The Dual Parameterization Approach to Optimal Least Square FIR Filter Design Subject to Maximum Error Constraints

Hai Huyen Dam, Kok Lay Teo, *Senior Member, IEEE*, Sven Nordebo, *Member, IEEE*, and Antonio Cantoni, *Fellow, IEEE*

***Abstract*—This paper is concerned with the design of linear-phase finite impulse response (FIR) digital filters for which the weighted least square error is minimized, subject to maximum error constraints. The design problem is formulated as a semi-infinite quadratic optimization problem. Using a newly developed dual parameterization method in conjunction with the Caratheodory's dimensional theorem, an equivalent dual finite dimensional optimization problem is obtained. The connection between the primal and the dual problems is established. A computational procedure is devised for solving the dual finite dimensional optimization problem. The optimal solution to the primal problem can then be readily obtained from the dual optimal solution. For illustration, examples are solved using the proposed computational procedure.**

***Index Terms*—FIR, least squares, quadratic optimization, quantized.**

I. Introduction

THE PEAK constrained weighted least square error (PCWLSE) filter design problem is concerned with the design of a linear-phase filter for which the weighted least square error with respect to a desired response is minimized, subject to constraints on the peak deviation from the desired response in the passband and stopband. This design problem is of interest since it has been found to have a good signal-to-noise ratio (SNR) with an acceptable error in the passband and stopband compared with filters designed according to the min-max and least square criteria [1]. The PCWLSE problem can be formulated as a semi-infinite quadratic optimization problem. The problem can then be solved using a multiple exchange algorithm [1]–[3] based on an approach similar to that used in the development of the Remez algorithm. This method appears to perform well in practice. However, there is no proof that the solution obtained satisfies the semi-infinite constraints.

An alternative approach for handling the semi-infinite quadratic optimization problem is to discretize the passband and the stopband using a dense grid. This gives rise to a standard quadratic optimization problem approximating the original problem. However, there is no guarantee that the solution obtained will satisfy the continuous constraints at points in between the grid points.

In this paper, we propose a new approach to solve the semi-infinite quadratic optimization problem that results from PCWLSE filter design problem. The approach has been applied to the envelop constrained filter design problem [4]. This approach is based on the dual parameterization method for convex semi-infinite programming developed in [5]. The approach leads to a new computational procedure. The solution obtained using the new approach can be shown to satisfy the continuous constraints at all points.

Another problem formulation relating to the PCWLSE problem is given in [6]. The problem is concerned with the design of a lowpass filter that maximizes the stopband attenuation subject to a given maximum passband ripple. The problem is solved by using the spectral factorization method.

The rest of the paper is organized as follows. In Section II, we formulate the PCWLSE filter design problem. The dual semi-infinite quadratic optimization problem is obtained in Section III and an algorithm for solving the problem is proposed. Design examples are given in Section IV.

II. Peak Constrained Weighted Least Square Error (PCWLSE)

The frequency response of a linear-phase finite impulse response (FIR) filter of length L and impulse response

$$\{h(0), h(1), \ldots, h(L-1)\}$$

can be written as

$$H(e^{j2\pi f}) = \sum_{n=0}^{L-1} h(n)e^{-j2\pi fn} = e^{-j2\pi f\frac{(L-1)}{2}} A(f).$$

The real response $A(f)$ is given by

$$A(f) = \phi^T(f)\mathbf{h}$$

Manuscript received February 8, 1999; revised February 28, 2000. This work was supported by the Cooperative Research Center for Broadband Telecommunications and Networking and the Australian Research Council. The associate editor coordinating the review of this paper and approving it for publication was Prof. Dr. Ir. Bart L. R. De Moor.

H. H. Dam is with Department of Telecommunications and Signal Processing, University of Karlskrona/Ronneby, Ronneby, Sweden on leave from the Australian Telecommunication Research Institute (ATRI), Curtin University of Technology, Perth, Australia.

K. L. Teo is with the Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.

S. Nordebo is with the Department of Telecommunications and Signal Processing, University of Karlskrona/Ronneby, Ronneby, Sweden.

A. Cantoni is with the Australian Telecommunication Research Institute (ATRI), Curtin University of Technology, Perth, Australia, and the University of Western Australia, Perth, Australia.

Publisher Item Identifier S 1053-587X(00)05983-3.

where the column vector $\mathbf{h}$ contains N impulse response coefficients with $N = (L+1)/2$ for odd values of L and $N = L/2$ for even values of L, and $\phi(f)$ is a column vector containing sine and cosine functions of f [7] (see Appendix A1).

The desired real response is given as

$$A_d(f) = \begin{cases} 1, & f \in \mathcal{P} \\ 0, & f \in \mathcal{S} \end{cases}$$

where $\mathcal{P}$ and $\mathcal{S}$ are passband and stopband, respectively. Without loss of generality, we concentrate on a lowpass filter with $\mathcal{P} = [0, f_p]$ and $\mathcal{S} = [f_s, 0.5]$. Highpass, bandpass, and bandstop filters can be handled similarly.

The cost function associated with the PCWLSE filter design problem is

$$e = W_p \int_0^{f_p} |A(f) - A_d(f)|^2\, df + W_s \int_{f_s}^{0.5} |A(f) - A_d(f)|^2\, df \tag{2.1}$$

where W_p and W_s are the weighting coefficients for the passband and stopband, respectively. Note that W_s is usually chosen much larger than W_p [1]. The cost function (2.1) reduces to

$$e = \frac{1}{2}\mathbf{h}^T \mathbf{\Phi} \mathbf{h} + \mathbf{p}^T \mathbf{h} + W_p f_p$$

where

$$\mathbf{\Phi} = 2W_p \int_0^{f_p} \phi(f)\phi^T(f)\, df + 2W_s \int_{f_s}^{0.5} \phi(f)\phi^T(f)\, df \tag{2.2}$$

and

$$\mathbf{p} = -2W_p \int_0^{f_p} \phi(f)\, df.$$

Let δ_p and δ_s denote the maximum passband and stopband errors, respectively. Then, the constraint in the passband is

$$|A(f) - A_d(f)| \le \delta_p, \quad \forall f \in [0, f_p] \tag{2.3}$$

whereas the constraint in the stopband is

$$|A(f)| \le \delta_s, \quad \forall f \in [f_s, 0.5]. \tag{2.4}$$

The PCWLSE filter design problem can now be stated as the following semi-infinite quadratic programming problem:

$$\begin{aligned} \min_{\mathbf{h}} \quad & \frac{1}{2}\mathbf{h}^T \mathbf{\Phi} \mathbf{h} + \mathbf{p}^T \mathbf{h} + W_p f_p \\ \text{subject to} \quad & 1 - \delta_p \le \phi^T(f)\mathbf{h} \le 1 + \delta_p, \quad \forall f \in [0, f_p] \\ & -\delta_s \le \phi^T(f)\mathbf{h} \le \delta_s, \quad \forall f \in [f_s, 0.5]. \end{aligned}$$

Let

$$B(f) = \begin{bmatrix} \phi^T(f) \\ -\phi^T(f) \end{bmatrix}$$

and

$$c_1(f) = \begin{bmatrix} 1 + \delta_p \\ -1 + \delta_p \end{bmatrix}, \quad c_2(f) = \begin{bmatrix} \delta_s \\ \delta_s \end{bmatrix}.$$

The above optimization problem can be written in the following standard form, which is referred to as the primal problem (P):

$$\begin{aligned} \min_{\mathbf{h}} \quad & \gamma(\mathbf{h}) \\ \text{subject to} \quad & g_1(\mathbf{h}, f) \le \mathbf{0}_2, \quad \forall f \in [0, f_p] \\ & g_2(\mathbf{h}, f) \le \mathbf{0}_2, \quad \forall f \in [f_s, 0.5] \end{aligned}$$

where

$$\gamma(\mathbf{h}) = \frac{1}{2}\mathbf{h}^T \mathbf{\Phi} \mathbf{h} + \mathbf{p}^T \mathbf{h} + W_p f_p$$

$$g_1(\mathbf{h}, f) = B(f)\mathbf{h} - c_1(f)$$

and

$$g_2(\mathbf{h}, f) = B(f)\mathbf{h} - c_2(f).$$

Clearly, $B(f) \in C([0, f_p] \bigcup [f_s, 0.5], \mathcal{R}^{2\times N}), c_1(f) \in C([0, f_p], \mathcal{R}^2)$, and $c_2(f) \in C([f_s, 0.5], \mathcal{R}^2)$, where $C([0, f_p] \bigcup [f_s, 0.5], \mathcal{R}^{2\times N})$ denotes the Banach space that consists of all continuous functions from

$$[0, f_p] \bigcup [f_s, 0.5] \text{ to } \mathcal{R}^{2\times N}$$

and equipped with the sup norm. $C([0, f_p], \mathcal{R}^2)$ and $C([f_s, 0.5], \mathcal{R}^2)$ are to be understood similarly.

III. Dual Parameterization Approach

The approach proposed in this paper for solving Problem (P) relies on adapting the dual parameterization method developed in [5]. The basis of the approach is to consider the dual problem with continuous constraints that can be reduced to a problem with finite constraints.

We assume that the following condition is satisfied for the PCWLSE filter design problem.

Assumption 3.1: There exists an $\mathbf{h}^o \in \mathcal{R}^N$ such as $g_1(\mathbf{h}^o, f) < \mathbf{0}_2$ for all $f \in [0, f_p]$ and $g_2(\mathbf{h}^o, f) < \mathbf{0}_2$ for all $f \in [f_s, 0.5]$.

Assumption 3.1 is also called Slater's constraints qualification. This condition is satisfied if the specified maximum allowable passband and stopband errors are properly chosen. If the condition is not satisfied, the approach developed in this paper is not applicable. This situation implies that the specified maximum allowable passband and/or stopband errors are too restrictive and, hence, should be relaxed.

Since the matrix $\mathbf{\Phi}$ defined by (2.2) is positive definite (for proof, see Appendix A2), the cost of the problem (P) is strictly convex in $\mathbf{h}$. Furthermore, the constraints are linear, and hence, $\mathbf{h}$ belongs to a convex set. Thus, problem (P) is a convex semi-infinite quadratic programming problem, and by Assumption 3.1, we have Theorem 3.1 [8].

Theorem 3.1: The problem (P) has a unique optimum solution.

We now derive the dual problem corresponding to problem (P). The Lagrangian for the problem (P) is

$$\mathcal{L}(\mathbf{h}, \Lambda_1, \Lambda_2) = \gamma(\mathbf{h}) + \int_0^{f_p} g_1^T(\mathbf{h}, f)\, d\Lambda_1(f) + \int_{f_s}^{0.5} g_2^T(\mathbf{h}, f) d\Lambda_2(f) \tag{3.5}$$

where

$$\Lambda_1(f) = [\lambda_{1,1}(f), \lambda_{1,2}(f)]^T \in M([0, f_p], \mathcal{R}^2)$$

and

$$\Lambda_2(f) = [\lambda_{2,1}(f), \lambda_{2,2}(f)]^T \in M([f_s, 0.5], \mathcal{R}^2)$$

are the Lagrange multipliers. $M([0, f_p], \mathcal{R}^2)$ (respectively, $M([f_s, 0.5], \mathcal{R}^2)$) is the dual space of $C([0, f_p], \mathcal{R}^2)$ (respectively, $C([f_s, 0.5], \mathcal{R}^2)$) representing by the space of all finite signed regular Borel measures on $[0, f_p]$ (respectively, $[f_s, 0.5]$). Note that the integration is to be understood as Lebesgue–Stieltjes integration with respect to measure induced by $\Lambda_1(f)$ (respectively, $\Lambda_2(f)$).

By virtue of the uniqueness of the primal optimum solution, the following theorem can be established from the Lagrangian duality theorem [9].

Theorem 3.2: Let $\mathbf{h}^*$ be the optimal solution of the primal problem (P). Then

$$\gamma(\mathbf{h}^*) = \max_{\Lambda_1 \geq \mathbf{0}_2; \Lambda_2 \geq \mathbf{0}_2} \left\{ \min_{\mathbf{h} \in R^N} \mathcal{L}(\mathbf{h}, \Lambda_1, \Lambda_2) \right\} \tag{3.6}$$

for $\Lambda_1 \in M([0, f_p], \mathcal{R}^2), \Lambda_2 \in M([f_s, 0.5], \mathcal{R}^2)$.

Clearly, the minimization problem

$$\min_{\mathbf{h} \in \mathcal{R}^N} \mathcal{L}(\mathbf{h}, \Lambda_1, \Lambda_2)$$

is a convex, unconstrained quadratic optimization problem. To obtain the optimal solution, we set the gradient of the function $\mathcal{L}(\mathbf{h}, \Lambda_1, \Lambda_2)$ to zero. This leads to the following set of equations:

$$\nabla_{\mathbf{h}} \mathcal{L}(\mathbf{h}, \Lambda_1, \Lambda_2) = \mathbf{\Phi h} + \mathbf{p} + \int_0^{f_p} B^T(f)\, d\Lambda_1(f) + \int_{f_s}^{0.5} B^T(f)\, d\Lambda_2(f) = 0. \tag{3.7}$$

Then, it follows that

$$\mathbf{h}(\Lambda_1, \Lambda_2) = -\mathbf{\Phi}^{-1} \left(\mathbf{p} + \int_0^{f_p} B^T(f)\, d\Lambda_1(f) + \int_{f_s}^{0.5} B^T(f)\, d\Lambda_2(f) \right). \tag{3.8}$$

Substituting (3.7) and (3.8) into the expression of the Lagrangian function $\mathcal{L}$, we obtain

$$\begin{aligned}\mathcal{L}(\Lambda_1, \Lambda_2) &= -\frac{1}{2} \left(\mathbf{p} + \int_0^{f_p} B^T(f)\, d\Lambda_1(f) + \int_{f_s}^{0.5} B^T(f) d\Lambda_2(f) \right)^T \mathbf{\Phi}^{-1} \left(\mathbf{p} + \int_0^{f_p} B^T(f) d\Lambda_1(f) \right. \\ &\quad \left. + \int_{f_s}^{0.5} B^T(f) d\Lambda_2(f) \right) - \int_0^{f_p} c_1^T(f)\, d\Lambda_1(f) \\ &\quad - \int_{f_s}^{0.5} c_2^T(f)\, d\Lambda_2(f) + W_p f_p. \end{aligned} \tag{3.9}$$

This gives rise to the following dual problem (D)

$$\begin{aligned} &\max_{\Lambda_1, \Lambda_2} \quad \mathcal{L}(\Lambda_1, \Lambda_2) \\ &\text{subject to} \quad \Lambda_1(f) \geq \mathbf{0}_2, \quad f \in [0, f_p] \\ &\qquad\qquad \Lambda_2(f) \geq \mathbf{0}_2, \quad f \in [f_s, 0.5]. \end{aligned}$$

The following theorem is obtained from Theorems 3.1 and 3.2.

Theorem 3.3: Let $(\Lambda_1^*, \Lambda_2^*)$ be a solution of the dual problem (D). Then, $\mathbf{h}^* = \mathbf{h}(\Lambda_1^*, \Lambda_2^*)$ given by (3.8) is the solution of the primal problem (P).

By Caratheodory's dimensional theorem [10], it is known that the dual problem (D) admits an optimal solution, which is a measure concentrating at no more than N points. This result is summarized in the following as a theorem.

Theorem 3.4: Let $\mathbf{h}^* \in \mathcal{R}^N$ be the optimal solution of problem (P). Then, there exists a solution of the dual problem (D), which is a measure concentrating at no more than N points.

Let m_1 and m_2 be the number of the supporting points in $[0, f_p]$ and $[f_s, 0.5]$), respectively, at which the dual optimal solution is concentrated. Let these supporting points be denoted by $\mathbf{f}_1 = [f_{1,1}, \ldots, f_{1,m_1}]$ and $\mathbf{f}_2 = [f_{2,1}, \ldots, f_{2,m_2}]$, where $f_{1,1}, \ldots, f_{1,m_1} \in [0, f_p]$ and $f_{2,1}, \ldots, f_{2,m_2} \in [f_s, 0.5]$. Then, the constraints of the primal problem (P) corresponding to the optimal solution are only achieved at these points. By Theorem 3.4, it is clear that $m_1 + m_2 \leq N$.

Let

$$\mathbf{\Lambda}_1 = [\Lambda_{1,1}, \Lambda_{1,2}, \ldots, \Lambda_{1,m_1}]$$

and

$$\mathbf{\Lambda}_2 = [\Lambda_{2,1}, \Lambda_{2,2}, \ldots, \Lambda_{2,m_2}]$$

be the dual variables associated with the supporting points $\mathbf{f}_1$ and $\mathbf{f}_2$, respectively. Note that $\Lambda_{i,j} \in \mathcal{R}^2, j = 1, \ldots, m_i; i = 1, 2$. Thus, it suffices to search for an optimal dual solution from $\mathcal{M}$, where $\mathcal{M}$ is a subset of $M([0, f_p], \mathcal{R}^2) \times [0, f_p] \bigcup M([f_s, 0.5], \mathcal{R}^2) \times [f_s, 0.5])$ consisting of elements characterized by $(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2)$. In other words, the dual problem (D) is equivalent to the one with $(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2)$ restricted to $\mathcal{M}$. On this basis, it follows that $\mathbf{h}(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2)$ obtained in (3.8) is reduced to

$$\mathbf{h}(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2) = -\mathbf{\Phi}^{-1} \left(\mathbf{p} + \sum_{i=1}^{m_1} B^T(f_{1,i}) \Lambda_{1,i} + \sum_{i=1}^{m_2} B^T(f_{2,i}) \Lambda_{2,i} \right).$$

Thus, the dual problem (D) is equivalent to the following finite-dimensional optimization problem (D1):

$$\begin{aligned} &\min_{\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2} \quad L(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2) \\ &\text{subject to} \quad \mathbf{\Lambda}_1 \geq \mathbf{0}_{2 \times m_1}, \quad \mathbf{\Lambda}_2 \geq \mathbf{0}_{2 \times m_2} \end{aligned}$$

where $\mathbf{\Lambda}_1 \geq \mathbf{0}_{2 \times m_1}$ (respectively, $\mathbf{\Lambda}_2 \geq \mathbf{0}_{2 \times m_2}$) denotes $\Lambda_{1,i} \geq \mathbf{0}_2, i = 1, \ldots, m_1$ (respectively, $\Lambda_{2,i} \geq \mathbf{0}_2, i = 1, \ldots, m_2$), and

$$L(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2) = \frac{1}{2}\left(\mathbf{p} + \sum_{i=1}^{m_1} B^T(f_{1,i})\Lambda_{1,i} + \sum_{i=1}^{m_2} B^T(f_{2,i})\Lambda_{2,i}\right)^T \mathbf{\Phi}^{-1}\left(\mathbf{p} + \sum_{i=1}^{m_1} B^T(f_{1,i})\Lambda_{1,i} + \sum_{i=1}^{m_2} B^T(f_{2,i})\Lambda_{2,i}\right) + \sum_{i=1}^{m_1} c_1^T(f_{1,i})\Lambda_{1,i} + \sum_{i=1}^{m_2} c_2^T(f_{2,i})\Lambda_{2,i} - W_p f_p.$$

The corresponding primal problem (P1) to the finite dimensional dual problem (D1) is

$$\begin{aligned} &\min_{\mathbf{h}} \quad \gamma(\mathbf{h}) \\ &\text{subject to} \quad g_1(\mathbf{h}, f_{1,i}) \leq \mathbf{0}_2, \quad i = 1, \ldots, m_1 \\ &\qquad\qquad\quad g_2(\mathbf{h}, f_{2,i}) \leq \mathbf{0}_2, \quad i = 1, \ldots, m_2. \end{aligned}$$

The following theorem establishes the relationship between the solution for the finite primal problem (P1) and the original primal problem (P).

Theorem 3.5: Let $\mathbf{h}^1$ be an optimum solution for (P1). If $\mathbf{h}^1$ satisfies the continuous constraints, then it is the optimum solution for (P).

Proof: See Appendix A3.

Note that the dual problem (D1) is a nonconvex optimization problem. Thus, there may exist local optima. However, by a close examination of the problem, we observe that the function L is convex in $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ corresponding to each given pair of $\mathbf{f}_1$ and $\mathbf{f}_2$. Thus, we propose to solve the problem (D1) as a two-level optimization problem

$$\min_{\mathbf{f}_1, \mathbf{f}_2} \left\{ \min_{\mathbf{\Lambda}_1 \geq \mathbf{0}, \mathbf{\Lambda}_2 \geq \mathbf{0}} L(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2) \right\}. \tag{3.10}$$

More specifically, for each $\mathbf{f}_1, \mathbf{f}_2$, define

$$F(\mathbf{f}_1, \mathbf{f}_2) = \min_{\mathbf{\Lambda}_1 \geq \mathbf{0}, \mathbf{\Lambda}_2 \geq \mathbf{0}} L(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2). \tag{3.11}$$

Since $L(\mathbf{\Lambda}_1, \mathbf{f}_1, \mathbf{\Lambda}_2, \mathbf{f}_2)$ is convex in $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ corresponding to each $\mathbf{f}_1$ and $\mathbf{f}_2$, it can be easily solved. We then view $F(\mathbf{f}_1, \mathbf{f}_2)$ as an optimization problem. It is known that F is nonsmooth with respect to $(\mathbf{f}_1, \mathbf{f}_2)$. However, it is found that optimization techniques such as the one given in the Matlab Optimization Toolbox using numerical gradient work rather well if the initial guess is close to the optimal solution. Finding a good initial guess is quite easy for our problem. More precisely, we discretize the frequency domain $[0, f_p]$ and $[f_s, 0.5]$ into a dense enough grid of points. Then, the problem (D) is approximated by a convex quadratic optimization problem subject to linear constraints. It is thus readily solvable. Clearly, if a dual variable is not equal to zero, then its corresponding frequency point $\mathbf{f}_{i,j}$ is an active point. These active points can be chosen as the initial frequency points $(\mathbf{f}_1^o, \mathbf{f}_2^o)$. This simple procedure enable us to find good initial points as well as to know approximately the number of active points and, hence, the dimensions m_1 and m_2.

The proposed computational procedure for solving the semi-infinite quadratic optimization problem (P) by using the dual method and Caratheodory's Theorem can be summarized as follows.

Computational Procedure:

- Step 1) Obtain a good initial guess $(\mathbf{f}_1^o, \mathbf{f}_2^o)$ for the set of active points for the constraints as suggested above.
- Step 2) Solve problem (D1) as a two-level optimization (3.10), where for each set of active points $(\mathbf{f}_1, \mathbf{f}_2)$, a convex quadratic optimization problem with a small number of constraints (3.11) is solved as the inner optimization problem. Optimization methods using numerical gradient such as the one given in the Matlab Optimization Toolbox is then used to solve the dual problem (3.10), starting from the initial active points $(\mathbf{f}_1^o, \mathbf{f}_2^o)$, as the outer optimization problem.
- Step 3) Let the solution obtained be denoted by $(\mathbf{h}^*, \mathbf{f}_1^*, \mathbf{f}_2^*)$. Check if the continuous constraints of the primal problem (P) are satisfied. If they are satisfied, it follows from Theorem 3.5 that $\mathbf{h}^*$ is the optimum solution of problem (P). If they are not satisfied, repeat Step 1 with a finer discretization to obtain a better initial guess for a set of active points. In practice, these discretization refining process will eventually produce an initial guess $(\mathbf{f}_1, \mathbf{f}_2)$, which is closed enough to the optimal solution of the outer optimization problem as defined in Step 2 to ensure the convergence to the optimal solution.

The multiple exchange algorithm and finitization methods can, in principle, give an arbitrarily accurate approximation to the primal problem (P). However, these approaches become excessively memory intensive as the grid spacing increases, especially when the number of filter coefficients is large. The dual parameterization approach, however, can obtain the global solution with only a small number of active points. Experience with a number of numerical examples shows that the procedure is relatively fast when the initial active point is chosen according to the above procedure. The solution obtained using the algorithm is the similar to those obtained by using the above two methods with a denser grid in the frequency domain.

IV. Design Examples

The first design example is for a lowpass FIR filter of length $L = 35$, where $f_p = 0.05$ and $f_s = 0.1$. The weightings are 1 in the passband and 1000 in the stopband.

Let DB_p and DB_s represent the passband and stopband ripples measured in decibels, where $\mathrm{DB}_p = 20\log_{10}\{(1+\delta_p)/(1-\delta_p)\}$, and $\mathrm{DB}_s = 20\log_{10}\delta_s$, respectively. The design requires $\mathrm{DB}_p = 1.0$ dB in the passband and $\mathrm{DB}_s \leq -40$ dB in the stopband.

The passband and the stopband are discretized into 256 equally spaced frequency points. Solve this discretized quadratic optimization problem for the problem (P) to obtain the corresponding dual variables. Identify those frequency points corresponding to nonzero dual variables. Choose $\mathbf{f}_1^o = [0.00, 0.033, 0.0497]$ and $\mathbf{f}_2^o = [0.10, 0.109, 0.135]$. Thus,

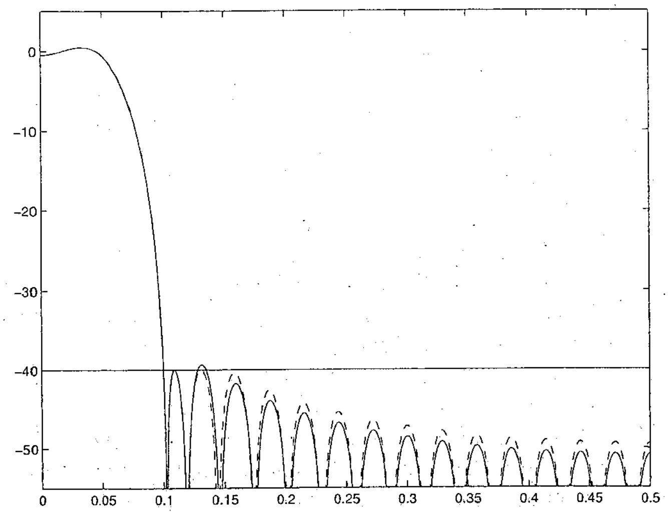


Fig. 1. Frequency responses for $L = 35, f_p = 0.05$ and $f_s = 0.1$. The frequency response for the initial filter (solid curve). The frequency response for the optimum filter (dashed curve).

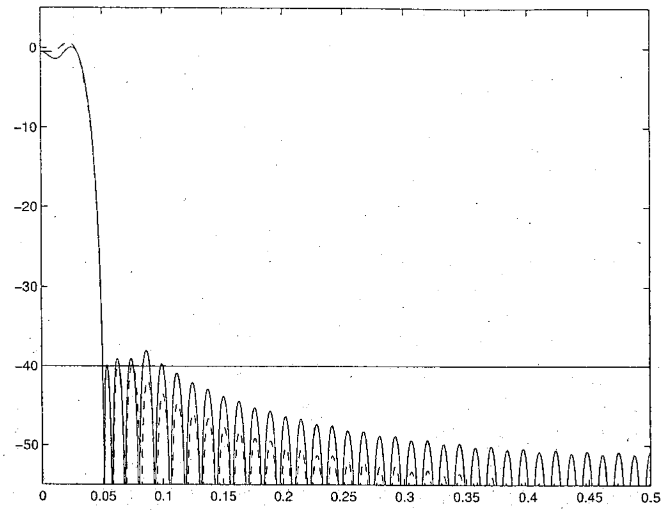


Fig. 2. Frequency responses for $L = 77, f_p = 0.03$ and $f_s = 0.05$. The frequency response for the initial filter (solid curve). The frequency response for the optimum filter (dashed curve).

$m_1 = 3$ and $m_2 = 3$. Note that the filter length $L = 35$ is much larger than the total number of active points. $(\mathbf{f}_1^o, \mathbf{f}_2^o)$ will be used as the initial guess for the Step 2 of the Computational Procedure.

The frequency response for the filter based on these active points is given by the solid curve in Fig. 1. These frequency points are used as the initial points for the Step 2 in the above computational procedure. The frequency response for the solution obtained from this procedure is given by the dashed curve in Fig. 1.

The straight line in Fig. 1 denotes the constraint in the stopband. The first filter does not satisfy the stopband constraints, whereas the second filter satisfies all the constraints. From Theorem 3.5, the second solution is the optimum solution for the semi-infinite quadratic optimization problem (P).

Another design example is a narrowband lowpass filter with $f_p = 0.03, f_s = 0.05$ and the filter length $L = 77$. The initial frequency points are chosen: $\mathbf{f}_1^o = [0, 0.02, 0.03]$ and $\mathbf{f}_2^o = [0.05, 0.054, 0.064, 0.072]$ with the corresponding $m_1 = 3, m_2 = 4$. Note that although the number of coefficients is almost twice the number in the first example, the number of active points is virtually the same (7). The frequency response based on this set of active points and the filter obtained from the Computational Procedure are given by the solid and dashed curves in Fig. 2, where the stopband constraint is given by the straight line. Obviously, the initial filter does not satisfy the constraints, whereas the second filter satisfies all the constraints.

V. Conclusion

A new method has been proposed to solve the semi-infinite quadratic program for the problem that arises from the PCWLSE filter design problem. The advantage of the method is that the original semi-infinite programming problem can be solved directly as a finite-dimensional optimization problem. In particular, only a few crucial active points are sufficient to give enough information for searching for the optimal solution. This method is preferred over methods based on discretization as there is no guarantee that the optimal solution will satisfy the continuous constraints of the original problem.

Appendix

A. FIR Filter Classes

There are two main cases of linear-phase FIR filters (with $h(i) = h(L - i - 1), \forall i$).

- The filter with odd length

$$\phi(f) = [2\cos(2\pi f((L-1)/2)), 2\cos(2\pi f((L-1)/2 - 1)) \\ , \ldots, 2\cos(2\pi f), 1]^T$$

and

$$\mathbf{h} = [h(0), h(1), \ldots, h((L-1)/2)]^T.$$

- The filter with even length

$$\phi(f) = [2\cos(2\pi f((L-1)/2)), 2\cos(2\pi f((L-1)/2 - 1) \\ , \ldots, 2\cos(\pi f)]^T$$

and

$$\mathbf{h} = [h(0), h(1), \ldots, h(L/2 - 1)]^T.$$

B. Matrix $\mathbf{\Phi}$ Is Positive Definite

We need to show that

$$\mathbf{h}^T \mathbf{\Phi} \mathbf{h} \geq 0, \quad \forall \mathbf{h} \tag{6.12}$$

and

$$\mathbf{h}^T \mathbf{\Phi} \mathbf{h} = 0 \quad \text{iff } \mathbf{h} = 0. \tag{6.13}$$

Equation (6.12) is true in view of the definition of the matrix $\mathbf{\Phi}$ given in (2.2). Hence, we only need to prove (6.13). Given $\mathbf{h}^T\mathbf{\Phi}\mathbf{h} = 0$, we have

$$2W_p \int_0^{f_p} \mathbf{h}^T\phi(f)\phi^T(f)\mathbf{h}\, df + 2W_s \int_{f_s}^{0.5} \mathbf{h}^T\phi(f)\phi^T(f)\mathbf{h}\, df = 0. \quad (6.14)$$

Since $W_p > 0, W_s > 0$ and $\mathbf{h}^T\phi(f)\phi^T(f)\mathbf{h} \geq 0$ for all vectors $\mathbf{h}$ and $f \in [0, f_p]\bigcup[f_s, 0.5]$, it follows from (6.14) that

$$\mathbf{h}^T\phi(f) = 0, \quad \forall f \in [0, f_p]\bigcup[f_s, 0.5]. \quad (6.15)$$

By [11, p. 137, Th. 2], it follows that

$$\{e^{j2\pi f(L-1)/2}, e^{j2\pi f((L-1/2)-1)}, \ldots, e^{j2\pi f}, 1\}$$

is linear independent on any open interval of $[0, f_p]\bigcup[f_s, 0.5]$. This clearly implies that the components of the vector $\phi(f)$ are linear independent functions on any open interval of $[0, f_p]\bigcup[f_s, 0.5]$. Thus, (6.15) holds iff $\mathbf{h} = 0$. Therefore, the matrix $\mathbf{\Phi}$ is positive definite.

C. Proof for Theorem 3.5

Let us define the feasible regions for the problems (P1) and (P) as follows:

$$\mathcal{H}^1 = \{\mathbf{h} : g_1(\mathbf{h}, f_{1,i}) \leq \mathbf{0}_2,\ i = 1, \ldots, m_1 \quad g_2(\mathbf{h}, f_{2,i}) \leq \mathbf{0}_2,\ i = 1, \ldots, m_2\}$$

and

$$\mathcal{H}^* = \{\mathbf{h} : g_1(\mathbf{h}, f) \leq \mathbf{0}_2,\ \forall f \in [0, f_p] \quad g_2(\mathbf{h}, f) \leq \mathbf{0}_2,\ \forall f \in [f_s, 0.5]\}.$$

Then, $\mathcal{H}^* \subset \mathcal{H}^1$. Denote by $\mathbf{h}^1$ and $\mathbf{h}^*$ the optimum solutions for the problem (P1) and (P), respectively. Since $\mathbf{h}^1$ is a feasible solution for the problem (P) and $\mathbf{h}^*$ is the optimum solution for (P), we have $\gamma(\mathbf{h}^1) \geq \gamma(\mathbf{h}^*)$. However, since $\mathcal{H}^* \subset \mathcal{H}^1$, it is clear that $\gamma(\mathbf{h}^*) \geq \gamma(\mathbf{h}^1)$. Thus, $\gamma(\mathbf{h}^*) = \gamma(\mathbf{h}^1)$. From Theorem 3.1, we have $\mathbf{h}^* = \mathbf{h}^1$, and hence, $\mathbf{h}^1$ is the optimum solution for the problem (P).

REFERENCES

[1] J. W. Adams, "FIR digital filters with least-squares stopbands subject to peak-gain constraints," *IEEE Trans. Circuits Syst.*, vol. 39, pp. 376–388, Apr. 1991.

[2] J. W. Adams and J. L. Sullivan, "Peak-constrainted least squares optimization," *IEEE Trans. Signal Processing*, vol. 46, pp. 306–321, Feb. 1998.

[3] J. W. Adams, "A new optimal window," *IEEE Trans. Signal Processing*, vol. 39, pp. 1753–1769, Aug. 1991.

[4] C. H. Tseng, K. L. Teo, A. Cantoni, and Z. Zang, "A dual approach to continuous-time enveloped-constrained filter design via orthonormal filters," *IEEE Trans. Circuits Syst. I*, vol. 46, pp. 1042–1054, 1999.

[5] S. Ito, Y. Liu, and K. L. Teo, "A dual parametrization method for convex semi-infinite programming," in *Proc. Int. Conf. Optim. Techn. Appl.*, vol. 2. Perth, Australia, June 1998, pp. 550–557.

[6] S. P. Wu, S. Boyd, and L. Vandenberghe, "FIR filter design via spectral factorization and convex optization," in *Applied and Computational Control, Signal and Circuits*. Boston, MA: Birkhauser, 1997, pp. 55–81.

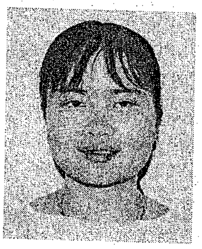
[7] T. W. Parks and C. S. Burrus, *Digital Filter Design*. New York: Wiley, 1987.

[8] T. Hoang, *Convex Analysis and Global Optimization*. Boston, MA: Kluwer, 1998.

[9] D. G. Luenberger, *Optimization by Vector Space Method.* New York: Wiley], 1969.

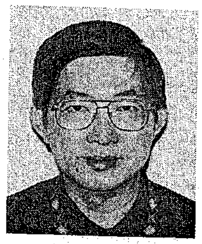
[10] R. T. Rockafella, *Convex Analysis* Princeton, NJ, 1970.

[11] E. Kreyszig, *Advanced Engineering Mathematics*. New York: Wiley, 1993.



Hai Huyen Dam was born in Vietnam in 1974. She received the B.S. degree (with first class honors) from Curtin University of Technology, Perth, Australia in 1996. She is pursuing the Ph.D. degree at the Australian Telecommunications Research Institute (ATRI) in optimal filter design.

Since May 1999, she has been on leave at the University of Karksrona/Ronneby, Ronneby, Sweden, as a Visiting Research Associate.

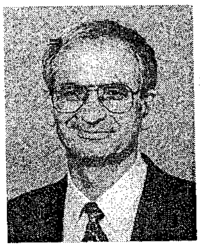


Kok Lay Teo (SM'87) received the Ph.D. degree in electrical engineering from the University of Ottawa, Ottawa, Ont., Canada in 1974.

He was with the Department of Applied Mathematics, University of New South Wales, Sydney, Australia, from 1974 to 1985, as a Lecturer and subsequently as a Senior Lecturer. He was an Associate Professor with the Department of Industrial and Systems Engineering, National University of Singapore from 1985 to 1987. He was an Associate Professor with the Department of Mathematics, The University of Western Australia, Nedlands, Australia, from 1988 to 1996. He was Professor of applied mathematics with Curtin University of Technology, Perth, Australia, from 1996 to 1998. He is currently the Chair Professor of applied mathematics and Head of the Department of Applied Mathematics at the Hong Kong Polytechnic University. He has delivered five keynote lectures, two fully funded invited lectures, and published four books and numerous journal and conference papers. The software package MISER3.2, which solves general constrained optimal control problems, was developed by the research team under his leadership. His research interests includes both the theoretical and practical aspects of optimal control and optimization and their applications, in particular, signal processing in telecommunications.

Sven Nordebo (M'95) received the M.S. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden, in 1989 and the Ph.D. degree in electrical engineering from Luleå University of Technology, Luleå, Sweden, in 1995.

Since 1996, he has been a Senior Lecturer with the University of Karlskrona/Ronneby, Ronneby, Sweden. During 1997 and 1998, he was with the Australian Telecommunications Research Institute (ATRI), Curtin University of Technology, Perth, Australia, working as a Post-doctoral Research Fellow. His research interests are in array signal processing, digital filter design, and optimization.



Antonio Cantoni (F'98) was born in Soliera, Italy, on October 30, 1946. He received the B.E. (with first class honors) and Ph.D. degrees from the University of Western Australia, Nedlands, in 1968 and 1972, respectively.

He was a Lecturer in computer science at the Australian National University, Canberra, in 1972. He joined the Department of Electrical and Electronic Engineering, the University of Newcastle, Shortland, Australia, in 1973, where he held the Chair of Computer Engineering until 1986. In 1987, he joined QPSX Communications Ltd., Perth, Australia, as Director of the Digital and Computer Systems Design Section for the development of the DQDB Metropolitan Area Network. From 1987 to 1990, he was also a Visiting Professor with the Department of Electrical and Electronic Engineering, University of Western Australia. From 1992 to 1997, he was the Director of the Australian Telecommunications Research Institute and Professor of Telecommunications at Curtin University of Technology, Perth. During this period, he was also the director of the Cooperative Research Centre for Broadband Telecommunications and Networking. He is currently Chief Technology Officer with Atmosphere Networks, Inc. and Professor of Telecommunications at the University of Western Australia. He is interested in adaptive signal processing, electronic system design, and networking and regularly acts as a consultant to industry in these areas.

Dr. Cantoni is a Fellow the Australian Academy of Technological Sciences and Engineering. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.