Global Optimum Design of Uniform FIR Filter Bank With Magnitude Constraints

C. Z. Wu, K. L. Teo, *Senior Member, IEEE*, V. Rehbock, and H. H. Dam

Abstract—The optimum design of a uniform finite impulse response filter bank can be formulated as a nonlinear semi-infinite optimization problem. However, this optimization problem is non-convex with infinitely many inequality constraints. In this paper, we propose a new hybrid approach for solving this highly challenging nonlinear, nonconvex semi-infinite optimization problem. In this approach, a gradient-based method is used in conjunction with a filled function method to determine a global minimum of the problem. This new hybrid approach finds an optimal result independent of the initial guess of the solution. The method is applied to some existing examples. The results obtained are superior to those obtained by other existing methods.

Index Terms—Constraint transcription, filled function, finite impulse response (FIR) filter bank, global optimization.

I. INTRODUCTION

THE rapid development of multirate digital signal processing is driven by emerging new applications. These include subband coding of speech, audio and video signals, multicarrier data transmission, fast transforms using digital filter banks, and discrete wavelet analysis for all types of signals [1]. In multirate digital signal processing, an analysis filter is used to divide the signal to be processed into subbands. They are then decimated according to the new bandwidth of the subbands. Since nonperfect filters are used within the filter bank, the decimation will cause aliasing of the subband signals. It is possible to cancel this aliasing through the design of a synthesis filter bank in such a way that the whole multirate chain yields no distortion. In other words, the total transfer function becomes a simple delay. This is referred to as the perfect reconstruction property (PR) [4].

PR decimated filter banks have been of great interest in subband coding [1]–[4]. However, these filter banks are less suitable for subband adaptive filtering since the PR property cannot be maintained when the subband signals are modified by filters with arbitrary magnitude and phase response (see [6] and [7]). Thus, aliasing may be caused in the reconstructed output of the subband adaptive filter. To overcome this problem, optimization

methods are often used in the design of filter banks. In [8], the design of a uniform discrete Fourier transform (DFT)-filter bank is considered. This problem is solved by a two-step optimization method. In the first step, the analysis filter bank is designed such that the aliasing term in each subband is minimized individually. On this basis, minimal aliasing is achieved at the output without aliasing cancellation. In the second step, the synthesis filter bank is designed to match the analysis filter bank in such a way that the analysis-synthesis response is optimized while all aliasing terms in the output signal are individually suppressed without aiming at aliasing cancellation. To design such a filter bank, it is necessary to solve two quadratic unconstrained optimization problems. Clearly, this method does not give rise to an optimum result since the design of such a filter bank is carried out by determining the analysis and synthesis filter banks separately.

The delay introduced by a filter bank has a critical impact in many practical applications [9]. Thus, there is an increasing interest in finding the minimum delay and the bounds on the delay and phase characteristics when the PR requirement is relaxed. In [10], the design of a uniformly modulated oversampled finite impulse response (FIR) filter bank with group delay specifications is considered. Accurate linear approximations for the phase and the group delay of the total filter bank are derived. Then, the design of this filter bank is formulated as a nonlinear optimization problem with group delay and amplitude distortion constraints. To solve this optimization problem, an iterative alternating method is applied. More specifically, the analysis filter is fixed when solving the synthesis filter, while the synthesis filter is fixed when solving the analysis filter. This alternating process is repeated iteratively until no improvement can be made. In each alternating iteration, a semi-infinite quadratic programming is solved via discretization method. This iterative alternating method has produced better results than those obtained in [8]. However, it does not, in general, produce an optimal solution for the overall system that is nonlinear and nonconvex with infinitely many inequality constraints. Therefore, there is a need to develop global optimization techniques.

In this paper, we develop a new hybrid method for solving this nonlinear, nonconvex semi-infinite optimization problem. We first use the constraint transcription method [11] to approximate the continuous inequality constraints by a sequence of integral constraints. Each of these integral constraints is then appended to the cost function as a penalty term. This give rise to a sequence of unconstrained optimization problems. Each of them is solvable by any gradient-based optimization algorithm. In [11], it is shown that if the penalty parameter is sufficiently large, the optimal solution of the unconstrained optimization problem satisfies the continuous inequality constraints of the original optimization problem. Thus, we obtain a sequence of approximate
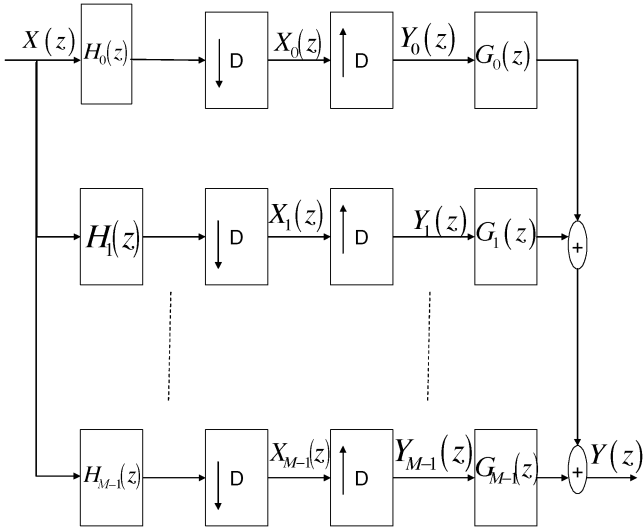
Fig. 1. Analysis and synthesis filter banks.

optimal solutions. In [11], it is further shown that if these optimal solutions are global, then they converge to the global optimal solution of the original nonlinear, nonconvex semi-infinite optimization problem. However, it is clear that the solutions obtained by using a gradient-based optimization algorithm are unlikely to be globally optimal. Therefore, we adopt a filled function method [13] to supplement the gradient-based algorithm. This filled function will be used to escape from a local basin in the search space and find an improved point. The original problem is then solved again with this improved point as an initial guess, leading to an improved local minimum. This process is repeated until a global minimum of the original problem is obtained. The new hybrid method finds the optimal result independent of the initial guess of the solution. This method is applied to some existing examples, and the results obtained are shown to be superior to those obtained by existing methods.

II. PROBLEM FORMULATION

A. Definition and Notations

An M-channel uniformly modulated FIR filter bank consists of a set of M branches. Each branch consists of an analysis filter $H_m(z)$, a decimator with decimation factor D, an interpolator with interpolation factor D, and a synthesis filter $G_m(z)$. Such a structure is depicted in Fig. 1.

The input signal $X(z)$ is filtered by M analysis filters $H_m(z)$, $0 \le m \le M - 1$, where M is the number of subbands. These analysis filters are modulated versions of the prototype analysis filter according to

$$H_m(z) = H(zW_M^m) \tag{1}$$

where $W_M = e^{-j2\pi/M}$. The analysis prototype filter is modeled as an FIR filter with real impulse response $\boldsymbol{h} = [h(0), \cdots, h(L_a - 1)]^T$. The corresponding transfer function is given by

$$H(z) = \boldsymbol{h}^T \phi_a(z) \tag{2}$$

where

$$\phi_a(z) = \left[1, z^{-1}, \cdots, z^{-(L_a - 1)}\right]^T.$$

Then, the output of each analysis filter is decimated by a factor $D, D \le M$. Thus, each subband signal $X_m(z)$ can be expanded as

$$X_m(z) = \frac{1}{D} \sum_{d=0}^{D-1} H\left(z^{\frac{1}{D}} W_M^m W_D^d\right) X\left(z^{\frac{1}{D}} W_D^d\right) \tag{3}$$

where $W_D = e^{-j2\pi/M}$. These signals are then interpolated by the same factor D and filtered by the synthesis filter bank $G_m(z)$. Similarly, the synthesis filters are modulated versions of the prototype synthesis filter according to

$$G_m(z) = G(zW_M^m) \tag{4}$$

where the synthesis prototype filter $G(z)$ is also modeled as an FIR filter with length L_s, that is

$$G(z) = \boldsymbol{g}^T \phi_s(z) \tag{5}$$

where

$$\boldsymbol{g} = [g(0), \cdots, g(L_s - 1)]^T$$

and

$$\phi_s(z) = \left[1, z^{-1}, \ldots, z^{-(L_s - 1)}\right]^T.$$

From above, the output $Y(z)$ can be expressed as

$$
\begin{aligned}
Y(z) &= \sum_{m=0}^{M-1} Y_m(z) G_m(z) \\
&= \frac{1}{D} \sum_{d=0}^{D-1} X(zW_D^d) \sum_{m=0}^{M-1} H(zW_M^m W_D^d) G(zW_M^m) \\
&= \frac{1}{D} X(z) \sum_{m=0}^{M-1} H(zW_M^m) G(zW_M^m) \\
&\quad + \frac{1}{D} \sum_{d=1}^{D-1} X(zW_D^d) \\
&\quad \cdot \sum_{m=0}^{M-1} H(zW_M^m W_D^d) G(zW_M^m).
\end{aligned}
\tag{6}
$$

The terms

$$\frac{1}{D} \sum_{m=0}^{M-1} H(zW_M^m W_D^d) G(zW_M^m) \tag{7}$$

where $1 \le d \le D - 1$, can be viewed as the transfer functions relating to the residual aliasing terms in the output signal. Similarly

$$T(z) = \frac{1}{D} \sum_{m=0}^{M-1} H(zW_M^m) G(zW_M^m) \tag{8}$$

can be viewed as the transfer function representing the desired output signal spectrum. Substituting (2) and (5) into (8), we obtain

$$T(z) = \boldsymbol{h}^T \Psi(z) \boldsymbol{g} \tag{9}$$

where

$$\Psi(z) = \frac{1}{D} \sum_{m=0}^{M-1} \phi_a(zW_M^m) \phi_s^T(zW_M^m). $$

Thus, the (n_1, n_2) element of $\Psi(z)$ can be expressed as

$$
\Psi_{n_1,n_2}(e^{jw}) \\
= \begin{cases} 0, & \text{if } mod(n_1+n_2-2, M) \neq 0 \\ \frac{M}{D} e^{-jw(n_1+n_2-2)}, & \text{if } mod(n_1+n_2-2, M) = 0 \end{cases} \tag{10}
$$

where $1 \leq n_1 \leq L_a$, $1 \leq n_2 \leq L_s$, and "mod" denotes the modulus.

B. Prototype Analysis and Synthesis Design by Optimization

Let the desired frequency response of the filter bank be denoted by

$$T_d(e^{jw}) = e^{-jw\tau_d} \tag{11}$$

where τ_d is a constant desired group delay. In [10], the design of such a filter bank is formulated as a nonlinear semi-infinite optimization problem given

$$\text{minimize } F(\boldsymbol{h}, \boldsymbol{g}) + \kappa G(\boldsymbol{h}) \tag{12}$$

subject to

$$\left| \boldsymbol{h}^T \Psi(e^{jw}) \boldsymbol{g} - e^{-jw\tau_d} \right| \leq \epsilon_T, \quad \forall w \in [0, \pi] \tag{13}$$

and

$$|e_\theta(w)| \leq \nu_\theta, \quad \forall w \in [0, \pi] \tag{14}$$

where

$$F(\boldsymbol{h}, \boldsymbol{g}) = \frac{c_1}{2\pi} \int_{-\pi}^{\pi} \sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \left| \boldsymbol{h}^T \Phi_{m,d}(e^{jw}) \boldsymbol{g} \right|^2 dw \tag{15}$$

is the residual aliasing

$$G(\boldsymbol{h}) = \boldsymbol{h}^T B \boldsymbol{h} \tag{16}$$

is the inband aliasing

$$e_\theta(w) = \theta(w) - \theta_d(w) = \boldsymbol{h}^T \Xi(w) \boldsymbol{g}$$

is the approximation for the phase error, and

$$
\Phi_{m,d}(e^{jw}) = \phi_a\left(e^{jw} W_M^m W_D^d\right) \phi_s^H\left(e^{jw} W_M^m\right) \\
B = \frac{c_2}{2\pi} \sum_{d=1}^{D-1} \int_{-\pi}^{\pi} \phi_a\left(e^{\frac{jw}{D}} W_D^d\right) \phi_a^H\left(e^{\frac{jw}{D}} W_D^d\right) dw \\
\Xi(w) = \mathcal{I}\left\{ \Psi(w) e^{jw\tau_d} \right\} \tag{17}
$$

where κ, c_1, c_2 are appropriate constants, ϵ_T, ν_θ are specified error constants, $(\cdot)^H$ denotes the Hermitian operation of a vector, and $\mathcal{I}\{\cdot\}$ denotes the imaginary part. Let this optimization problem be referred to as Problem (P). In this paper, we choose $c_1 = c_2 = 1$. Calculating the integrals in (17) and (15), the (t, s) element of B obtained is

$$B_{t,s} = \frac{D\varphi(t,s) \sin\left(\frac{(t-s)\pi}{D}\right)}{\pi(t-s)} \tag{18}$$

while

$$
F(\boldsymbol{h}, \boldsymbol{g}) = M \sum_{t=1}^{La} \sum_{s=1}^{Ls} \sum_{k=1}^{La} h(t)g(s)h(k) \\
\cdot g(t+s-k) \varkappa_{[1,L_s]}(t+s-k)\varphi(t,s) \\
\varphi(t,s) = \begin{cases} D-1, & \text{if } mod(t-s, D) = 0 \\ -1, & \text{else} \end{cases} \tag{19}
$$

and

$$\varkappa_{[1,L_s]}(l) = \begin{cases} 1, & \text{if } 1 \leq l \leq L_s \\ 0, & \text{else} \end{cases}. $$

III. SOLUTION STRATEGY

Note that Problem (P) is a nonlinear nonconvex semi-infinite optimization problem. This optimization problem is very difficult to solve and many local minima may exist. We develop a new hybrid method for solving Problem (P). First, the constraint transcription method [11] is applied to approximate the continuous inequality constraints by a sequence of integral constraints as detailed below. The constraints (13) and (14) are, respectively, equivalent to the following constraints:

$$G_1(\boldsymbol{x}, w) = \left| \boldsymbol{h}^T \Psi(e^{jw}) \boldsymbol{g} - e^{-jw\tau_d} \right|^2 - \epsilon_T^2 \leq 0, \quad \forall w \in [0, \pi] \tag{20}$$

and

$$G_2(\boldsymbol{x}, w) = |e_\theta(w)|^2 - \nu_\theta^2 \leq 0, \quad \forall w \in [0, \pi] \tag{21}$$

where $\boldsymbol{x} = (\boldsymbol{h}^T, \boldsymbol{g}^T)^T \in \mathbb{R}^{L_a + L_s}$. These continuous inequality constraints (20) and (21) are equivalent to

$$\tilde{G}_i(\boldsymbol{x}) = \int_0^\pi \max\{G_i(\boldsymbol{x}, w), 0\} \, dw = 0, \quad i = 1, 2. \tag{22}$$

Since the function $\max\{z, 0\}$ is not differentiable at $z = 0$, the function $\tilde{G}_i(\boldsymbol{x})$ is, in general, nonsmooth. Consequently, gradient-based optimization routines could not be used to handle optimization problems involving such equality constraints. Thus, as in [11], we construct the following smoothing approximation:

$$\tilde{G}_{i,\varepsilon}(\boldsymbol{x}, w) = \begin{cases} 0, & \text{if } G_i(\boldsymbol{x}, w) < -\varepsilon \\ (G_i(\boldsymbol{x}, w) + \varepsilon)^2 / 4\varepsilon, & \text{if } -\varepsilon \leq G_i(\boldsymbol{x}, w) \leq \varepsilon. \\ G_i(\boldsymbol{x}, w), & \text{if } G_i(\boldsymbol{x}, w) > \varepsilon \end{cases} \tag{23}$$

Denote

$$\hat{G}_{i,\varepsilon}(\boldsymbol{x}) = \int_0^{\pi} \tilde{G}_{i,\varepsilon}(\boldsymbol{x}, w) dw. \qquad (24)$$

Define an approximate version of Problem (P) as

$$\min f_{\varepsilon,\gamma}(\boldsymbol{x}) \qquad (25)$$

where

$$f_{\varepsilon,\gamma}(\boldsymbol{x}) = F(\boldsymbol{h}, \boldsymbol{g}) + \kappa G(\boldsymbol{h}) + \gamma \sum_{i=1}^{2} \hat{G}_{i,\varepsilon}(\boldsymbol{x}) \qquad (26)$$

and γ is a penalty constant. Let this problem be referred to as Problem $(\mathrm{P}_{\varepsilon,\gamma})$ and let $\mathcal{F}$ be the feasible region of Problem (P), i.e.,

$$\mathcal{F} = \left\{ \boldsymbol{x} \in \mathbb{R}^{L_a + L_s} : G_i(\boldsymbol{x}, w) \le 0, \ \forall w \in [0, \pi], \ i = 1, 2 \right\}. \qquad (27)$$

Define

$$\overset{\circ}{\mathcal{F}} = \left\{ \boldsymbol{x} \in \mathbb{R}^{L_a + L_s} : G_i(\boldsymbol{x}, w) < 0, \ \forall w \in [0, \pi], \ i = 1, 2 \right\}. \qquad (28)$$

We assume that the following conditions are satisfied.

A1) $\overset{\circ}{\mathcal{F}}$ is nonempty.

A2) For any $\hat{\boldsymbol{x}} \in \mathcal{F}$, there exists an $\bar{\boldsymbol{x}} \in \overset{\circ}{\mathcal{F}}$ such that $\alpha \bar{\boldsymbol{x}} + (1-\alpha)\hat{\boldsymbol{x}} \in \overset{\circ}{\mathcal{F}}$ for all $\alpha \in (0, 1]$.

Theorem 3.1: Consider Problem (P) and assume that A1) and A2) are satisfied. Then, for any $\varepsilon > 0$, there exists a $\gamma(\varepsilon) > 0$ such that for any $\gamma > \gamma(\varepsilon)$, any solution of Problem $(\mathrm{P}_{\varepsilon,\gamma})$ satisfies the continuous inequality constraints of Problem (P).

Proof: The proof is given in Appendix I. ∎

Furthermore, we have the following.

Theorem 3.2: Assume that the conditions given in Theorem 3.1 are satisfied. Let $\boldsymbol{x}^* = \left((\boldsymbol{h}^*)^T, (\boldsymbol{g}^*)^T\right)^T$ be a global optimal solution of Problem (P) and let $\boldsymbol{x}^*_{\varepsilon,\bar{\gamma}(\varepsilon)} = \left((\boldsymbol{h}^*_{\varepsilon\bar{\gamma}(\varepsilon)})^T, (\boldsymbol{g}^*_{\varepsilon,\bar{\gamma}(\varepsilon)})^T\right)^T$ be a global solution of Problem $(\mathrm{P}_{\varepsilon,\bar{\gamma}(\varepsilon)})$, where $\bar{\gamma}$ is chosen sufficiently large such that $\boldsymbol{x}^*_{\varepsilon,\bar{\gamma}(\varepsilon)} \in \mathcal{F}$. Then

$$\lim_{\varepsilon \to 0} f_{\varepsilon,\bar{\gamma}(\varepsilon)}\left(\boldsymbol{x}^*_{\varepsilon,\gamma}\right) = F(\boldsymbol{h}^*, \boldsymbol{g}^*) + \kappa G(\boldsymbol{h}^*, \boldsymbol{g}^*). \qquad (29)$$

Proof: The proof is given in Appendix II. ∎

In practice, we only consider a finite sequence in ε. By Theorem 3.1, the choice of $\bar{\gamma}(\varepsilon)$ for each ε in this sequence is finite. Numerical experiments confirm this fact.

We are now in a position to present an algorithm to search for a minimizer of Problem $(\mathrm{P}_{\varepsilon,\gamma})$.

Algorithm 3.1:

Step 1) Set $\varepsilon = 0.1$, $\gamma = 10$, $\varepsilon_0 = 10^{-8}$, $\gamma_0 = 10^8$.

Step 2) Solve Problem $(\mathrm{P}_{\varepsilon,\gamma})$ by a quasi-Newton method and let the solution obtained be denoted by $\boldsymbol{x}^*_{\varepsilon,\gamma}$.

Step 3) Check the feasibility of $G_i(\boldsymbol{x}, w) \le 0$, $i = 1, 2$, at a sufficiently dense set of points in $[0, \pi]$. If these constraints are satisfied, go to Step 4). Otherwise, set $\gamma = 10\gamma$. If $\gamma > \gamma_0$, stop. There is no feasible point found. Otherwise, go to Step 2).

Step 4) Set $\varepsilon = \varepsilon/10$. If $\varepsilon < \varepsilon_0$, go to Step 2). Otherwise, stop.

Remark 3.1: In Step 2) to Step 3) of Algorithm 3.1, it follows from Theorem 3.1 that the increment of γ for each $\varepsilon > 0$ only needs to be carried out a finite number of times.

From (17), (19), and (25), we see that Problem $(\mathrm{P}_{\varepsilon,\gamma})$ is non-convex. Clearly, Algorithm 3.1 may not yield a global optimal solution. Thus, strictly speaking, Theorem 3.2 is not applicable. To overcome this difficulty, we develop a filled function method to search for a global optimal solution in Step 2) of Algorithm 3.1. Suppose that $\boldsymbol{x}^*_{\varepsilon,\gamma}$ is a minimizer obtained by a gradient-based method in Step 2) of Algorithm 3.1. We introduced the following filled function (see [13]):

$$\begin{aligned}
p\left(\boldsymbol{x}, \boldsymbol{x}^*_{\varepsilon,\gamma}, \mu, \rho\right) = & f_{\varepsilon,\gamma}\left(\boldsymbol{x}^*_{\varepsilon,\gamma}\right) - \rho \left\|\boldsymbol{x} - \boldsymbol{x}^*_{\varepsilon,\gamma}\right\|^2 \\
& - \min\left[f_{\varepsilon,\gamma}\left(\boldsymbol{x}^*_{\varepsilon,\gamma}\right), f_{\varepsilon,\gamma}(\boldsymbol{x})\right] \\
& + \mu\left\{\max\left[0, f_{\varepsilon,\gamma}(\boldsymbol{x}) - f_{\varepsilon,\gamma}\left(\boldsymbol{x}^*_{\varepsilon,\gamma}\right)\right]\right\}^2 \quad (30)
\end{aligned}$$

where μ, ρ are properly chosen small parameters. The proposed global algorithm makes use of the filled function (30) as follows. Initially, a gradient based optimization method is applied to determine a local minimizer $\boldsymbol{x}^*_{\varepsilon,\gamma}$ of Problem $(\mathrm{P}_{\varepsilon,\gamma})$. To allow the algorithm to move away from the current local minimum and its associated basin, we then choose appropriate values for μ and ρ and apply a gradient-based optimization method to the filled function $p(\boldsymbol{x}, \boldsymbol{x}^*_{\varepsilon,\gamma}, \mu, \rho)$ defined by (30). Suppose that this results in the points $\boldsymbol{x}_{\varepsilon,\gamma}$, where $f_{\varepsilon,\gamma}(\boldsymbol{x}_{\varepsilon,\gamma}) < f_{\varepsilon,\gamma}(\boldsymbol{x}^*_{\varepsilon,\gamma})$. The algorithm has then escaped from a local minimum and its associated basin, and we can return to minimizing $f_{\varepsilon,\gamma}$ with a gradient-based method, using $\boldsymbol{x}_{\varepsilon,\gamma}$ as a starting point to determine another local minimum. The cycle is then repeated as often as necessary. For more details, see Remark 3.2.

Algorithm 3.2:

Step 1) Starting from an initial point $\boldsymbol{x}^0$, $f_{\varepsilon,\gamma}(\boldsymbol{x})$ is minimized by Algorithm 3.1. Let the local minimizer obtained be denoted as $\boldsymbol{x}^*_{\varepsilon,\gamma}$.

Step 2) Set μ_0, ρ_0. Let $\bar{\mu}$ and $\bar{\rho}$ be two appropriately chosen small positive numbers. Let $\varkappa$ be a large number and let $S = \{\boldsymbol{x}^{0,1}, \ldots, \boldsymbol{x}^{0,m}\}$ such that $\boldsymbol{x}^{0,k} \in N(\boldsymbol{x}^*_{\varepsilon,\gamma}, \delta_k) \setminus N(\boldsymbol{x}^*_{\varepsilon,\gamma}, \delta_{k-1})$, where $\delta_k > \delta_{k-1}$ and where $N(\boldsymbol{x}^*_{\varepsilon,\gamma}, \delta_k)$ denotes a neighborhood of $\boldsymbol{x}^*_{\varepsilon,\gamma}$ with center $\boldsymbol{x}^*_{\varepsilon,\gamma}$ and radius δ_k.

Step 3) Set $\mu = \mu_0$, $\rho = \rho_0$, $i = 1$.

Step 4) Set $\boldsymbol{x} = \boldsymbol{x}^{0,i}$. Use $\boldsymbol{x}$ as the initial point for the minimization of $p(\boldsymbol{x}, \boldsymbol{x}^*_{\varepsilon,\gamma}, \mu, \rho)$ by a modified quasi-Newton method as described below. If an $\tilde{\boldsymbol{x}}$ is obtained such that $f_{\varepsilon,\gamma}(\tilde{\boldsymbol{x}}) < f_{\varepsilon,\gamma}(\boldsymbol{x}^*_{\varepsilon,\gamma})$, then let $\boldsymbol{x}^0 = \tilde{\boldsymbol{x}}$ and goto Step 1). If $\|\boldsymbol{x}\| \ge \varkappa$, stop.

Step 5) Set $i = i + 1$. If $i < m$, goto Step 4). Otherwise, set $\mu_0 = \mu/10$. If $\mu_0 > \bar{\mu}$, goto Step 3). Otherwise, set $\rho_0 = \rho/10$. If $\rho_0 \leq \bar{\rho}$, stop. The $\boldsymbol{x}_{\varepsilon,\gamma}^*$ obtained is a global minimizer of $f_{\varepsilon,\gamma}(\boldsymbol{x})$. Otherwise, goto Step 3).

In Step 4) of Algorithm 3.2, we set $\|\boldsymbol{x}\| \geq \varkappa$ to be a stopping criterion because $f_{\varepsilon,\gamma}(\boldsymbol{x})$ will be very large in the domain $\|\boldsymbol{x}\| \geq \varkappa$. In Algorithm 3.2, the solution $\boldsymbol{x}_{\varepsilon,\gamma}^*$ obtained is regarded as a globally optimal solution of Problem (P).

Remark 3.2: In the minimization process of the filled function in Step 4) of Algorithm 3.2, we do not use the quasi-Newton method directly. Instead, in each iteration, a direction along which the values of $f_{\varepsilon,\gamma}(\boldsymbol{x})$ and $p(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho)$ are both reduced is the first choice. More specifically, this search direction is chosen as

$$D_1 = \frac{-\nabla f_{\varepsilon,\gamma}(\boldsymbol{x})}{\|\nabla f_{\varepsilon,\gamma}(\boldsymbol{x})\|} - \frac{\nabla p\left(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho\right)}{\|\nabla p\left(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho\right)\|}.$$

If this direction does not lead to reduced values of both $f_{\varepsilon,\gamma}(\boldsymbol{x})$ and $p(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho)$, then we will choose $D_2 = -\nabla p(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho)$ as the search direction. In this way, a new point $\tilde{\boldsymbol{x}}$ can be found such that either both $f_{\varepsilon,\gamma}(\boldsymbol{x})$ and $p(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho)$ are reduced simultaneously or $p(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho)$ alone is reduced. Once $\tilde{\boldsymbol{x}}$ is obtained, we compute the value $f_{\varepsilon,\gamma}(\tilde{\boldsymbol{x}})$ and compare it with $f_{\varepsilon,\gamma}(\boldsymbol{x}_{\varepsilon,\gamma}^*)$. If $f_{\varepsilon,\gamma}(\tilde{\boldsymbol{x}}) < f_{\varepsilon,\gamma}(\boldsymbol{x}_{\varepsilon,\gamma}^*)$, then we go to Step 1) and use $\tilde{\boldsymbol{x}}$ as a new initial point to minimize $f_{\varepsilon,\gamma}(\boldsymbol{x})$. Otherwise, this process is repeated until a point located in a lower basin is found. If this process fails to find a point in a lower basin, we return to Step 2) and use another starting point to repeat this process. If all the stating points in S have been used and no better initial point is found, we may increase the number of points in S and continue with the process using these new points in S. We will regard the solution obtained as a global optimal solution if the process does not give rise to a point in a lower basin after all the points in S have been utilized. Clearly, we may miss finding the global optimizer unless the set S is very dense, but this becomes impractical even for problems of moderate size. There is clearly a tradeoff. Our extensive simulation studies suggest that good quality optimal solutions can often be obtained without having to take the number of points in S to be excessively large. The questions of how the points in the set S to be chosen remains open. It is an important and challenging question for future research.

Remark 3.3: The design of a uniform FIR filter bank with magnitude constraints is formulated as an optimization problem with continuous inequality constrains in [10]. An alternating method is developed to solve this optimization problem. In this alternating method, the synthesis prototype filter is solved as a quadratic optimization problem with the preassigned analysis prototype filter. Then, using the obtained synthesis prototype filter, the analysis prototype filter is determined by solving a corresponding linear quadratic optimization problem. In each of these optimization processes, the continuous inequality constraints are discretized into finitely many inequality constraints. This alternating process is repeated until the reduced cost is smaller than a preassigned positive real number. Since Problem (P) is nonlinear optimization problem with continuous inequality

constraints, the solution obtained by this alternating method is likely to be a local minimizer. As for all local optimization methods, the initial guess is critically important for the quality of the solution to be obtained. This is a motivation for the development of the hybrid method presented in this paper.

IV. NUMERICAL EXPERIMENTS

In this section, we will apply the proposed method to some numerical examples and compare the results with those obtained by existing methods.

Consider the design of the analysis and synthesis prototype filters. The transfer function of the overall filter bank is given by $20 \log_{10}(\epsilon_T) = -30[\text{dB}]$. We wish to obtain an optimum design such that the aliasing levels are minimized subject to a constraint on the distortion level of the transfer function. Let $\kappa = M$ and $\nu_\theta = 0.04$. The total delay for the filter bank is chosen as $\tau_d = L_a/2$. The performance of the filter bank for different cases is compared in terms of the relative inband aliasing and the relative residual aliasing measures, which are, respectively, defined by

$$M_i(\boldsymbol{h}) = \frac{\sum_{d=1}^{D-1} \int_{-\pi}^{\pi} \left| H\left(e^{\frac{jw}{D}} W_D^d\right) \right|^2 dw}{\int_{-\pi}^{\pi} \left| H\left(e^{\frac{jw}{D}}\right) \right|^2 dw} \tag{31}$$

and

$$M_r(\boldsymbol{h}, \boldsymbol{g}) = \frac{\sum_{d=1}^{D-1} \sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \left| H\left(e^{jw} W_M^m W_D^d\right) G(e^{jw} W_M^m) \right|^2 dw}{\sum_{m=0}^{M-1} \int_{-\pi}^{\pi} \left| H(e^{jw} W_M^m) G(e^{jw} W_M^m) \right|^2 dw}. \tag{32}$$

In this section, we will use a quasi-Newton method with limited memory developed in [14] to look for local minimizers of the function $f_{\varepsilon,\gamma}(\boldsymbol{x})$ and its modified version for the corresponding filled function $p(\boldsymbol{x}, \boldsymbol{x}_{\varepsilon,\gamma}^*, \mu, \rho)$. All the examples are performed in Compaq Visual Fortran on a computer with Intel Core 2 × 2.4G CPU and 2G RAM. In our computation, the integration approximated by Simpson's rule with 512 grid points.

For the first case, let $M = 16$, $D = M/2$, and $L_a = L_s = 2M$. The initial coefficients are all set to 0.01. For the parameters in Algorithm 3.2, we set $\delta = 0.1$ and $m = 32$. All the initial points in Algorithm 3.2 are generated randomly. We first use a gradient-based method to find its local minimizer. Then, we construct a corresponding filled function. After 25 iterations, a better initial point is found. With this point as a new initial guess, we minimize the original function, leading to a better minimizer. However, with this minimizer, we test all the 32 points in the set S, but no better initial point is found. To make sure we have found a global minimizer, we introduce an additional 28 points in the set S. However, no improved point is found. Thus, the current minimizer is regarded as a global minimizer. This process takes 320.1 s. The coefficients of the analysis and synthesis prototype filters are given in Table I. If we increase the number of points in the set S to 256, the computational time is increases to 3341.5 s. Using the alternating method in [10] with the stopband of the prototype filters being set to π/M and the initial analysis filters being set to the solution obtained from the method in [8] to solve this problem, it takes 26 s. If this solution is adopted as

TABLE I
ANALYSIS AND SYNTHESIS PROTOTYPE FILTER COEFFICIENTS WITH $M = 16$, $D = M/2$, $La = Ls = 2M$

n	$h(n)$	n	$h(n)$	n	$g(n)$	n	$g(n)$
0	0.0065433	16	0.06984	0	0.077432	16	0.2906
1	0.0113329	17	0.060582	1	0.13355	17	0.13969
2	0.017508	18	0.050606	2	0.20709	18	0.0024552
3	0.024969	19	0.040504	3	0.29569	19	-0.11548
4	0.033572	20	0.030684	4	0.39896	20	-0.20593
5	0.043028	21	0.021686	5	0.50991	21	-0.26823
6	0.052985	22	0.013941	6	0.62844	22	-0.29688
7	0.063028	23	0.0076463	7	0.74579	23	-0.29504
8	0.070789	24	0.0019425	8	0.78507	24	-0.30247
9	0.078256	25	-0.0014336	9	0.83696	25	-0.26228
10	0.083948	26	-0.0034993	10	0.85415	26	-0.21691
11	0.08746	27	-0.004505	11	0.83584	27	-0.16838
12	0.088503	28	-0.004601	12	0.77825	28	-0.12285
13	0.08699	29	-0.0041766	13	0.6864	29	-0.082461
14	0.082852	30	-0.0033768	14	0.55952	30	-0.051611
15	0.076246	31	-0.0024986	15	0.40683	31	-0.030522

TABLE II
THE RESULTS WITH $M = 16$, $D = M/2$, $La = Ls = 2M$

	$M_i(\boldsymbol{h}, \boldsymbol{g})$	$M_r(\boldsymbol{h}, \boldsymbol{g})$
[8]	−20.0781	−24.3550
[10]	−29.1993	−33.1888
our proposed method	−32.3665	−33.7584

an initial guess, our method takes an additional 117.5 s to obtain a global optimal solution. If we use the solution obtained in [8] as the initial guess, then one call of the filled function yields the global minimizer. This process takes 265 s. Our results and those obtained by other methods are compared in Table II. The results solved by our method and the method in [10] with different initial conditions are presented in Table III.

From Table II, we see that our results are superior to those obtained by other methods, including the iterative alternating method [10]. The reason is quite clear. The iterative alternating method does not produce a global optimal solution, as the optimization problem is nonconvex. Figs. 2 and 3 plot the frequency responses of the analysis and synthesis prototype filters, respectively. Figs. 4 and 5 plot the corresponding magnitude error and phase error, respectively. If we try to solve this problem with all initial coefficients being set to 0.1, then after two calls of the filled function, the global minimizer is obtained. This process takes 612 s.

Table IV shows the results for the case when $M = 16$, $D = M/2$, $L_a = Ls = 4M$ with 32 different starting points.

The frequency responses of the analysis and synthesis prototype filters are depicted in Fig. 6. The phase error for this case is depicted in Fig. 7. X

From these results, we see that our method obtains better results than those obtained by the iterative alternating method developed in [10]. Furthermore, the iterative alternating method may be sensitive to the initial guess of $\boldsymbol{h}$. This can be seen from Table III. To illustrate this further, let us reconsider the case when $M = 16$, $D = M/2$, $L_a = Ls = 4M$ with 32 starting points. Let $\bar{\boldsymbol{h}}$ be the optimal solution obtained by our method. Choose two initial guesses of $\boldsymbol{h}$ as follows: $\boldsymbol{h}_1 = 1/2 * \bar{\boldsymbol{h}}$ and $\boldsymbol{h}_2 = 2 * \bar{\boldsymbol{h}}$. We run the iterative alternating method with these two initial guesses, and the results obtained are depicted in Fig. 8. We can clearly see that these two results are signif-

icantly different. For our method, after two calls of the filled function, we obtain the same result. V

The remaining seven cases in [10, Table II] are solved using our method with the solutions obtained in [10] taken as our initial guesses and 32 starting points. Better results are obtained for five of the seven cases, where the number of calling of the filled function is one for each of these five cases. The results for the remaining two cases are the same as those obtained in [10]. These results and those obtained by existing methods are presented in Table V.

Our method is not effective for the critically sampled case. This is because the inband aliasing is considered in (12). The inband aliasing is related to the frequency selectivity and will be large if the taps of the prototype filters are small by our method for this case. How to balance them and design an efficient method for this case is an interesting future research problem. At the same time, our method is not effective for the PR filter bank design. In the design of PR filter bank, a special structure is constructed in such a way that the aliasing effects are cancelled. In our formulation, the residual cost is expressed as the sum of squared terms. Thus, no cancellation will take place. All of the drawbacks are also shared by the method in [10]. To design these types of filter banks, the methods developed in [2], [3], and [16] can be applied.

V. CONCLUSION

In this paper, we have developed a new computational algorithm for the optimum design of an FIR filter bank. This problem is formulated as a nonlinear and nonconvex semi-infinite optimization problem. A constraint transcription method and the concept of a penalty function are used to approximate this nonlinear nonconvex semi-infinite optimization as a sequence of unconstrained optimization problems. The quasi-Newton method with limited memory is used in conjunction with a filled function method to solve each of these approximated unconstrained optimization problems. Solutions obtained by our method are superior to all those obtained by other methods. Furthermore, our method is independent of the initial guess of the solution. The method can be readily applied to the optimum design of uniform FIR filter bank with group delay specifications. This extension is straightforward.

TABLE III
THE RESULTS FOR DIFFERENT INITIAL CONDITIONS WITH $M = 16, D = M/2, La = Ls = 2M$

| Different initial conditions | $\mathcal{M}_i(h, g)$ | | $\mathcal{M}_r(h, g)$ | | cost (12) | |
	our method	method in [10]	our method	method in [10]	our method	method in [10]
All with 0.01	-32.3665	-26.2193	-33.7584	-33.0750	-30.7520	-29.5247
All with 0.1	-32.3665	-26.2203	-33.7584	-33.4316	-30.7520	-29.5233
Solution in [8]	-32.3665	-33.7533	-33.7584	-28.7225	-30.7520	-19.9626
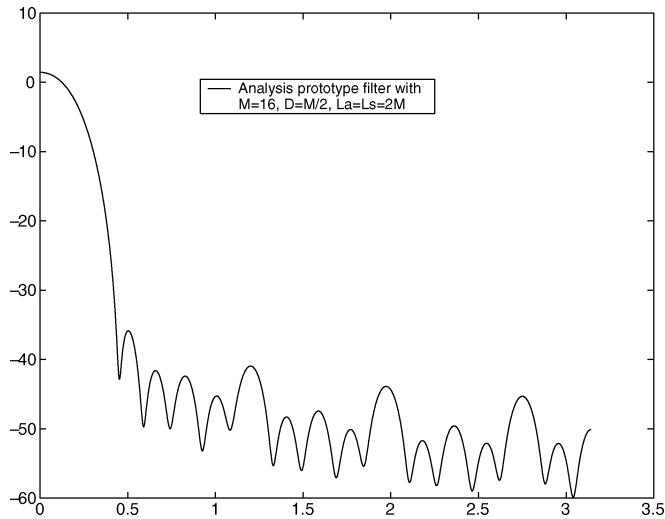
Fig. 2. The frequency response of analysis filter with $M = 16, D = 8, L_a = L_s = 32$.
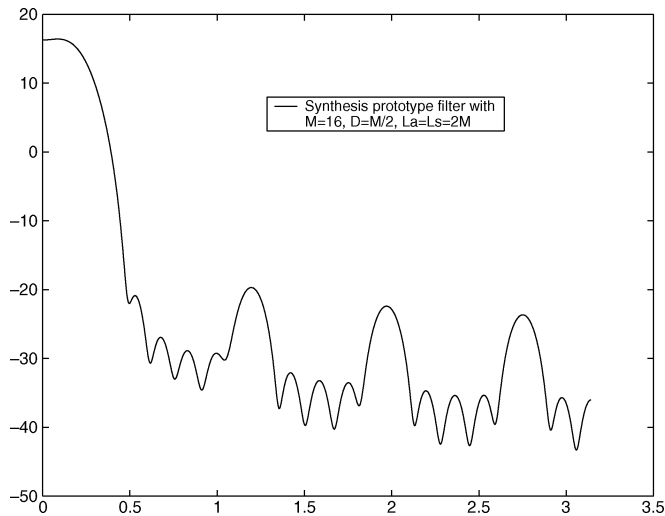

Fig. 3. The frequency response of synthesis filter with $M = 16, D = 8, L_a = L_s = 32$.
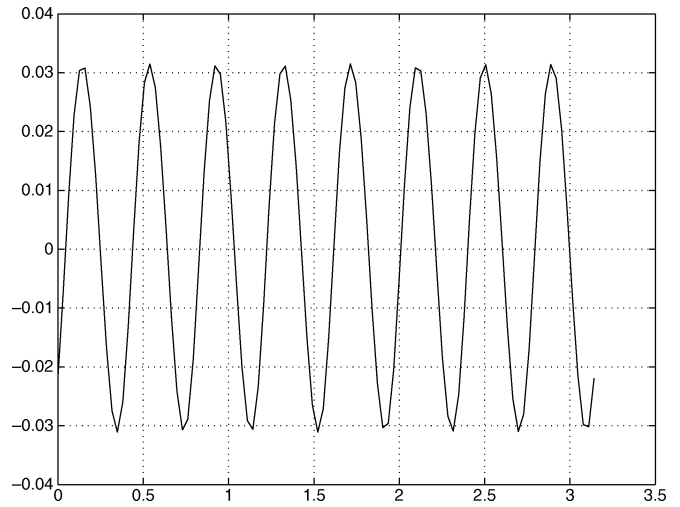

Fig. 4. The magnitude error for the case when $M = 16, D = 8, L_a = L_s = 32$.
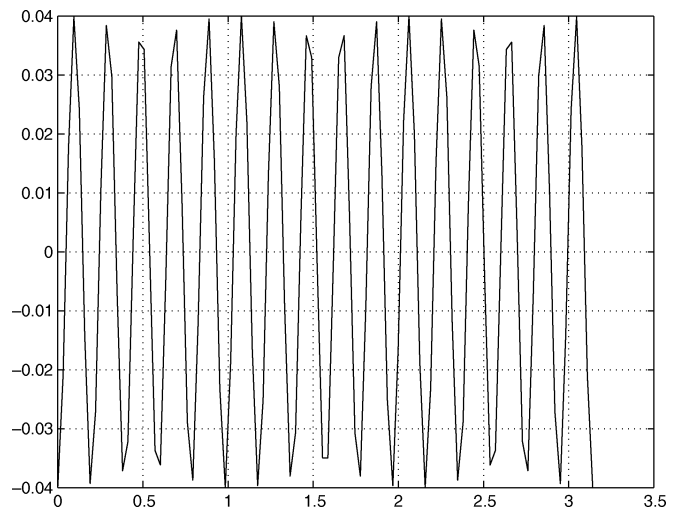

Fig. 5. The phase error for the case when $M = 16, D = 8, L_a = L_s = 32$, $\nu_\theta = 0.04$.

APPENDIX I
PROOF OF THEOREM 3.1

Define

$$\mathcal{F}_\varepsilon = \left\{ \boldsymbol{x} \in \mathbb{R}^{L_a + L_s} : G_i(\boldsymbol{x}, w) \leq -\varepsilon \right.$$
$$\forall w \in [0, \pi], \; i = 1, 2\}$$
$$= \left\{ \boldsymbol{x} \in \mathbb{R}^{L_a + L_s} : \hat{G}_{i,\varepsilon}(\boldsymbol{x}) = 0, \; i = 1, 2 \right\}. \quad (33)$$

TABLE IV
THE RESULTS WITH $M = 16, D = M/2, La = Ls = 4M$

	$M_i(h, g)$	$M_r(h, g)$
[8]	-34.6964	-38.5248
[10]	-63.9694	-66.9775
Our proposed method	-63.9889	-68.7237

TABLE V
THE RESULTS OBTAINED BY OUR METHOD AND THOSE OBTAINED BY EXISTING METHODS

| M	D	L_a	$\mathcal{M}_i(h,g)$			$\mathcal{M}_r(h,g)$		
			[8]	*[10]*	*our method*	*[8]*	*[10]*	*our method*
32	16	2M	-19.1502	-28.1670	-31.2417	-23.2661	-32.7133	-32.7211
32	16	4M	-33.9531	-62.7587	-65.2094	-37.8805	-66.9370	-67.3108
64	32	2M	-18.4164	-27.6148	-27.6148	-22.6857	-32.4970	-32.4970
64	32	4M	-33.2612	-61.9921	-62.0015	-37.6073	-67.2469	-67.3412
16	4	2M	-47.6116	-87.2924	-87.9231	-51.4305	-92.9208	-93.2613
32	8	2M	-47.0397	-83.8481	-87.5832	-50.7785	-90.5891	-90.6354
64	16	2M	-46.0754	-81.0713	-81.0713	-49.8074	-89.8934	-89.8934
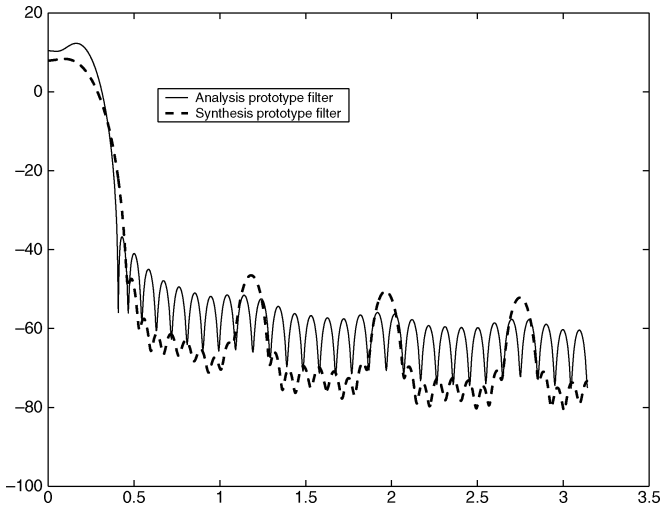
Fig. 6. The frequency responses of analysis and synthesis filters for the case when $M = 16$, $D = 8$, $L_a = L_s = 64$.
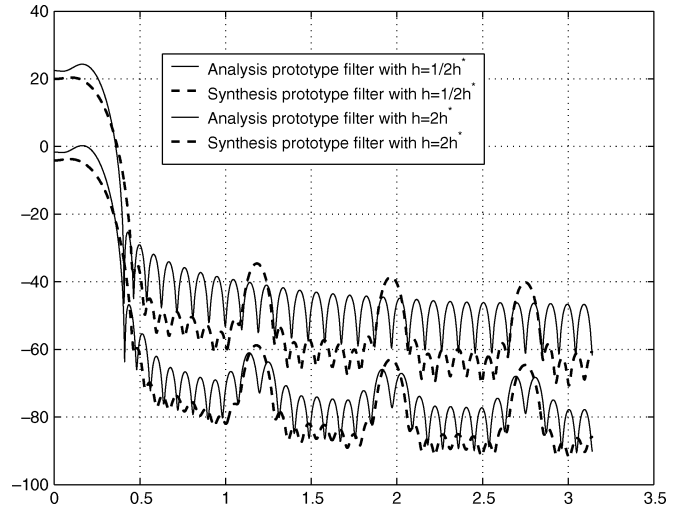

Fig. 8. The frequency responses of analysis and synthesis filters obtained by the iterative alternating method with two different initial guesses.
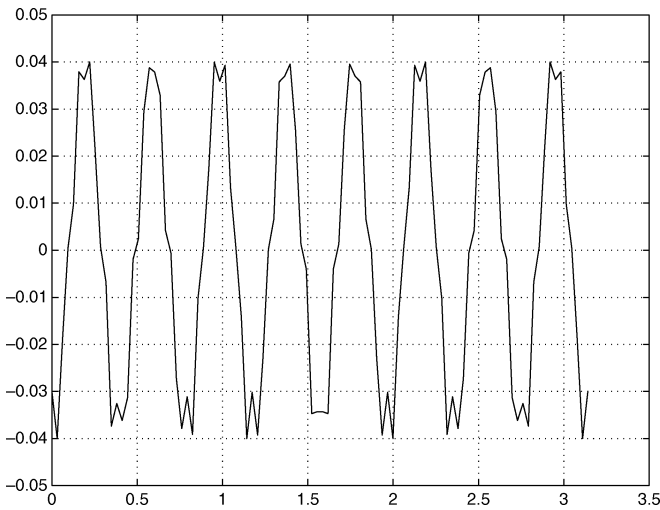

Fig. 7. The phase error for the case when $M = 16$, $D = 8$, $L_a = L_s = 64$.

Let $\boldsymbol{x}_{\varepsilon,\gamma}^* = \left((\boldsymbol{h}_{\varepsilon,\gamma}^*)^T, (\boldsymbol{g}_{\varepsilon,\gamma}^*)^T\right)^T$ be an optimal solution of Problem $(P_{\varepsilon,\gamma})$. We have

$$f_{\varepsilon,\gamma}\left(\boldsymbol{x}_{\varepsilon,\gamma}^*\right) \leq f_{\varepsilon,\gamma}(\boldsymbol{x}) \tag{34}$$

for all $\boldsymbol{x} = (\boldsymbol{h}^T, \boldsymbol{g}^T)^T \in \mathbb{R}^{L_a + L_s}$. Let $\boldsymbol{x}_\varepsilon \in \mathcal{F}_\varepsilon$ be fixed. Then, it follows from (33) that $\hat{G}_{i,\varepsilon}(\boldsymbol{x}_\varepsilon) = 0$ for $i = 1, 2$. Note that $F(\boldsymbol{h}, \boldsymbol{g}) + \kappa G(\boldsymbol{h}) \geq 0$ for all $\boldsymbol{h}$ and $\boldsymbol{g}$. We have

$$\gamma \sum_{i=1}^{2} \hat{G}_{i,\varepsilon}\left(\boldsymbol{x}_{\varepsilon,\gamma}^*\right) \leq f_{\varepsilon,\gamma}\left(\boldsymbol{x}_{\varepsilon,\gamma}^*\right) \leq f_{\varepsilon,\gamma}(\boldsymbol{x}_\varepsilon).$$

That is

$$\sum_{i=1}^{2} \hat{G}_{i,\varepsilon}\left(\boldsymbol{x}_{\varepsilon,\gamma}^*\boldsymbol{x}\right) \leq f_{\varepsilon,\gamma}(\boldsymbol{x}_\varepsilon)/\gamma.$$

By Theorem 2.3 [15], there exists an $\gamma(\varepsilon) > 0$, such that for all $\gamma > \gamma(\varepsilon)$, $\boldsymbol{x}_{\varepsilon,\gamma}^* \in \mathcal{F}$. This completes the proof.

APPENDIX II
PROOF OF THEOREM 3.2

By Assumption 2), there exists an $\bar{\boldsymbol{x}} \in \overset{\circ}{\mathcal{F}}$ such that $\boldsymbol{x}_\alpha = \alpha\bar{\boldsymbol{x}} + (1-\alpha)\boldsymbol{x}^* \in \overset{\circ}{\mathcal{F}}$ for all $\alpha \in (0, 1]$. From (19) and (16), it follows that for any $\delta_1 > 0$, there exists an $\alpha_1 \in (0, 1]$, such that for all $\alpha \in (0, \alpha_1)$, we have

$$F(\boldsymbol{h}^*, \boldsymbol{g}^*) + \kappa G(\boldsymbol{h}^*) \leq F(\boldsymbol{h}_\alpha, \boldsymbol{g}_\alpha) + \kappa G(\boldsymbol{h}_\alpha)$$
$$\leq F(\boldsymbol{h}^*, \boldsymbol{g}^*) + \kappa G(\boldsymbol{h}^*) + \delta_1. \tag{35}$$

Choose $\alpha_2 = \alpha_1/2$. Clearly, $\boldsymbol{x}_{\alpha_2} \in \overset{\circ}{\mathcal{F}}$. Thus, there exists a $\delta_2 > 0$ such that $\max_{0 \le \pi \le w} G_i(\boldsymbol{x}_{\alpha_2}, w) < -\delta_2$, $i = 1, 2$. Thus, $\hat{G}_{i,\delta_2}(\boldsymbol{x}_{\alpha_2}) = 0$, $i = 1, 2$. By the definition of $\boldsymbol{x}^*_{\varepsilon,\gamma}$, if we choose $\varepsilon = \delta_2$, then

$$f_{\varepsilon,\bar{\gamma}(\varepsilon)}\left(\boldsymbol{x}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}\right) \le f_{\varepsilon,\bar{\gamma}(\varepsilon)}\left(\boldsymbol{x}_{\alpha_2}\right) = F\left(\boldsymbol{h}_{\alpha_2}, \boldsymbol{g}_{\alpha_2}\right) + \kappa G\left(\boldsymbol{h}_{\alpha_2}\right)$$

where for each $\varepsilon > 0$, $\bar{\gamma}(\varepsilon)$ is chosen such that the constraints (20) and (21) are satisfied. Thus

$$F\left(\boldsymbol{h}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}, \boldsymbol{g}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}\right) + \kappa G\left(\boldsymbol{h}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}\right) \le F\left(\boldsymbol{h}_{\alpha_2}, \boldsymbol{g}_{\alpha_2}\right) + \kappa G\left(\boldsymbol{h}_{\alpha_2}\right).$$

By (35) and the fact that $\boldsymbol{x}^*$ is an optimal solution of Problem (P), we obtain

$$\begin{aligned} F(\boldsymbol{h}^*, \boldsymbol{g}^*) + \kappa G(\boldsymbol{h}^*) &\le F\left(\boldsymbol{h}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}, \boldsymbol{g}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}\right) + \kappa G\left(\boldsymbol{h}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}\right) \\ &\le f_{\varepsilon,\bar{\gamma}(\varepsilon)}\left(\boldsymbol{x}^*_{\varepsilon,\bar{\gamma}(\varepsilon)}\right) \\ &\le F\left(\boldsymbol{h}_{\alpha_2}, \boldsymbol{g}_{\alpha_2}\right) + \kappa G\left(\boldsymbol{h}_{\alpha_2}\right) \\ &\le F(\boldsymbol{h}^*, \boldsymbol{g}^*) + \kappa G(\boldsymbol{h}^*) + \delta_1. \end{aligned}$$

Letting $\varepsilon \to 0$ and noting that $\delta_1 > 0$ is arbitrary, (29) is satisfied. This completes the proof.

REFERENCES

[1] N. J. Fliege, *Multirate Digital Signal Processing*. West Sussex, U.K.: Wiley, 1994.

[2] T. Q. Nguyen and P. P. Vaidyanathan, "Two-channel perfect-reconstruction FIR QMF structures which yield linear-phase analysis and synthesis filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 676–690, 1989.

[3] B. R. Horng and A. N. Willson, Jr., "Lagrange multiplier approaches to the design of two-channel perfect-reconstruction linear-phase FIR filter banks," *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 364–374, 1992.

[4] P. P. Vaidyanathan, *Multirate Systems and Fiterbanks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[5] T. E. Tuncer and T. Q. Nguyen, "General analysis of two-band QMF banks," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 544–548, 1995.

[6] Q.-G. Liu, B. Champagne, and D. K. C. Ho, "Simple design of over-sampled uniform DFT-filter banks with applications to subband acoustic echo cancellation," *IEEE Trans. Signal Process.*, pp. 831–847, 2000.

[7] M. Harteneck, S. Weiss, and R. W. Stewart, "Design of near perfect reconstruction oversampled filter banks for subband adaptive filters," *IEEE Trans. Circuits Syst.*, vol. 46, pp. 1081–1085, 1999.

[8] J. M. de Haan, N. Grbic, I. Classon, and S. Norholm, "Filter bank design for subband adaptive microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 1, pp. 14–23, 2003.

[9] , S. L. Gay and J. Benesty, Eds., *Acoustic Signal Processing for Telecompunication*. Norwell, MA: Kluwer, 2000.

[10] H. H. Dam, S. Nordholm, and A. Cantoni, "Uniform FIR filterbank optimization with group delay specifications," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4249–4260, 2005.

[11] K. L. Teo, V. Rehbock, and L. S. Jennings, "A new computational algorithm for functional inequality constrained optimization problems," *Automatica*, vol. 29, no. 3, pp. 789–792, 1993.

[12] D. G. Luenburger, *Linear and Nonlinear Programming*, 2nd ed. New York: Addison-Wesley, 1984.

[13] L. S. Zhang, C. K. Ng, D. Li, and W. W. Tian, "A new filled function method for global optimization," *J. Global Optim.*, vol. 28, pp. 17–43, 2004.

[14] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, L-BFGS-B: A limited memory FORTRAN code for solving bound constrained optimization problems, Electrical Engineering and Computer Science Dept., Northwestern Univ., Tech. Rep. NAM-11, 1994.

[15] L. S. Jennings and K. L. Teo, "A computational algorithm for functional inequality constrained optimization problems," *Automatica*, vol. 26, pp. 371–375, 1990.

[16] T. Q. Nguyen, "Near-perfect-reconstruction pseudo-QMF banks," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 65–76, 1994.
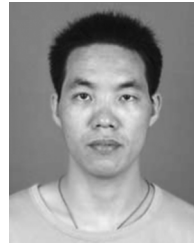
C. Z. Wu received the bachelor's degree from Anhui Normal University, China, in 2001 and the Ph.D. degree from Zhongshan University, China, in 2006.

He joined Chongqing Normal University as a Lecturer in 2006. At the same time, he was a Visiting Research Fellow with Curtin University of Technology, Perth, Australia. His main interests include both theoretical and practical aspects of optimization and optimal control and their applications in signal processing.

K. L. Teo (M'74–SM'87) received the B.Sc. degree in telecommunications engineering from Ngee Ann Technical College, Singapore, and the M.A.Sc and Ph.D. degrees in electrical engineering from the University of Ottawa, Ottawa, ON, Canada.

He was with the Department of Applied Mathematics, University of New South Wales, Australia; the Department of Industrial and Systems Engineering, National University of Singapore, Singapore; and the Department of Mathematics, University of Western Australia, Australia. In 1996, he joined the Department of Mathematics and Statistics, Curtin University of Technology, Perth, Australia, as a Professor of applied mathematics. He then became Chair Professor of Applied Mathematics and Head of the Department of Applied Mathematics at Hong Kong Polytechnic University, China, from 1999 to 2004. He is currently a Professor of applied mathematics and Head of the Department of Mathematics and Statistics, Curtin University of Technology. He has published five books and more than 350 journal papers. The software package MISER3.3 for solving general constrained optimal control problems was developed by the research team under his leadership. He is the Founding Editor-in-Chief of the *Journal of Industrial and Management Optimization.* He has been Editor-in-Chief of *Dynamics of Continuous, Discrete and Impulsive Systems, Series B* since September 2007. He is a Regional Editor of *Nonlinear Dynamics and Systems Theory.* In addition, he is in the editorial board of 14 international journals. He was a Co-Guest Editor of 12 special journal issues and a Coeditor of two books. His research interests include both the theoretical and practical aspects of optimal control and optimization and their practical applications in areas that include signal processing in telecommunications and financial portfolio optimization.

V. Rehbock received the Ph.D. degree in applied mathematics from the University of Western Australia, Australia, in 1994.

Since 1995, he has been with the Department of Mathematics and Statistics, Curtin University of Technology, Perth, Australia. His main interests are in the modelling and numerical solution of dynamic optimization problems, particularly those involving variable switching times. He has coauthored more than 60 refereed publications in the areas of optimization and optimal control.

H. H. Dam received the bachelor's degree (first-class honors) and Ph.D. degree (with distinction) from Curtin University of Technology, Perth, Australia, in 1996 and 2001, respectively.

From 2001 to 2007, she was a Reseach Fellow/Senior Research Fellow with the Western Australian Telecommunications Research Institute. Since 2006, she has been a Senior Lecturer with the Department of Mathematics and Statistics, Curtin University of Technology. Her research interests are adaptive array processing, optimization, equalization, and filter design.