Optimal Design of Frequency-Response-Masking Filters Using Semidefinite Programming

Wu-Sheng Lu, *Fellow, IEEE,* and Takao Hinamoto, *Fellow, IEEE*

Abstract—Since Lim's 1986 paper on the frequency-response-masking (FRM) technique for the design of finite-impulse response digital filters with very small transition widths, the analysis and design of FRM filters has been a subject of study. In this paper, a new optimization technique for the design of various FRM filters is proposed. Central to the new design method is a sequence of linear updates for the design variables, with each update carried out by semidefinite programming. Algorithmic details for the design of basic and multistage FRM filters are presented to show that the proposed method offers a unified design framework for a variety of FRM filters. Design simulations are included to illustrate the proposed algorithms and to evaluate the design performance in comparison with that of several existing methods.

Index Terms—Frequency-response-masking filters, optimal design, semidefinite programming.

I. INTRODUCTION

SINCE the publication of [1], the frequency-response-masking (FRM) technique for the design of finite-impulse response (FIR) digital filters with very narrow transition bands has been a subject of study [2]–[9]. As a result, in many cases, it has become the method of choice primarily because of the considerably reduced realization complexity it offers compared with other available options [5], [8].

As illustrated in Fig. 1(a), a *basic* FRM *filter* involves a linear-phase prototype filter $H_a(z)$ up-sampled by M, a pair of linear-phase masking filters $\{H_{ma}(z), H_{mc}(z)\}$, and a delay line that, together with the prototype filter, helps form a linear-phase complementary pair $\{H_a, H_c\}$ [1]. For additional reduction of realization complexity, the prototype filter itself may be realized with a basic FRM filter, yielding a *multistage* FRM *filter* [1], [5], see Fig. 1(b) for a two-stage FRM filter structure. Given an up-sampling factor, lengths of the subfilters involved, and passband/stopband edges, the design of a basic or multistage FRM filter is usually carried out by *separately* designing the subfilters [1], [5], [6]. As such, the FRM filter obtained is only suboptimal. In [9], a two-step optimization technique for the optimal design of basic FRM filters is proposed. In the first step of the method, an initial FRM filter is designed by alternately optimizing the prototype filter and masking filters. The second step

then refines the design using an algorithm in [10]. Although the optimal design of multistage FRM filters was not addressed in [9], the method in [9] appears to be, at least in principle, applicable to such a design. However, as the number of filter stages increases, a great many of the subfilters need to be optimized alternately in the first step of the method, making the design process increasingly involved. In this paper, we present a rather different optimization technique in which the set of filter coefficients of *all* subfilters is treated as a single design vector and an optimal basic FRM filter is designed through a sequence of linear updates for the design variables, with each update carried out in a semidefinite programming (SDP) framework. As will be demonstrated by design examples, starting with a reasonable initial design, the proposed algorithm converges to an optimized design with performance considerably better than that of [1] and comparable with that of [9]. An advantage of the proposed design methodology is that with straightforward modifications it can be readily extended to multistage FRM filters. We shall elaborate on this point in Sections IV and V with algorithmic details for the class of two-stage FRM filters and simulation results.

The second issue to be addressed in this paper is the optimal design of FRM filters with *reduced passband group delay*. Linear-phase FIR filters have constant group delay in the entire frequency band, but for a filter with very narrow transition width, the group delay can be exceedingly large, a property not desirable in many applications. Note that although linear-phase FRM filters have been successful in reducing realization complexity, its group delay is even larger than that of the direct-form FIR filter with the same approximation accuracy. For a linear-phase FRM filter with a large up-sampling factor M, its large group delay is dominantly contributed by the prototype filter. Therefore, if the prototype filter has a nonlinear phase response with a reduced passband group delay, say d, and if the delay line [the lower-left block in Fig. 1(a)] is accordingly modified to z^{-dM}, then the filter is expected to have its passband group delay reduced by $M[0.5(N-1) - d]$ where N is the length of the prototype filter. Hence, the reduction in group delay can be significant especially when M is large. In this paper, we pursue this idea and show that, by a joint optimization of the entire set of subfilters, the prototype as well as masking filters all contribute to minimizing the fluctuation in the reduced passband group delay. Furthermore, we extend the design method to the multistage FRM filters with reduced group delay.

The paper is organized as follows. Section II gives a brief overview of the FRM filtering and some basic elements of SDP that we need in the rest of the paper. Section III describes the core of the design methodology based on which our algorithms are developed in Sections IV and V, respectively. Illustrative
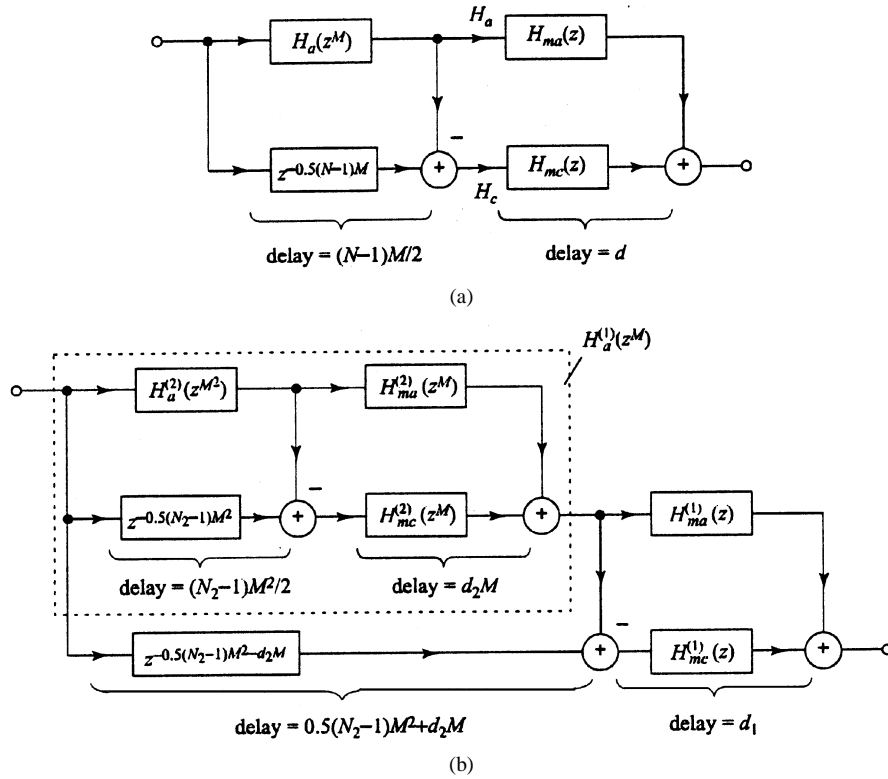
Fig. 1. (a) Basic FRM filter structure and (b) Two-stage FRM filter.

examples will be presented for each class of FRM filters for performance demonstration and comparison.

In the rest of the paper, boldfaced characters denote matrices and vectors, I_r represents the identity matrix of dimension r, $\mathrm{diag}\{a_1, \ldots, a_n\}$ denotes a diagonal matrix, $\mathrm{diag}\{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n\}$ denotes a block-diagonal matrix with matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n$ as its diagonal blocks, ω_p and ω_a denote normalized passband and stopband edges, respectively, and Ω denotes the normalized base frequency band, i.e., $\Omega = \{\omega\colon -\pi \leq \omega \leq \pi\}$.

II. PRELIMINARIES

A. Brief Overview of FRM Filters

As can be seen from Fig. 1(a), the prototype filter and its complement are upsampled by M, yielding sparse filter coefficients and a reduced transition width for each. They are then connected in cascade to a pair of frequency-response masking filters in order to approximate a desired sharp frequency response.

The transfer functions of the subfilters in a basic FRM filter are denoted by

$$H_a(z) = \sum_{k=0}^{N-1} h_k z^{-k}$$

$$H_{ma}(z) = \sum_{k=0}^{N_a-1} h_k^{(a)} z^{-k}$$

$$H_{mc}(z) = \sum_{k=0}^{N_c-1} h_k^{(c)} z^{-k}. \qquad (1)$$

If all subfilters have linear phase responses with N_a and N_c either both even or both odd and $(N-1)M$ even, then the FRM filter has a linear phase response with the group delay

$$D = 0.5(N-1)M + \max\{0.5(N_a - 1), 0.5(N_c - 1)\}$$

provided that a delay line of appropriate length is cascaded with either $H_{ma}(z)$ or $H_{mc}(z)$. Let us consider the design of a low-pass filter as an example. Given sampling factor M, normalized passband edge ω_p and stopband edge ω_a, a reasonable *initial* design of subfilters $H_a(z)$, $H_{ma}(z)$, and $H_{mc}(z)$ can be obtained by using a standard method [11] to design three low-pass filters whose passband and stopband edges are determined as follows [1].

1) For $H_a(z)$, the passband edge θ and stopband edge ϕ are given by

$$\theta = \omega_p M - 2m\pi \qquad (2a)$$

$$\phi = \omega_a M - 2m\pi \qquad (2b)$$

$$m = \lfloor \omega_p M / 2\pi \rfloor \qquad (2c)$$

where $\lfloor x \rfloor$ denotes the largest integer less than x, or by

$$\theta = 2m\pi - \omega_a \pi \qquad (3a)$$

$$\phi = 2m\pi - \omega_p \pi \qquad (3b)$$

$$m = \lceil \omega_a M / 2\pi \rceil \qquad (3c)$$

where $\lceil x \rceil$ denotes the smallest integer larger than x, depending on which set of $\{\theta, \phi\}$ satisfies $0 < \theta < \phi < \pi$.

2) If (2) is used to determine the values of θ and ϕ, then the passband and stopband edges of $H_{ma}(z)$ are given by $(2m\pi + \theta)/M$ and $[2(m+1)\pi - \phi]/M$, respectively, and

the passband and stopband edges of $H_{mc}(z)$ are given by $(2m\pi - \theta)/M$ and $(2m\pi + \phi)/M$, respectively.

3) If (3) is used to determine the values of θ and ϕ, then the passband and stopband edges of $H_{ma}(z)$ are given by $[2(m-1)\pi + \phi]/M$ and $(2m\pi - \theta)/M$, respectively, and the passband and stopband edges of $H_{mc}(z)$ are given by $(2m\pi - \phi)/M$ and $(2m\pi + \theta)/M$, respectively.

B. SDP

SDP is a relatively new optimization methodology which is primarily concerned with minimizing a linear or convex quadratic objective function subject to linear-matrix inequality (LMI) type constraints that depend on the design variables affinely [12]. The class of SDP problems most relevant to the design problems of interest is expressed as

$$\text{minimize} \quad \boldsymbol{c}^T \boldsymbol{x} \tag{4a}$$

$$\text{subject to:} \quad \boldsymbol{F}(\boldsymbol{x}) \succeq \boldsymbol{0} \tag{4b}$$

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 + \sum_{i=1}^{n} x_i \boldsymbol{F}_i \tag{4c}$$

where $\boldsymbol{x} = [x_1 \cdots x_n]^T$, $\boldsymbol{F}_i \in R^{m \times m}$ for $i = 0, 1, \ldots, n$ are known symmetric matrices and $\boldsymbol{F}(\boldsymbol{x}) \succeq \boldsymbol{0}$ denotes that $\boldsymbol{F}(\boldsymbol{x})$ is positive semidefinite at $\boldsymbol{x}$. Note that the constraint matrix $\boldsymbol{F}(\boldsymbol{x})$ in (4) is *affine* with respect to $\boldsymbol{x}$. SDP includes both linear and quadratic programming as its special cases and represents a subclass of convex programming that covers many optimization problems encountered in various engineering disciplines. Many interior-point methods which have proven efficient for linear programming have recently been extended to SDP [12], [13] and efficient software implementation of various SDP algorithms are available. In particular we mention the LMI Control Toolbox [14], SeDuMi [15], and SDPT3 Toolbox [16], all of which work with MATLAB.

III. OPTIMIZATION METHODOLOGY

In this section, we describe the core of the optimization technique based on which the design algorithms for various FRM filters will be developed in the subsequent sections. As such, our description will be given in a setting more general than each individual algorithm in Sections IV and V.

Let $H_d(\omega)$ be a desired real-valued or complex-valued function of frequency variable ω, and $H(\omega, \boldsymbol{x})$ be a real-valued or complex-valued function of ω, which depends on a real-valued parameter vector $\boldsymbol{x} \in R^{n \times 1}$. We seek to find a vector $\boldsymbol{x}^*$ that solves the weighted minimax optimization problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \left\{ \underset{\omega \in \Omega}{\text{maximize}} \, W(\omega)|H(\omega, \boldsymbol{x}) - H_d(\omega)| \right\} \tag{5}$$

where $W(\omega) \geq 0$ is a weighting function. With different interpretations for $H(\omega, \boldsymbol{x})$ and $H_d(\omega)$, the problem in (5) covers many minimax design problems for digital filters, including those to be addressed in Sections IV and V.

Let η be an upper bound of $W^2(\omega)|H(\omega, \boldsymbol{x}) - H_d(\omega)|^2$ on Ω. As the first step of the optimization we convert the problem in (5) into a constrained minimization problem

$$\text{minimize} \quad \eta \tag{6a}$$

$$\text{subject to:} \quad W^2(\omega)|H(\omega, \boldsymbol{x}) - H_d(\omega)|^2 \leq \eta, \qquad \text{for } \omega \in \Omega. \tag{6b}$$

Suppose we have a reasonable initial point $\boldsymbol{x}_0$ to start the design, and we are now in the kth iteration. For a nonlinear and smooth $H(\omega, \boldsymbol{x})$ in a vicinity of $\boldsymbol{x}_k$, we can write

$$H(\omega, \boldsymbol{x}_k + \boldsymbol{\delta}) = H(\omega, \boldsymbol{x}_k) + \boldsymbol{g}_k^T(\omega)\boldsymbol{\delta} + o(\|\boldsymbol{\delta}\|)$$

where $\boldsymbol{g}_k(\omega)$ is the gradient of $H(\omega, \boldsymbol{x})$ with respect to $\boldsymbol{x}$ and evaluated at $\boldsymbol{x}_k$. Hence, provided that $\|\boldsymbol{\delta}\|$ is small, with $\boldsymbol{x} = \boldsymbol{x}_k + \boldsymbol{\delta}$ we have

$$W^2(\omega)|H(\omega, \boldsymbol{x}) - H_d(\omega)|^2$$
$$\approx W^2(\omega) \left| \boldsymbol{g}_k^T(\omega)\boldsymbol{\delta} + [H(\omega, \boldsymbol{x}_k) - H_d(\omega)] \right|^2$$
$$= \left[\boldsymbol{g}_{rk}^T(\omega)\boldsymbol{\delta} + e_{rk}(\omega) \right]^2 + \left[\boldsymbol{g}_{ik}^T(\omega)\boldsymbol{\delta} + e_{ik}(\omega) \right]^2 \tag{7}$$

where $\boldsymbol{g}_{rk}(\omega)$ and $\boldsymbol{g}_{ik}(\omega)$ are the real and imaginary parts of $W(\omega)\boldsymbol{g}_k(\omega)$, respectively, and

$$e_{rk}(\omega) = W(\omega)[H_r(\omega, \boldsymbol{x}_k) - H_{rd}(\omega)]$$
$$e_{ik}(\omega) = W(\omega)[H_i(\omega, \boldsymbol{x}_k) - H_{id}(\omega)]$$

with $H_r(\omega, \boldsymbol{x})$, $H_i(\omega, \boldsymbol{x}_k)$, $H_{rd}(\omega)$, and $H_{id}(\omega)$ being the real and imaginary parts of $H(\omega, \boldsymbol{x}_k)$ and $H_d(\omega)$, respectively. From (6) and (7), it follows that an approximate solution of (6) in the kth iteration can be obtained by solving the following problem:

$$\text{minimize} \quad \eta \tag{8a}$$

$$\text{subject to:} \quad \left(\boldsymbol{g}_{rk}^T\boldsymbol{\delta} + e_{rk} \right)^2 + \left(\boldsymbol{g}_{ik}^T\boldsymbol{\delta} + e_{ik} \right)^2 \leq \eta, \qquad \text{for } \omega \in \Omega \tag{8b}$$

$$\|\boldsymbol{\delta}\|^2 \leq b \tag{8c}$$

where b is a prescribed bound to control the magnitude of $\boldsymbol{\delta}$, and for notation simplicity, the dependence of $\boldsymbol{g}_{rk}$, $\boldsymbol{g}_{ik}$, e_{rk}, and e_{ik} on ω has been omitted.

By using linear algebraic arguments, it can be readily shown that the constraint in (8b) holds if and only if

$$\hat{\boldsymbol{\Phi}}_k(\eta, \boldsymbol{\delta}, \omega) = \begin{bmatrix} \eta & \boldsymbol{g}_{rk}^T\boldsymbol{\delta} + e_{rk} & \boldsymbol{g}_{ik}^T\boldsymbol{\delta} + e_{ik} \\ \boldsymbol{g}_{rk}^T\boldsymbol{\delta} + e_{rk} & 1 & 0 \\ \boldsymbol{g}_{ik}^T\boldsymbol{\delta} + e_{ik} & 0 & 1 \end{bmatrix}_{3 \times 3} \succeq \boldsymbol{0}$$
$$\text{for } \omega \in \Omega \tag{9}$$

and that the constraint in (8c) holds if and only if

$$\boldsymbol{\Psi}(\boldsymbol{\delta}) = \begin{bmatrix} b & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \boldsymbol{I}_n \end{bmatrix} \succeq \boldsymbol{0}. \tag{10}$$

If we treat the upper bound η as an additional design variable and define an augmented vector as $\boldsymbol{u} = [\eta \quad \boldsymbol{\delta}^T]^T$, then the objective function in (8a) can be expressed as $\eta = \boldsymbol{c}^T \boldsymbol{u}$ with

$c = [1 \quad 0 \quad \cdots \quad 0]^T$, and the problem in (8) can be formulated as

$$\text{minimize} \quad c^T u \tag{11a}$$

$$\text{subject to:} \quad G_k(u) \succeq 0 \tag{11b}$$

where

$$G_k(u) = \text{diag}\{\Phi_k(\eta, \delta), \Psi(\delta)\}$$

$$\Phi_k(\eta, \delta) = \text{diag}\left\{\hat{\Phi}_k(\eta, \delta, \omega_1), \ldots, \hat{\Phi}_k(\eta, \delta, \omega_K)\right\}$$

and $\Omega_d = \{\omega_1, \ldots, \omega_K\} \subset \Omega$ is a set of dense grid points in the frequency bands of interest. Note that matrix $G_k(u)$ in (11b) is affine with respect to vector u and the problem in (11) is, therefore, an SDP problem.

Having solved the problem in (11) for a minimizer

$$u_k^* = \begin{bmatrix} \eta_k^* \\ \delta_k^* \end{bmatrix}$$

vector δ_k^* is used to update x_k as

$$x_{k+1} = x_k + \delta_k^*.$$

The iteration continues until $\|\delta_k^*\|$ becomes insignificant compared to a prescribed tolerance.

The original problem in (5) and, equivalently, the problem in (6) are highly nonlinear and nonconvex optimization problems. As such, the above method, if it converges, only provides a local minimizer for the problem. Among other things, the performance of such a local solution depends largely on how the initial point is chosen. Fortunately, for FRM filter designs, a technique that generates a reasonably good initial point is available, see [1] and Section II-A. Concerning the convergence of the method, although a rigorous proof is presently not available, in our simulations when the method was applied to design a variety of FRM filters, we had not detected a single failure of convergence. One might attribute the success of the proposed method to three factors: 1) the global convergence of each subproblem in (11) when an interior-point convex programming algorithm is applied; 2) the use of constraint (8c) that validates the key approximation in (7); and 3) the use of a good initial point.

Another related issue is the convergence rate or, in a more general term, the computational efficiency. From the above description of the method, it is quite clear that the computational efficiency is determined by how efficient each individual SDP problem in (11) is solved and how many linear updates are needed to reach a minimizer of (6). For the former, most of the algorithms that are presently available for solving the SDP problem (11) are so-called polynomial-time algorithms, meaning that the amount of computations required is bounded by a polynomial of the data size [13]. Consequently, the computational complexity for problem (11) is affordable for today's computing devices even for designing high-order FRM filters, and it will increase only moderately when the size of the problem increases. For the latter, with a given bound b in constraint (8c), the number of updates needed depends on how far the initial point is from the minimizer. In the context

of FRM filter design, with an initial point generated by the method in [1], the number of updates required is typically in the range of 10 to 30.

It should also be pointed out that although (11) is merely an *approximation* of (6), as the iteration continues and the local minimizer gets closer, the increment vector δ obtained by solving (11) gradually shrinks in magnitude and within a limited number of iterations it eventually becomes such a value that the updated solution point is practically the same as the true minimizer.

In summary, we have described a method for minimax optimization of an objective function that is frequently encountered in filter design problems and is allowed to be highly nonlinear. The method proposed here accomplishes the optimization through a sequence of linear updates where each update is solvable in an SDP setting. The usefulness of this methodology will be demonstrated in the next two sections where various FRM filter design problems are addressed.

We conclude this section with a remark on the scale of the SDP problem in (11). In the general case, as addressed above where both $H(\omega, x)$ and $H_d(\omega)$ are complex-valued functions, the dimension of the LMI constraint in (11b) is $3K+n+1$ where n is the number of design variables and K is the total number of grid points in Ω_d. If, however, both $H(\omega, x)$ and $H_d(\omega)$ are real-valued functions (which is indeed the case encountered in the design of *linear-phase* FRM filters), then the gradient $g(\omega, x)$ is also real-valued and the matrix $\hat{\Phi}_k(\eta, \delta, \omega)$ in (9) becomes a 2×2 matrix and the constraint in (9) becomes

$$\hat{\Phi}_k(\eta, \delta, \omega) = \begin{bmatrix} \eta & g_k^T \delta + e_k \\ g_k^T \delta + e_k & 1 \end{bmatrix}_{2 \times 2} \succeq 0, \quad \text{for } \omega \in \Omega \tag{12}$$

where $e_k = W(\omega)[H(\omega, x_k) - H_d(\omega)]$. The size of matrix $G_k(u)$ in (11b) in this case is reduced to $2K + n + 1$. In a typical FRM filter design problem, the "do-not-care" region is very small, hence the grid points in set Ω_a need to be distributed over almost the entire base band Ω and K is likely in the range of 5×10^2 to 10^3. Thus, for an FRM filter of moderate order, the number of constraints can be in the range of 1.5×10^3 to 3×10^3 (in a variety of FRM filter designs that we have conducted, it was found that as long as K is in the above range, changing K from one value to another does not lead to significant changes in the performance of the FRM filter). Fortunately, a robust SDP solver (such as those available in the MATLAB LMI Control Toolbox [14]) today can easily handle a SDP problem with the size of $G_k(u)$ in the range of several thousand on a Pentium PC without numerical difficulties.

IV. OPTIMIZATION OF FRM FILTERS: BASIC AND MULTISTAGE STRUCTURES

Throughout the section, all subfilters are assumed to have linear-phase responses, and the lengths of the masking filters at any given stage are either both even or both odd.

A. Basic FRM Filters

1) Frequency Response and Its Gradient: The reader is referred to the structure in Fig. 1 and (1). Without loss of gener-

ality, the FRM filter can be treated as a zero-phase FIR filter, and the frequency response of the FRM filter is then given by

$$H(\omega, \boldsymbol{x}) = [\boldsymbol{a}^T \boldsymbol{c}(\omega)] [\boldsymbol{a}_a^T \boldsymbol{c}_a(\omega) - \boldsymbol{a}_c^T \boldsymbol{c}_c(\omega)] + \boldsymbol{a}_c^T \boldsymbol{c}_c(\omega) \quad (13)$$

where

$$\boldsymbol{a} = \begin{cases} [h_{(N-1)/2} \ 0.5h_{(N+1)/2} \ \cdots \ 0.5h_{N-1}]^T, & \text{if } N \text{ odd} \\ 0.5[h_{N/2} \ \cdots \ h_{N-1}]^T, & \text{if } N \text{ even} \end{cases}$$

$$\boldsymbol{c}(\omega) = \begin{cases} [1 \ \cos M\omega \ \cdots \ \cos[(N-1)M\omega/2]]^T, & \text{if } N \text{ odd} \\ [\cos(M\omega/2) \ \cdots \ \cos[(N-1)M\omega/2]]^T, & \text{if } N \text{ even} \end{cases}$$

$$\boldsymbol{a}_a = \begin{cases} \left[h_{(N_a-1)/2}^{(a)} \ 0.5h_{(N_a+1)/2}^{(a)} \cdots 0.5h_{N_a-1}^{(a)}\right]^T, & \text{if } N_a \text{ odd} \\ 0.5\left[h_{N_a/2}^{(a)} \ \cdots \ h_{N_a-1}^{(a)}\right]^T, & \text{if } N_a \text{ even} \end{cases}$$

$$\boldsymbol{c}_a(\omega) = \begin{cases} [1 \ \cos\omega \ \cdots \ \cos[(N_a-1)\omega/2]]^T, & \text{if } N_a \text{ odd} \\ [\cos(\omega/2) \ \cdots \ \cos[(N_a-1)\omega/2]]^T, & \text{if } N_a \text{ even} \end{cases}$$

$$\boldsymbol{a}_c = \begin{cases} \left[h_{(N_c-1)/2}^{(c)} \ 0.5h_{(N_c+1)/2}^{(c)} \cdots 0.5h_{N_c-1}^{(c)}\right]^T, & \text{if } N_c \text{ odd} \\ 0.5\left[h_{N_c/2}^{(c)} \ \cdots \ h_{N_c-1}^{(c)}\right]^T, & \text{if } N_c \text{ even} \end{cases}$$

$$\boldsymbol{c}_c(\omega) = \begin{cases} [1 \ \cos\omega \ \cdots \ \cos[(N_c-1)\omega/2]]^T, & \text{if } N_c \text{ odd} \\ [\cos(\omega/2) \ \cdots \ \cos[(N_c-1)\omega/2]]^T, & \text{if } N_c \text{ even} \end{cases}$$

and the design variables are put together as parameter vector

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{a}_a \\ \boldsymbol{a}_c \end{bmatrix}.$$

The group delay of the FRM filter is given by

$$D = \frac{(N-1)M}{2} + d \quad (14)$$

where $d = \max((N_a - 1)/2, (N_c - 1)/2)$, and the gradient of $H(\omega, \boldsymbol{x})$ with respect to $\boldsymbol{x}$ is given by

$$\boldsymbol{g}(\omega, \boldsymbol{x}) = \begin{bmatrix} y(\omega)\boldsymbol{c}(\omega) \\ [\boldsymbol{a}^T \boldsymbol{c}(\omega)]\boldsymbol{c}_a(\omega) \\ [1 - \boldsymbol{a}^T \boldsymbol{c}(\omega)]\boldsymbol{c}_c(\omega) \end{bmatrix} \quad (15)$$

where $y(\omega) = \boldsymbol{a}_a^T \boldsymbol{c}_a(\omega) - \boldsymbol{a}_c^T \boldsymbol{c}_c(\omega)$.

2) Desired Frequency Response and Weighting Function: For the sake of presentation clarity, we consider the case of designing a low-pass FRM filter with up-sampling factor M, normalized passband edge ω_p and stopband edge ω_a. The desired $H_d(\omega)$ in this case becomes

$$H_d(\omega) = \begin{cases} 1, & \text{for } 0 \le \omega \le \omega_p \\ 0, & \text{for } \omega_a \le \omega \le \pi \end{cases} \quad (16)$$

and a staircase weighting function

$$W(\omega) = \begin{cases} 1, & \text{for } 0 \le \omega \le \omega_p \\ w, & \text{for } \omega_a \le \omega \le \pi \\ 0, & \text{elsewhere} \end{cases} \quad (17)$$

is chosen, where w is a positive scalar to weigh the stopband relative to the passband.

3) Initial Design: Given parameters M, N, N_a, N_c, ω_p, and ω_a, a reasonable initial design can be obtained by designing low-pass $H(z)$, $H_{ma}(z)$, and $H_{mc}(z)$ as discussed in Section II-A. It is important to stress that although (as will be demonstrated by simulations shortly) the optimized $H(z)$ does not at all look like a low-pass filter, the initial design prepared here worked flawlessly in a variety of FRM designs we have attempted.

4) Placement of Grid Points and Bound b: Our design practice has indicated that relatively denser grid points should be placed in the regions near the band edges in both passband and stopband so as to avoid using unnecessarily large number of total grid points. We recommend that about 25% of the grid points be placed in the 10% of that band nearest to the band edge.

As expected, the value of bound b in constraint (8c) is taken to be proportional to the dimension of vector $\boldsymbol{x}$, namely, $b = \gamma n$ where n denotes the dimension of $\boldsymbol{x}$ and γ is a constant factor. It was found in our simulations that the norm constraint (8c) worked effectively when the value of γ was in the range of $[0.005, 0.05]$.

5) Design Examples:

Example 1: The design is a linear-phase low-pass FRM filer with the same design parameters as the first example in [1], i.e., $N = 45$, $N_a = 41$, $N_c = 33$, $M = 9$, $\omega_p = 0.6\pi$, and $\omega_a = 0.61\pi$. The weight was set to $w = 1$, bound b in (8c) was set to $b = 0.005n$ ($n = 61$ in this design), and the total number of grids was $K = 900$. In this case the optimization algorithm in Section III handles 62 variables with 1862 constraints. With ten iterations, the algorithm converges to an FRM filter with the amplitude response of its subfilters $H_a(z^M)$, $H_{ma}(z)$, and $H_{mc}(z)$ shown in Fig. 2(a) and (b), respectively, and the amplitude response of the FRM filter and its passband ripples shown in Fig. 2(c) and (d), respectively. The maximum passband ripple was found to be 0.0674 dB and the minimum stopband attenuation was 42.25 dB. By comparison, the passband ripple and stopband attenuation of the design in [1] were 0.0896 dB and 40.96 dB, respectively.

As can be seen from Fig. 2, the masking filters $H_{ma}(z)$ and $H_{mc}(z)$ resulted from the joint optimization remain low pass with very similar passband widths, but the optimized prototype filter $H_a(z^M)$ is not at all a low-pass filter. Note that $H_a(z^M)$ has a sharp drop down precisely at the passband edge [normalized to 0.3 in Fig. 2(a)].

Example 2: This is a linear-phase FRM filer with the same design parameters as the example discussed in [9], i.e., $N = 123$, $N_a = 56$, $N_c = 78$, $M = 21$, $\omega_p = 0.4\pi$, and $\omega_a = 0.402\pi$. The scalar weight was set to $w = 12$ and bound $b = 0.025n$ with $n = 129$. A total number of 1000 grid points was used, hence the SDP-based optimization method handles in this case 130 design variables and 2130 constraints. It took the proposed algorithm 15 iterations to converge to an FRM filter with the amplitude responses of its subfilters and the overall filter shown in Fig. 3. The maximum passband deviation was found to be 0.0855 dB and the minimum stopband attenuation achieved was 60.93 dB. By comparison, the maximum passband ripple and minimum stopband attenuation achieved by the design in
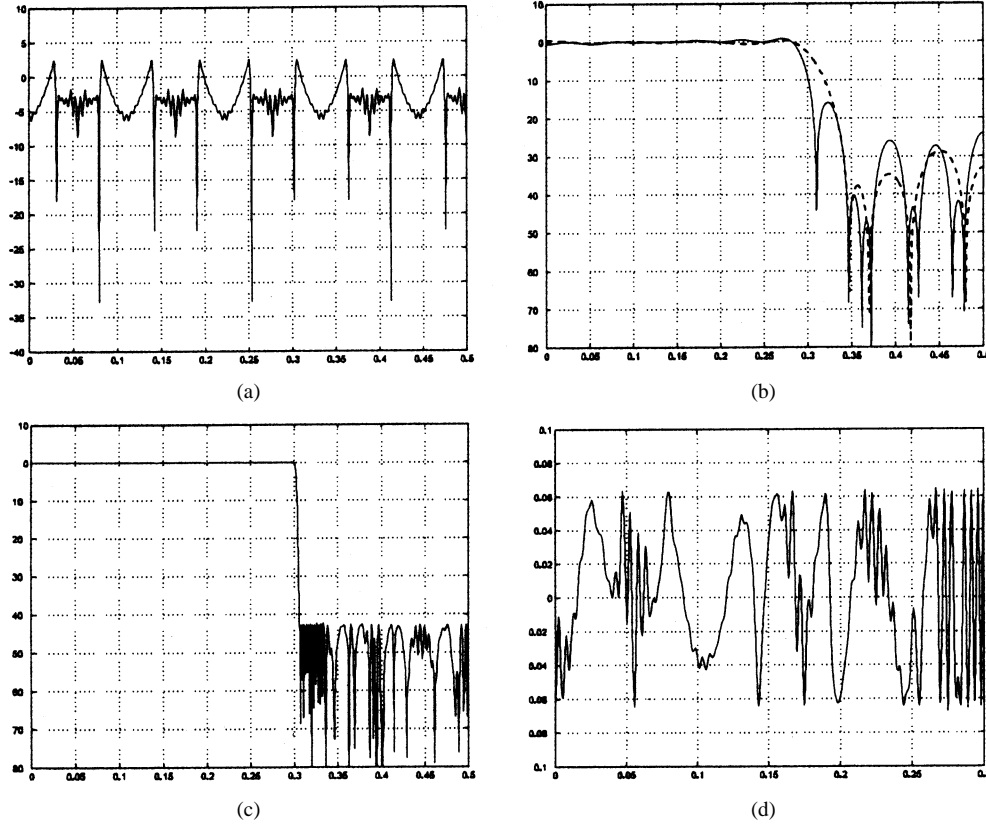
Fig. 2.   Amplitude responses of (a) prototype filter $H_a(z^9)$; (b) masking filters $H_{ma}(z)$, and $H_{mc}(z)$; (c) FRM filter; and (d) passband ripples of the FRM filter, all in decibels.

[9] were 0.0864 dB ($\delta_p = 0.01$) and 60 dB ($\delta_s = 0.001$), respectively.

B. Multistage FRM Filters

The algorithm in Section III also applies to multistage FRM filters. For the sake of notation simplicity, in what follows we focus our attention on the two-stage case and we refer the reader to the filter structure in Fig. 1(b). However, as we proceed it should become apparent as how a similar treatment can be used for the design of FRM filters with more than two stages.

1) Frequency Response and Its Gradient: Since the subfilters involved have linear phase responses, they will be treated as zero-phase filters and the frequency response of the two-stage FRM filter is given by

$$H(\omega, \boldsymbol{x}) = H_a^{(1)}(\omega) \left[\boldsymbol{a}_{1a}^T \boldsymbol{c}_{1a}(\omega) - \boldsymbol{a}_{1c}^T \boldsymbol{c}_{1c}(\omega) \right] + \boldsymbol{a}_{1c}^T \boldsymbol{c}_{1c}(\omega)$$

(18a)

where

$$H_a^{(1)}(\omega) = \left[\boldsymbol{a}_2^T \boldsymbol{c}_2(\omega) \right] \left[\boldsymbol{a}_{2a}^T \boldsymbol{c}_{2a}(\omega) - \boldsymbol{a}_{2c}^T \boldsymbol{c}_{2c}(\omega) \right] + \boldsymbol{a}_{2c}^T \boldsymbol{c}_{2c}(\omega).$$

(18b)

The vectors $\boldsymbol{a}_2, \boldsymbol{a}_{ia}, \boldsymbol{a}_{ic}$ (for $i = 1, 2$) in (18) are the coefficient vectors for filters $H_a^{(2)}(z), H_{ma}^{(i)}(z), H_{mc}^{(i)}(z)$, respectively, and are defined in a way similar to vectors $\boldsymbol{a}, \boldsymbol{a}_a$, and $\boldsymbol{a}_c$ in (13), respectively. The cosine vectors for the first-stage masking filters, namely $\boldsymbol{c}_{1a}(\omega)$ and $\boldsymbol{c}_{1c}(\omega)$, are defined in the same way as $\boldsymbol{c}_a(\omega)$ and $\boldsymbol{c}_c(\omega)$ in (13); and the cosine vectors for the second-stage subfilters, i.e., $\boldsymbol{c}_2(\omega), \boldsymbol{c}_{2a}(\omega)$, and $\boldsymbol{c}_{2c}(\omega)$ are defined in a way

similar to $\boldsymbol{c}(\omega), \boldsymbol{c}_a(\omega)$, and $\boldsymbol{c}_c(\omega)$ in (13) except that each ω is replaced by $M\omega$. The parameter vector $\boldsymbol{x}$ in (18) is defined by

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a}_2 \\ \boldsymbol{a}_{2a} \\ \boldsymbol{a}_{2c} \\ \boldsymbol{a}_{1a} \\ \boldsymbol{a}_{1c} \end{bmatrix}.$$

(19)

The group delay of the FRM filter is given by

$$D^{(2)} = \frac{(N_2 - 1)M^2}{2} + d_2 M + d_1$$

(20)

where N_2 is the length of $H_a^{(2)}(z)$ and $d_i = \max((N_{ai} - 1)/2, (N_{ci} - 1)/2)$ for $i = 1, 2$ with N_{ai} and N_{ci} being the lengths of the masking filters $H_{ma}^{(i)}(z)$ and $H_{mc}^{(i)}(z)$, respectively. Using (18), we compute the gradient of $H(\omega, \boldsymbol{x})$ with respect to $\boldsymbol{x}$ as

$$\boldsymbol{g}(\omega, \boldsymbol{x}) = \begin{bmatrix} y^{(1)}(\omega) y^{(2)}(\omega) \boldsymbol{c}_2(\omega) \\ \left[\boldsymbol{a}_2^T \boldsymbol{c}_2(\omega) \right] y^{(1)}(\omega) \boldsymbol{c}_{2a}(\omega) \\ \left[1 - \boldsymbol{a}_2^T \boldsymbol{c}_2(\omega) \right] y^{(1)}(\omega) \boldsymbol{c}_{2c}(\omega) \\ H_a^{(1)}(\omega) \boldsymbol{c}_{1a}(\omega) \\ \left[1 - H_a^{(1)}(\omega) \right] \boldsymbol{c}_{1c}(\omega) \end{bmatrix}$$

(21a)

where

$$y^{(1)}(\omega) = \boldsymbol{a}_{1a}^T \boldsymbol{c}_{1a}(\omega) - \boldsymbol{a}_{1c}^T \boldsymbol{c}_{1c}(\omega)$$

(21b)

$$y^{(2)}(\omega) = \boldsymbol{a}_{2a}^T \boldsymbol{c}_{2a}(\omega) - \boldsymbol{a}_{2c}^T \boldsymbol{c}_{2c}(\omega)$$
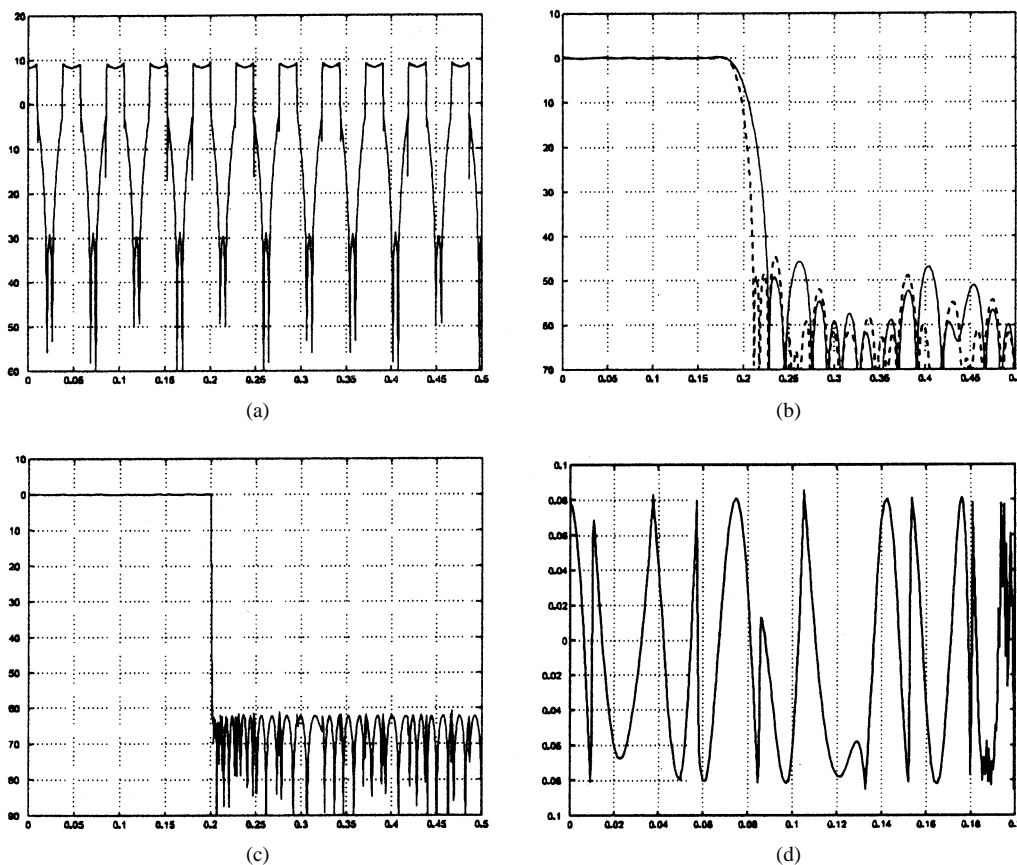
(21c)

Fig. 3. Amplitude responses of (a) prototype filter $H_a(z^{21})$; (b) masking filters $H_{ma}(z)$, and $H_{mc}(z)$; (c) FRM filter; and (d) passband ripples of the FRM filter, all in decibels.

and $H_a^{(1)}(\omega)$ is defined in (18b).

2) Initial Design: Again, we consider a low-pass filter design. Given parameters M, N_2, N_{a2}, N_{c2}, N_{a1}, N_{c1}, ω_p, and ω_a, an initial design for a two-stage FRM filter can be readily obtained in a two-stage manner as follows.

i) Use parameters M, N_1, N_{a1}, N_{c1}, ω_p, and ω_a, where N_1 is an arbitrary positive integer for the reason that will become apparent shortly, to obtain an initial design of $H_a^{(1)}(z)$, $H_{ma}^{(1)}(z)$, and $H_{mc}^{(1)}(z)$ by the method in Section II-A. In doing so, we also obtain the values of θ and ϕ as the passband and stopband edges for $H_a^{(1)}(z)$.

ii) Now, denote $\hat{\omega}_p = \theta$ and $\hat{\omega}_a = \phi$ and use parameters M, N_2, N_{a2}, N_{c2}, $\hat{\omega}_p$, and $\hat{\omega}_a$ to obtain an initial design of $H_a^{(2)}(z)$, $H_{ma}^{(2)}(z)$, and $H_{mc}^{(2)}(z)$ by the same method (in Section II-A). The coefficients of $H_a^{(2)}(z)$, $H_{ma}^{(2)}(z)$, $H_{mc}^{(2)}(z)$, $H_{ma}^{(1)}(z)$, and $H_{mc}^{(1)}(z)$ are now used to form the initial parameter vector $\boldsymbol{x}_0$ as in (19). As we can see, filter $H_a^{(1)}(z)$ was not involved in $\boldsymbol{x}_0$, but just used as an entity to which the values of θ and ϕ become meaningful parameters, and this is why parameter N_1 can be chosen arbitrarily.

As can be expected in an m-stage FRM filter design, an initial design can be generated by repeatedly applying the method in Section II-A m times and extract the coefficients of the subfilters obtained to form vector $\boldsymbol{x}_0$.

3) Design Example: For illustration purposes, we applied the optimization technique in Section III to design a two-stage FRM filter with $N_2 = 27$, $N_{a2} = 13$, $N_{c2} = 27$, $N_{a1} = 15$, $N_{c1} = 23$, $M = 4$, $\omega_p = 0.6$, and $\omega_a = 0.61$. With $w = 1$, $K = 700$, and $b = 0.005n$ (with $n = 68$), it took the proposed algorithm 18 iterations to converge to an FRM filter. The amplitude response of various subfilters are shown in Fig. 4. The maximum deviation in the passband was 0.032 dB and the minimum stopband attenuation was 48.77 dB. From Fig. 4(a) and (b), we see that the prototype $H_a^{(2)}(z^{16})$ as well as the masking filters $H_{ma}^{(2)}(z^4)$ and $H_{mc}^{(2)}(z^4)$ do not act like low-pass filters and it seems that they can be obtained only through a joint optimization of all subfilters. From Fig. 4(c) and (d), we observe that the masking filters $H_{ma}^{(1)}(z)$ and $H_{mc}^{(1)}(z)$ remain low pass in nature but the "prototype" filter $H_a^{(1)}(z^4)$ synthesized by the second stage of FRM filtering [see Fig. 1(b)] is not. Note the sharp drop-down in $H_a^{(1)}(z^4)$ precisely at the passband edge (normalized to 0.3 in the figure). Finally, we remark that although the design resulted in a slightly more economical FRM filter compared to the one in Example 1, the two-stage filter offers better performance in both passband and stopband as is evident from Fig. 4(e) and (f). Nevertheless, the performance gain is achieved at the cost of an increased group delay $D^{(2)} = 271$ compared to the group delay $D = 218$ for the filter in Example 1.
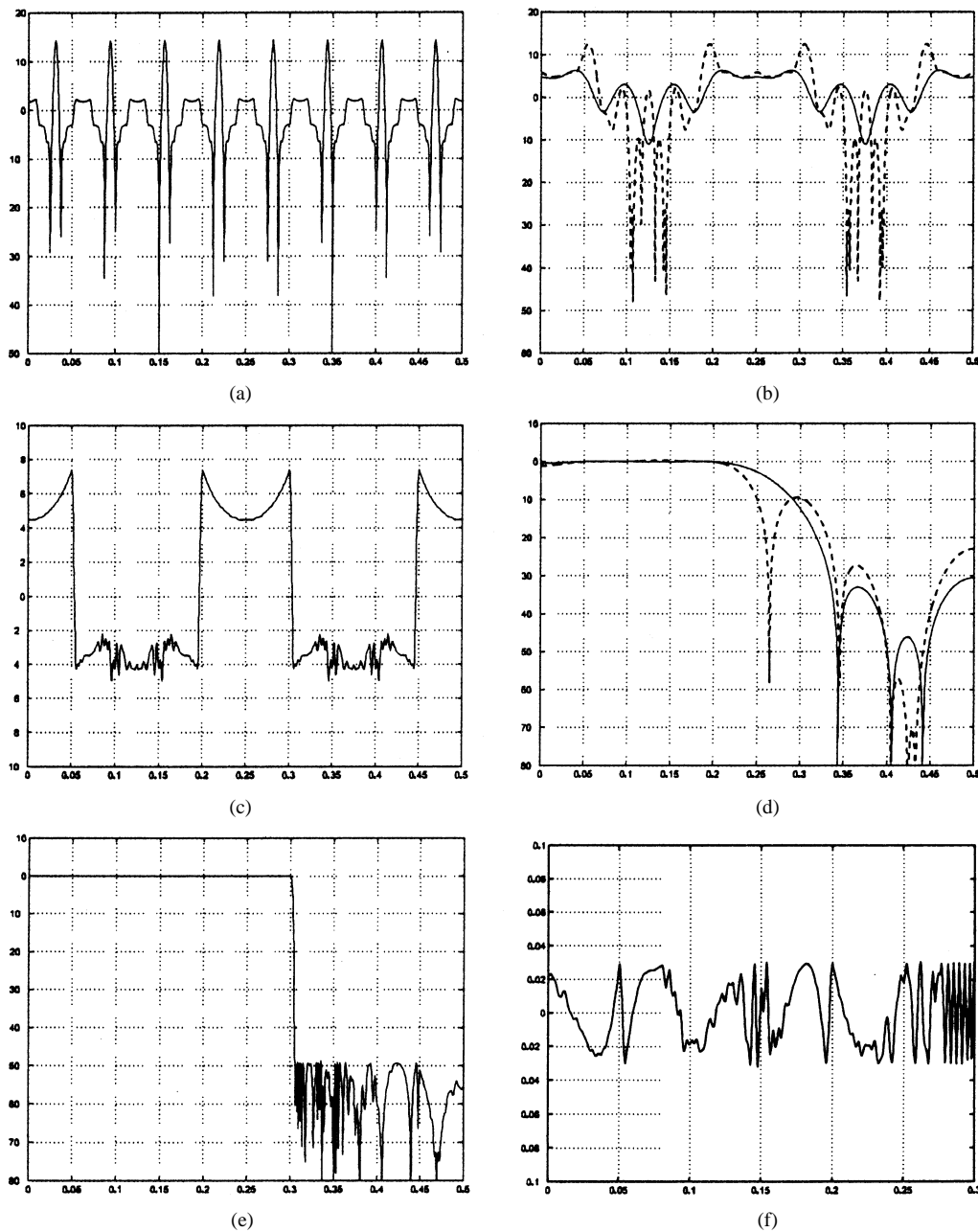
Fig. 4. Amplitude responses of (a) prototype filter $H_a^{(2)}(z^{16})$; (b) masking filters $H_{ma}^{(2)}(z^4)$, and $H_{mc}^{(2)}(z^4)$; (c) "prototype" $H_a^{(1)}(z^4)$; (d) masking filters $H_{ma}^{(1)}(z)$ and $H_{mc}^{(1)}(z)$; (e) FRM filter; and (f) passband ripples of the FRM filter, all in decibels.

V. OPTIMIZATION OF FRM FILTERS WITH REDUCED GROUP DELAY

As can be seen from (14) and (20), the group delay of an FRM filter is dominantly contributed by the prototype filter, and it is even more so as the number of stages increases. For this reason, throughout the section, it is assumed that the prototype filter is the only filter with a *nonlinear* phase response in the entire FRM filter structure regardless of the number of stages the FRM filter has. The treatment adopted here, however, is such that it can be readily extended to an FRM filter where other subfilters are allowed to have nonlinear phase responses. Such an extension is made possible because of the *joint* optimization methodology we adopt in that all subfilters participate in achieving the design goal.
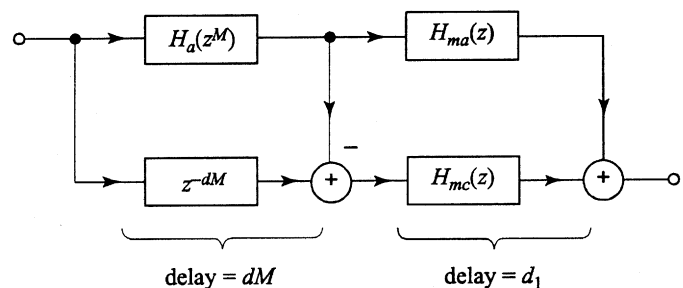

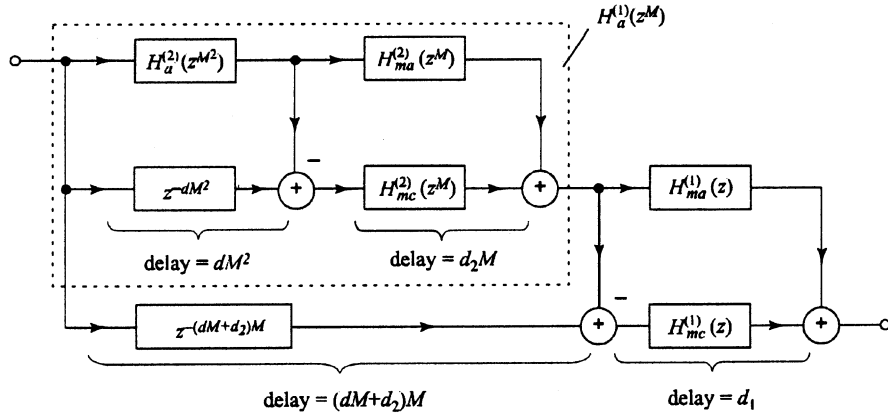Fig. 5. Basic FRM filter with reduced group delay $dM + d_1$.

Fig. 6. Two-stage FRM filter with reduced group delay.

A. Basic FRM Filters

1) Frequency Response and Its Gradient: A basic FRM filter with passband group delay $D_r = dM + d_1$ is illustrated in Fig. 5, where the prototype filter has the frequency response

$$H_a(e^{jM\omega}) = \sum_{k=0}^{N-1} h_k e^{-jkM\omega}. \tag{22}$$

Assuming the group delays of masking filters $H_{ma}(z)$ and $H_{mc}(z)$ have been equalized to $d_1 = \max((N_a - 1)/2, (N_c - 1)/2)$, the frequency response of the FRM filter can be expressed as

$$e^{-jD_r\omega} H(\omega, \boldsymbol{x}) \tag{23a}$$

with

$$H(\omega, \boldsymbol{x}) = \tilde{H}_a(\omega)\big[\boldsymbol{a}_a^T \boldsymbol{c}_a(\omega) - \boldsymbol{a}_c^T \boldsymbol{c}_c(\omega)\big] + \boldsymbol{a}_c^T \boldsymbol{c}_c(\omega) \tag{23b}$$

$$\tilde{H}_a(\omega) = \boldsymbol{h}^T[\boldsymbol{c}(\omega) + j\boldsymbol{s}(\omega)] \tag{23c}$$

$$\boldsymbol{h} = [h_0 \ \ h_1 \ \cdots \ h_{N-1}]^T \tag{23d}$$

$$\boldsymbol{c}(\omega) = [\cos M\,d\omega \ \ \cos M(d-1)\omega \cdots \ \cos M(d-N+1)\omega]^T \tag{23e}$$

$$\boldsymbol{s}(\omega) = [\sin M\,d\omega \ \ \sin M(d-1)\omega \cdots \ \sin M(d-N+1)\omega]^T \tag{23f}$$

and $\boldsymbol{a}_a$, $\boldsymbol{a}_c$, $\boldsymbol{c}_a(\omega)$, and $\boldsymbol{c}_c(\omega)$ defined in (13). Note that because of $\tilde{H}_a(\omega)$ in (23c), $H(\omega, \boldsymbol{x})$ in (23b) is a complex-valued function, where parameter vector $\boldsymbol{x}$ is defined as

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{a}_a \\ \boldsymbol{a}_c \end{bmatrix}.$$

Now if the value of d is strictly less than $(N-1)/2$, then the first factor in (23a) with $D_r = dM + d_1$ represents a reduced group delay provided that the second factor in (23a), $H(\omega, \boldsymbol{x})$, best approximates the zero-phase desired frequency response $H_d(\omega)$ in (16).

Two remarks before we proceed to compute the gradient of $H(\omega, \boldsymbol{x})$: First, obviously the algorithm in Section III applies to the current design problem but the number of constraints increases to $3K + n + 1$ since $H(\omega, \boldsymbol{x})$ is now complex-valued; second, in an effort of minimizing the maximum error $\max |H(\omega, \boldsymbol{x}) - H_d(\omega)|$, from (23b) it is quite clear that all subfilters will jointly participate in that effort.

The gradient of $H(\omega, \boldsymbol{x})$ is also complex-valued and is given by $\boldsymbol{g}(\omega, \boldsymbol{x}) = \boldsymbol{g}_r(\omega, \boldsymbol{x}) + j\boldsymbol{g}_i(\omega, \boldsymbol{x})$ where

$$\boldsymbol{g}_r(\omega, \boldsymbol{x}) = \begin{bmatrix} y(\omega)\boldsymbol{c}(\omega) \\ [\boldsymbol{h}^T \boldsymbol{c}(\omega)]\boldsymbol{c}_a(\omega) \\ [1 - \boldsymbol{h}^T \boldsymbol{c}(\omega)]\boldsymbol{c}_c(\omega) \end{bmatrix} \tag{24a}$$

$$\boldsymbol{g}_i(\omega, \boldsymbol{x}) = \begin{bmatrix} y(\omega)\boldsymbol{s}(\omega) \\ [\boldsymbol{h}^T \boldsymbol{s}(\omega)]\boldsymbol{c}_a(\omega) \\ -[\boldsymbol{h}^T \boldsymbol{s}(\omega)]\boldsymbol{c}_c(\omega) \end{bmatrix} \tag{24b}$$

$$y(\omega) = \boldsymbol{a}_a^T \boldsymbol{c}_a(\omega) - \boldsymbol{a}_c^T \boldsymbol{c}_c(\omega). \tag{24c}$$

2) Initial Design: The initial design of the masking filters $H_{ma}(z)$ and $H_{mc}(z)$ remains the same as in Section IV-A. As well, one can use the formulas given in Section II-A to predict the passband and stopband edges θ and ϕ. However, at this point one needs a low-pass filter $H_a(z)$ with $\omega_p = \theta$, $\omega_a = \phi$ and a reduced passband group delay d (strictly less than $(N-1)/2$). A reasonably good initial $H_a(z)$ is the weighted least-squares solution that minimizes

$$\int_0^\pi W(\omega) \left| H_a(e^{j\omega}) - \tilde{H}_d(\omega) \right|^2 d\omega \tag{25}$$

where

$$W(\omega) = \begin{cases} 1, & \text{for } \omega \in [0, \theta) \\ w, & \text{for } \omega \in [\phi, \pi] \\ 0, & \text{elsewhere} \end{cases}$$

and

$$\tilde{H}_d(\omega) = \begin{cases} e^{-jd\omega}, & \text{for } \omega \in [0, \theta) \\ 0, & \text{for } \omega \in [\phi, \pi] \text{ and elsewhere.} \end{cases}$$

It can be shown that the objective function in (25) is a strictly convex quadratic function with a Toeplitz type Hessian matrix. Consequently, the least-square solution can be computed efficiently by solving a Toeplitz system of linear equations [17].

B. Multistage FRM Filters

1) Frequency Response and Its Gradient: The design method in this case will be illustrated using a two-stage low-pass FRM filter with reduced group delay shown in Fig. 6,

where only prototype filter $H_a^{(2)}(z^{M^2})$ has a nonlinear phase response and its frequency response is given by

$$H_a^{(2)}\left(e^{jM^2\omega}\right) = \sum_{k=0}^{N_2-1} h_k e^{-jkM^2\omega}. \tag{26}$$

The frequency response of the two-stage FRM filter can be expressed as

$$e^{-jD_r^{(2)}\omega}H(\omega, \boldsymbol{x}) \tag{27a}$$

where

$$D_r^{(2)} = dM^2 + d_2 M + d_1 \tag{27b}$$

$$H(\omega, \boldsymbol{x}) = \tilde{H}_a^{(1)}(\omega)\left[\boldsymbol{a}_{1a}^T \boldsymbol{c}_{1a}(\omega) - \boldsymbol{a}_{1c}^T \boldsymbol{c}_{1c}(\omega)\right] + \boldsymbol{a}_{1c}^T \boldsymbol{c}_{1c}(\omega) \tag{27c}$$

$$\tilde{H}_a^{(1)}(\omega) = \tilde{H}_a^{(2)}(\omega)\left[\boldsymbol{a}_{2a}^T \boldsymbol{c}_{2a}(\omega) - \boldsymbol{a}_{2c}^T \boldsymbol{c}_{2c}(\omega)\right] + \boldsymbol{a}_{2c}^T \boldsymbol{c}_{2c}(\omega) \tag{27d}$$

$$\tilde{H}_a^{(2)}(\omega) = \boldsymbol{h}_2^T[\boldsymbol{c}_2(\omega) + j\boldsymbol{s}_2(\omega)] \tag{27e}$$

$$\boldsymbol{h}_2 = [h_0 \ \ h_1 \ \ \cdots \ \ h_{N_2-1}]^T \tag{27f}$$

$$\boldsymbol{c}_2(\omega) = [\cos M^2 d\omega \ \ \cos M^2(d-1)\omega \ \cdots \ \cos M^2(d-N_2+1)\omega]^T \tag{27g}$$

$$\boldsymbol{s}_2(\omega) = [\sin M^2 d\omega \ \ \sin M^2(d-1)\omega \ \cdots \ \sin M^2(d-N_2+1)\omega]^T \tag{27h}$$

vectors $\boldsymbol{a}_{ia}, \boldsymbol{a}_{ic}, \boldsymbol{c}_{ia}(\omega)$, and $\boldsymbol{c}_{ic}(\omega)$ for $i = 1, 2$ are the same as in (18), and vector $\boldsymbol{x}$ is defined by

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{h}_2 \\ \boldsymbol{a}_{2a} \\ \boldsymbol{a}_{2c} \\ \boldsymbol{a}_{1a} \\ \boldsymbol{a}_{1c} \end{bmatrix}.$$

With a value of d strictly less than $(N_2 - 1)/2$, (27a) represents a frequency response with reduced group delay $D_r^{(2)}$ provided that $H(\omega, \boldsymbol{x})$ in (27c) approximates the zero-phase desired frequency response $H_d(\omega)$ in (16). To proceed, we compute the gradient of $H(\omega, \boldsymbol{x})$ as $\boldsymbol{g}(\omega, \boldsymbol{x}) = \boldsymbol{g}_r(\omega, \boldsymbol{x}) + j\boldsymbol{g}_i(\omega, \boldsymbol{x})$ where

$$\boldsymbol{g}_r(\omega, \boldsymbol{x}) = \begin{bmatrix} y^{(1)}(\omega)y^{(2)}(\omega)\boldsymbol{c}_2(\omega) \\ \left[\boldsymbol{h}_2^T \boldsymbol{c}_2(\omega)\right]y^{(1)}(\omega)\boldsymbol{c}_{2a}(\omega) \\ \left[1 - \boldsymbol{h}_2^T \boldsymbol{c}_2(\omega)\right]y^{(1)}(\omega)\boldsymbol{c}_{2c}(\omega) \\ \tilde{Y}_a^{(1)}(\omega)\boldsymbol{c}_{1a}(\omega) \\ \left[1 - \tilde{Y}_a^{(1)}(\omega)\right]\boldsymbol{c}_{1c}(\omega) \end{bmatrix} \tag{28a}$$

$$\boldsymbol{g}_i(\omega, \boldsymbol{x}) = \begin{bmatrix} y^{(1)}(\omega)y^{(2)}(\omega)\boldsymbol{s}_2(\omega) \\ \left[\boldsymbol{h}_2^T \boldsymbol{s}_2(\omega)\right]y^{(1)}(\omega)\boldsymbol{c}_{2a}(\omega) \\ -\left[\boldsymbol{h}_2^T \boldsymbol{s}_2(\omega)\right]y^{(1)}(\omega)\boldsymbol{c}_{2c}(\omega) \\ \left[\boldsymbol{h}_2^T \boldsymbol{s}_2(\omega)\right]y^{(2)}(\omega)\boldsymbol{c}_{1a}(\omega) \\ -\left[\boldsymbol{h}_2^T \boldsymbol{s}_2(\omega)\right]y^{(2)}(\omega)\boldsymbol{c}_{1c}(\omega) \end{bmatrix} \tag{28b}$$

$$\tilde{Y}_a^{(1)}(\omega) = \left[\boldsymbol{h}_2^T \boldsymbol{c}_2(\omega)\right]y^{(2)}(\omega) + \boldsymbol{a}_{2c}^T \boldsymbol{c}_{2c}(\omega) \tag{28c}$$

and $y^{(1)}(\omega), y^{(2)}(\omega)$ are defined by (21b) and (21c).

2) Initial Design: Initial designs of the two pairs of masking filters $\{H_{ma}^{(1)}(z), H_{mc}^{(1)}(z)\}$ and $\{H_{ma}^{(2)}(z), H_{mc}^{(2)}(z)\}$ can be obtained in the same way as described in Section IV-B-2. During that process, the passband and stopband edges of prototype filter $H_a^{(2)}(z)$ can also be identified. One can then perform a least-square design of $H_a^{(2)}(z)$ with passband group delay d by minimizing the convex quadratic objective function in (25), see the discussion in Section V-A-2.

3) Design Example: To illustrate the design method, we apply the optimization method in Section III to design a two-stage FRM filter with $N_2 = 27$, $N_{a2} = 13$, $N_{c2} = 27$, $N_{a1} = 15$, $N_{c1} = 23$, $M = 4$, $d = 9$, $\omega_p = 0.6\pi$, and $\omega_a = 0.61\pi$. The use of $d = 9$ represents a group delay reduction of $[0.5(N_2 - 1) - d] \times M^2 = 64$, a 24% reduction in group delay from its linear-phase counterpart. With $w = 1$, $b = 0.01n$ (with $n = 68$), and $K = 700$, it took the proposed algorithm 24 iterations to converge to a two-stage FRM filter. The amplitude responses of the subfilters as well as the overall FRM filter, and its passband ripples in magnitude and group delay are depicted in Fig. 7. The maximum passband amplitude ripple was 0.04 dB and the minimum stopband attenuation was 46.01 dB. The relative deviation in passband delay was 0.0132. Compared with its linear-phase counterpart which was the design discussed in Section IV-B, slight performance degradation in terms of passband ripple and stopband attenuation were observed, a cost for having a considerable reduction in group delay.

VI. CONCLUDING REMARKS

In this paper, we have attempted to lay out a methodology for optimal design of various FRM filters. To conclude, we highlight several features of the design method as follows: a) it is an optimization method that treats the coefficients of all subfilters as a single set of design variables regardless of the number of stages the FRM filter has. As a result, the joint optimization of all participating subfilters leads to improved design performance; b) it is a method based on SDP which is a special class of convex programming equipped with efficient interior-point solvers for large scale problems, and as such the proposed design algorithms are able to handle designs of high order, multistage FRM filters with linear or nonlinear phase responses; c) it is a method in which various types of FRM filters can be designed in a unified manner. Consequently, the coding of the design algorithms is substantially simplified, with a set of core codes in common plus small size routines to fit the set of core codes into a specific class of FRM filters; and d) it is a method that works well as long as it starts with a reasonable initial point. Finally, we stress that the class of FRM filters with reduced group delay offers the designer an additional option for the tradeoff between realization complexity, performance, and system delay.
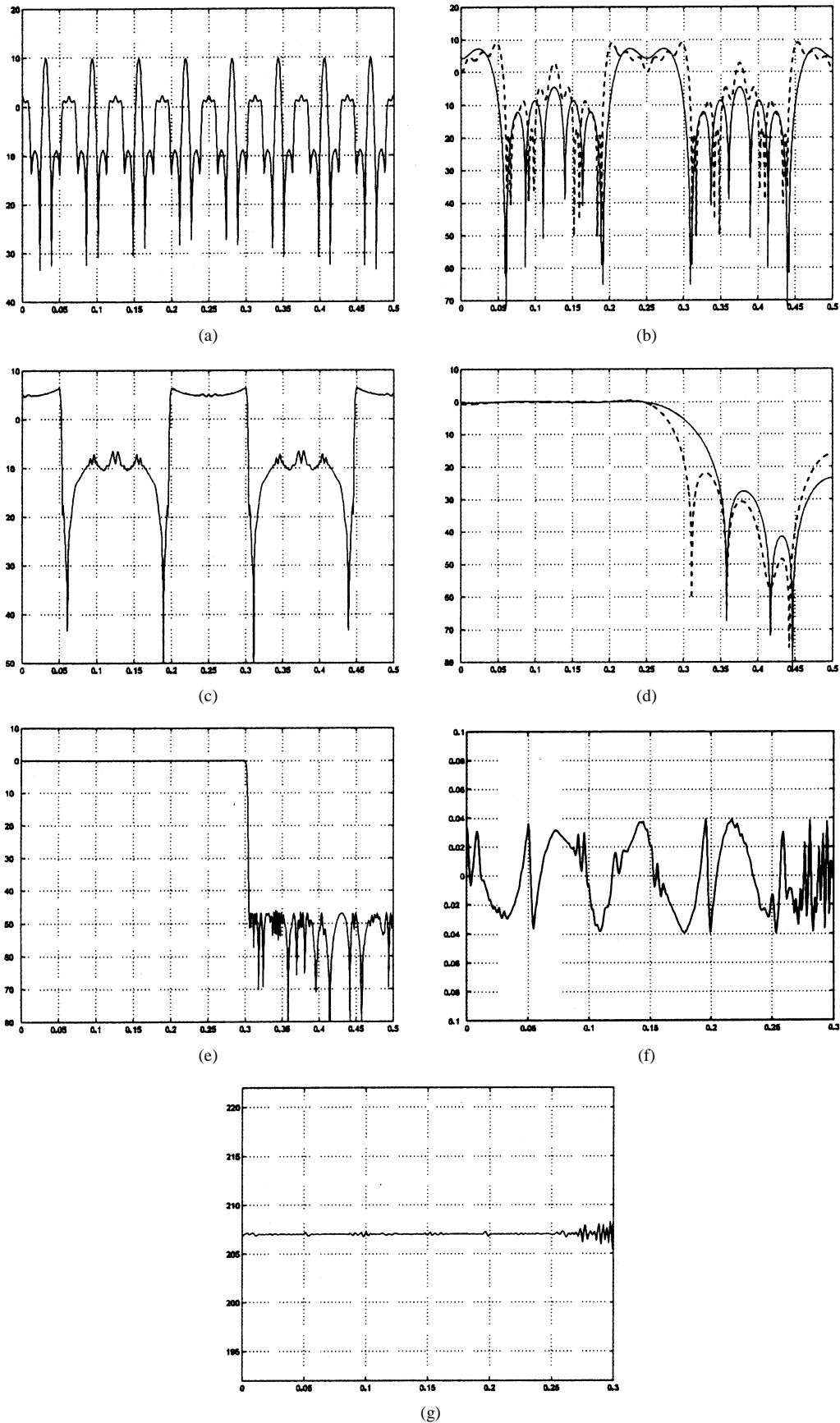
Fig. 7. Amplitude responses of: (a) prototype filter $H_a^{(2)}(z^{16})$; (b) masking filters $H_{ma}^{(2)}(z^4)$ and $H_{mc}^{(2)}(z^4)$; (c) prototype filter $H_a^{(1)}(z^4)$; (d) masking filters $H_{ma}^{(1)}(z)$ and $H_{mc}^{(1)}(z)$; (e) FRM filter; (f) passband amplitude ripples of the FRM filter; and (g) passband group delay of FRM filter.

REFERENCES

[1] Y. C. Lim, "Frequency-response masking approach for the synthesis of sharp linear phase digital filters," *IEEE Trans. Circuits Syst.*, vol. 33, pp. 357–364, Apr. 1986.

[2] R. Yang, B. Liu, and Y. C. Lim, "A new structure of sharp transition FIR filters using frequency-response masking," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 955–966, Aug. 1988.

[3] G. Rajan, Y. Neuvo, and S. K. Mitra, "On the design of sharp cutoff wide-band FIR filters with reduced arithmetic complexity," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1447–1454, Nov. 1988.

[4] T. Saramaki and A. T. Fam, "Subfilter approach for designing efficient FIR filters," in *Proc. 1988 ISCAS*, pp. 2903–2915.

[5] Y. C. Lim and Y. Lian, "The optimum design of one- and two-dimensional FIR filters using the frequency response masking technique," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 88–95, Feb. 1993.

[6] Y. C. Lim and Y. Lian, "Frequency-response masking approach for digital filter design: Complexity reduction via masking filter factorization," *IEEE Trans. Circuits Syst. II*, vol. 41, pp. 518–525, Aug. 1994.

[7] T. Saramaki, Y. C. Lim, and R. Yang, "The synthesis of half-band filter using frequency-response marking technique," *IEEE Trans. Circuits Syst. II*, vol. 42, pp. 58–60, Jan. 1995.

[8] M. G. Bellanger, "Improved design of long FIR filters using the frequency masking technique," in *Proc. 1996 ICASSP*, pp. 1272–1275.

[9] T. Saramaki and H. Johansson, "Optimization of FIR filters using frequency-response masking approach," in *Proc. 2001 ISCAS*, vol. II, pp. 177–180.

[10] S. R. K. Dutta and M. Vidyasagar, "New algorithms for constrained minimax optimization," *Math. Progr.*, vol. 13, pp. 140–155, 1977.

[11] A. Antoniou, *Digital Filters: Analysis, Design, and Applications*, 2nd ed. New York: McGraw-Hill, 1993.

[12] L. Vanderberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, pp. 49–95, 1996.

[13] H. Wolkowicz, R. Saigal, and L. Vanderberghe, Eds., *Handbook on Semidefinite Programming*. Norwell, MA: Kluwer Academic, 2000.

[14] P. Gahinet, A. Nemirovski, A. J. Laub, and M. Chilali, *Manual of LMI Control Toolbox*. Natick, MA: The MathWorks, Inc., 1995.

[15] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11–12, pp. 625–653, 1999. [Online]. Available: http://fewcal.kub.nl/sturm/software/sedumi.html.

[16] R. H. Tütüncü, K. C. Toh, and M. J. Todd. (2001, Aug.) SDPT3—A MATLAB software package for semidefinite–quadratic-linear programming, version 3.0. [Online]. Available: http://www.math.nus.edu.sg/~mattohkc/.

[17] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1989.
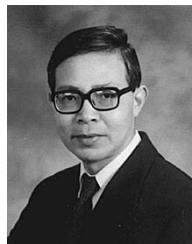
Wu-Shen Lu (F'99) received the undergraduate degree in mathematics from Fudan University, Shanghai, China, in 1964, and the M.S. degree in electrical engineering, and Ph.D. degree in control science from University of Minnesota, Minnesota, in 1983 and 1984, respectively.

He was a Post-Doctoral Fellow at the University of Victoria, Victoria, BC, Canada in 1985 and a Visiting Assistant Professor at the University of Minnesota, in 1986. Since 1987, he has been with the University of Victoria, where he is currently a Professor. His teaching and research interests are in the areas of digital signal processing and application of optimization methods. He is the coauthor of *Two-Dimensional Digital Filters* (New York: Marcel Dekker, 1992). He was an Associate Editor of the *Canadian Journal of Electrical and Computer Engineering* in 1989, and its Editor from 1990 to 1992. He is presently an Associate Editor for the *International Journal of Multidimensional Systems and Signal Processing*.

Dr. Lu served as an Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II from 1993 to 1995, and for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I from 1999 to 2001. He is a Fellow of the Engineering Institute of Canada.

Takao Hinamoto (M'77–SM'84–F'01) received the B.E. degree from Okayama University, Okayama, Japan, in 1969, the M.E. degree from Kobe University, Kobe, Japan, in 1971, and the Dr. Eng. degree from Osaka University, Osaka, Japan, in 1977, all in electrical engineering.

From 1972 to 1988, he was with the Faculty of Engineering, Kobe University. From 1979 to 1981, he was on leave from Kobe University as visiting member of staff in the Department of Electrical Engineering, Queen's University, Kingston, ON, Canada. During 1988–1991, he was Professor of electronic circuits in the Faculty of Engineering, Tottori University, Tottori, Japan. Since January 1992, he has been Professor of electronic control in the Department of Electrical Engineering, Hiroshima University, Hiroshima, Japan. His research interests include digital signal Processing, system theory, and control engineering. He has published more than 270 papers in these areas and is the co-editor and co-author of *Two-Dimensional Signal and Image Processing* (Tokyo, Japan: SICE, 1996).

Dr. Hinamoto served as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II from 1993 to 1995, and presently serves as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I. He also served as Chair of the 12th Digital Signal Processing (DSP) Symposium held in Hiroshima in November 1997, sponsored by the DSP Technical Committee of IEICE. He was the Guest Editor of the special section of DSP in the August 1998 issue of the *IEICE Transactions on Fundamentals*. Since 1995, he has been a member of the steering committee of the IEEE Midwest Symposium on Circuits and Systems, and since 1998, a member of the Digital Signal Processing Technical Committee in the IEEE Circuits and Systems Society. He served as a member of the Technical Program Committee for ISCAS'99. From 1993 to 2000, he served as a senator or member of the Board of Directors in the Society of Instrument and Control Engineers (SICE), and from 1999 to 2001, he was Chair of the Chugoku Chapter of SICE. He played a leading role in establishing the Hiroshima Section of IEEE, and served as the Interim Chair of the section. He is a recipient of the IEEE Third Millennium Medal.