A Discrete Filled Function Method for the Design of FIR Filters With Signed-Powers-of-Two Coefficients

Zhi Guo Feng and Kok Lay Teo, *Senior Member, IEEE*

***Abstract*—In this paper, we consider the optimal design of finite-impulse response (FIR) filters with coefficients expressed as sums of signed powers-of-two (SPT) terms, where the normalized peak ripple (NPR) is taken as the performance measure. This problem is formulated as a mixed-integer programming problem. Based on a transformation between two different integer spaces and the computation of the optimal scaling factor for a given set of coefficients, this mixed integer programming problem is transformed into an equivalent integer programming problem. Then, an efficient algorithm based on a discrete filled function is developed for solving this equivalent problem. For illustration, some numerical examples are solved.**

***Index Terms*—Discrete filled function, finite-impulse response (FIR) filter, signed powers-of-two (SPT).**

I. Introduction

DIGITAL filters with coefficients expressed as sums of signed powers-of-two (SPT) have been widely studied in the literature due to their ease in implementation. The optimal design of finite-impulse response (FIR) filters with SPT terms is a mixed-integer optimization problem and can be transformed into an integer programming problem. Several optimization methods are now available in the literature to deal with this class of problems. For example, mixed integer linear programming (MILP) is used in [1]–[3], some stochastic methods, such as simulated annealing (SA) and genetic algorithm (GA) are applied in [4]–[6], respectively. In [7], a polynomial-time algorithm is presented based on the relationship between the coefficients and its frequency response. An improved version of this algorithm is reported in [8]. Searching method, such as those reported in [9] and [10], are based on trellis or enumeration search. Other methods or techniques are reported in [11]–[15].

The concept of a filled function was first introduced in [16] for global optimization with continuous variables. It searches for a global minimizer among the local minimizers by means of a function, which is called a filled function. A discrete filled function method was developed in [17] for solving a discrete global optimization problem. It searches for a local minimizer by a local search method. Then, a discrete filled function is constructed, from which a better local minimizer, if it exists, is obtained. In this paper, we formulate the problem and transform it into an equivalent integer programming problem. We then develop an efficient algorithm, by incorporating a procedure for choosing initial points, a discrete steepest descent method, and a discrete filled function.

The rest of this paper is organized as follows. In Section II, we present the problem formulation. In Section III, we transform the original problem into an equivalent integer programming problem. An algorithm is then developed to solve this problem in Section IV. Several examples are solved in Section V and the numerical results obtained are compared with those obtained by other methods.

Manuscript received November 30, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Hakan Johansson. This paper was supported by the Australian Research Council under an ARC Grant.

Z. G. Feng is with College of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China, and also with the Department of Mathematics and Statistics, Curtin University of Technology, Perth 6845, Australia (e-mail: z.feng.scholar@gmail.com).

K. L. Teo is with the Department of Mathematics and Statistics, Curtin University of Technology, Perth 6845, Australia (e-mail: k.l.teo@curtin.edu.au).

Digital Object Identifier 10.1109/TSP.2007.901164

II. Problem Formulation

Ignoring the linear phase term, the frequency response of a linear phase FIR filter is given by

$$H(\omega) = h^{\mathsf{T}} C(\omega) \tag{2.1}$$

where $h = (h(0), \ldots, h(m))^{\mathsf{T}}$, and m, which depends on whether the filter length L is odd or even, is defined by $m = (L-1)/2$ if L is odd and $m = (L-2)/2$ if L is even. $C(\omega)$ is an appropriate cosine function vector.

Suppose that the wordlength is b -bit. Then, the coefficients $h(n), n = 0, \ldots, m,$ are expressed as

$$h(n) = \sum_{i=1}^{b} s_{i,n} 2^{-i} \tag{2.2}$$

where $s_{i,n} \in \{-1, 0, 1\}$. Let N_1 denote the total allowable number of the SPT terms used. Then, we have the constraint

$$\sum_{i=1}^{b} \sum_{n=0}^{m} |s_{i,n}| \le N_1. \tag{2.3}$$

In some applications, the following constraints may also be required:

$$\sum_{i=1}^{b} |s_{i,n}| \le N_2, \quad n = 0, \ldots, m \tag{2.4}$$

where N_2 denotes the allowable number of SPT terms for each coefficient $h(n)$.

The normalized peak ripple (NPR) is given by

$$E(h, v) = \max_{\omega \in \mathcal{F} = \mathcal{P} \cup \mathcal{S}} W(\omega) \left| D(\omega) - H(\omega)/v \right|, \quad v \in \mathbb{R}^{+} \tag{2.5}$$

where $\mathbb{R}^+ = (0, +\infty)$, υ is a real scaling factor, $\mathcal{P}$ and $\mathcal{S}$ are, respectively, the passband and the stopband regions, and $W(\omega)$ is the frequency weighting function given by

$$W(\omega) = \begin{cases} 1, & \omega \in \mathcal{P} \\ \gamma(\omega), & \omega \in \mathcal{S} \end{cases} \tag{2.6}$$

while $\gamma(\omega)$ is a function on $\mathcal{S}$ and $D(\omega)$ is the target response given by

$$D(\omega) = \begin{cases} 1, & \omega \in \mathcal{P} \\ 0, & \omega \in \mathcal{S}. \end{cases} \tag{2.7}$$

$\mathcal{F} = \mathcal{P} \cup \mathcal{S} \subseteq [0, (1/2)]$ is a compact subset.

The optimal filter design problem may now be formulated formally as follows.

Problem 1: Find a pair (s, υ), where $s = \{s_{i,n}, i = 1, \ldots, b, n = 0, \ldots, m\}$ and $\upsilon \in \mathbb{R}^+$, such that the normalized peak ripple (2.5) is minimized, subject to the constraint (2.3) or the constraints (2.3) and (2.4).

III. Problem Transformation

Let us first construct a transformation between two integer spaces, which is then used to convert Problem 1 to an equivalent integer programming problem.

Ignoring the constraints (2.3) and (2.4), it is easy to see that the set of all $h(n)$ is $\{2^{-b}i, i = 1 - 2^b, \ldots, 2^b - 1\}$, for each n. That is, for any $\bar{h} \in \{2^{-b}i, i = 1 - 2^b, \ldots, 2^b - 1\}$ there exists a vector $\bar{s} = (\bar{s}_i)$, $\bar{s}_i \in \{-1, 0, 1\}$, such that

$$\bar{h} = \sum_{i=1}^{b} \bar{s}_i 2^{-i}. \tag{3.1}$$

Then, when the wordlength is taken as b -bit, the number of SPT terms for $\bar{h}$ is defined as

$$P_h(\bar{h}, b) = \min_{\bar{s}} \sum_{i=1}^{b} |\bar{s}_i|$$
$$s.t. \quad \bar{h} = \sum_{i=1}^{b} \bar{s}_i 2^{-i}. \tag{3.2}$$

We now introduce an integer vector $x = (x(0), \ldots, x(m))^\top$ such that

$$x = 2^b h. \tag{3.3}$$

Then, for each n, $x(n) \in \{1 - 2^b, \ldots, 2^b - 1\}$. Let $\mathcal{X}$ denote the space of all such x. Thus, when the wordlength is taken as b -bit, the number of SPT terms for $\bar{x}$ can be defined by $P_x(\bar{x}, b) = P_h(2^{-b}\bar{x}, b)$.

Clearly, the total number of SPT terms for all coefficients and the number of SPT terms for each coefficient are bounded by

$$\sum_{n=0}^{m} P_x(x(n), b) \le N_1 \tag{3.4}$$
$$P_x(x(n), b) \le N_2, \quad n = 0, \ldots, m. \tag{3.5}$$

The frequency response (2.1) is equivalent to

$$X(\omega) = 2^{-b} x^\top C(\omega) \tag{3.6}$$

and the normalized peak ripple (NPR) (2.5) is equivalent to

$$E(x, \upsilon) = \max_{\omega \in \mathcal{F} = \mathcal{P} \cup \mathcal{S}} W(\omega) \left| D(\omega) - X(\omega)/\upsilon \right|, \quad \upsilon \in \mathbb{R}^+. \tag{3.7}$$

For any $x \in \mathcal{X}$, we can find an optimal real scaling factor, which is denoted by $\upsilon(x)$, that is

$$\min_{x,\upsilon} E(x, \upsilon) = \min_{x} E(x, \upsilon(x)). \tag{3.8}$$

To deal with the constraints, we use the penalty function method. Let Q_1 and Q_2 denote two sufficiently large positive real numbers, the constraints (3.4) and (3.5) are replaced with

$$g_1(x) = \max\left\{0, Q_1\left(\sum_{n=0}^{m} P_x(x(n), b) - N_1\right)\right\} \tag{3.9}$$

and

$$g_{2,n}(x) = \max\{0, Q_2(P_x(x(n), b) - N_2)\}, \quad n = 0, \ldots, m. \tag{3.10}$$

Then, the objective function becomes

$$\bar{E}(x) = 20 * \log_{10} E(x, \upsilon(x)) + g_1(x) + \sum_{n=0}^{m} g_{2,n}(x) \tag{3.11}$$

where the error function is in dB scale.

Thus, Problem 1 can be stated equivalently as follows.

Problem 2: Find an $x \in \mathcal{X}$ such that $\bar{E}(x)$, which is defined by (3.11), is minimized.

Remark 1: There are several methods available in the literature for computing $P_x(\bar{x}, b)$ and converting an integer into signed digit code. The method we apply here is a recursive function method, which is given in Appendix A.

Remark 2: To compute $\upsilon(x)$, we suppose that

$$\min_{\omega \in \mathcal{P}} X(\omega) > 0. \tag{3.12}$$

Then, $\upsilon(x)$ is given by

$$\upsilon(x) = \max\{\upsilon_1, \upsilon_2\} \tag{3.13}$$

where

$$\upsilon_1 = \left(\min_{\omega \in \mathcal{P}} X(\omega) + \max_{\omega \in \mathcal{P}} X(\omega)\right)/2 \tag{3.14}$$
$$\upsilon_2 = \min_{\omega \in \mathcal{P}} X(\omega) + \max_{\omega \in \mathcal{S}} |\gamma(\omega) X(\omega)|. \tag{3.15}$$

The proof of this result is similar to that given in [12] or [18].

Remark 3: If condition (3.12) is not satisfied, then the NPR value is larger and hence x cannot, in general, be an optimal solution. For this case, it is not necessary to compute $\upsilon(x)$ and we set $E(x, \upsilon(x))$ to a sufficiently large value.

IV. Computational Algorithm

To solve Problem 2, we develop a two-step algorithm. The first step is a local search, while the second step is a global search.

A. Local Search

The steepest descent algorithm that we are proposing starts from a point in $\mathcal{X}$. Then, by searching over its neighborhood, we select the point which produces the largest reduction in the value of the objective function (3.11). Note that x is an $(m+1)$ -dimensional integer variable, its neighborhood is defined in the following.

Definition 1: For any $x \in \mathcal{X}$, the neighborhood of x is defined by

$$\mathcal{N}(x) = \{x, x \pm e_i : i = 0, 1, \ldots, m\}$$

where e_i is the i th unit vector (the $(m+1)$ -dimensional vector with the i th component equal to one and all other components equal to zero).

If we have found a point which minimizes the objective function (3.11) over its neighborhood, then the local search stops and the point obtained is called a local minimizer. The precise definition of local minimizers is given as follows.

Definition 2: A point x^* is called a local minimizer of $\bar{E}$ over $\mathcal{X}$ if $\bar{E}(x^*) \leq \bar{E}(x) \quad \forall x \in \mathcal{X} \cap \mathcal{N}(x^*)$.

Based on the previous two definitions given, we present a discrete steepest descent algorithm to search for a local minimizer as follows as an algorithm.

Algorithm 1

1) Start from an initial point $x^* = x_0$. Compute the objective function value $\bar{E}$. Set $k = 1$.
2) For each point $x \in \mathcal{X} \cap \mathcal{N}(x^*) \setminus \{x^*\}$, where $A_1 \setminus A_2 = \{x \in A_1 : x \notin A_2\}$, compute the corresponding objective function value $\bar{E}(x)$. Suppose x^k is such that $\bar{E}(x^k)$ is the minimum. If $\bar{E}(x^k) \geq \bar{E}(x^*)$, then x^* is a local minimizer of $\bar{E}$ and stop, else goto Step 3.
3) Set $x^* = x^k$ and $k = k + 1$. Goto Step 2.

B. Global Search

With Algorithm 1, we can find a local minimizer from any initial point. But for this problem, there exist many local minimizers and not all of them are useful in practice. Thus, we shall derive a discrete filled function method to search for a global solution.

We introduce the following function based on the one constructed in [17]:

$$F_{\mu,\rho}(x; x^*) = \mu \left[\bar{E}(x) - \bar{E}(x^*)\right]^2 - \rho \|x - x^*\|^2, \quad \text{if} \quad \bar{E}(x) \geq \bar{E}(x^*) \tag{4.1}$$

where $\|\cdot\|$ denotes the usual Euclidean norm. When $\rho > 0$, $0 < \mu < \rho/K$ (K is a sufficiently large real number), (4.1) is called a discrete filled function.

It is not necessary to define the function F when $\bar{E}(x) < \bar{E}(x^*)$. For in this case, if the discrete steepest descent method is used directly with x as the initial point, we will obtain a local minimizer, which is better than x^* with reference to the objective function (3.11).

The search according to this discrete filled function (4.1) takes place as follows. With the starting point x^*, the $\mu[\bar{E}(x) - \bar{E}(x^*)]^2$ term favors a solution with lower objective function value while the $-\rho\|x - x^*\|^2$ term favors a solution far away from x^*. Combining the effects, the discrete filled function favors a solution whose objective function value is not too much greater than that of x^* and at a considerable distance away from x^*. The idea behind the search strategy is to direct its search towards the direction with the least increase in the objective function value. This idea is somewhat similar to that proposed in [12].

To address the situation when we fail to find a point x such that $\bar{E}(x) < \bar{E}(x^*)$ using the discrete filled function (4.1) as the objective function, we choose a positive integer number n_s. When the number of searching steps is greater than n_s, we stop and return the optimal solution. This is because for the problem considered in this paper, we can obtain the optimal infinite precision solution (continuous solution) by the Remez exchange algorithm [19]. Since the optimal solution should be in the neighborhood of the continuous solution, it means that the optimal solution, if it exists, is "far" away from the continuous solution when the number of searching steps is greater than n_s. This situation is unlikely to occur.

C. Initial Points

Now, we can devise an efficient computational method, by combining the discrete steepest descent method and the discrete filled function method, to search for the optimal solution. But first, let us find some "good" initial points. Suppose that the infinite precision solution of Problem 1 is $\{\hat{h}(n), n = 0, 1, \ldots, m\}$ corresponding to $\upsilon = 1$. Let $\{\hat{x}(n) = 2^b \hat{h}(n), n = 0, 1, \ldots, m\}$ be the rounded values of the infinite precision coefficients scaled by 2^b. Define

$$\bar{x}_0 = g_3(\bar{x}) = \begin{cases} [\bar{x} + 0.5], & \text{if } |\bar{x}| \leq 2^b - 1 \\ 2^b - 1, & \text{if } \bar{x} > 2^b - 1 \\ 1 - 2^b, & \text{if } \bar{x} < 1 - 2^b \end{cases} \tag{4.2}$$

where $[\bar{x}]$ denotes the largest integer less than or equal to $\bar{x}$. Then, the initial point $x_0 \in \mathcal{X}$ nearest to $\hat{x}$ is given by

$$x_0 = \left(g_3\left(\hat{x}(0)\right), \ldots, g_3\left(\hat{x}(m)\right)\right)^{\mathsf{T}}. \tag{4.3}$$

The scaling factor υ takes its value from the interval $[\upsilon_l, \upsilon_u]$, where υ_l and υ_u are given lower and upper bounds for υ. Then, for each $\upsilon \in [\upsilon_l, \upsilon_u]$, the corresponding infinite precision solution is $\upsilon\hat{x}$ and the nearest point to $\upsilon\hat{x}$ in $\mathcal{X}$ can be found by using (4.3), which is then used as the initial point x_0.

In fact, although there is an infinite number of υ between υ_l and υ_u, not all the values of υ will produce distinct sets of coefficients by using (4.3). Each distinct set of coefficients corresponds to an interval of values of υ. When υ crosses the end points of these intervals, the set of coefficients changes abruptly from one to another. Let all these end points be arranged in the ascending order and let them be denoted as

$$\left\{\upsilon^{(i)}, i = 0, \ldots, m_\upsilon\right\}. \tag{4.4}$$

Then, for each $i = 0, \ldots, m_\upsilon - 1$, $x_0 = (x_0(0), \ldots, x_0(m))$ takes the same value for any $\upsilon \in (\upsilon^{(i)}, \upsilon^{(i+1)})$. Thus, all the initial points can be computed when we choose one of υ in $(\upsilon^{(i)}, \upsilon^{(i+1)})$, for each $i = 0, \ldots, m_\upsilon - 1$.

The computation of υ^i is given in Appendix B.

D. Algorithm

We present the following algorithm to solve Problem 2.

Algorithm 2

1) Initialize the parameters μ, ρ, n_s, Q_1, Q_2.
2) Calculate the infinite precision solution $\bar{h}$ corresponding to $v = 1$ by using Remez exchange algorithm [19]. Compute $\bar{x} = 2^b \bar{h}$. Then, for each $v \in [v_l, v_u]$, the infinite precision solution corresponding to v is $v\bar{x}$. Compute all the initial points $\mathcal{X}_0 = \{x_0^k, k = 1, 2, \ldots, m_v\}$ nearest to $v\bar{x}$, $v \in [v_l, v_u]$. Set $k = 1$.
3) Select $x_0 = x_0^k \in \mathcal{X}_0$ as the initial point. Let $J(k)$ denote the optimal value when x_0^k is used as the initial point. Set $J(k)$ to a sufficiently large value.
4) Apply Algorithm 1 to obtain a local minimizer of the objective function $\bar{E}$ with the initial point x_0. Let the local minimizer be denoted as x^*. Modify the optimal value as $J(k) = \bar{E}(x^*)$.
5) Set x^* as the initial point, that is, $x_0 = x^*$, and apply Algorithm 1 to search for a point better than the current local minimizer x^*, with the discrete filled function F used as the objective function.
6) Let l be the number of searching steps in Step 5. If a point x is found such that $\bar{E}(x) < \bar{E}(x^*)$ when $l \leq n_s$ during Step 5, then stop searching. Set the initial point as $x_0 = x$ and goto Step 4. Else, stop searching and return the optimal value $J(k)$. Hence, when the initial point is x_0^k, the optimal value $J(k)$ is found. Set $k = k + 1$.
7) If each point in $\mathcal{X}_0$ has been used as initial point, then goto Step 8, else select $x_0 = x_0^k \in \mathcal{X}_0$ as the initial point and goto Step 4.
8) Compare the optimal value $J(k)$ for every initial point x_0^k and select the best. Stop.

V. Simulation Results

The proposed method has been used to solve many examples. The results obtained are consistently favorable when compared with results obtained by other methods. In this section, we only present the results obtained for three examples. The computation was performed in Compaq Visual Fortran double precision. The coefficients are set to be

$$\mu = 10^{-6}, \rho = 1, n_s = 30, Q_1 = Q_2 = 5.$$

Example 1: The filter length is 71 with the normalized passband and stopband edge frequencies of 0.11 and 0.137. The ripple weighting factor is $W = 1$ and the wordlength is 8.

We set $N_1 = 51$. The solution obtained is

$$\begin{aligned} h = (&2^{-7}, 2^{-7}, 2^{-8}, 0, -2^{-6}, -2^{-6}, -2^{-7}, 2^{-8}, 2^{-6}, 2^{-6},\\ &2^{-7}, -2^{-8}, -2^{-6} - 2^{-8}, -2^{-5}, -2^{-6}, 2^{-7},\\ &2^{-5}, 2^{-5} + 2^{-7}, 2^{-5} - 2^{-8}, -2^{-7}, -2^{-5} - 2^{-7},\\ &-2^{-4}, -2^{-5} - 2^{-7}, 2^{-7}, 2^{-4}, 2^{-3} - 2^{-5} - 2^{-7}, 2^{-4},\\ &-2^{-8}, -2^{-3} + 2^{-5} - 2^{-7}, -2^{-3} - 2^{-5},\\ &-2^{-3}, 2^{-7}, 2^{-2} - 2^{-6}, 2^{-1} - 2^{-6}, 2^{-1} + 2^{-2}\\ &-2^{-4} - 2^{-8}, 2^{-1} + 2^{-2}).\end{aligned}$$

TABLE I
RESULTS FOR EXAMPLE 1

Method	NPR(dB)/SPT
[7]	-36.54/59
[9]	-36.52/51
proposed method	-37.25/51

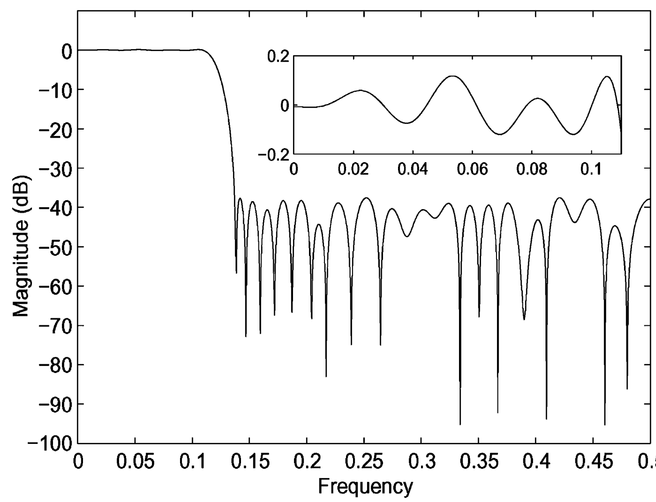


Fig. 1. Frequency response of the filter considered in Example 1.

TABLE II
RESULTS FOR EXAMPLE 2

Method	PBR	SBA(dB)	SPT
[11]	0.17	-30.3	20
[7]	0.16	-33.8	19
[8]	0.15	-31.3	17
[9]	0.17	-30.7	16
	0.16	-34.4	17
	0.13	-31.3	17
	0.10	-33.5	19
	0.13	-34.8	19
	0.11	-35.4	20
proposed method	0.17	-31.2	16
	0.16	-34.4	17
	0.13	-31.6	17
	0.10	-34.1	19
	0.13	-36.9	19
	0.11	-35.4	20

PBR: passband ripple; SBA: stopband attenuation.

Our results and those obtained by other methods are given in Table I and Fig. 1.

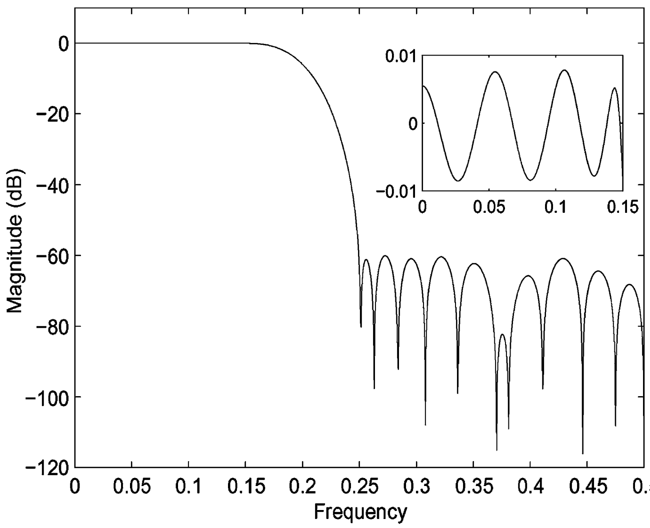
Example 2: The filter length is 28 with the normalized passband and stopband edge frequencies of 0.128 and 0.2048. In order to compare our results with other known results, the objective function (2.5) is replaced by

$$E(h, v) = \begin{cases} \max_{\omega \in \mathcal{S}} W(\omega) \left| \frac{H(\omega)}{v} \right|, & \text{when } \max_{\omega \in \mathcal{P}} W(\omega) \left| 1 - \frac{H(\omega)}{v} \right| \leq E_{\text{pbr}} \\ +\infty, & \text{when } \max_{\omega \in \mathcal{P}} W(\omega) \left| 1 - \frac{H(\omega)}{v} \right| > E_{\text{pbr}} \end{cases}$$

where E_{pbr} is a given value for the limitation of the passband ripple. Results obtained and those obtained by other methods are presented in Table II.

TABLE III
RESULTS FOR EXAMPLE 3

Method	NPR(dB)	SPT	Filter Length	N_2
[10]	-50.07	58	28	4
	-50.12	60	28	3
	-60.13	80	34	4
proposed method	-50.23	56	28	4
	-50.14	60	28	3
	-60.15	74	34	4



$h(0) = 2^{-9}$ $\quad$ $h(1) = 2^{-9}$

$h(2) = -2^{-9} - 2^{-10}$ $\quad$ $h(3) = -2^{-7} + 2^{-10} + 2^{-12}$

$h(4) = 0$ $\quad$ $h(5) = 2^{-6} - 2^{-9}$

$h(6) = 2^{-6} - 2^{-8} - 2^{-11}$ $\quad$ $h(7) = -2^{-6} + 2^{-10} - 2^{-12}$

$h(8) = -2^{-5} - 2^{-11}$ $\quad$ $h(9) = 0$

$h(10) = 2^{-4} - 2^{-7} - 2^{-9} + 2^{-11}$ $\quad$ $h(11) = 2^{-5} + 2^{-7} + 2^{-8} - 2^{-12}$

$h(12) = -2^{-4} + 2^{-7} - 2^{-9}$ $\quad$ $h(13) = -2^{-3} - 2^{-11}$

$h(14) = 0$ $\quad$ $h(15) = 2^{-2} + 2^{-4} + 2^{-8}$

$h(16) = 2^{-1} + 2^{-4} + 2^{-5} + 2^{-9}$

Fig. 2. Frequency response of the filter considered in Example 3 with filter length 34.

Example 3: The normalized passband and stopband edge frequencies are 0.15 and 0.25 and the ripple weighting factor is $W = 1$. The coefficient wordlength is 12. We consider the filter length in three cases: 28, 34, and 38, and the maximum allowed number N_2 of SPT terms per coefficient is taken as 3 or 4. The respective results and those obtained by other methods are given in Table III and Fig. 2.

VI. CONCLUSION

In this paper, we developed a computational method based on a discrete filled function, a discrete steepest descent method and a procedure for choosing initial points for the design of FIR filters with coefficients expressed as sums of signed powers-of-two terms. From our numerical studies, we observe that the proposed method is highly effective and efficient.

APPENDIX A

A. *Computation of* $P_x(\bar{x}, b)$

Function $(P_x(\bar{x}, b))$

1) If $|\bar{x}| \geq 2^b$, then $P_x(\bar{x}, b) = +\infty$. Stop and return $P_x(\bar{x}, b)$.
2) If $\bar{x} = \pm 1$, then $P_x(\bar{x}, b) = 1$. Stop and return $P_x(\bar{x}, b)$.
3) If $|\bar{x}| > 1$ and $\bar{x}$ is odd, then call the functions $P_x(\bar{x} - 1, b)$ and $P_x(\bar{x} + 1, b)$, respectively. Then, let $P_x(\bar{x}, b) = 1 + \min\{P_x(\bar{x} - 1, b), P_x(\bar{x} + 1, b)\}$. Stop and return $P_x(\bar{x}, b)$.
4) If $\bar{x}$ is even, then call the function $P_x(\bar{x}/2, b - 1)$. Then, $P_x(\bar{x}, b) = P_x(\bar{x}/2, b - 1)$. Stop and return $P_x(\bar{x}, b)$.

B. *Converting an Integer Into Signed Digit Code*

Let $\Lambda(\bar{x}, b)$ denote the minimal number of signed powers of two terms of $\bar{x}$ when the wordlength is taken as b -bit. Generally, $\Lambda(\bar{x}, b)$ is not unique (see [20]).

Function $(\Lambda(\bar{x}, b))$

1) If $|\bar{x}| \geq 2^b$, then $\Lambda(\bar{x}, b) = \emptyset$. Stop.
2) If $\bar{x} = 1$, then $\Lambda(\bar{x}, b) = \{2^0\}$. If $\bar{x} = -1$, then $\Lambda(\bar{x}, b) = \{-2^0\}$. Stop and return $\Lambda(\bar{x}, b)$.
3) If $|\bar{x}| > 1$ and $\bar{x}$ is odd, then call the functions $P_x(\bar{x} - 1, b)$, $P_x(\bar{x} + 1, b)$, and $\Lambda(\bar{x} - 1, b)$, $\Lambda(\bar{x} + 1, b)$. If $P_x(\bar{x}, b) = 1 + P_x(\bar{x} + 1, b)$, then let $\Lambda(\bar{x}, b) = \Lambda(\bar{x} + 1, b) \cup \{-2^0\}$, else if $P_x(\bar{x}, b) = 1 + P_x(\bar{x} - 1, b)$, then let $\Lambda(\bar{x}, b) = \Lambda(\bar{x} - 1, b) \cup \{2^0\}$. Stop and return $\Lambda(\bar{x}, b)$.
4) If $\bar{x}$ is even, then call the function $\Lambda(\bar{x}/2, b - 1)$. Suppose that $\Lambda(\bar{x}/2, b - 1) = \{s_{i_k} 2^{i_k}, k = 1, \ldots\}$, then let $\Lambda(\bar{x}, b) = \{s_{i_k} 2^{i_k + 1}, k = 1, \ldots\}$. Stop and return $\Lambda(\bar{x}, b)$.

APPENDIX B
PROCEDURE FOR CHOOSING INITIAL POINTS

From (4.2), we see that for each $n = 0, \ldots, m$, if $\upsilon' \hat{x}(n)$ and $\upsilon'' \hat{x}(n)$ belong to a same interval in

$$I = \{(-\infty, 0.5 - 2^b), (2^b - 0.5, +\infty), (i - 0.5, i + 0.5), \\ i = 1 - 2^b, \ldots, 2^b - 1\}$$

then their corresponding $x_0(n)$ take the same values. On the other hand, if $\upsilon' \hat{x}(n)$ and $\upsilon'' \hat{x}(n)$ belong to different intervals in I, then their corresponding $x_0(n)$ take different values.

If $\hat{x}(n) = 0$, then $\upsilon \hat{x} = 0 \in (-0.5, 0.5)$ for any υ and we obtain $x_0(n) = 0$.

If $\hat{x}(n) > 0$, then $\upsilon \hat{x}(n) \in [\upsilon_l \hat{x}(n), \upsilon_u \hat{x}(n)]$. Hence, for any υ such that $\upsilon \hat{x}(n)$ in the same interval in $[\upsilon_l \hat{x}(n), \upsilon_u \hat{x}(n)] \cap I$, $x_0(n)$ takes the same value. Then, the critical values of $v\hat{x}(n)$ are the end points of all the intervals in $[v_l\hat{x}(n), v_u\hat{x}(n)] \cap I$, that is

$$B \cap [v_l\hat{x}(n), v_u\hat{x}(n)]$$

where

$$B = \{0.5 - 2^b, \ldots, 0.5, 1.5, \ldots, 2^b - 0.5\}.$$

Thus, the critical values of v are $B_1 \cap [v_l, v_u]$, where

$$B_1 = \left\{ \frac{0.5 - 2^b}{\hat{x}(n)}, \ldots, \frac{0.5}{\hat{x}(n)}, \frac{1.5}{\hat{x}(n)}, \ldots, \frac{2^b - 0.5}{\hat{x}(n)} \right\}.$$

Similarly, if $\hat{x}(n) < 0$, the critical values of v are obtained as $B_1 \cap [v_l, v_u]$.

Then, for each $n = 0, \ldots, m$, we mix all these critical values of v up in the ascending order. Thus, we obtain (4.4).

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their valuable comments and suggestions.

REFERENCES

[1] Y. C. Lim and S. R. Parker, "FIR filter design over a discrete powers-of-two coefficient space," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-31, no. 3, pp. 583–591, Jun. 1983.
[2] Y. C. Lim, "Design of discrete-coefficient-value linear phase FIR filters with optimum normalized peak ripple magnitude," *IEEE Trans. Circuits Syst.*, vol. 37, no. 12, pp. 1480–1486, Dec. 1990.
[3] O. Gustafsson, H. Johansson, and L. Wanhammar, "An MILP approach for the design of linear-phase FIR filters with minimum number of signed-power-of-two terms," in *Proc. Euro. Conf. Circuit Theory Des.*, 2001, pp. 217–220.
[4] N. Benvenuto, M. Marchesi, and A. Uncini, "Applications of simulated annealing for the design of special digital filters," *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 323–332, Feb. 1992.
[5] R. Cemes and D. Ait-Boudaoud, "Genetic approach to design of multiplierless FIR filters," *Electron. Lett.*, vol. 29, pp. 2090–2091, 1993.
[6] P. Persson, S. Nordebo, and I. Claesson, "Design of discrete coefficient FIR filters by a fast entropy-directed deterministic annealing algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 1006–1014, Mar. 2005.
[7] D. Li, J. Song, and Y. C. Lim, "A polynomial-time algorithm for designing digital filters with power-of-two coefficients," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1993, pp. 84–87.
[8] C. L. Chen, K. Y. Khoo, and A. N. Willson, "An improved polynomial-time algorithm for designing digital filters with power-of-two coefficients," in *Proc. IEEE Int. Symp. Circuits Syst.*, 1995, pp. 223–226.
[9] T. Ciloglu, "Design of FIR filters for low implementation complexity," *Electron. Lett.*, vol. 46, pp. 577–584, 1999.
[10] C. L. Chen and A. N. Willson, "A trellis search algorithm for the design of FIR filters with signed-powers-of-two coefficients," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 46, no. 1, pp. 29–39, Jan. 1999.
[11] D. Ait-Boudaoud and R. Cemes, "Modified sensitivity criterion for the design of powers-of-two FIR filters," *Electron. Lett.*, vol. 29, pp. 1467–1469, 1993.
[12] T. Ciloglu, "An efficient local search method guided by gradient information for discrete coefficient FIR filter design," *Signal Process.*, vol. 82, no. 10, pp. 1337–1350, 2002.
[13] D. E. Quevedo and G. C. Goodwin, "Moving horizon design of discrete coefficient FIR filters," *IEEE Trans. Signal Process.*, vol. 53, no. 6, pp. 2262–2267, Jun. 2005.
[14] T. Ciloglu, "New initialization methods for discrete coefficient FIR filter design with coefficient scaling and the use of scale factor in the design process," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 796–800, Feb. 2006.
[15] W. R. Lee, L. Caccetta, K. L. Teo, and V. Rehbock, "Optimal design of complex FIR filters with arbitrary magnitude and group delay responses," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1617–1628, May 2006.
[16] R.-P. Ge, "A filled function method for finding a global minimizer of a function of several variables," *Math. Program.*, vol. 46, pp. 191–204, 1990.
[17] C. K. Ng, L. S. Zhang, D. Li, and W. W. Tian, "Discrete filled function method for discrete global optimization," *Comput.Opt. Appl.*, vol. 31, pp. 87–115, 2005.
[18] W. R. Lee, V. Rehbock, K. L. Teo, and L. Caccetta, "A global approach to discrete-coefficient-value FIR filter design with optimum NPWR," WACEIO, Dept. Math. Stat., Curtin Univ. Technol., Tech. Rep. 1/02, 2002.
[19] T. W. Parks and J. H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuit Syst.*, vol. CAS-19, no. 2, pp. 189–194, Mar. 1972.
[20] Y. C. Lim, R. Yang, D. Li, and J. Song, "Signed power-of-two term allocation scheme for the design of digital filters," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 35, pp. 529–530, 1999.

Zhi Guo Feng was born in Guangdong, China, in 1977. He received the B.S. degree in mathematics and the Ph.D. degree in operational research and control theory from Sun Yat-Sen University, Guangzhou, China, in 2000 and 2006, respectively.

He is currently a visiting Research Associate with Curtin University of Technology, Perth, Australia. Since 2006, he has been with the College of Mathematics and Computer Science, Chongqing Normal University, Chongqing, China. His research interests include theory and application of optimization and optimal control in areas, such as sensor scheduling, filter design, and hybrid systems.

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

From 1999 to 2004, he was a Chair Professor of Applied Mathematics and the Head of the Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong. He is currently Chair of Applied Mathematics and Head of Department of Mathematics and Statistics of Curtin University of Technology, Perth, Australia. He has published six books and over 300 journal and conference papers. The software package, MISER3.2, for solving general constrained optimal control problems was developed by the research team under his leadership. His research interests include both the theoretical and practical aspects of optimal control and optimization, and their practical applications in areas, such as in signal processing in telecommunications, and financial portfolio optimization.

Dr. Teo is Editor-in-Chief of the *Journal of Industrial and Management Optimization* and Regional Editor of *Nonlinear Dynamics and Systems Theory*. He also serves as an Associate Editor of an additional 14 international journals, including *Automatica*; *Journal of Global Optimization*; *Discrete and Continuous Dynamic Systems (Series A and Series B)*; *Optimization and Engineering*.