{\mathbf{R}}_{k-1}$ and $\mathbf{\Lambda}_{k-1}$ is a diagonal matrix consisting of the eigenvalues of $\overline{\mathbf{R}}_{k-1}$. In addition, suppose that $\mathbf{\Lambda}_{k-1} = \text{diag}\left(\lambda_{k-1,0}, \lambda_{k-1,1}, \ldots, \lambda_{k-1,\rho-1}\right)$ and that the eigenvalues have been ordered in decreasing order, i.e. $\lambda_{k-1,0} \geq \lambda_{k-1,1} \geq \cdots \geq \lambda_{k-1,\rho-1}$. Then, the solution to the optimization problem of (20) is given to be [2],

$$\sigma^2_{\mathbf{w},k} = \sum_{i=0}^{L-1} \lambda_{k-1,i}$$

which occurs iff we have,

$$\mathbf{a} = \mathbf{V}_{k-1}\mathbf{b}$$

where $\mathbf{b}$ is a $\rho \times L$ matrix of the form,

$$\mathbf{b} = \begin{bmatrix} \mathbf{B} \\ \mathbf{0}_{(\rho-L)\times L} \end{bmatrix} \qquad (21)$$

and $\mathbf{B}$ is any $L \times L$ square unitary matrix. In other words, the optimal $\mathbf{a}$ is a such that its columns are unitary combinations of the first L eigenvectors of $\overline{\mathbf{R}}_{k-1}$. Once an optimal $\mathbf{a}$ has been found, the corresponding optimal $\widehat{\mathbf{f}}_k$ can be found using $\widehat{\mathbf{f}}_k = \mathbf{U}_{k-1}\mathbf{a}$. As computing the optimal synthesis filter matrix $\widehat{\mathbf{f}}_k$ requires the eigendecomposition of a particular matrix, it follows that the original linearized optimization problem of (16) is an eigenfilter type problem [7]. The simulation results presented in the next section shown the merit of the proposed iterative eigenfilter method.

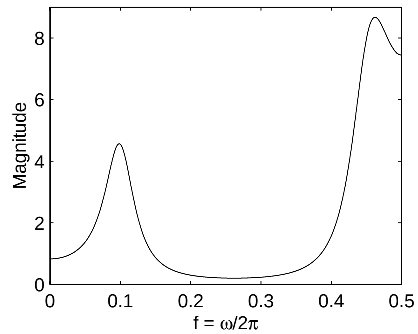

Fig. 2. Input power spectral density $S_{xx}(e^{j\omega})$.

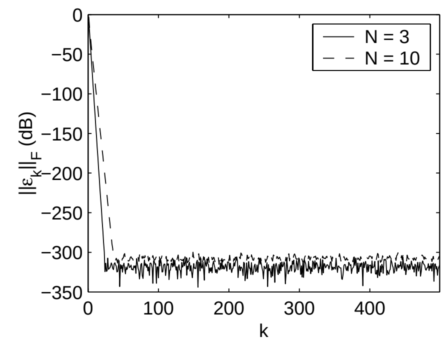

Fig. 3. Orthonormality error $||\boldsymbol{\epsilon}_k||_F$ vs. the iteration index k ($L = 1$, $M = 7$).

5. SIMULATION RESULTS

To test the proposed iterative eigenfilter algorithm, we chose the input process $x(n)$ to be a real WSS autoregressive process of order 4 (AR(4)) whose power spectrum $S_{xx}(e^{j\omega})$ is plotted in Fig. 2. For all of the simulation results here, we used $M = 7$, i.e. the block size of the filter bank used was chosen to be 7. Also, the matrix $\mathbf{B}$ from (21) was chosen to be $\mathbf{I}_L$ for all examples.

We first considered the design of a single channel of the overdecimated system (i.e. $L = 1$). The observed error in orthonormality using the iterative eigenfilter method is shown in Fig. 3 in dB for two values of orders, namely $N = 3$ and 10. In order to observe the behavior of our algorithm, we ran it for 500 iterations and opted not to choose a stopping threshold value δ_T from (18). As can be seen from Fig. 3, proposed method is indeed converging toward an orthonormal solution for both cases of filter orders. The error $||\boldsymbol{\epsilon}_k||_F$ saturates at around -300 dB for both cases, most likely due to quantization effects as a result of finite precision arithmetic.

To gauge the performance of the algorithm, a plot of the magnitude squared response of the resulting synthesis filter $F_0(z)$ from Fig. 1(a) is shown in Fig. 4 for the two orders $N = 3$ and 10, along with that of the first filter of the infinite order PCFB. As can be seen, both FIR filters have a response close to that of the ideal compaction filter. Furthermore, the higher order filter offers a better approximation than the lower order one, in line with intuition.

For a quantitative measure of the performance of the algorithm, we opted to calculate the *compaction gain* of the designed filters. This quantity is defined as [8],

$$G_{\text{comp}} \triangleq \frac{\frac{1}{2\pi}\int_0^{2\pi} \left|F_0(e^{j\omega})\right|^2 S_{xx}(e^{j\omega})\, d\omega}{\sigma_x^2}$$

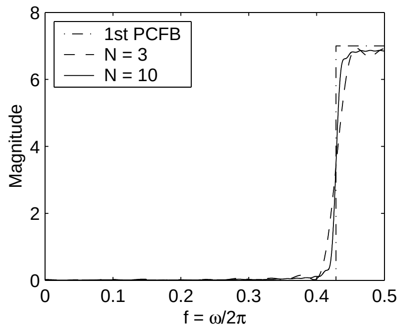

Fig. 4. Magnitude squared responses of the designed FIR synthesis filter $F_0(z)$ along with the first filter of the infinite order PCFB ($L = 1$, $M = 7$).

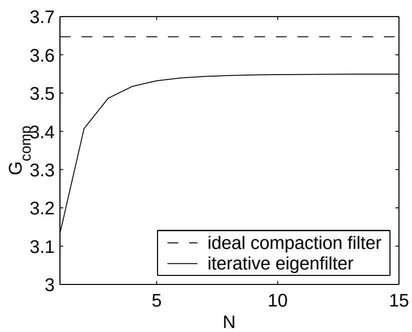

Fig. 5. Compaction gain G_{comp} vs. the filter order parameter N.

The ideal compaction filter maximizes this quantity over all filters satisfying the required orthonormality condition of (1) [8]. A plot of the observed compaction gain as a function of the filter order parameter N is shown in Fig. 5. Though the compaction gain increases monotonically as N increases, it appears to saturate well below the ideal compaction gain. At this time, it is not known why this phenomenon occurs. Despite this, however, for small orders, the observed compaction gain is reasonably large.

To further test the algorithm, we then considered the design of two channels of the overdecimated system (i.e. $L = 2$) and fixed the order to be $N = 10$. The observed error in orthonormality (in dB) as a function of iteration is shown in Fig. 6. As before, it can be seen that the algorithm is converging to an orthonormal solution. The magnitude squared responses of the designed synthesis filters $F_0(z)$ and $F_1(z)$ are shown in Fig. 7 along with those of the first two filters of the infinite order PCFB. From this, it is clear that the proposed algorithm is yielding filters close to the ideal compaction filters of the infinite order PCFB, as desired.

6. CONCLUDING REMARKS

An iterative eigenfilter method for designing overdecimated FIR compaction filter banks was proposed and shown to be useful in simulations. Future research includes finding ways to guarantee the convergence toward an orthonormal solution and improving the performance of the algorithm in terms of compaction gain.

7. REFERENCES

[1] S. Akkarakaran and P. P. Vaidyanathan, "Filterbank optimization with convex objectives and the optimality of principal component forms," *IEEE Trans. Signal Processing*, 49(1):100-114, Jan. 2001.

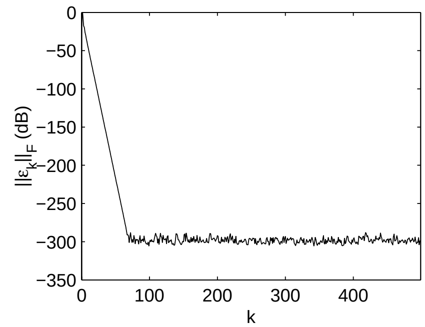

Fig. 6. Orthonormality error $||\boldsymbol{\epsilon}_k||_F$ vs. the iteration index k ($L = 2$, $M = 7$, $N = 10$).

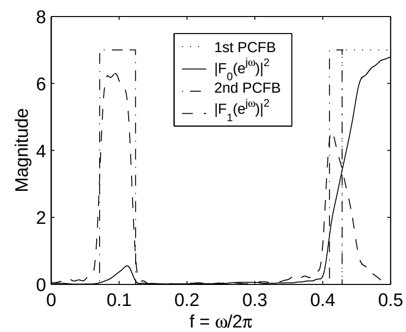

Fig. 7. Magnitude squared responses of the designed FIR synthesis filters $F_0(z)$ and $F_1(z)$ along with the first two filters of the infinite order PCFB ($L = 2$, $M = 7$, $N = 10$).

[2] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, U.K., 1985.

[3] A. Kıraç and P. P. Vaidyanathan, "Theory and design of optimum FIR compaction filters," *IEEE Trans. Signal Processing*, 46(4):903-919, Apr. 1998.

[4] P. Moulin and M. K. Mıhçak, "Theory and design of signal-adapted FIR paraunitary filter banks," *IEEE Trans. Signal Processing*, 46(4):920-929, Apr. 1998.

[5] M. K. Tsatsanis and G. B. Giannakis, "Principal component filter banks for optimal multiresolution analysis," *IEEE Trans. Signal Processing*, 43(8):1766-1777, Aug. 1995.

[6] J. Tuqan and P. P. Vaidyanathan, "A state space approach to the design of globally optimal FIR energy compaction filters," *IEEE Trans. Signal Processing*, 48(10):2822-2838, Oct. 2000.

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

[8] P. P. Vaidyanathan, "Theory of optimal orthonormal subband coders," *IEEE Trans. Signal Processing*, 46(6):1528-1543, Jun. 1998.

[9] P. P. Vaidyanathan, T. Q. Nguyen, Z. Doğanata and T. Saramäki, "Improved technique for design of perfect reconstruction FIR QMF banks with lossless polyphase matrices," *IEEE Trans. Acoust., Speech, Signal Processing*, 37(7):1042-1056, Jul. 1989.

[10] B. Xuan and R. H. Bamberger, "FIR principal component filter banks," *IEEE Trans. Signal Processing*, 46(4):930-940, Apr. 1998.