A Novel Systematic Approach for Synthesizing Multiplication-Free Highly-Selective FIR Half-Band Decimators and Interpolators

Tapio Saramäki and Juha Yli-Kaakinen
Institute of Signal Processing,
Tampere University of Technology
P. O. Box 553, FI-33101 Tampere, Finland
email: {ts, ylikaaki}@cs.tut.fi

***Abstract*— This paper discloses a systematic approach to generating for sampling rate applications multiplication-free linear-phase finite-impulse response half-band filters based on an article by Saramäki, Karema, Ritoniemi, and Tenhunen. This article counts on the following facts. First, the transfer function of such filters is expressible as a sum of the terms $(1/2)z^{-M}$ and $G(z^2)$, where M is odd and is the order of $G(z)$. Second, $G(z)$ is constructed using identical copies of the same sub-filter that are properly interconnected with the aid a few additional adders and tap coefficients. The first step is to quantize in a simple manner the additional tap coefficient values to be a few powers-of-two terms. The second step is to find the sub-filter coefficient values in the same representation forms. This can be accomplished using a very simple quantization technique because the ripple values for the sub-filter are very huge compared to the overall filter. In addition, efficient structures for implementing both the resulting decimators and interpolators are given. An example in the above-mentioned article is included for illustrating the efficiency of the proposed multiplication-free structures compared with the conventional direct-form designs.**

I. Introduction

Linear-phase finite-impulse response (FIR) half-band filters play a very important role, due to their many attractive properties, in many digital signal processing applications including, among others, signal analysis, instrumentation, and systems where sampling rate conversion is needed. First of all, a low-pass half-band filter transfer function of order $2M$ with M odd is expressible as (see, e.g., [1])

$$H(z) = \sum_{n=0}^{2M} h[n]z^{-n} = \frac{1}{2}z^{-M} + G(z^2), \tag{1a}$$

where

$$G(z) = \sum_{n=0}^{M} g[n]z^{-n}, \quad g[M-n] = g[n]. \tag{1b}$$

This means that $h[M] = 1/2$ and $h[M+2r] = 0$ for $r = \pm1, \pm2, \ldots, \pm M/2$, that is, $h[M]$ is equal to 2^{-1}, being realizable just by one shift operation, and every second coefficient value around this coefficient is automatically zero-valued. Second, constructing $H(z)$ in the above manner guarantees that the resulting filter has a linear-phase response that is of great importance in many applications. Third, when the goal is to split the input signal into output low-pass and high-pass channels with exactly linear-phase responses, the corresponding high-pass filter is obtained from the low-pass counterpart by simply replacing the summation in Eq. (1a) between $G(z^2)$ and $(1/2)z^{-M}$ by the subtraction, as shown in Fig. 1 (see, e.g., [1]).

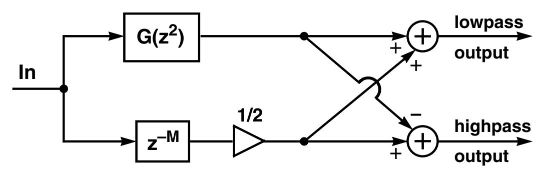

Fig. 1. Efficient implementation for a complementary half-band filter pair.

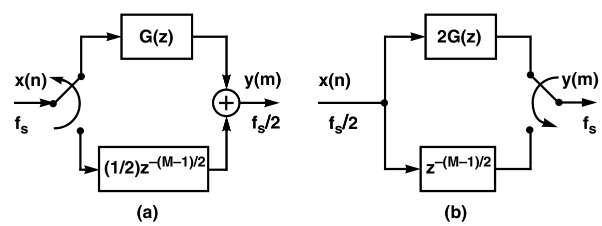

Fig. 2. Efficient commutative implementations of a half-band linear-phase filter for (a) the 2-to-1 decimator and (b) the 1-to-2 interpolator.

Efficients commutative structures [2] of using the above half-band filter as a 2-to-1 decimator and a 1-to-2 interpolator are shown in Figs. 2(a) and 2(b), respectively. In most overall sampling-rate alteration systems, these filters serve as sub-blocks in generating the overall system (see, e.g., [3]).

The key drawback of using the above half-band decimators and interpolators is the fact that in cases, where a high attenuation and high filter orders are required, it is not possible to express all the coefficient values of $G(z)$ in Fig. 2 as a few powers-of-two terms. This is very crucial when implementing the filter as a very large-scale integration (VLSI) circuit because in these cases there is a need to use a very costly multiplier element. This element increases significantly the silicon area and power consumption as well as decreases the maximal achievable sampling rate. On the other hand, when all the coefficients are expressible as a few powers-two terms, they can be implemented as a sequence of shifts and adds and/or subtracts. The shifts are often hardwired and, therefore, essentially free. Thus, only a few adders and/or subtracters are required for implementing each coefficient.

The main idea of this paper is to disclose the secret of designing multiplication-free half-band filters discussed without giving any synthesis details in [3]. This article has received a great interest (see, e.g., [4]–[9]). The outline of this paper is as follows. Section II shows how to construct $G(z)$ in the overall half-band transfer function [cf. Eqs. (1a)–(1b)] as a tapped cascaded interconnection of the same properly constructed sub-filter, as shown in [3] without detailed explanations. Furthermore, the zero-phase frequency response connection between the overall filter and the sub-filter as well as an efficient implementation structures for decimation and interpolations purposes are given. Section III shows a simple design scheme for determining the additional tap coefficient as a few powers of two. Section IV concentrates on how to achieve the results of the paper [3]. Finally, Section V gives the concluding remarks for improving the approach proposed in [3].

II. Proper Generation of $G(z)$ in Terms of Identical Sub-Filters

In [1], [10], [11], it has been shown how to properly generate even-order linear-phase FIR filters having an even symmetrical impulse response as a tapped cascaded interconnection of identical even-order sub-filters. For the present filters, the design is different due to the fact that the order of $G(z)$, as mentioned in the Introduction, should be odd in order to generate a proper linear-phase FIR overall half-band function [cf. Eqs. (1a)–(1b)]. As proposed in [3], a promising alternative in this case is to construct $G(z)$ as follows:

$$G(z) = \sum_{l=0}^{L} a_l z^{-(L-l)K} [F(z)]^{2l+1}, \tag{2a}$$

where

$$F(z) = \sum_{n=0}^{K} f[n] z^{-n}, \quad f[K-n] = f[n]. \tag{2b}$$

Here, K is odd and the impulse response of $F(z)$ has an even symmetry. Most importantly, the delay of each term in the summation of Eq. (2a) is $(2L+1)K/2$, indicating that the resulting $G(z)$ has an even symmetric impulse response and $(2L+1)K$, the overall order, is odd, as is desired. Only the term with $l = L$ in Eq. (2a) has nonzero-valued impulse-response coefficients for $n = 0, 1, \ldots, (2L+1)K$. For the remaining terms in Eq. (2a), there are lK zero-valued impulse response coefficient values in both the beginning and the end of the above-mentioned interval.

The first and second subfigures in Fig. 3 show the decimator and interpolator structures for the overall half-band filter in the $L = 3$ case. These are commutative structures [2], where the delay terms are shared for reducing their use in the implementation.

The zero-phase frequency response of $H(z)$, as given by Eq. (1a), where $G(z)$ is described by Eqs. (2a) and (2b), is expressible as

$$H(\omega) = 1/2 + G(2\omega), \tag{3a}$$

where

$$G(\omega) = \sum_{l=0}^{L} a_l [F(\omega)]^{2l+1} \tag{3b}$$

with

$$F(\omega) = \sum_{n=0}^{(K-1)/2} f\left[\frac{K-1}{2} - n\right] \cos\left[\left(n + \frac{1}{2}\right)\omega\right]. \tag{3c}$$

TABLE I
Connections Between the Coefficients a_l and b_l

$b_0 = a_0 + a_1 + a_2 + a_3$	$b_3 = a_1 + 10a_2 + 35a_3$	$b_6 = 7a_3$
$b_1 = a_0 + 3a_1 + 5a_2 + 7a_3$	$b_4 = 5a_2 + 35a_3$	$b_7 = a_3$
$b_2 = 3a_1 + 10a_2 + 21a_3$	$b_5 = a_2 + 21a_3$	

TABLE II
Proper Selections of the Values of the Tap Coefficients in the $L = 1$, $L = 2$, and $L = 3$ Cases

$L = 1$	$a_0 = 2^0 - 2^{-2}$, $a_1 = -2^{-2}$, $a_2 = 0$, $a_3 = 0$
$L = 2$	$a_0 = 2^0 - 2^{-4}$, $a_1 = -2^{-1} - 2^{-3}$, $a_2 = 2^{-2} - 2^{-4}$, $a_3 = 0$
$L = 3$	$a_0 = 2^0 + 2^{-3} - 2^{-5}$, $a_1 = -2^0 - 2^{-3} + 2^{-5}$, $a_2 = 2^{-1} + 2^{-3} + 2^{-5}$, $a_3 = -2^{-3} - 2^{-5}$

III. Simple Design Technique for Determining the Additional Tap Coefficient to be Expressible as a Few Powers of Two

This section describes the simple technique proposed in [3] for determining the additional tap coefficients to have the desired representation forms. Only the cases with $L = 1$, $L = 2$, and $L = 3$ are under consideration.

For $L = 3$,

$$G(\omega) = a_0 F(\omega) + a_1 [F(\omega)]^3 + a_2 [F(\omega)]^5 + a_3 [F(\omega)]^7. \tag{4}$$

The overall design is based on the following facts. First, if $G(\omega)$ oscillates within the limits $1/2 \pm \delta$ on $[0, 2\omega_p]$, then $H(\omega)$ oscillates within $1 \pm \delta$ $(\pm\delta)$ on $[0, \omega_p]$ $([\pi - \omega_p, \pi])$ (see, e.g., [1]). Second, it is assumed that $F(\omega)$ stays within the limits $1 - \epsilon_1$ and $1 + \epsilon_2$ on $[0, 2\omega_p]$. Third, the problem is to find the values of the tap coefficients a_l such that they have the desired representation forms and the resulting values of ϵ_1 and ϵ_2 guarantee the desired performances for $G(z)$ and $H(z)$.

This problem can be solved by first assuming that the value of $F(\omega)$ is $1 + \epsilon$, where ϵ is either positive or negative and then studying the corresponding value of $G(\omega)$. According to Eq. (4), this value is expressible as

$$\begin{aligned} G(\omega) &= a_0[1+\epsilon] + a_1[1+\epsilon]^3 + a_2[1+\epsilon]^5 + a_3[1+\epsilon]^7 \\ &= b_0 + b_1\epsilon + b_2\epsilon^2 + b_3\epsilon^3 + b_4\epsilon^4 + b_5\epsilon^5 + b_6\epsilon^6 + b_7\epsilon^7, \end{aligned} \tag{5}$$

where the connections between the coefficients a_l and b_l are given in Table I.

For $L = 1$, $L = 2$, and $L = 3$, the goal is to express $G(\omega)$ in the following form:

$$G(\omega) = 1/2 + \Delta, \tag{6}$$

where the maximum possible number of the lowest powers of ϵ become zero. This implies the selection of the coefficients such that b_0 becomes equal to $1/2$. The second goal in the $L = 1$, $L = 2$, and $L = 3$ cases is achieved by selecting the values of the values of the coefficients a_l as shown in Table II. Table III, in turn, shows the resulting relations between Δ and ϵ. In addition, the values of ϵ_1 and ϵ_2 guaranteeing the arrival at the half-band filter with $\delta = 10^{-6}$ (a 120-dB attenuation) are shown in this table. The resulting minimum and maximum values of $G(\omega)$ are denoted by $1 - \Delta_1$ and $1 + \Delta_2$. These values of Δ_1 and Δ_2 are also included in this table.

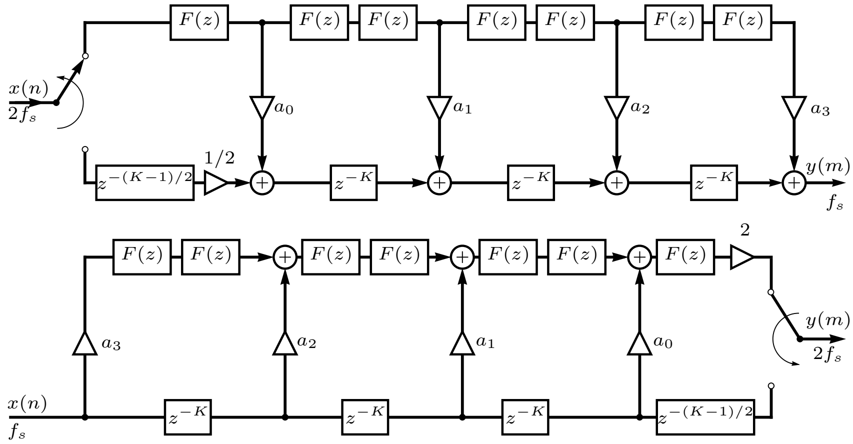

Fig. 3. Implementation of the proposed half-band filters for the sampling rate alteration by a factor of two in the $L = 3$ case.

TABLE III
CONNECTION BETWEEN Δ AND ϵ IN THE CASE OF TABLE II AS WELL AS THE VALUES OF ϵ_1, ϵ_2, Δ_1, AND Δ_2

$L=1$	$\Delta = -(3/4)\epsilon^2 - (1/4)\epsilon^3$ $\epsilon_1 = 0.001\,154\,92$, $\epsilon_2 = 0.001\,154\,47$, $\Delta_1 = 10^{-6}$, $\Delta_2 = 0$
$L=2$	$\Delta = (5/4)\epsilon^3 + (15/16)\epsilon^4 + (3/16)\epsilon^5$ $\epsilon_1 = 0.009\,304\,83$, $\epsilon_2 = 0.009\,261\,74$ $\Delta_1 = \Delta_2 = 10^{-6}$
$L=3$	$\Delta = -(35/16)\epsilon^4 - (21/8)\epsilon^5 - (35/32)\epsilon^6 - (5/32)\epsilon^7$ $\epsilon_1 = 0.026\,208\,62$, $\epsilon_2 = 0.025\,802\,80$ $\Delta_1 = 10^{-6}$, $\Delta_2 = 0$

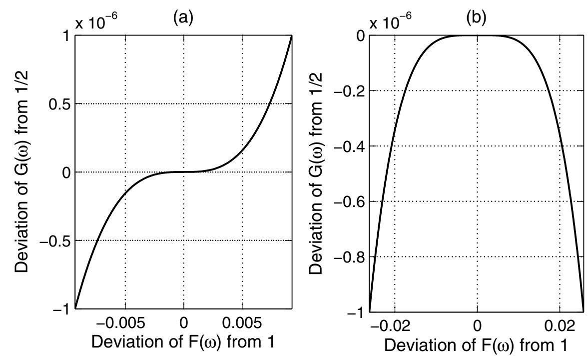

Fig. 4. The relations of the deviation of $G(\omega)$ from 1/2 to deviation of $F(\omega)$ from 1. (a) $L = 2$. (b) $L = 3$.

Figure 4, in turn, shows $1/2 - G(\omega)$ as a function of $1 - F(\omega)$ in the $L = 2$ and $L = 3$ cases. As seen from Table II and Fig. 4, $\Delta_2 = 0$ for both $L = 1$ and $L = 3$, even though it is allowed to achieve the value of 10^{-6}. For $L = 1$, using $a_1 = -2^{-2} + 2^{-20}$, instead of the value in Table I, gives rise to $\epsilon_1 = 0.001\,612\,49$, $\epsilon_2 = 0.001\,615\,44$, $\Delta_1 = 10^{-6}$, and $\Delta_2 = 0.94 \cdot 10^{-6}$. Similarly, when using $a_3 = -2^{-3} - 2^{-5} + 2^{-20}$, instead of the value in Table I results in $\epsilon_1 = 0.030\,263\,24$, $\epsilon_2 = 0.031\,314\,13$, and $\Delta_1 = \Delta_2 = 10^{-6}$. Hence, in both cases, the above-mentioned change of one tap coefficient value makes the values of both ϵ_1 and ϵ_2 larger. This means milder criteria for $F(z)$, thereby decreasing the minimum odd order to meet the resulting criteria with few powers-of-two representation forms for the filter coefficients.

IV. ILLUSTRATIVE EXAMPLE

In [3], an efficient multiplication-free decimator has been generated without any general multipliers after a sigma-delta modulator so that the overall anatog-to-digital converter has a 20-bit resolution. In this decimator, there are three linear-phase half-band filters. This example concentrates on the design of the one having the most stringent criteria.

It is desired to design a half-band decimator so that the sampling rate reduction ratio is two, the output-sampling rate is 44.1 kHz, and the components aliasing into the band from 0 Hz to 20 kHz are attenuated at least 120 dB. In this case, the problem is to design $G(z)$ such that the deviation of its zero-phase frequency response $G(\omega)$ from 1/2 is at most 10^{-6} in the passband with edge angle being $2\omega_p = [20/(44.1/2)]\pi = 0.907\,029\,48\pi$.

According to the above discussion, for $L = 3$, $a_0 = 2^0 + 2^{-3} - 2^{-5}$, $a_1 = -2^0 - 2^{-3} + 2^{-5}$, $a_2 = 2^{-1} + 2^{-3} + 2^{-5}$, and $a_3 = -2^{-3} - 2^{-5} + 2^{-20}$ are proper selections for the additional tap coefficients. In this case, the problem is to design an odd-order $F(z)$ such that $F(\omega)$ stays within the limits $1 - \epsilon_1$ and $1 + \epsilon_2$ on $[0, 2\omega_p]$ with $\epsilon_1 = 0.030\,263\,24$ and $\epsilon_2 = 0.031\,314\,13$.

The design can be performed with the aid of the Remez multiple exchange algorithm using a single band $[0, 2\omega_p]$. The desired function is $1+(1/2)(\epsilon_2-\epsilon_1) = 1.000\,525\,45$, whereas the allowable deviation is $1+\epsilon_2-1.000\,525\,45 = 0.030\,788\,68$. These criteria are well met by $F(z)$ of order 21. The impulse-response coefficient values of $F(z)$ having at most three powers-of-two terms are shown in Table IV.

TABLE IV
COEFFICIENT VALUES FOR THE MULTIPLICATION-FREE SUB-FILTER IN THE ILLUSTRATIVE EXAMPLE

$f[0] = f[21] = 2^{-6}$	$f[6] = f[15] = +2^{-4} - 2^{-8}$
$f[1] = f[20] = -2^{-6} + 2^{-8}$	$f[7] = f[14] = -2^{-4} - 2^{-6} - 2^{-8}$
$f[2] = f[19] = 2^{-6}$	$f[8] = f[13] = 2^{-3} - 2^{-8}$
$f[3] = f[18] = 2^{-7}$	$f[9] = f[12] = -2^{-2} + 2^{-5} + 2^{-7}$
$f[4] = f[17] = 2^{-5}$	$f[10] = f[11] = +2^{-1} + 2^{-3} + 2^{-7}$
$f[5] = f[16] = -2^{-4} + 2^{-6} + 2^{-8}$	

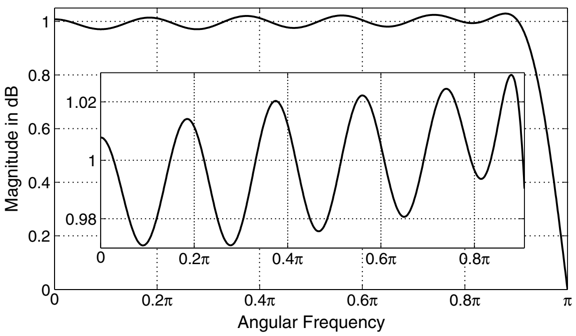

Fig. 5. Magnitude response as well as the passband details for the optimized finite-precision $F(z)$ in Illustrative Example.

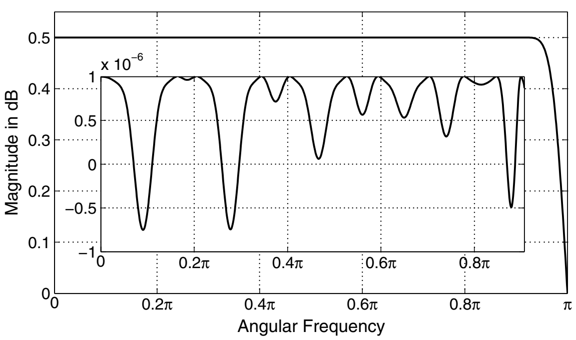

Fig. 6. Magnitude response as well as the passband details for the optimized finite-precision $G(z)$ in the Illustrative Example.

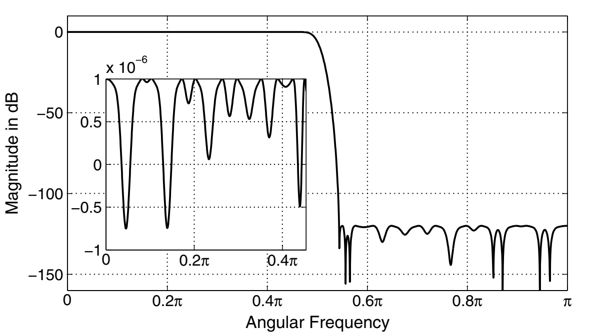

Fig. 7. Magnitude response as well as the passband details for the optimized finite-precision half-band FIR filter $H(z)$ in the Illustrative Example.

These values have been obtained by first rounding the coefficients of $F(z)$ to eight fractional bits. This rounding gives $f[10] = f[11] = 163 \cdot 2^{-8}$ and $f[9] = f[12] = -53 \cdot 2^{-8}$, which are not expressible as three powers-of-two terms, whereas the remaining coefficient values have the desired representation forms as shown in Table IV. Therefore, $f[10] = f[11]$ and $f[9] = f[12]$ are rounded to the nearest three powers-of-two representations, giving the values shown in Table IV. In this case, the number of adders and/or subtracters needed to implement all the coefficient values is 19.

Figures 5, 6, and 7 show the characteristics of $F(z)$, $G(z)$, and the resulting overall half-band FIR filter $H(z)$ with passband and stopband edges at $\omega_p = 0.453\,514\,74\pi$ and $\omega_s = \pi - \omega_p = 0.546\,485\,26\pi$ and having at least a 120-dB stopband attenuation.

For comparing the proposed design with the direct-form half-band filters, the coefficient values of these filters with various orders were rounded to the minimum number of fractional bits to meet the criteria. Table V summarizes the results by showing the minimum value of M, half the filter order, required for 23, 24, 25, and 26 fractional bit representations. When using rounding, 23 factional bits is the limit for achieving the given criteria.

TABLE V

MINIMUM VALUE OF M, HALF THE FILTER ORDER, TO MEET THE STATED CRITERIA WITH VARIOUS NUMBERS OF FRACTIONAL BITS

Number of fractional bits	22	23	24	25	26
M, Half the filter order	–	91	87	87	85

V. CONCLUDING REMARKS

This paper disclosed the design technique proposed in [3]. Future work is devoted to improving this design scheme as follows. First, for the given values of L and δ, there is a need to generate a synthesis scheme for expressing the additional tap coefficients a_l as a few powers-of-two terms such that $1/2 - G(\omega)$ as a function of $1 - F(\omega)$ [cf. Fig. 4] stays in a wide range of the values of $1 - F(\omega)$ roughly within the limits $\pm\delta$ and oscillates around zero $L+1$ times. This makes the criteria for $F(z)$ significantly milder. Second, a more sophisticated algorithm should be applied for quantizing the coefficients of $F(z)$.

REFERENCES

[1] T. Saramäki, "Finite impulse response filter design," in *Handbook for Digital Signal Processing*, S. K. Mitra and J. F. Kaiser, Eds. New York: John Wiley and Sons, 1993, ch. 4, pp. 155–277.

[2] L. Rabiner and D. Crochiere, *Multirate Digital Signal Processing*. Prentice-Hall, 1983.

[3] T. Saramäki, T. Karema, T. Ritoniemi, and H. Tenhunen, "Multiplier-free decimator algorithms for superresolution oversampled converters," in *Proc. IEEE Int. Symp. Circuits Syst.*, New Orleans, LA, May 1–3 1990, pp. 3275–3278.

[4] B. Leung, "The oversampling technique for analog to digital conversion: A tutorial overview," *J. Analog Integrated Circuits, Signal Process.*, vol. 1, no. 1, Mar. 1991.

[5] L. Lin and T. Aboulnasr, "Adaptive signal processing in subbands using sigma-delta modulation technique," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, Adelaide, SA, Australia, Apr. 19–22 1994, pp. 169–172.

[6] V. Liberali, S. Brigati, F. Francesconi, and F. Maloberti, "Progress in high-speed and high-resolution CMOS data converters," in *Proc. 1995 20th Int. Conf. Microelectron.*, vol. 1, Nis, Serbia, Sept. 12–14 1995, pp. 19–28.

[7] K. Ichige and M. Kamada, "An approximation for discrete B-splines in time domain," *IEEE Signal Processing Lett.*, vol. 4, no. 3, pp. 82–84, Mar. 1997.

[8] L. Presti, "Efficient modified-sinc filters for sigma-delta A/D converters," *IEEE Trans. Circuits Syst. II*, vol. 47, no. 11, pp. 1204–1213, Nov. 2000.

[9] T. Asahi, K. Ichige, and R. Ishii, "A new formulation for discrete box splines reducing computational cost and its evaluation," *IEICE Trans. Fund.*, vol. E84-A, no. 3, pp. 884–892, Mar. 2001.

[10] T. Saramäki, "Design of FIR filters as a cascaded tapped interconnection of identical subfilters," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 1011–1029, Sept. 1987.

[11] ——, "A systematic technique for designing highly selective multiplier-free FIR filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, Singapore, June 11–14 1991, pp. 484–487.