On the Design of High-Complexity Cosine-Modulated Transmultiplexers Based on the Frequency-Response Masking Approach

Miguel B. Furtado, Jr., *Student Member, IEEE*, Paulo S. R. Diniz, *Fellow, IEEE*, Sergio L. Netto, *Senior Member, IEEE*, and Tapio Saramäki, *Fellow, IEEE*

***Abstract*—Two efficient techniques exploiting the frequency-response masking (FRM) approach are proposed in order to make it feasible to design prototype filters for highly selective nearly perfect-reconstruction cosine-modulated transmultiplexers and filter banks (CMTs and CMFBs) having a very large number of channels. In these design schemes, the number of unknowns is drastically reduced when compared with the corresponding techniques for designing direct-form prototype filters. Furthermore, in the proposed techniques, the main figures of merits, that is, the intersymbol interference and the interchannel interference for CMTs and the overall and aliasing distortions for CMFBs are taken into account in a controlled manner. In order to speed up the convergence of these two optimization techniques, simplifications for computing the resulting nonlinear constraints and the corresponding gradient vectors are proposed. They differ from each other in the sense that the first and second ones utilize the frequency-domain and time-domain constraints for controlling the figures of merit, respectively. Combining these two techniques results in numerically efficient algorithms for designing optimized CMTs (or CMFBs) with a reduced computational complexity (number of arithmetic operations per output sample), particularly when both branches of the FRM structure are required. Design examples are included illustrating the efficiency of the design methods and the high performance of the resulting CMT structures.**

***Index Terms*—Cosine-modulation, filter banks, frequency-response masking (FRM), optimization, transmultiplexers.**

I. Introduction

During the past few years, the area of digital communications has experienced a growing interest on the research of filter banks, due to their inherent subband processing capability. Such structures, when applied to multiple input/output digital communications systems, are known as transmultiplexers (TMUXes) [3]. Currently, the most widely used TMUX structure is the orthogonal frequency division multiplexing (OFDM) [21]. The OFDM system features a low computational complexity and an easy equalization process, when using a proper cyclic prefix. However, such a prefix may greatly reduce the OFDM payload, thereby motivating the research for generating alternative TMUX structures such as the cosine-modulated transmultiplexer (CMT) [20], [14]. In practice, modern multicarrier communications systems may require subchannels with a very small amount of overlap, known in the literature as roll-off, and possibly a large number of subbands such that each subchannel has flat response, allowing a simple equalization process. Such a demand for small roll-off and/or large number of subbands turns the design of CMT prototype filters very difficult, because the number of coefficients involved in the optimization process tends to be prohibitively large.

This paper proposes efficient design procedures for CMTs, enabling the design of high-order optimized prototype filters meeting demanding prescribed specifications. For this purpose, two optimization techniques are proposed, exploiting the frequency-response masking (FRM) approach [12] for optimizing such prototype filters. In addition, some analytical simplifications for the computation of the constraints and corresponding gradient vectors are provided. These crucial simplifications are employed by the subsequent optimization procedure aiming at the minimization of the associated intersymbol and interchannel interferences (ISI and ICI) for CMTs or the overall and aliasing distortions for the corresponding cosine-modulated filter banks (CMFBs). A particular case where the ISI is inherently zero is also presented. The first and the second proposed optimization techniques use the frequency-domain and the time-domain constraints, respectively, for properly controlling the above-mentioned figures for CMTs and CMFBs.

The proper combination of these two optimization schemes results in a numerically efficient algorithm for designing optimized CMTs or CMFBs with reduced computational complexity (number of arithmetic operations per output sample), particularly when both branches of the FRM structure are employed. The method is suitable for cases where the prototype filter length extends to some thousands of coefficients, enabling one to design extremely selective multirate systems with a very large number of subbands, even more than one thousand. Furthermore, the designed prototype filter can be used not only with CMFB, but with SMFB [2] and MDFT [11] filter banks as well.

Optimizing such systems using direct-form (DF) prototype filters would take an extremely long time and would require a huge amount of memory resources in order to finalize the task. The first proposed building-block optimization technique using

Manuscript received December 8, 2004; revised March 7, 2005. This work was supported by the Academy of Finland, Finnish Centre of Excellence Program (2000–2005), under Project 44876.This paper was recommended by Associate Editor M. Chakraborty.

M. B. Furtado, P. S. R. Diniz, and S. L. Netto are with the Electrical Engineering Program, COPPE/Federal University of Rio de Janeiro, Rio de Janeiro, RJ 21941-972, Brazil (e-mail: furtado@lps.ufrj.br; diniz@lps.ufrj.br; sergioln@lps.ufrj.br).

T. Saramäki is with the Institute of Signal Processing, Tampere University of Technology, FIN-33101 Tampere, Finland (e-mail: ts@cs.tut.fi).

Digital Object Identifier 10.1109/TCSI.2005.853919

frequency domain constraints can be regarded as an efficient extension of the optimization technique described in [7] in the sense that instead of using the DF prototype filter, the FRM approach is utilized for generating it. The second optimization scheme, in turn, is an extension of the technique described in [10]. The original synthesis scheme has been aimed at generating CMTs with zero ISI or CMFBs with no overall distortion. In the modified version, these figures of merit are allowed to get any predetermined values.

This paper is organized as follows. Sections II and III review the general basic concepts and ideas behind CMTs and the FRM approach, respectively. Furthermore, Section II shows how the evaluation of the figures of merit for both CMTs and CMFBs can be significantly reduced based on the use of the techniques described in [7]. These reductions will be utilized in the following sections when developing the aforementioned numerically efficient algorithm for designing optimized CMTs or CMFBs with reduced computational complexity, especially for designing multirate systems with a very small roll-off factor and a huge number of channels. In addition to reviewing the basic concepts of the FRM approach for constructing the prototype filters for multirate systems, Section III considers also how to properly select the parameters of the FRM-based prototype filter in order to arrive at satisfactory overall multirate systems. Section IV states the optimization problems for both CMTs and CMFBs in order to determine the optimal FRM-based prototype filter in the least-mean-square and minimax senses subject to the given allowable figures of merit. This section also introduces the objective function and presents closed-form expressions for its first-order derivatives to accelerate the optimization procedure, and considers a few alternatives to solve the resulting problem in a convenient manner. Section V formulates the figures of merit in terms of both the time- and frequency-domain constraints and presents their respective closed-form derivatives. Section VI introduces an efficient procedure, based on the results of Section V, for the design of FRM-based CMTs. Section VII presents some practical CMT design examples, illustrating the numerical efficiency of the proposed optimization procedure and the high performance of the resulting CMT structure. Finally, Section VIII summarizes the main results presented in this paper.

II. Cosine-Modulated Transmultiplexers

CMTs are a well-known class of multirate systems, where all the analysis and synthesis subfilters can be easily generated by just modulating a low-pass prototype filter with the transfer function as given by

$$H_p(z) = \sum_{n=0}^{N_p} h_p(n) z^{-n} \tag{1}$$

as follows [14], [20]:

$$\begin{aligned} H_m(z) &= \sum_{n=0}^{N_p} h_p(n) c_{m,n} z^{-n} \\ F_m(z) &= \sum_{n=0}^{N_p} h_p(n) \tilde{c}_{m,n} z^{-n} \end{aligned} \tag{2}$$

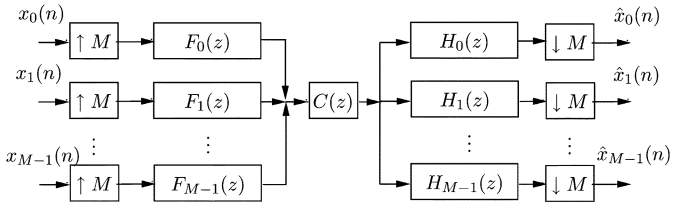


Fig. 1. M-channel maximally decimated TMUX system.

for $m = 0, 1, \ldots, (M-1)$ and for $n = 0, 1, \ldots, N_p$, where

$$\begin{aligned} c_{m,n} &= 2\cos[(2m+1)(n - N_p/2)\pi/(2M) + (-1)^m \pi/4] \\ \tilde{c}_{m,n} &= 2\cos[(2m+1)(n - N_p/2)\pi/(2M) - (-1)^m \pi/4]. \end{aligned} \tag{3}$$

This modulated filter bank possesses the property that all filtering operations in both the analysis and synthesis banks can be efficiently implemented with a reduced number of operations per output sample [20].

The maximally decimated M-channel CMT system is depicted in Fig. 1. At the transmitter (synthesis bank) of this structure, each input signal is first interpolated and filtered and, then, the resulting signals on each branch are added to form a single signal for transmission over a given channel $C(z)$. At the receiver (analysis bank), the output signal of the channel is split, with the aid of filtering and decimation, back into M-channels in order to generate the desired M output signals. In the sequel, in order to simplify the design of CMT systems, it is assumed that the channel response is unitary $(C(z) \equiv 1)$ or a pure delay.

By taking into account the aforementioned simplification, the general relation that describes the transfer functions of the TMUX system is

$$\hat{\boldsymbol{x}}(z^M) = \frac{1}{M} \boldsymbol{T}(z^M) \boldsymbol{x}(z^M) \tag{4}$$

where

$$\hat{\boldsymbol{x}}(z) = [\hat{X}_0(z) \quad \hat{X}_1(z) \quad \cdots \quad \hat{X}_{M-1}(z)]^T \tag{5}$$

$$\boldsymbol{x}(z) = [X_0(z) \quad X_1(z) \quad \cdots \quad X_{M-1}(z)]^T \tag{6}$$

$$[\boldsymbol{T}(z^M)]_{ab} = \sum_{i=0}^{M-1} H_a\left(zW_i^{-1}\right) F_b\left(zW_i^{-1}\right) \tag{7}$$

for $a, b = 0, 1, \ldots, (M-1)$, and $W_i = e^{((j2\pi i)/M)}$. It is worth mentioning that the letter j is used to define the imaginary part of a complex number, as opposed to i, which is just used for indexing. The matrix $\boldsymbol{T}(z^M)$ is the so-called transfer matrix whose elements, $[\boldsymbol{T}(z^M)]_{ab}$, represent the transfer functions between the interpolated input a and the interpolated (undecimated) output b.

In a TMUX system, one would be interested in estimating the total ISI and ICI figures of merit, which are given by [1]

$$\text{ISI} = \max_a \left\{ \sum_n [\delta(n-d) - t_{aa}(n)]^2 \right\} \tag{8}$$

$$\text{ICI} = \max_{a,\omega} \left\{ \sum_{b=0, b\neq a}^{M-1} |[\boldsymbol{T}(e^{j\omega})]_{ab}|^2 \right\} \tag{9}$$

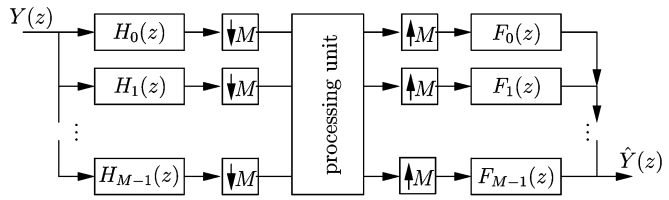


Fig. 2. M-channel maximally decimated filter bank.

where $\delta(n)$ is the ideal impulse, d is a properly chosen delay, $t_{aa}(n)$ is the impulse response for the ath subchannel, and the term $[\boldsymbol{T}(e^{j\omega})]_{ab}$, as given by (7), is the crosstalk between the ath and bth subchannels.

The maximally decimated M-channel TMUX system is a filter bank where the analysis and synthesis blocks are switched in order to form a system with M input/output channels [20], [6]. The reciprocal is true, in such a manner that the filter bank can be built by first omitting the transmission channel $C(z)$ in Fig. 1, and switching both analysis and synthesis banks, resulting in the M-channel maximally decimated filter bank structure of Fig. 2. Consequently, the figures of merit of a filter bank are closely related to those of the corresponding TMUX. Therefore, when designing a TMUX, it is enough to consider input-output relation of the corresponding filter bank as given by

$$\hat{Y}(z) = \frac{1}{M}\left[T_0(z)Y(z) + \sum_{i=1}^{M-1} T_i(z)Y(zW_i)\right]. \qquad (10)$$

The first term in (10), $T_0(z)$, is the overall transfer function and must be the unique term in an alias-free design, which includes the perfect reconstruction (PR) filter bank as a particular case. The second term, involving the remaining $T_i(z)$, includes the aliasing transfer functions, which quantify the influences in a given band from all other bands. These terms are expressible as

$$T_0(z) = \sum_{m=0}^{M-1} F_m(z)H_m(z) \qquad (11)$$

$$T_i(z) = \sum_{m=0}^{M-1} F_m(z)H_m\left(zW_i^{-1}\right) \qquad (12)$$

for $i = 1, 2, \ldots, M-1$.

The aforementioned figures of merit, namely the overall and aliasing distortions, are always used when designing maximally decimated filter banks with constrained distortions, and are of high computational complexity. However, the evaluation of these figures of merit can be significantly simplified as described in [7], where it was shown that the functions $T_i(z)$ can be evaluated convolving the prototype filter with its complex modulated version as follows:

$$T_i(z) = \mathcal{Z}\left\{\left(W_i^n h_p(n) * h_p(n)\right)\gamma(n)\right\} \qquad (13)$$

for $i = 0, \ldots, (M-1)$, where

$$\gamma(n) = \begin{cases} 0, & (N_p - n) \neq 2Mc \\ 2M(-1)^c, & (N_p - n) = 2Mc, \end{cases} \qquad c \in \mathbb{Z}. \qquad (14)$$

Due to the symmetry $T_i(z) = T_{M-i}(z)$, the expressions for the functions $T_i(z)$ may be determined only for $i = 0, \ldots, \lfloor M/2 \rfloor$, where $\lfloor x \rfloor$ denotes the integer part of x. If the prototype filter is symmetric (linear-phase) and is of order $N_p = (2KM - 1)$, then the functions $T_i(z)$ can be written as

$$T_i(z) = z^{-N_p} 2M \left[a_i(N_p) + \sum_{l=1}^{K-1} a_i(N_p - 2Ml)(-1)^l (z^{2Ml} + z^{-2Ml})\right] \qquad (15)$$

where the coefficients $a_i(n)$ are defined according to the following equation

$$\mathcal{Z}\left\{W_i^n h_p(n) * h_p(n)\right\} = \sum_{n=0}^{2N_p} a_i(n) z^{-n} \qquad (16)$$

that is, the coefficients $a_i(n)$ can be generated by convolving the impulse response $h_p(n)$ with its complex exponentially modulated version $W_i^n h_p(n)$.

Each impulse response $t_i(n) = \mathcal{Z}^{-1}\{T_i(z)\}$, for $i = 0, \ldots, \lfloor M/2 \rfloor$, is evaluated as follows: First, for a given prototype filter $h_p(n)$ of order N_p, set $L = \lceil (2N_p - 1)/M \rceil M$, where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x, and evaluate the L-point DFT of $h_p(n)$ padded with zeros, which will be called $H_p(k)$. Second, evaluate $H_p^i(k) = H_p(k - L/Mi)$ by just right circular shifting the components of $H_p(k)$. Third, the coefficients $a_i(n)$ result from the inverse DFT of the product $H_p(k)H_p^i(k)$. Finally, what remains is to multiply $a_i(n)$ by $\gamma(n)$, resulting in the desired $t_i(n)$.

TABLE I
COMPUTATIONAL COMPLEXITY FOR THE STANDARD AND SIMPLIFIED FORMULATIONS

stage	standard	simplified
CM	$4KM$	-
PM	$KM(2KM+1) \times M$	$KM(K+1)$
Total	$2KM^2(2KM+1+\frac{3}{M})$	$2KM(K+1)$

Table I shows the number of floating-point multiplications associated with (12) and (15), which represent the standard and the simplified formulations, respectively. It was assumed that the prototype filter is of order $N_p = (2KM - 1)$ and possesses the linear phase property, facts which will be considered throughout this paper and must be emphasized. The entries CM and PM stand for cosine modulation and polynomial multiplication, respectively. Clearly, if $K \gg 1$ and $M \gg 1$, then the total number of multiplications approximate $4K^2M^3$ and $2K^2M$ for the standard and simplified formulations, respectively, leading to a reduction factor of $r = (1/(2M^2))$ in favor of the simplified formulations, which were detailed in [7].

Using the facts that the computational gain for evaluating the functions $T_i(z)$ is specially impressive when the number of bands M is high, and the usual requirement of CMTs with a large M, which brings high selectivity for multicarrier transmissions, this paper makes use of the simplified formulations stated before when designing both CMFBs and CMTs.

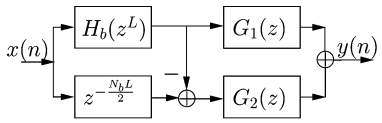


Fig. 3. FRM structure.

III. Cosine-Modulated Transmultiplexers With Frequency-Response Masking

The FRM approach uses a complementary pair of interpolated linear-phase FIR filters. The base filter $H_b(z)$ with group-delay $((N_b)/2)$, and its complementary version $z^{-N_bL/2} - H_b(z^L)$ are interpolated by a factor L, to form sharp transition bands at the cost of introducing multiple images (passbands) in both responses. These repetitive bands are then filtered out by the so-called positive and negative masking filters, denoted by $G_1(z)$ and $G_2(z)$, respectively, and added together to compose the desired overall transfer function as given by [12]

$$H_f(z) = H_b(z^L)G_1(z) + (z^{-N_bL/2} - H_b(z^L))G_2(z) \quad (17)$$

which will be used as the prototype filter $H_p(z)$ in the proposed CMTs and CMFBs. This is because the aim of this contribution is to design filters with sharp transition band and wide bandwidth and/or a large number of subchannels. In such cases, first, the use of the above prototype transfer function when properly constructed in a manner to be described later, instead of its direct-form counterpart, results in implementations of CMTs and CMFBs with a considerably reduced arithmetic complexity at the expense of a slightly reduced performance of the overall filter bank, as shown in [8]. Second, most importantly, the FRM structure enables one to use very fast optimization techniques, requiring just a small amount of memory for designing CMTs and CMFBs having the aforementioned requirements, as will be seen later in this paper. In extreme cases with very large number of channels and a very narrow transition band, it is either impossible to optimize the DF prototype filter or designing it takes an enormous amount of time. It is worth emphasizing that if the bandwidth is required to be narrow, the lower branch, where the polynomial multiplication by $G_2(z)$ takes place, is discarded. Fig. 3 depicts the general FRM structure for generating the prototype filter for both CMTs and CMFBs.

As shown in [4], an efficient FRM-CMT structure or, equivalently, an FRM-CMFB structure can be derived provided that the FRM interpolation factor is expressible as

$$L = 2K_aM + M/K_b \quad (18)$$

where $K_a \geq 0$ and $K_b > 0$ are integers. In such a case, using solely the upper branch in the FRM scheme, the mth analysis filter can be written as

$$H_m(z) = \sum_{q=0}^{Q-1} \left[z^{-Lq} H'_{b1q}(-z^{LQ}) \times \sum_{i=0}^{2M-1} c_{m,(n+\frac{M}{K_b}q)} z^{-i} E'_i(-z^{2M}) \right] \quad (19)$$

where

$$H'_{b1q}(z) = \sum_{k=0}^{K_c-1} (-1)^{K_aq} h_b(kQ+q) z^{-k} \quad (20)$$

for $q = 0, 1, \ldots, (Q-1)$. Here $Q = 2K_b$ is the number of polyphase components for the FRM base filter

$$E'_i(z) = \sum_{k=0}^{K_d-1} g_1(2kM+i) z^{-k} \quad (21)$$

for $i = 0, 1, \ldots, (2M-1)$, where $g_1(n)$ is the impulse response of the masking filter of order N_{g_1}, and K_c is related to N_b and Q as $(N_b+1) = QK_c$. Equation (19) leads to the so-called FRM-CMT efficient structure described in [4], which was developed not only for the case where the upper branch of the FRM structure is used, but also for the case where both branches are employed. The details of the structure are outside the scope of this paper. Nevertheless, the FRM-based prototype filter can be mapped into a DF FIR prototype filter, in such a manner that the traditional fast implementation of CMTs is explored [20], [2]. In this case, there is an additional gain in the computational complexity reduction during the optimization task, due to a reduced number of unknowns to be optimized.

Depending on the application at hand, TMUXes with large number of subchannels are necessary, thereby demanding the optimization of very long CMT prototype filters. Such optimization problems may not even be feasible due to the computational complexity involved. Another issue is the need for a high computational storage capacity since efficient optimization algorithms, as the sequential quadratic programming (SQP) [13] used in this paper, require the computation of a Hessian matrix (its inverse or an approximation to it), whose size is about the square of the number of parameters being optimized. To overcome these problems, filter banks designed with the FRM [12] approach were proposed in [4], [9], and [17], resulting in a remarkable reduction in the number of parameters to be optimized. In these works, the price paid for using an FRM prototype filter with a least squares (LS) (or minimax) characteristic is a slightly higher stopband energy (or stopband ripple) when compared to the DF realization, whenever this alternative is feasible, for the same overall filter order. This is the motivation to employ FRM-based prototype filters in this paper.

IV. Optimal Design of FRM-CMTS: Objective Function

Standard optimization goals for the CMT prototype filter are to minimize the objective functions

$$E_2 = \int_{\omega_r}^{\pi} |H_p(e^{j\omega})|^2 \, d\omega$$
$$E_\infty = \max_{\omega \in [\omega_r, \pi]} |H_p(e^{j\omega})| \quad (22)$$

which correspond to the total energy and the maximum magnitude value in the filter's stopband, respectively. Moreover, the 3-dB attenuation frequency and the stopband edge are given by [20]

$$\omega_{3\text{ dB}} \approx \frac{\pi}{2M}$$
$$\omega_r = \frac{(1+\rho)\pi}{2M} \quad (23)$$

respectively, where ρ controls the amount of overlap between adjacent bands.

The optimization of the CMT must be conveniently carried out by keeping the aliasing distortions and the overall distortion of the corresponding CMFB under control. This is because

these distortions can be evaluated in a simplified manner and are the figures of merit directly related to the ICI and ISI, respectively, as explained in Section II. Consequently, the following constraints are introduced:

$$1-\delta_1 \le |T_0(e^{j\omega})| \le 1+\delta_1 \tag{24}$$

$$|T_i(e^{j\omega})| \le \delta_2, \qquad \text{for } \omega \in [0,\pi] \tag{25}$$

for $i = 1, 2, \ldots, (M-1)$.

Traditional filter design algorithms [5] are not suitable to solve the general CMT optimization problem due to the required nonlinear constraints. In such cases, the problem of finding an optimized solution can be solved with a modified objective function F, that combines the original objective function $E_{\mathfrak{p}}$ [$\mathfrak{p} = 2, \infty$, as in (22)] with a weighted set of constraints, such that

$$F = E_{\mathfrak{p}} + \boldsymbol{\lambda}^T \boldsymbol{c} \tag{26}$$

where $\boldsymbol{\lambda}$ is the vector of constraint weights, and $\boldsymbol{c}$ is the vector of constraints, that is

$$\boldsymbol{\lambda} = [\lambda_0 \quad \lambda_1 \quad \cdots \quad \lambda_{N_c}]^T \tag{27}$$

$$\boldsymbol{c} = [c_0 \quad c_1 \quad \cdots \quad c_{N_c}]^T \tag{28}$$

where N_c is the overall number of constraints.

The minimization of $E_{\mathfrak{p}}$ in (26) can be performed using the WLS-Chebyshev method for the design of CMTs, as described in [7]. The WLS-Chebyshev method is specially suitable for the design of filters with the minimax criterion, and is based on successive approximations of the L_∞ norm by means of weighted L_2 optimizations. Alternatively, the second algorithm of Dutta and Vidyasagar (DV) can be employed in a similar manner as presented in [19]. The performance is very similar in both designs, but the WLS-Chebyshev algorithm usually converges faster, in approximately half the amount of time required by the DV algorithm, as observed in our simulations, the reason it will be employed in this paper.

The optimization is carried out by quadratic programming (QP) algorithms [13], which whenever possible employ the closed-form first and second derivatives of F to simplify its implementation and improve the performance. A second-order cone programming approach is another choice to perform the optimization [22]. The modified objective function F is optimized by minimizing $E_{\mathfrak{p}}$ and satisfying predetermined constraints, as in (26). With the QP algorithms, the constraint weights are set by the designer. Another alternative is to use a sequential quadratic programming (SQP) algorithm, which optimally sets the weights of the constraints based on the method of Lagrange multipliers with the Kuhn-Tucker conditions [13].

In the WLS-Chebyshev design, the minimization of $E_{\mathfrak{p}}$ is performed by successively minimizing WLS-like objective functions, given in each iteration by

$$F = \sum_{\omega'_j \in [\omega_s, \pi]} W_k^2(\omega_{j'}) f_{j'}^2 \Delta_\omega + \boldsymbol{\lambda}^T \boldsymbol{c} \tag{29}$$

with

$$f_{j'} = |H_p(e^{j\omega_{j'}})|, \qquad j' = 1, 2, \ldots, J' \tag{30}$$

which is evaluated over a discrete set of frequencies $\omega_{j'}$. In (29), $W_k(\omega_{j'})$ is the weighting function at the kth iteration and Δ_ω is the given frequency grid interval. Further information regarding the updating procedure of the weighting function and the algorithm details can be found in [7].

Now, since the objective function $E_{\mathfrak{p}}$ can be written as a sum of squares, it is straightforward to derive its first-order derivatives with respect to the FRM subfilters. Using (29) enables one to write

$$E_{\mathfrak{p}} = \sum_{\omega'_j \in [\omega_s, \pi]} W_k^2(\omega_{j'}) f_{j'}^2 \Delta_\omega \tag{31}$$

$$\frac{\partial E_{\mathfrak{p}}}{\partial h(n)} = 2 \sum_{\omega'_j \in [\omega_s, \pi]} W_k^2(\omega_{j'}) f_{j'} \Delta_\omega \frac{\partial f_{j'}}{\partial h(n)} \tag{32}$$

where $h(n)$ may be $h_b(n), g_1(n)$, or $g_2(n)$, and

$$\begin{aligned}
\frac{\partial f_{j'}}{\partial h_b(n)} &= 2c_1 \text{sign}\left(A_{H_b}(\omega_{j'}L)\right) \\
&\quad \times \cos\left[(n - N_b/2)\omega_{j'}L\right] \left|A_{G_1}(\omega_{j'}) - A_{G_2}(\omega_{j'})\right|, \\
&\quad 0 \le n \le \left\lfloor \frac{N_b}{2} \right\rfloor \\
\frac{\partial f_{j'}}{\partial g_1(n)} &= 2c_2 \text{sign}\left(A_{G_1}(\omega_{j'})\right) \\
&\quad \times \cos\left[(n - N_g/2)\omega_{j'}\right] \left|A_{H_b}(\omega_{j'}L)\right|, \\
&\quad 0 \le n \le \left\lfloor \frac{N_g - 1}{2} \right\rfloor \\
\frac{\partial f_{j'}}{\partial g_2(n)} &= 2c_2 \text{sign}\left(A_{G_2}(\omega_{j'})\right) \\
&\quad \times \cos\left[(n - N_g/2)\omega_{j'}\right] \left|1 - A_{H_b}(\omega_{j'}L)\right|, \\
&\quad 0 \le n \le \left\lfloor \frac{N_g - 1}{2} \right\rfloor
\end{aligned} \tag{33}$$

where

$$\text{sign}(x) = \begin{cases} 1, & \forall x \ge 0 \\ -1, & \forall x < 0 \end{cases} \tag{34}$$

$$c_1 = \begin{cases} 1/2, & \text{if } n = N_b/2 \quad (N_b \text{ even}) \\ 1, & \text{otherwise} \end{cases}$$

$$c_2 = \begin{cases} 1/2, & \text{if } n = N_g/2 \quad (N_g \text{ even}) \\ 1, & \text{otherwise} \end{cases} \tag{35}$$

with $N_{g_1} = N_{g_2} = N_g$. Furthermore, the description of the frequency response of the linear-phase filters $H_b(z^L), G_1(z)$, and $G_2(z)$ in terms of their phase and zero-phase (amplitude) responses are

$$\begin{aligned}
H_b(e^{j\omega L}) &= e^{-j\omega N_b L/2} A_{H_b(\omega L)} \\
G_1(e^{j\omega}) &= e^{-j\omega N_g/2} A_{G_1(\omega)} \\
G_2(e^{j\omega}) &= e^{-j\omega N_g/2} A_{G_2(\omega)}.
\end{aligned} \tag{36}$$

With (31) and (32), along with the auxiliary (33), (34), (35), and (36), one can implement a fast routine for accurately evaluating the objective function and its derivatives, which will lead to a fast convergence for the optimization process.

V. Optimal Design of FRM-CMTs: Constraint Evaluation

This section considers how to formulate the constraints involved in the objective function given previously so that the resulting optimization of FRM-CMTs becomes computationally efficient. As shown in this section, this goal is achieved by defining these constraints either in the time or frequency

domains. Depending on the problem at hand, one of the following four alternative combinations can be explored, namely, frequency-domain (FD) inequality or equality constraints, or time-domain (TD) inequality or equality constraints. In the dedicated literature, FD inequality constraints are most frequently used [8], [19], [20]. TD equality constraints were apparently first employed in [15], and subsequently in [10], with a different treatment. To the best of our knowledge, TD inequality constraints were not treated so far in the literature, but are exploited in this paper in an efficient manner.

A. FD Constraints for the FRM-CMT Design

Both the overall distortion and the aliasing distortions of the overall filter bank can be controlled in the frequency domain by stating the following constraints

$$\hat{c}_0(\omega_j) = \left|\left|T_0(e^{j\omega_j})\right| - 1\right| - \delta_1 \leq 0 \tag{37}$$

$$\hat{c}_i(\omega_j) = \left|T_i(e^{j\omega_j})\right| - \delta_2 \leq 0 \tag{38}$$

for $i = 1, 2, \ldots, \lfloor M/2 \rfloor$ and $\omega_j \in [0, \pi/M]$ for $j = 1, 2, \ldots, J$. Selecting all grid points ω_j to be within $[0, \pi/M]$ is due to the fact that the functions $T_i(e^{j\omega})$ are periodic with the periodicity equal to π/M, as will be seen later. The overall number of the resulting constraints is $J\lfloor M/2 \rfloor$, depending on J, the number of grid points in use. Having a small (large) value for J results in a poor (good) control on the actual overall and aliasing errors and in an optimization algorithm with a low (high) complexity. It has been observed in [8] that selecting $J = 40K$ provides a proper compromise, leading to $40K\lfloor M/2 \rfloor$ constraints.

What is left is to convert the aforementioned $40K\lfloor M/2 \rfloor$ constraints of (37) and (38) into the constraints in the modified objective function, as given by (28). To do this in an efficient and convenient manner, it is beneficial to first express the constraints in the following forms:

$$c_i = \sum_{\omega_j \in [0, \frac{\pi}{M}] \setminus \hat{c}_i > 0} \hat{c}_i^2(\omega_j). \tag{39}$$

In this case, only those constraints $\hat{c}_i$ being greater than zero are taken into account, thereby reducing the computational complexity. Second, the computational burden of evaluating the above equation can be significantly reduced by exploiting (15). This equation enables one to write all functions $T_i(e^{j\omega})$ in the following computationally efficient form

$$T_i(e^{j\omega}) = e^{-j(\omega N_p + \frac{\pi i}{M})} A_{T_i}(\omega) \tag{40}$$

where

$$A_{T_i}(\omega) = 2Me^{j\frac{\pi i}{M}} \left[a_i(N_p) + 2\sum_{l=1}^{K-1} a_i(N_p - 2Ml) \cos\left[(2M\omega + \pi)l\right] \right]. \tag{41}$$

On the right-hand side (RHS) of (41), the variable ω is multiplied by $2M$, making the trigonometric function $A_{T_i}(\omega)$ periodic with the periodicity equal to π/M.

In the FRM-CMT structure, the prototype filter $H_p(z)$ is implemented using $H_f(z)$ as given by (17), and the optimization problem entails on finding a base filter, a positive masking filter (upper branch), and a negative masking filter (lower branch) that optimize F. In such a case, $T_i(z)$ must consider the entire FRM structure given by (17). In the TD, the coefficients $a_i(n)$ in (15) are given by

$$\begin{aligned} a_i(N_p - 2Mk) &= \sum_{\tau=0}^{N_p - 2Mk} e^{\frac{j2\pi i\tau}{M}} h_p(N_p - 2Mk - \tau) h_p(\tau) \\ &= \sum_{\tau=0}^{N_p - 2Mk} e^{\frac{j2\pi i\tau}{M}} h_p(2Mk + \tau) h_p(\tau) \end{aligned} \tag{42}$$

where the prototype filter coefficients $h_p(n)$ are expressed as

$$\begin{aligned} h_p(n) = \sum_{\nu=0}^{\lfloor n/L \rfloor} &(h_b(\nu) g_1(n - L\nu) \\ &- h_b(\nu) g_2(n - L\nu)) + g_2(n - N_b L/2). \end{aligned} \tag{43}$$

Based on (39), in order to determine the constraint derivatives, one can write

$$\frac{\partial c_i}{\partial h_p(n')} = 2 \sum_{\omega_j \in [0, \frac{\pi}{M}] \setminus \hat{c}_i > 0} \hat{c}_i(\omega_j) \frac{\partial \hat{c}_i(\omega_j)}{\partial h_p(n')} \tag{44}$$

where

$$\frac{\partial \hat{c}_i(\omega_j)}{\partial h_p(n')} = \begin{cases} \operatorname{sign}\left(\hat{c}_0(\omega_j) A_{T_0(\omega_j)}\right) \frac{\partial A_{T_0}(\omega_j)}{\partial h_p(n')}, & i = 0 \\ \operatorname{sign}(A_{T_i}(\omega_j)) \frac{\partial A_{T_i}(\omega_j)}{\partial h_p(n')}, & i \neq 0 \end{cases} \tag{45}$$

with

$$\begin{aligned} &\frac{\partial A_{T_i}(\omega_j)}{\partial h_p(n')} \\ &\quad = 8M \cos \frac{\pi i(2n' + 1)}{M} \\ &\qquad \times \left[h_p(n') + \sum_{l=1}^{K-1} [h_p(n' + 2Ml) + h_p(n' - 2Ml)] \right. \\ &\qquad \left. \times \cos\left(2M\omega_j + \pi\right)l) \right]. \end{aligned} \tag{46}$$

Here, $h_p(n') = 0$ for $n' < 0$ and $n' > N_p$.

Finally, as the proposed technique deals with the FRM structure, the constraint derivatives with respect to the coefficients of the subfilters must be analytically formulated, that is

$$\frac{\partial c_i}{\partial h(n)} = \sum_{n'=0}^{N_p} \frac{\partial c_i}{\partial h_p(n')} \frac{\partial h_p(n')}{\partial h(n)} \tag{47}$$

where $h(n)$ could be either $h_b(n), g_1(n)$, or $g_2(n)$. Moreover

$$\begin{aligned} \frac{\partial h_p(n')}{\partial h_b(n)} &= c_1[g_1(n' - Ln) + g_1(n' - L(N_b - n)) \\ &\qquad - g_2(n' - Ln) + g_2(n' - L(N_b - n))] \\ \frac{\partial h_p(n')}{\partial g_1(n)} &= c_2 \left[h_b\left(\frac{n' - n}{L}\right) + h_b\left(\frac{n' + n - N_g}{L}\right) \right] \\ \frac{\partial h_p(n')}{\partial g_2(n)} &= -c_2 \left[h_b\left(\frac{n' - n}{L}\right) + h_b\left(\frac{n' + n - N_g}{L}\right) \right] + c_3 \end{aligned} \tag{48}$$

where $h_b(n) = 0$ for noninteger values of n as well as for $n < 0$ and $n > N_b$, whereas $g_1(n) = 0$ and $g_2(n) = 0$ for $n < 0$ and

$n > N_g$ in the $N_g = N_{g_1} = N_{g_2}$ case. The constants c_1 and c_2 are given by (35), whereas c_3 is given by

$$c_3 = \begin{cases} 1, & \text{if } (n'-n) = N_b L/2 \text{ or } (n'-n) = N_b L/2 + N_g \\ 0, & \text{otherwise} \end{cases}$$

for $n' = 0, 1, \ldots, (N_p - 1)/2$. Here, it is assumed, without much loss of generality, that the impulse-response coefficients of the prototype filter satisfies the following symmetry condition: $h_p(n') = h_p(N_p - n')$.

B. TD Constraints for the FRM-CMT

Nyquist filters are a class of FIR filters whose impulse-response coefficients satisfy the following conditions [16], [18]

$$\begin{aligned} h_n(N) &\neq 0 \\ h_n(N + lM) &= 0, \qquad l = \pm 1, \pm 2, \ldots \end{aligned} \tag{49}$$

for a given integer N and an interpolation factor M. These filters are also known as Mth-band filters and may have a linear phase. For later use, $H_n^{(M)}(z)$ denotes the transfer function of a $2N$th-order Mth-band (Nyquist) filter with the impulse-response coefficients $h_n(n)$ satisfying, in addition to the conditions of (49), $h_n(2N - n) = h_n(n)$, thereby resulting in a linear-phase FIR filter. Such a class of filters is useful, for example, in digital communications where the data symbols are upsampled by M and filtered by $H_n^{(M)}(z)$ [18]. In these cases, there will be no ISI in the received sequence provided that both the transmitter and receiver are synchronized and the channel is of the form $C(z) \equiv cz^{-d}$ with c and z^{-d} being a constant and the overall delay, respectively. Hence, the receiver has only to downsample the received information, thereby enabling one to recover the original sequence in a very straightforward manner. If the sampling period for the upsampled signal is T_s, then the data symbols will be spaced in time by $T = MT_s$. There are many alternatives for designing the Nyquist filter for the channel bandwidth $W > 1/2T$, including both the single-stage and multistage implementations of $H_n^{(M)}(z)$.

It has been verified independently in [9] and [15] that CMFBs with zero overall distortion and, consequently, CMTs with null ISI are directly achieved by designing the prototype transfer function $H_p(z)$ such that it satisfies $[H_p(z)]^2 \equiv H_n^{(2M)}(z)$, that is, the cascade of the prototype filter transfer function with itself is a $2M$th-band filter. Achieving this property in the case of linear-phase FRM-based prototype filter $H_p(z) \equiv H_f(z)$, with $H_f(z)$ as given by (17), can be stated in terms of the proper TD equality constraints. Before stating these constraints, it is worth recalling the following facts. First, the impulse-response coefficients $h_p(n)$ are related in the case where $H_p(z) \equiv H_f(z)$, for a given L, to the impulse-response coefficients of the linear-phase base filter transfer function $H_b(z)$ of order N_b, as well as the two linear-phase masking filter transfer functions $G_1(z)$ and $G_2(z)$ of order N_g, denoted by $h_b(n)$, $g_1(n)$, and $g_2(n)$, respectively, through (43). Second, in this case, the prototype filter order is $N_p = (LN_b + N_g)$, that should be expressible as $N_p = (2KM - 1)$, with K being an integer in order to be able to generate highly selective prototype filters. Under these circumstances, the desired conditions are satisfied by requiring that the impulse-response coefficients $h_n(n)$ of $H_n^{(2M)}(z) = [H_p(z)]^2$, with the resulting order $N_n = 2N_p = 2(2KM - 1)$, satisfy the following properties

$$h_n(N_p - 2Ml) = \begin{cases} 1/(2M), & \text{if } l = 0 \\ 0, & \text{if } l = 1, \ldots, K-1 \end{cases} \tag{50}$$

where $1/(2M)$ is the value of the central impulse-response coefficient, according to the definition of a linear-phase FIR $2M$th-band filter [18], whereas the coefficients $h_n(n)$ and $h_p(n)$ are related as follows:

$$h_n(m) = \sum_{\tau=0}^{m} h_p(m-\tau) h_p(\tau) \tag{51}$$

for $m = 0, 1, \ldots, N_p$, where $h_p(n) = 0$ for $n < 0$ and $n > N_p$.

The idea of imposing TD constraints for ISI-free filters was treated in [18] and [15], where [15] focused on the DF prototype filter for CMFBs and the former developed a cascaded structure to design a general Mth-band filter. The building block for the Nyquist filter treated here is shown in Fig. 4.

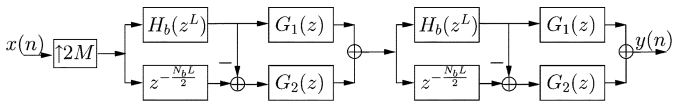


Fig. 4. $2M$th-band Nyquist filter composed by FRM subfilters.

Given the nonlinear equality constraints in (52)[1], or, for conformity with Section V-A, the constraints

$$\begin{aligned} c_l &= h_n(N_p - 2Ml) \\ &= \begin{cases} 1/(2M), & \text{if } l = 0 \\ 0, & \text{if } l = 1, \ldots, K-1 \end{cases} \end{aligned} \tag{52}$$

one can determine the first-order derivatives, required by an SQP-like optimization algorithm. Once more, by considering a generic filter $h(n)$, which can be replaced by $h_b(n)$, $g_1(n)$, or $g_2(n)$, it is possible to write

$$\frac{\partial h_n(N_p - 2Ml)}{\partial h(n)} = \sum_{n=0}^{N_p} \frac{\partial h_n(N_p - 2Ml)}{\partial h_p(n')} \frac{\partial h_p(n')}{\partial h(n)} \tag{53}$$

for $l = 0, \ldots, K-1$, where

$$\frac{\partial h_n(N_p - 2Ml)}{\partial h_p(n)} = 2(h_p(n + 2Ml) + h_p(n - 2Ml)). \tag{54}$$

With the aid of (35), (48), and (49), all needed first-order derivatives are determined.

The direct benefit of using the TD constraints in (52) is that their overall computational complexity depends only on K, which tends to be much smaller than M. The same does not apply for problems considering FD constraints, as in Section V-A, and in [9]. Therefore, when optimizing filter banks with a large number of bands, one should consider using those TD constraints. Furthermore, as this technique leads to TMUXes with zero ISI, the distortion is only dependent on the ICI, which has about the same magnitude as the maximal ripple in the stopband of the magnitude response of the prototype filter [10], [15], [19]. On the other hand, equality constraints may be very restrictive and by allowing a small distortion in the ISI, the respective ICI can be reduced. In

[1]The coefficients of $H_n(z)$ have a quadratic dependence with respect to the FRM-subfilter coefficients.

this manner, instead of employing time-equality constraints as before, one should use time-inequality constraints, by imposing that the magnitude of the coefficients in the TD must be less than a given tolerance, that is

$$c_l = \begin{cases} |h_n(N_p - 2Ml) - 1/(2M)|, & \text{if } l = 0 \\ |h_n(N_p - 2Ml)|, & \text{if } l = 1, \ldots, K-1 \end{cases} \leq \delta. \quad (55)$$

In this case, the value of δ should be selected in such a manner that the overall distortion of the system remains practically the same as in the FD case. The resulting impulse-response coefficients $h_n(N_p - 2Ml)$ may differ from the ideal values, as given by (50), and are, therefore, denoted by $\bar{h}_n(N_p - 2Ml)$. A small increment δ, which can be constrained in a suitable dynamic range, determines the differences between the resulting and ideal coefficient values. In the worst case, the resulting errors $e(l) = \bar{h}_n(N_p - 2Ml) - h_n(N_p - 2Ml)$ for $l = 0, 1, \ldots, (K-1)$ add constructively in such a way that, in the frequency domain, there is a gain of $(2K-1)$. This is due to the fact that the number of the error terms $e(l)$ is $(2K-1)$. Based on this fact, δ should be properly chosen in the following range

$$\frac{\delta_1}{2K-1} \leq \delta \leq \delta_1 \quad (56)$$

where δ_1 is the desired overall distortion, as in (37). Once δ has been appropriately determined, the optimization algorithm can be carried out.

Equation (55) is very suitable to be employed in SQP-based routines, but the same does not apply to the QP-based ones because the nonlinear constraints c_l, for $l = 0, \ldots, (K-1)$, must be explicitly inserted in the modified objective function, as given in (26). This task is accomplished, for example, by writing the modified objective function as

$$F = E_{\mathfrak{p}} + \lambda \sum_{l \in [0, K-1] \setminus c_l > 0} c_l \quad (57)$$

where, in this case, just the active constraints will be considered in each iteration of the optimization process, and λ is a scalar set by the designer to control the relevance of the constraints. In fact, this modified objective function was tested in QP-based algorithms and the attained results were very similar to the ones provided by the SQP-based algorithms, which try to set the constraint weight vector $\boldsymbol{\lambda}$ in an optimal manner. Both the QP and SQP routines tested are available in the optimization toolbox of MATLAB [23] under the names *fminunc* and *fmincon*, respectively.

VI. Proposed Coarse-to-Fine Optimization Procedure

The previous section considered how to use properly both FD and TD constraints for optimizing CMFBs and CMTs with prescribed allowable errors. It has been observed that the use of FD (TD) constraints leads to routines being characterized by a higher (lower) computational complexity and a higher (lower) accuracy of the final result. Therefore, in order to arrive at a routine finding quickly a very accurate solution, it is beneficial to properly combine the benefits of the routines based on the use of the two aforementioned constraints. A good alternative is to use the following coarse-to-fine optimization routine: Given the basic filter bank specifications, namely, the number of bands M, the roll-off factor ρ, the overlapping factor K, one is able to evaluate the order of the resulting prototype as $N_p = (2KM - 1)$. The set of specifications is closed when the maximum allowable aliasing and overall distortions, δ_1 and δ_2, respectively, are taken into account. Given the whole set of specifications, find the optimized solution using the following three-step procedure.

1) Choose between QP and SQP routines, which are suitable for the optimization of the filter bank under consideration. Design an FRM low-pass prototype filter satisfying the set of all specifications but the maximum allowable aliasing and overall distortions.
2) Optimize efficiently the filter bank with TD constraints, as in Section V-B, using an appropriate choice for δ, given the maximum overall distortion δ_1. Depending on the problem under consideration, more than 500 iterations may be required.
3) Use the prototype filter coefficients of Step 2 as the initial solution for this step. The optimization is now carried out in the FD, with the same optimization routine used in Step 2. The overall number of iterations in this step is considerably low. Usually, less than 50 iterations are required.

In the above procedure, Step 2 accounts for the coarse design of the filter bank to attain the prescribed specifications, whereas Step 3 is responsible for the fine adjustment of the resulting prototype filter, in such a manner that the initial specifications are met, and that the objective function is conveniently minimized.

VII. Design Examples

Some design examples are considered in this section to clarify the use of the techniques described in this paper. Comparing the examples included in this section enables one to decide when or where employment of the FD or TD constraints is most appropriate to reduce the amount of computational complexity. Furthermore, in some applications, the choice of one of them may be mandatory. For example, if the synthesis of a TMUX with negligible roll-off and high stopband attenuation is under consideration, then the ICI may be already extremely small, and only the ISI has to be taken into account when performing the optimization. Hence, for conveniently solving such an optimization problem, using the TD constraints is the best choice, since the resulting ISI will automatically have a very small magnitude.

In order to validate the proposed optimization techniques, two examples are presented: in the first example, CMTs with a small roll-off and a small number of bands (relatively large bandwidth) are designed; the second one requires a small roll-off and a large number of subbands, leading to a prototype with very narrow bandwidth. In the first case, both the LS and minimax criteria for designing the prototype filter subject the given constraints are under consideration, whereas in the second case, only the LS criterion is used.

In both examples, the initial solutions were set as FRM filters designed with the technique provided in [12]. Such initial prototype filters usually present very poor figures of merit for a filter bank, that is, both overall and aliasing distortions are considerable.

TABLE II
SETTINGS FOR FRM SUBFILTERS AND OVERALL PROTOTYPE IN EXAMPLE 1

Parameter	Direct-form	FRM
N_b	–	30
N_1	–	199
N_2	–	199
N_p	639	639
L	–	18
$\mathcal{N}$	320 (*)	116
$\mathcal{M}$	$\approx$ 819.2 kB (**)	$\approx$ 107.6 kB

(*) 16 unknowns are necessary if the designed prototype provides the PR, that is, both ISI and ICI are zero.

(**) If the problem is PR, the size of the Hessian matrix reduces to 204.8 kB.

TABLE III
RESULTING FIGURES OF MERIT FOR DIRECT-FORM AND FRM-BASED CMT DESIGNS IN EXAMPLE 1

Figures of Merit	DF		FRM	
	LS	Minimax	LS	Minimax
d_1	0.01	0.01	0.01	0.01
d_2	0.7×10^{-4}	10.0×10^{-4}	0.8×10^{-4}	20.0×10^{-4}
ISI (dB)	-43.0	-43.4	-44.1	-43.0
ICI (dB)	-79.5	-53.4	-75.0	-49.0
E_2	0.2×10^{-7}	3.3×10^{-7}	0.3×10^{-7}	10.0×10^{-7}
E_∞ (dB)	-51.7	-61.7	-47.0	-58.0

The orders of the FRM subfilters $H_b(z), G_1(z)$, and $G_2(z)$, and the interpolation factor L were chosen such that the resulting subfilters had a reasonable reduction in their orders subject to the constraint that the overall order is restricted to be equal to $N_p = (2KM - 1)$. This is the usual restriction when designing DF prototype filters since this order selection leads to filter banks with the highest selectivity.

A. Example 1: Wide-Band Design

In this example, the requirements are set in such a manner that the prototype filter presents a relatively wide bandwidth but a very sharp transition band. Therefore, it is enough to choose a small number of bands and a reduced roll-off factor. The specifications are as follows:

$$\begin{aligned} M &= 8 \text{ (number of bands)} \\ \rho &= 0.1 \text{ (rolloff factor)} \\ K &= 40 \text{ (overlapping factor)} \\ \delta_1 &\leq 0.01 \text{ (maximum overall distortion)} \\ \delta_2 &\leq 0.02 \text{ (maximum aliasing distortion).} \end{aligned}$$

In Table II, N_b is the base filter order, N_1 is the masking filter order, N_p is the overall filter order, and L corresponds to the FRM interpolation factor. In addition, $\mathcal{N}$ indicates the number of parameters being optimized and $\mathcal{M}$ specifies the amount of memory, in bytes (considering 8 bytes to represent a real number), required for the storage of a single Hessian matrix.

Table III presents the figures of merit for both LS and minimax designs. For comparison purposes, the DF CMT was designed, consisting of a simple FIR filter and providing a lower bound on the minimal values attainable for each figure of merit, that is, the DF structure generally leads to the best results. The optimization algorithm used was based on SQP, provided by the function `fmincon` under MATLAB [23], to minimize the LS and minimax objective functions as given by (22). A reasonable number of points in the discrete frequency grid was found to be $20N_p$. Table III clearly shows that both DF and FRM-based prototypes performed similarly, whereas the second optimization procedure required substantially less computational efforts. In Table III, d_1 and d_2, are, respectively, the attained maximum overall and aliasing distortions after optimization.

The strategy adopted in the design was to perform the optimization taking into account only the desired overall distortion δ_1, and checking whether the resulting aliasing distortion d_2 was within the prescribed margin δ_2 or not. If not, the optimization routine must be run again, considering the effects of the aliasing distortion. In other words, the constraints c_i in (39) must be evaluated not only for $i = 0$, but also for $i = 1, \ldots, \lfloor M/2 \rfloor$.

Fig. 5(a) and (b) depicts the magnitude response of the FMR-based prototype filters, optimized according to the LS and Minimax optimality criteria, respectively.

Alternatively, using the same specifications given before, one can use TD inequality constraints to design such a prototype filter. In this case, the ISI will no longer be zero, but the ICI and the stopband energy and attenuation will be reduced. The specifications are the same as in Example 1. Redesigning the prototype filters for the DF and FRM systems using the LS and minimax optimality criteria, we get the results in Table IV. The constraints were closely met with an error of order 10^{-12}.

The results in Table IV approach the ones obtained from the optimization carried out with FD constraints, depicted in Table III, but with a reduced computational burden because there is no need to evaluate the overall distortion over a dense frequency grid. Over many designs performed with a wide set of specifications, the computational time required when using TD constraints was about one third of that required when using FD constraints to satisfy the same specifications, which are usually considered in [8], [15], [19], and [20]. Furthermore, in order to exploit the benefits and eliminate the main drawbacks of both classes of designs, they should be used together, resulting in a kind of coarse-to-fine optimization procedure. In this procedure, it is useful to design first the prototype filter is to meet TD specifications with a reduced computational burden, and then to further optimize the prototype filter under FD constraints to make fine adjustments in the overall distortions.

Fig. 6(a) and (b) shows the magnitude responses of the resulting prototype filters after optimization with TD inequality constraints.

B. Example 2: Narrow-band Design

Now, a very challenging example with a large number of bands is presented. In this case, the proposed method clearly shows one of its advantages against the traditional DF. The specifications to be met are the following:

$$\begin{aligned} M &= 2048 \text{ (number of bands)} \\ \rho &= 0.5 \text{ (rolloff factor)} \\ K &= 8 \text{ (overlapping factor)} \\ \delta_1 &\leq 0.01 \text{(maximum overall distortion)} \\ \delta_2 &\leq 0.01 \text{(maximum aliasing distortion).} \end{aligned}$$

The specifications are constructed such that the prototype filter has both a narrow bandwidth and transition band. This

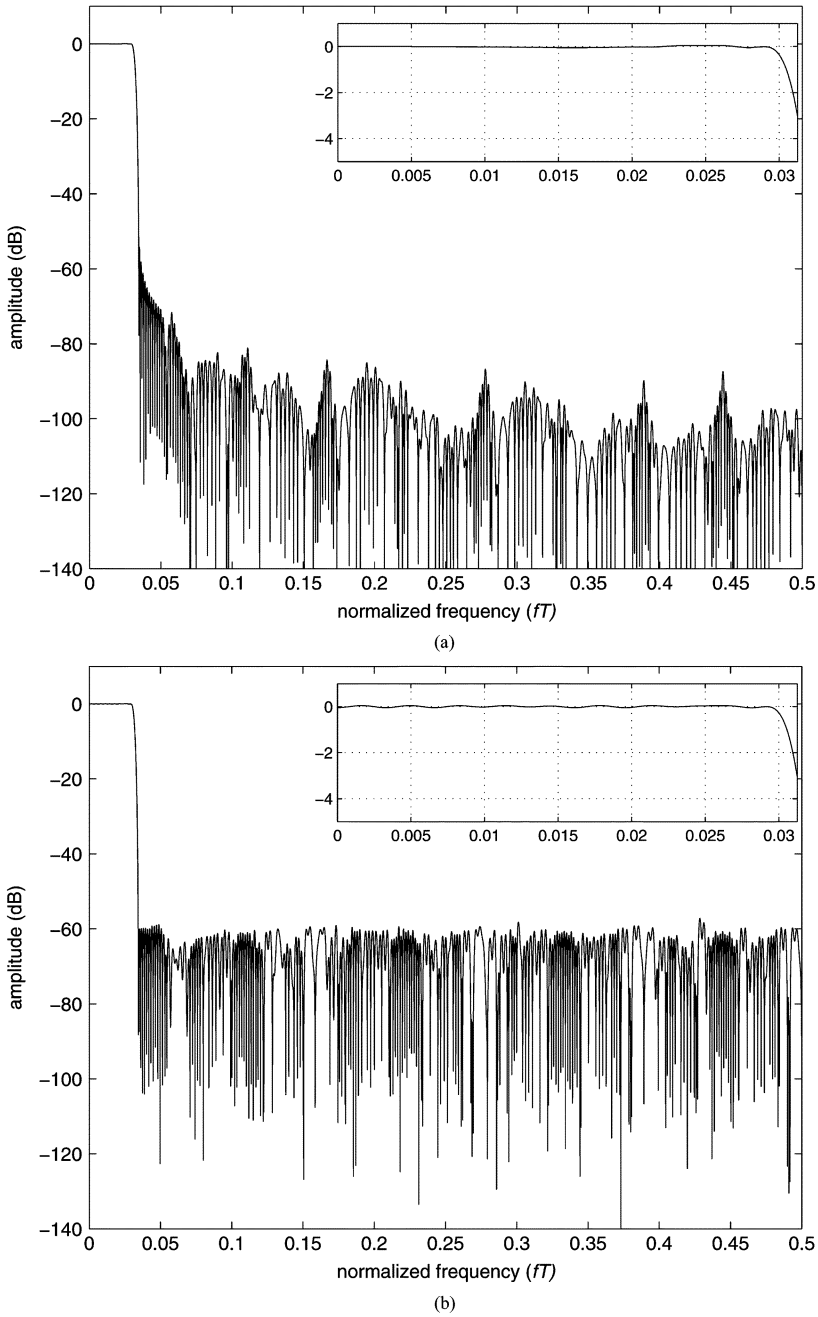


Fig. 5. Magnitude responses and passband details for the FRM prototype filters optimized using FD constraints in Example 1. (a) LS design. (b) Minimax design.

TABLE IV
RESULTING FIGURES OF MERIT FOR DF AND FRM-BASED CMT DESIGNS IN EXAMPLE 1, WHEN USING TD INEQUALITY CONSTRAINTS

Figures of Merit	DF		FRM	
	LS	Minimax	LS	Minimax
d_1	0.01	0.01	0.01	0.01
d_2	0.4×10^{-4}	2.6×10^{-4}	2.1×10^{-4}	26.0×10^{-4}
ISI (dB)	-49.8	-44.0	-48.7	-45.7
ICI (dB)	-86.6	-55.5	-70.0	-46.2
E_2	0.4×10^{-7}	2.6×10^{-7}	1.1×10^{-7}	24.0×10^{-7}
E_∞ (dB)	-46.8	-55.6	-41.0	-51.0

TABLE V
SETTINGS FOR THE FRM SUBFILTERS AND OVERALL PROTOTYPE IN EXAMPLE 2

Parameter	Direct-form	FRM
N_b	–	62
N_1	–	1024
N_2	–	–
N_p	32767	32767
L	–	512
$\mathcal{N}$	16384 (*)	544
$\mathcal{M}$	$\approx 2,150$ MB (**)	≈ 2.37 MB

(*) 8192 unknowns are necessary if the designed prototype provides the PR, that is, both ISI and ICI are zero.

(**) If the problem is PR, the size of the Hessian matrix reduces to 536.9 MB.

leads to a situation where only the upper branch of the FRM filter is necessary [10], since there is no need to increase the bandwidth of the resulting prototype filter with filtered images of the complementary filter resulting in the lower branch of Fig. 4. Moreover, since the order of the prototype filter is given

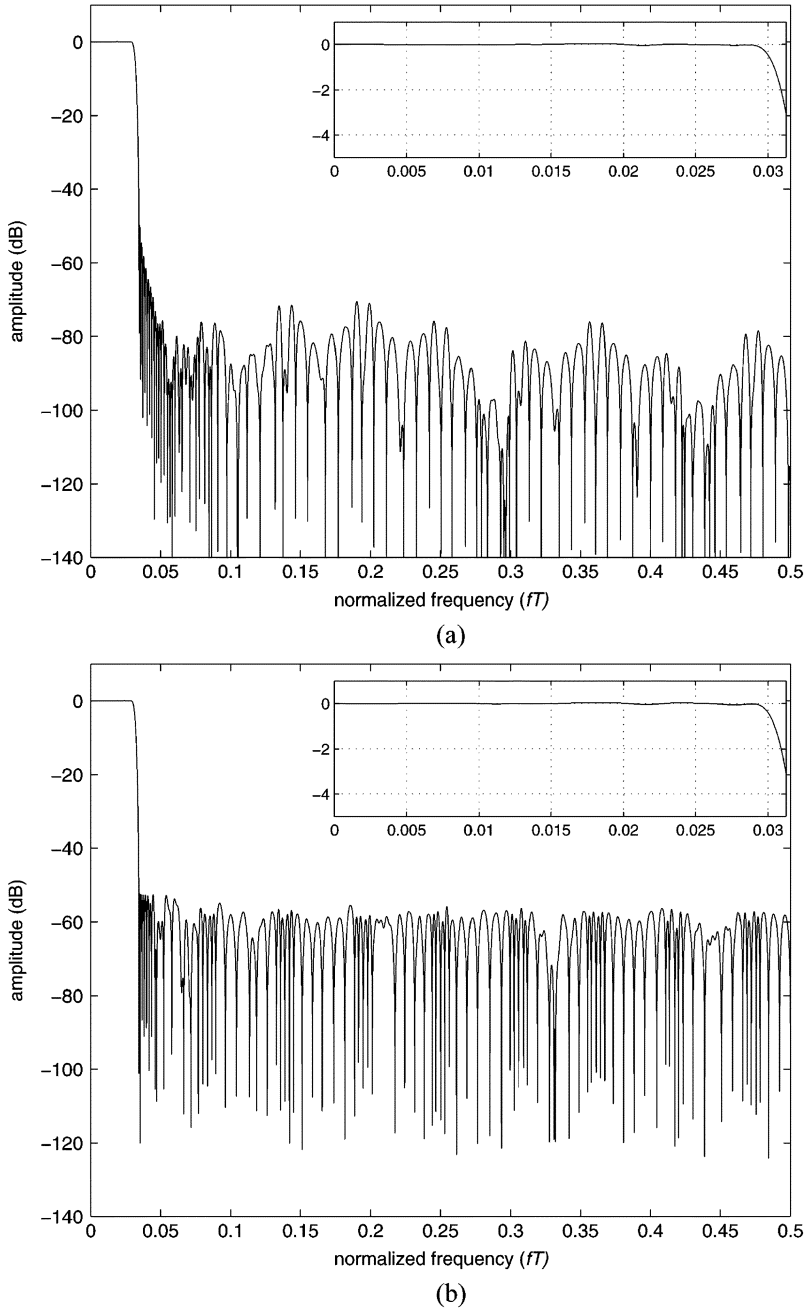


Fig. 6. Magnitude responses and passband details for the FRM prototype filter optimized using TD constraints in Example 1. (a) LS design. (b) Minimax design.

TABLE VI
RESULTING FIGURES OF MERIT FOR THE FRM-BASED CMT DESIGN IN EXAMPLE 2 WITH THE LS OPTIMALITY CRITERION

Figure of Merit	**Step**	
	coarse	**fine**
d_1	0.01	0.01
d_2	5.0×10^{-3}	2.0×10^{-3}
ISI (dB)	-43.2	-43.1
ICI (dB)	-42.6	-46.0
E_2	3.9×10^{-9}	1.0×10^{-9}
E_∞ (dB)	-47.0	-50.0

by $N_p = (2KM - 1) = 32\,767$, optimization techniques based on QP and SQP routines cannot be directly applied, as they require the storage of a Hessian matrix, which will be, in this case, extremely large. The available characteristics of both DF and FRM-based CMTs structures for this example are depicted in Table V, along with the number of coefficients to be optimized and the memory requirements during the optimization stage for the storage of the Hessian matrix. The DF structure in this example is not feasible because its design would require an enormous amount of memory.

The coarse-to-fine optimization procedure described in Section VI with the LS optimality criterion was used in this example, to reduce the computational requirements because the specifications led to a very computationally intensive problem.

Fig. 7(a) and (b) shows the resulting FRM-CMT prototype filter and a small number of subbands of the resulting

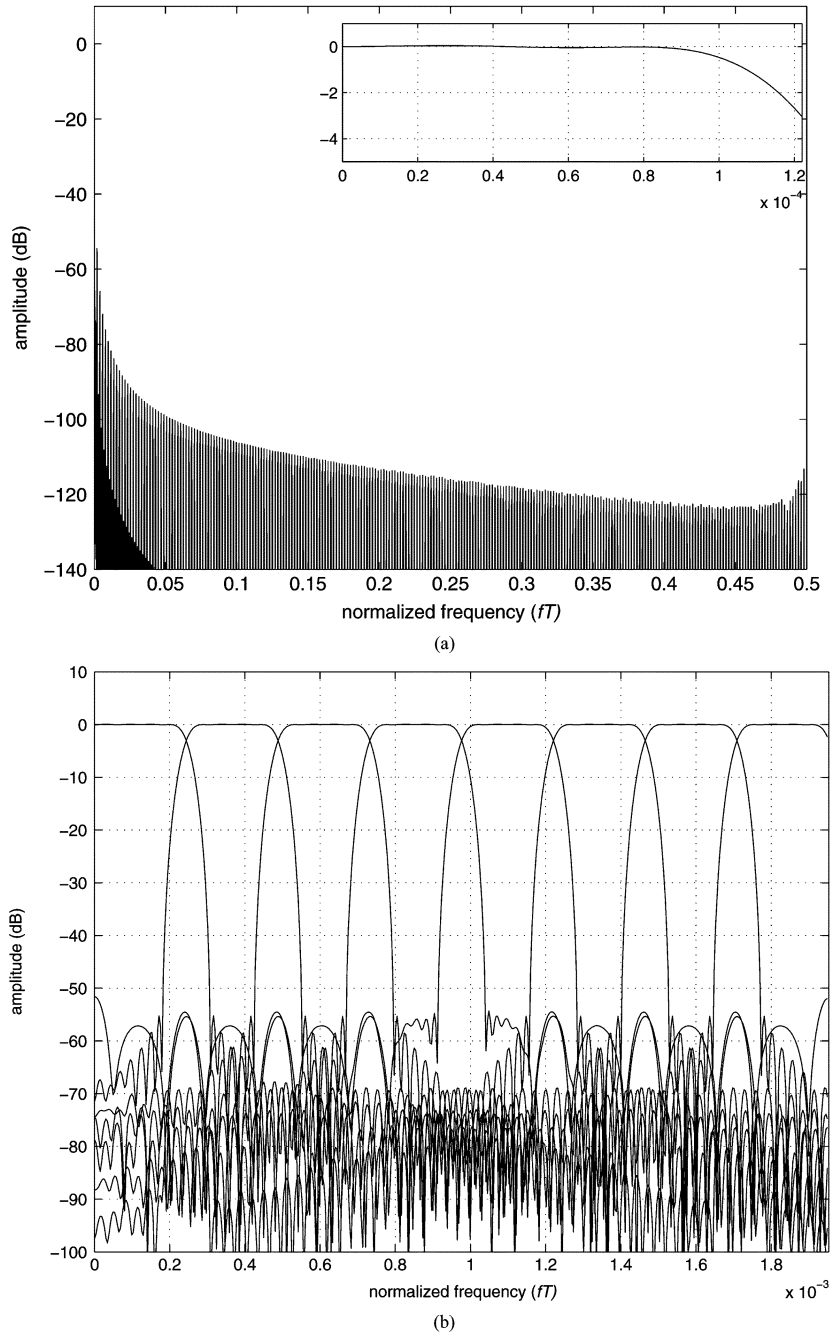


Fig. 7. Resulting FRM-CMT in Example 2. (a) Magnitude response and passband detail. (b) Some lowest subbands of the FRM-CMT: 8 out of 2048.

TMUX, respectively, after the fine adjustment performed in the coarse-to-fine procedure.

Table VI depicts the result of the optimized FRM-CMT subject to the LS optimality criterion, for both coarse and fine adjustment steps. It is worth mentioning that, in both examples here, the specifications for the CMTs were very restrictive but illustrative of the low complexity of the proposed FRM-CMT when designing lengthy prototype filters.

VIII. Conclusion

This paper has proposed a complete optimization method for the design of CMTs and CMFBs, based on the FRM approach. For this matter, we introduced in a general and unified framework, a new and complete set of expressions for FD and TD constraints, which are computationally efficient. The appropriate use of simplified constraint expressions allows the design of very sophisticated filter banks which could not be designed otherwise. One of the main contributions of this work is to indicate when each kind of constraint should be used in order to simplify the optimization process. Overall, the tools provided in this work enable one to generate highly selective filter banks with large number of subbands by optimizing a reduced number of parameters, and using a numerically efficient procedure.

The TD constraints were treated in a very new fashion, where nonlinear inequality constraints were considered, presenting the

advantage of a reduced computational complexity when compared to the traditional FD constraints, and leading to an adjustable tradeoff between the overall and aliasing distortions and the ISI and ICI, for CMFBs and CMTs, respectively. The employment of a coarse-to-fine procedure for the design of both FRM-CMTs and FRM-CMFBs is also proposed, starting with TD constraints and finely adjusting the optimized filter with FD constraints.

REFERENCES

[1] J. Alhava and A. Viholainen, "Coefficient quantization in perfect-reconstruction cosine-modulated filter banks," in *Proc. Europ. Signal Processing Conf.*, Tampere, Finland, Sep. 2000, pp. 1747–1750.

[2] J. Alhava, A. Viholainen, and M. Renfors, "Efficient implementation of complex exponentially modulated filter banks," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. IV, May 2003, pp. 157–160.

[3] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, pp. 5–14, May 1990.

[4] P. S. R. Diniz, L. C. R. de Barcellos, and S. L. Netto, "Design of high-resolution cosine-modulated transmultiplexers with sharp transition band," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1278–1288, May 2004.

[5] P. S. R. Diniz, E. A. B. da Silva, and S. L. Netto, *Digital Signal Processing: System Analysis and Design*. Cambridge, U.K.: Cambridge University Press, 2002.

[6] N. J. Fliege, *Multirate Digital Signal Processing*. Chichester, U.K.: Wiley, 1994.

[7] M. B. Furtado Jr., P. S. R. Diniz, and S. L. Netto, "Numerically efficient optimal design of cosine-modulated filter banks and transmultiplexers with peak-constrained least-squares behavior," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 3, pp. 597–608, Mar. 2005.

[8] ——, "Optimization techniques for cosine-modulated filter banks based on the frequency-response masking approach," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. IV, Bangkok, Thailand, May 2003, pp. 13–16.

[9] M. B. Furtado, P. S. R. Diniz, and S. L. Netto, "Optimized prototype filter based on the FRM approach for cosine-modulated filter banks," *Circuits, Syst., Signal Process.*, vol. 22, no. 2, pp. 193–210, Mar./Apr. 2003.

[10] M. B. Furtado Jr., P. S. R. Diniz, S. L. Netto, and T. Saramäki, "Time-domain constraints for the design of cosine-modulated and modified DFT filter banks with large number of bands and zero intersymbol interference," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. III, Vancouver, BC, Canada, May 2004, pp. 189–192.

[11] T. Karp and N. J. Fliege, "Modified DFT filter banks with perfect reconstruction," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 46, no. 11, pp. 1404–1414, Nov. 1999.

[12] Y. C. Lim, "Frequency-response masking approach for the synthesis of sharp linear phase digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-33, no. 4, pp. 357–364, Apr. 1986.

[13] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Menlo Park, CA: Addison-Wesley, 1989.

[14] H. S. Malvar, *Signal Processing with Lapped Transforms*. Boston, MA: Artech House, 1991.

[15] T. Q. Nguyen, "Near-perfect-reconstruction pseudo-QMF banks," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 65–76, Jan. 1994.

[16] J. G. Proakis, *Digital Communications*, 3rd ed: McGraw-Hill, 1995.

[17] L. Rosenbaum, P. Lowenborg, and M. Johansson, "Cosine and sine modulated FIR filter banks utilizing the frequency-response masking approach," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. III, May 2003, pp. 882–885.

[18] T. Saramäki and Y. Neuvo, "A class of FIR Nyquist (Nth-band) filters with zero intersymbol interference," *IEEE Trans. Circuits Syst.*, vol. CAS-34, no. 10, pp. 1182–1190, Oct. 1987.

[19] T. Saramäki, "A generalized class of cosine-modulated filter banks," in *Proc. TICSP Workshop Transforms and Filter Banks*, Tampere, Finland, Jun. 1998, pp. 336–365.

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

[21] R. D. J. van Nee, R. Prasad, and R. Van Nee, *OFDM for Wireless Multimedia Communications*. Reading, MA: Artech House, 2000.

[22] W.-S. Lu, T. Saramäki, and R. Bregovic, "Design of practically perfect-reconstruction cosine-modulated filter banks: A second-order cone programming approach," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 3, pp. 552–563, Mar. 2004.

[23] *MATLAB Optimization Toolbox: User's Guide*, 1997.

Miguel B. Furtado Jr. (S'02) was born in Itaperuna, RJ, Brazil. He received the B.Sc. degree (*cum laude*) in electrical engineering from the Federal University of Rio de Janeiro (UFRJ), Brazil, in 2001, and the M.Sc. degree in electrical engineering from COPPE/UFRJ, in 2002. He is currently pursuing the D.Sc. degree on electrical engineering at COPPE/UFRJ.

His research interest is on signal processing, communications, information theory, and optimization. During 2004, he worked as a technical consultant for the Instituto Nokia de Desenvolvimento Tecnologico (INDT, Nokia Institute of Technology Development). He is currently with Petrobras, a Brazilian oil company.

Paulo S. R. Diniz (S'80–M'81–SM'92–F'00) was born in Niterói, Brazil. He received the electronics engineering degree (*cum laude*) from the Federal University of Rio de Janeiro (UFRJ), Brazil, in 1978, the M.Sc. degree from COPPE/UFRJ in 1981, and the Ph.D. from Concordia University, Montreal, Canada, in 1984, all in electrical engineering.

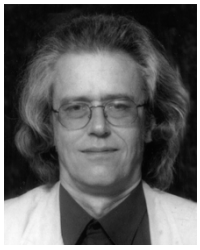
Since 1979 he has been with the Department of Electronic Engineering, UFRJ. He has also been with the Program of Electrical Engineering (Graduate Studies Dept.), COPPE/UFRJ, since 1984, where he is presently a Professor. He served as Undergraduate Course Coordinator and as Chairman of the Graduate Department. He is one of the three senior researchers and coordinators of the National Excellence Center in Signal Processing. From January 1991 to July 1992, he was a visiting Research Associate with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada. He also holds a Docent position with Helsinki University of Technology, Finland. From January 2002 to June 2002, he was a Melchor Chair Professor with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN. His teaching and research interests are in analog and digital signal processing, adaptive signal processing, digital communications, wireless communications, multirate systems, stochastic processes, and electronic circuits. He has published several refereed papers in some of these areas and authored *Adaptive Filtering: Algorithms and Practical Implementation* (Boston, MA: Kluwer Academic, Second ed., 2002) and *Digital Signal Processing: System Analysis and Design* (Cambridge, U.K.: Cambridge University Press, 2002) (with E. A. B. da Silva and S. L. Netto).

Dr. Diniz was the Technical Program Chair of the 1995 MWSCAS held in Rio de Janeiro, Brazil. He has been on the technical committee of several international conferences including ISCAS, ICECS, EUSIPCO, and MWSCAS. He has served Vice President for Region 9 of the IEEE Circuits and Systems Society and as Chairman of the DSP technical committee of the same Society. He was elected a Fellow of the IEEE for "for fundamental contributions to the design and implementation of fixed and adaptive filters and electrical engineering education." He has served as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: ANALOG AND DIGITAL SIGNAL PROCESSING from 1996 to 1999, IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 to 2002, and the *Circuits, Systems and Signal Processing Journal* from 1998 to 2002. He was a distinguished lecturer of the IEEE Circuits and Systems Society for the year 2000 to 2001. In 2004 he served as distinguished lecturer of the IEEE Signal Processing Society and received the 2004 Education Award of the IEEE Circuits and Systems Society. He has also received the Rio de Janeiro State Scientist award from the Governor of Rio de Janeiro state.

Sergio L. Netto (SM'04) was born in Rio de Janeiro, Brazil. He received the B.Sc. degree (*cum laude*) from the Federal University of Rio de Janeiro (UFRJ) in 1991, the M.Sc. degree from the COPPE/UFRJ in 1992, and the Ph.D. degree from the University of Victoria, BC, Canada, in 1996, all in electrical engineering.

Since 1997, he has been an Associate Professor with the Department of Electronics and Computer Engineering, UFRJ, and, since 1998, with the Program of Electrical Engineering, COPPE/UFRJ. His research interests lie in the fields of digital signal processing, adaptive signal processing, and speech processing. He is the coauthor (with P. S. R. Diniz and E. A. B. da Silva) of *Digital Signal Processing: System Analysis and Design* (Cambridge, U.K., Cambridge University Press, 2002).

Dr. Netto is an Associate Editor of *Circuits, Systems and Signal Processing*. He has served as the Vice-President for Region 9 of the IEEE Circuits and Systems Society in 2002 and 2003.



Tapio Saramäki (M'98–SM'01–F'02) was born in Orivesi, Finland, on June 12, 1953. He received the Diploma Engineer (with honors) and Doctor of Technology (with honors) degrees in electrical engineering from the Tampere University of Technology (TUT), Tampere, Finland, in 1978 and 1981, respectively.

Since 1977, he has held various research and teaching positions at TUT, where he is currently a Professor of Signal Processing and a Docent of Telecommunications (a scientist having valuable knowledge for both the research and education at the corresponding laboratory). He is also a Cofounder and a System-Level Designer of VLSI Solution Oy, Tampere, Finland, originally specializing in VLSI implementations of sigma-delta modulators and analog and digital signal processing algorithms for various applications. He is also the President of Aragit Oy Ltd., Tampere, Finland, which was founded by four TUT professors, specializing in various services for the industry, including the application of information technology to numerous applications. In 1982, 1985, 1986, 1990, and 1998, he was a Visiting Research Fellow (Professor) with the University of California, Santa Barbara, in 1987 with the California Institute of Technology, Pasadena, and in 2001 with the National University of Singapore, Singapore. His research interests are in digital signal processing, especially filter and filter bank design, VLSI implementations, and communications applications, as well as approximation and optimization theories. He has written more than 250 international journal and conference articles, various international book chapters, and holds three worldwide-used patents.

Dr. Saramäki received the 1987 Guillemin–Cauer Award for the Best Paper of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, as well as two other Best Paper awards. In 2004, he was awarded the honorary membership of the A. S. Popov Society for Radio-Engineering, Electronics, and Communications (the highest membership grade in the society and the 80th honorary member since 1945) for "great contributions to the development of DSP theory and methods and great contributions to the consolidation of relationships between Russian and Finnish organizations." He is a founding member of the Median-Free Group International. He was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: ANALOG AND DIGITAL SIGNAL PROCESSING (2000–2001), and is currently an Associate Editor of *Circuits, Systems, and Signal Processing* (2003–2008). He was also a Distinguished Lecturer of the IEEE Circuits and Systems Society (2002–2003) and the Chairman of the IEEE Circuits and Systems DSP Technical Committee (May 2002–May 2004).