Generalized Comb Decimation Filters for $\Sigma\Delta$ A/D Converters: Analysis and Design

Massimiliano Laddomada, *Member, IEEE*

***Abstract*—This paper addresses the design of generalized comb decimation filters, proposing some novel decimation schemes tailored to $\Sigma\Delta$ modulators. We present a mathematical framework to optimize the proposed decimation filters in such a way as to increase the $\Sigma\Delta$ quantization noise (QN) rejection around the so called folding bands, frequency intervals whose QN gets folded down to baseband because of the decimation process. Comparisons are given in terms of passband drop and selectivity with respect to classic comb filters with orders ranging from 3 to 6. As far as the practical implementation of the proposed filters is concerned, we present two different architectures, namely a recursive and a nonrecursive implementation, the latter of which constitutes the basis for realizing multiplier-less generalized comb filter (GCF) realizations. We propose a mathematical framework for evaluating the sensitivity of GCFs to the approximation of the multipliers embedded in the filter architectures. The considerations deduced from the sensitivity analysis, pave the way to an optimization algorithm useful for approximating the multipliers with power-of-2 coefficients.**

***Index Terms*—Analog–digital (A/D) converter, cascade integrator-comb (CIC) filters, comb, decimation, decimation filter, power-of-2, sinc filters, $\Sigma\Delta$.**

I. Introduction and Problem Formulation

COMB filters are efficient anti-aliasing rate-conversion filters widely used for decimation down to four times the Nyquist rate in $\Sigma\Delta$ analog–digital (A/D) converters [1], [2].

Consider, as reference, the decimation architecture shown in Fig. 1, whereby an analog input signal with maximum frequency f_x is sampled by a $\Sigma\Delta$ A/D converter of order B with a rate f_s much greater than $2f_x$. The oversampling ratio ρ is defined as $\rho = f_s/2f_x$. We define $f_c = f_x/f_s = 1/2\rho$ as the normalized maximum frequency contained in the input signal, so that the digital signal at the input of the first decimation filter has frequency components belonging to the range $f_d \in [-f_c, f_c]$ (where f_d denotes the digital frequency).

From a practical point of view, the decimation of an oversampled signal is accomplished with two (or more) stages of decimation, such as the one shown in Fig. 1, whereby the first stage is a comb filter of order N decimating by a factor D, while the second stage comprises a finite-impulse response (FIR) filter that decimates by ν and provides the required selectivity on the sampled signal at baseband [3]. In this setting, we have $\rho = D \cdot \nu$.

Manuscript received July 8, 2006; revised October 19, 2006. This work has been supported by the CAPANINA under project FP6-IST-2003-506745, as part of the EU VI Framework Programme. This paper was recommended by Associate Editor J. S. Chang.

The author is with the Dipartimento di Elettronica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy (e-mail: laddomada@polito.it).

Digital Object Identifier 10.1109/TCSI.2007.895528

It is also known that the order N of the comb filter has to be greater or equal to $B+1$, whereby $B \geq 1$ is the order of the $\Sigma\Delta$ modulator [1], [2].

When dealing with a multistage decimator for $\Sigma\Delta$ converters, particular care must be devoted to the design of the first decimation stage. In fact, the frequency response of the antialiasing filter employed in the first decimation stage must attenuate the quantization noise (QN) falling inside the so called folding bands, i.e., the frequency ranges, pictorially shown in Fig. 2, defined as $[(k/D) - f_c; (k/D) + f_c]$ with $k = 1, \ldots, \lfloor D/2 \rfloor$ if D is even, and $k = 1, \ldots, \lfloor (D-1)/2 \rfloor$ for D odd (for conciseness, the set of values assumed by k will be denoted as K_k throughout the paper). The reason is that the $\Sigma\Delta$ QN falling inside these frequency bands, will fold down to baseband because of the sampling rate reduction by D in the first decimation stage, irremediably affecting the signal resolution after the multistage decimation chain. This issue is especially important for the first decimation stage, since the QN folding down to baseband has not been previously attenuated. To this end, the first stage of decimation in the multistage architecture shown in Fig. 1, is usually accomplished by employing a comb decimation filter, which provides an inherent antialising function by placing its zeros just in the middle of each folding band. We recall that the transfer function of a Nth-order comb filter can be expressed as [4]

$$H_{C_N}(z) = \left(\frac{1}{D}\frac{1-z^{-D}}{1-z^{-1}}\right)^N = \frac{1}{D^N}\prod_{i=1}^{D-1}\left(1 - z^{-1}e^{j\frac{2\pi}{D}i}\right)^N \tag{1}$$

where D is the desired decimation factor.

For completeness, Fig. 2 also shows the frequency range $[0; f_c]$ where the useful signal bandwidth falls, along with the so called *don't care* bands, i.e., the frequency ranges whose QN will be rejected by the selective filter placed at the end of the downconversion chain.

The aim of this paper is at increasing the $\Sigma\Delta$ QN rejection around the folding bands with respect to a classic Nth-order comb filter. A first attempt on this direction was presented in [5], [6], whereby a third-order modified sinc filter was proposed. Other works somewhat related to the topic addressed in this paper are [7]–[12]. In [7] and [8], the authors proposed computational efficient decimation filter architectures based on comb decimation filters. Reference [9] proposed the use of decimation sharpened filters embedding comb filters, while [10] addressed the design of a new two-stage sharpened comb decimator. Also related to the previous work is [11], whereby the authors proposed novel decimation schemes for $\Sigma\Delta$ A/D converters based on Kaiser and Hamming sharpened filters, then generalized in [12] for higher order decimation filters.

1549-8328/$25.00 © 2007 IEEE

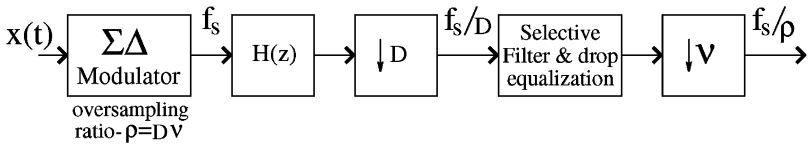

Fig. 1. General architecture of a two-stage decimation chain for ΣΔ A/D converters.

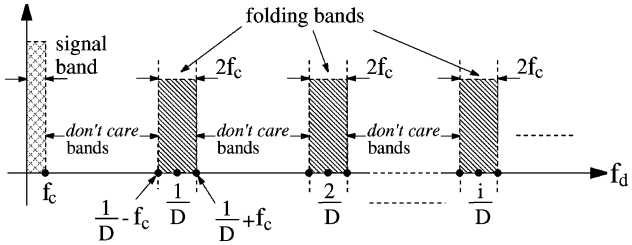

Fig. 2. Pictorial representation of the frequency intervals that should be carefully considered for the design of the first decimation stage.

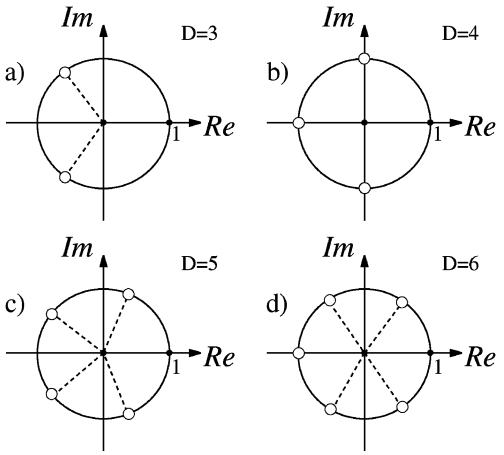

Fig. 3. Zero locations in the z-plane of a classic comb filter for various values of the decimation factor D.

The rest of the paper is organized as follows. In Section II, we derive the transfer functions of generalized comb filters (GCFs). Section III presents the main performance metrics of GCFs. Section IV derives a recursive implementation of GCFs, while Section V proposes a nonrecursive architecture suitable for power-of-2 decimation factors. In Section VI, we present a sensitivity analysis of the proposed nonrecursive filters, the considerations of which are at the basis of the optimization algorithm proposed in Section VII. Section VIII presents two sample filter designs, and presents a brief analysis of the computational complexity of the proposed filters. In Section IX, we compare the proposed filters with two architectures proposed in literature. Finally, Section X draws the conclusions.

II. Derivation of GCFs

Before presenting the main idea of this paper, let us examine the zero positions of the classic comb filter in (1). For $N = 1$, $H_{C_1}(z)$ presents zeros in the positions $z_k = e^{j(2\pi/D)i}$, $\forall i = 1, \ldots, D-1$, in the z-plane, as shown in Fig. 3 for some specific values of D. Note that the zero in $z = 1$ is cancelled by the pole $(1 - z^{-1})$ in (1). For $N > 1$, zero locations do not change; the only difference is that each zero is now of order N. With this background, let us answer the following question. How can we increase the ΣΔ QN rejection around the folding bands with respect to a classic comb filter?

We can synthesize a GCF by placing the zeros of a classic Nth-order comb filter in such a way as to span the folding bands $[(k/D) - f_c; (k/D) + f_c]$, $\forall k \in K_k$. Let us first observe that any complex zero $z = e^{j(2\pi/D)k}$ (with $k = 1, \ldots, (D/2) - 1$, if D is even, and $k = 1, \ldots, (D-1)/2$, for D odd) should be coupled with its complex conjugate, in order for the transfer function to have real coefficients. This in turn suggests the use of a second-order basic cell of the type $(1 - z^{-1}e^{+j(2\pi/D)k})(1 - z^{-1}e^{-j(2\pi/D)k})$, which simplifies to $(1 - 2\cos((2\pi/D)k)z^{-1} + z^{-2})$, with zeros placed at the digital frequencies $\pm(k/D)$, with k as specified above. This basic second-order cell can be generalized by rotating its zeros of an angle α_p, as shown in Fig. 4, obtaining the following second-order cell $(1 - z^{-1}e^{+j(2\pi/D)k - j\alpha_p})(1 - z^{-1}e^{-j(2\pi/D)k + j\alpha_p})$, which simplifies to $(1 - 2\cos((2\pi/D)k - \alpha_p)z^{-1} + z^{-2})$, with zeros placed at the digital frequencies $+(k/D) - (\alpha_p/2\pi)$ and $-(k/D) + (\alpha_p/2\pi)$, as shown in Fig. 4. Based on the considerations drawn above on the folding bands, a convenient choice for α_p is $\alpha_p = q_p 2\pi f_c$, with $q_p \in [-1, +1]$: this solution is such that the pair of zeros falls inside the relative folding band.

The next step is the generalization of the second-order cell discussed above. To this end, upon noting that a Nth-order comb filter should place N zeros in each folding band, the transfer function of a Nth-order GCF can be defined after defining the following two functions:

$$H_1(z) = \prod_{i=1}^{D_M} \prod_{p=1}^{N} \left(1 - 2\cos\left(\frac{2\pi}{D}i - \alpha_p\right) z^{-1} + z^{-2}\right)$$

with $D_M = (D/2) - 1$, for D even, and $D_M = (D-1)/2$, for D odd, and

$$\begin{aligned} H_2(z, \alpha_n) &= (1 - z^{-1}e^{+j\pi - j\alpha_n})(1 - z^{-1}e^{-j\pi + j\alpha_n}) \\ &= 1 - 2\cos(\pi - \alpha_n)z^{-1} + z^{-2}. \end{aligned}$$

With this setup, the transfer function of a Nth-order GCF, $H_{\mathrm{GCF}_N}(z)$, can be defined as

$$\begin{cases} \dfrac{1}{H_{o,\mathrm{ev},1}} H_1(z) \displaystyle\prod_{n=1}^{\lfloor \frac{N}{2} \rfloor} H_2(z, \alpha_{n+N}), & D \text{ even},\ N \text{ even} \\ \dfrac{(1 + z^{-1})}{H_{o,\mathrm{ev},2}} H_1(z) \displaystyle\prod_{n=1}^{\lfloor \frac{N}{2} \rfloor} H_2(z, \alpha_{n+N}), & D \text{ even},\ N \text{ odd} \\ \dfrac{1}{H_{o,\mathrm{od}}} H_1(z), & D \text{ odd} \end{cases} \tag{2}$$

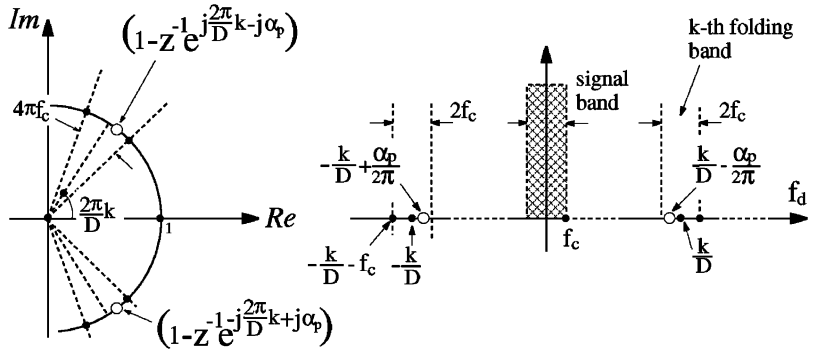

Fig. 4. Zero locations of a generalized second-order filter cell. Zeros are displayed in the z-plane and in the frequency axis in order to highlight their effect in both domains.

whereby $H_{o,\mathrm{ev},1}$, $H_{o,\mathrm{ev},2}$, and $H_{o,\mathrm{od}}$ are suitable normalization constants to be chosen in such a way as to have $H_{\mathrm{GCF}_N}(z)|_{z=1}=1$, and $\lfloor\cdot\rfloor$ is the floor function. Notice that, for D even and N odd, the polynomial $(1+z^{-1})$ places a first-order zero in the digital frequency 1/2, or, equivalently, in $z=-1$ in the z-domain. These zeros are only present for D even and N odd, as clearly depicted in Fig. 3.

By imposing $z=1$ in (2), it is simple to obtain the following normalization factors:

$$\begin{aligned} H_{o,\mathrm{ev},1} &= \prod_{i=1}^{D_M}\prod_{p=1}^{N}\left[2-2\cos\left(\frac{2\pi}{D}i-\alpha_p\right)\right] \\ &\quad\cdot\prod_{n=1}^{\lfloor\frac{N}{2}\rfloor}\left[2+2\cos(\alpha_{n+N})\right], \\ H_{o,\mathrm{ev},2} &= 2\cdot\prod_{i=1}^{D_M}\prod_{p=1}^{N}\left[2-2\cos\left(\frac{2\pi}{D}i-\alpha_p\right)\right] \\ &\quad\cdot\prod_{n=1}^{\lfloor\frac{N}{2}\rfloor}\left[2+2\cos(\alpha_{n+N})\right], \\ H_{o,\mathrm{od}} &= \prod_{i=1}^{D_M}\prod_{p=1}^{N}\left[2-2\cos\left(\frac{2\pi}{D}i-\alpha_p\right)\right] \end{aligned} \tag{3}$$

which assure unity gain in baseband. Finally, note that N should be chosen larger than the $\Sigma\Delta$ modulator order B [2]. The choice of the decimation factor D follows the same rules as for classic comb filters. We will discuss this issue when dealing with the drop performances of GCFs.

Notice that GCFs are linear-phase filters since they are composed by linear-phase basic filters. In fact, the second-order cells constituting $H_1(z)$, and $H_2(z)$ where it applies, are linear-phase filters.

The problem of finding the optimal values for the q_p-factors, defining the zero positions across the folding bands is addressed in the next section.

A. Optimal Choice of the Zero Positions

It is possible to optimize the rotation parameters q_p in (2), by maximizing the attenuation undergone by the QN of a Bth-order $\Sigma\Delta$ modulator around the folding bands, as a function of the q_p-factors. The $\Sigma\Delta$ QN power falling inside the folding bands can be defined as [1]

$$P_{\mathrm{qn}}=\sum_{k\in K_k}\int_{\frac{k}{D}-f_c}^{\frac{k}{D}+f_c}\left|H_{\mathrm{GCF}_N}(f_d)\right|^2 S_B(f_d)df_d \tag{4}$$

where $S_B(f_d)$, the power spectral density of the $\Sigma\Delta$ QN, can be expressed as $S_B(f_d)=S_e(f_d)\cdot[2\sin(\pi f_d)]^{2B}$. Note that (4) is valid only if the noise transfer function (NTF) of the modulator is maximally flat, i.e., it does not contain stabilizing poles. In higher order modulators this requires multi-bit feedback structures.

In the previous relation $S_e(f_d)=\Delta^2/12f_s$ is the spectral density of the sampled noise under the hypothesis of representing the QN as a white noise [1], Δ is the quantization level of the quantizer contained in the $\Sigma\Delta$ modulator [1], and f_s is the $\Sigma\Delta$ sampling rate.

Optimization has been accomplished by exhaustive evaluation of (4) for any $q_p\in[-1,+1]$ with steps of $\Delta q_p=10^{-2}$, and, then, by selecting the set of q_p's yielding the minimum value of P_{qn} in (4).

Results of this optimization are shown in Table I for various filter orders, N. In addition, we show the gain G, i.e., the extra attenuation received by the quantization noise because of the GCF with respect to classical comb filters. We found experimentally that the optimal values of q_p can be practically considered independent upon both the decimation factor D and the normalized maximum frequency f_c of the input signal. As we will see in Section IV, this behavior highly simplifies the decimation filter design. In the following, where not otherwise specified, we consider the optimal zero rotations (OZRs) specified in Table I. Notice that OZRs shown in each column in Table I, are independent of the folding bands in that the $q_i's$, $\forall i=1,\ldots,N$, represent the OZRs of the N zeros in any folding band (i.e., the ones belonging to $H_1(z)$), while the $q_i's$, $\forall i=N+1,\ldots,N+\lfloor N/2\rfloor$, are the OZRs around $z=-1$, i.e., the ones imposed by $H_2(z)$ where it applies.

Finally, notice that, roughly speaking, the values of G in Table I indicate that any additional rotated zero assures about 5 dB of QN rejection with respect to a classic comb filter.

TABLE I
OPTIMAL PARAMETERS OF GCF DECIMATION FILTERS

N	3	4	5	6
q_1	-0.79	-0.35	+0.55	+0.95
q_2	0.0	+0.35	+0.93	+0.675
q_3	+0.79	-0.88	-0.55	+0.25
q_4	+0.79	+0.88	-0.93	-0.25
q_5	-	+0.88	0.0	-0.675
q_6	-	+0.35	+0.55	-0.95
q_7	-	-	+0.93	+0.95
q_8	-	-	-	+0.675
q_9	-	-	-	+0.25
$G-[dB]$	$\sim$8	$\sim$13	$\sim$18	$\sim$ 23

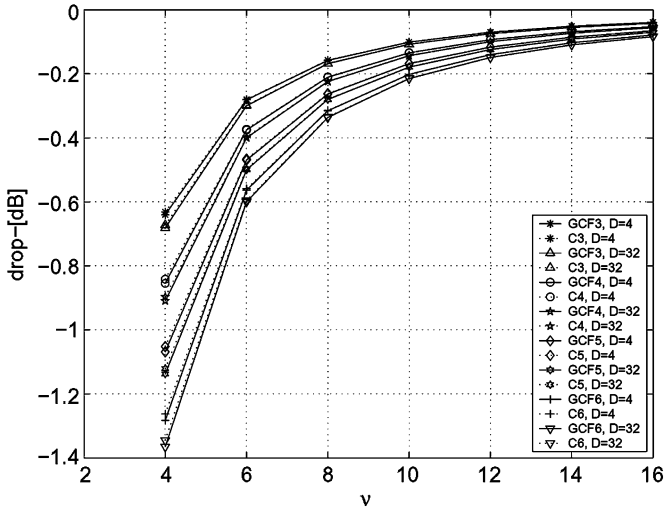

Fig. 5. Passband drops of filters $H_{\mathrm{GCF}_i}(f_d)$ and $H_{C_i}(f_d)$, with $i = 3, \ldots, 6$, when the first stage in Fig. 1 operates a decimation by D, as a function of the residual decimation ν.

III. PERFORMANCES OF GCFS

This section focuses on the performances of GCFs. The main performance measures [1], [2] considered for comparing decimation filters are the passband drop, defined as $d(f_c) = |H(f_c)/H(0)|$, and the selectivity ϕ, defined as $\phi = |H(f_c)/H((1/D) - f_c)|$, whereby $H(f_d)$ is the specific decimation filter frequency response considered, $f_c = 1/2D\nu$ is the normalized maximum frequency contained in the input signal, and $((1/D) - f_c)$ is the lower edge of the first folding band. OZRs specified in Table I have been used for filter performance evaluation.

Passband drop indicates the maximum attenuation of the actual designed filter at the edge of the useful signal bandwidth with respect to an ideal low-pass filter; this is the amount of amplitude distortion on the sampled signal that should be equalized after decimation. Selectivity, ϕ, measures the ability of the decimation filter to suppress high-frequency noise components accompanying the useful signal. Performance comparisons between GCF and classic comb filters will be accomplished in terms of both selectivity and passband drop.

Fig. 5 compares the passband drops of filters $H_{\mathrm{GCF}_i}(f_d)$ and $H_{C_i}(f_d)$, with $i = 3, \ldots, 6$, when the first stage in Fig. 1 operates a decimation by D, as a function of the residual decimation ν. Note that, practically speaking, passband drops of GCFs are very similar to the ones guaranteed by classic comb filters with the same order.

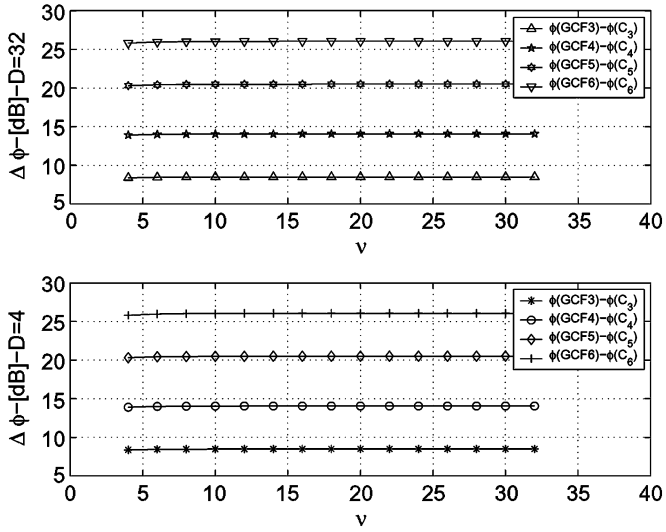

Fig. 6. Difference between the selectivities of filters $H_{\mathrm{GCF}_i}(f_d)$ and $H_{C_i}(f_d)$, with $i = 3, \ldots, 6$, when the first stage in Fig. 1 operates a decimation by D, as a function of the residual decimation ν.

Let us turn the attention on the selectivity of the proposed decimation filters. Fig. 6 shows, for any $i = 3, \ldots, 6$, the difference between the selectivities of filters $H_{\mathrm{GCF}_i}(f_d)$ and $H_{C_i}(f_d)$, when the first stage in Fig. 1 operates a decimation by D, as a function of the residual decimation ν. Note that GCFs have superior performances with respect to classic comb filters in terms of selectivity. The reason resides in the zero rotation accomplished in the first folding bands: the leftmost zero in the first folding band has the effect to reduce the values taken on by the frequency responses $H_{\mathrm{GCF}_i}(f_d)$, with $i = 3, \ldots, 6$, for $f_d \sim (1/D) - f_c$, with respect to the values $H_{C_i}((1/D) - f_c)$.

IV. RECURSIVE IMPLEMENTATION OF GCFS

This section deals with recursive implementations of GCFs. For conciseness we take as a reference example the implementation of filter $H_{\mathrm{GCF}_3}(z)$, and D equal to 4, but the considerations that follow can be easily applied to the design of any other GCF.

Consider the transfer function in (2) for $N = 3$ and $D = 4$

$$\begin{aligned} H_{\mathrm{GCF}_3}(z)_{|_{D=4}} &= \frac{1}{H_{o,\mathrm{ev},2}} \cdot (1 + z^{-1}) \\ &\quad \cdot \left[\prod_{p=1}^{3} \left(1 - z^{-1} e^{+j\frac{\pi}{2} - j\alpha_p}\right) \left(1 - z^{-1} e^{-j\frac{\pi}{2} + j\alpha_p}\right) \right] \\ &\quad \cdot \left[(1 - z^{-1} e^{+j\pi - j\alpha_4})(1 - z^{-1} e^{-j\pi + j\alpha_4}) \right]. \end{aligned} \tag{5}$$

Notice that the order N identifies the number of zeros falling in each folding band, while D identifies the number of folding bands since it defines the set K_k.

Zeros of $H_{\mathrm{GCF}_3}(z)$ in (5) are placed on the unity circle in the z-plane, as shown in Fig. 7 for the OZRs in Table I, i.e., for $\alpha_1 = -\alpha_3$, $\alpha_3 = \alpha_4 = 0.79 \cdot 2\pi f_c$, and $\alpha_2 = 0$. For simplicity, let us define $\alpha = \alpha_3 = \alpha_4$, and $\alpha_1 = -\alpha$.

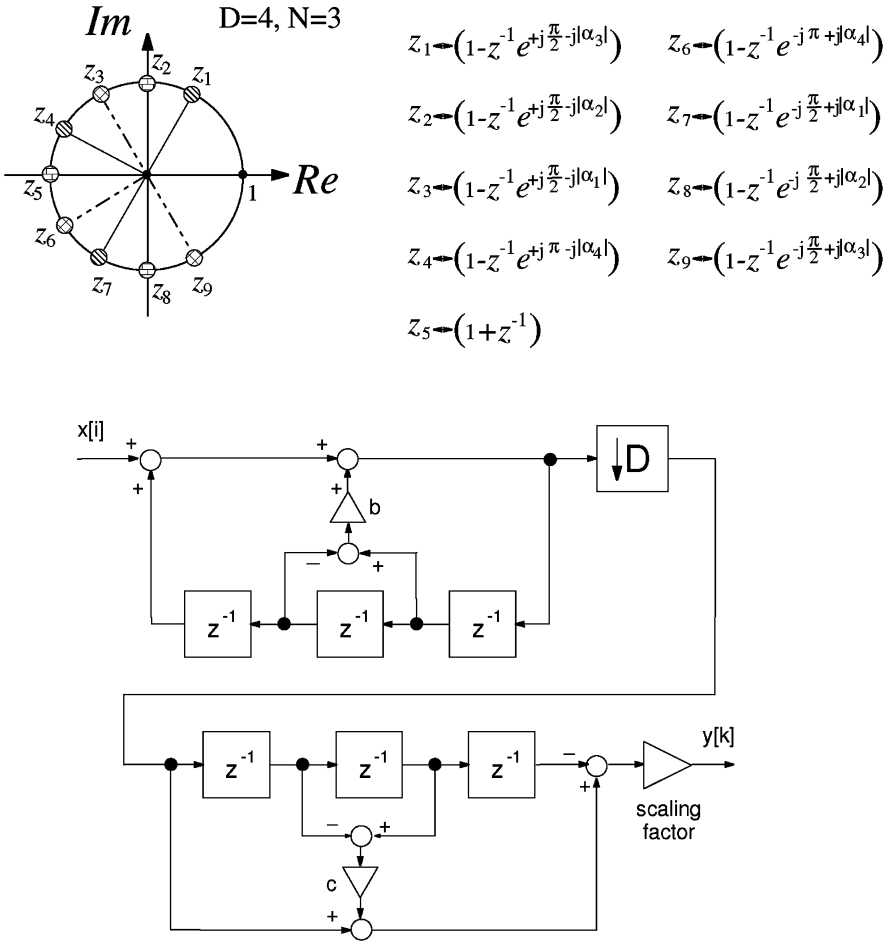

Fig. 7. Upper subplot shows the zero locations in the z-plane of filter $H_{\mathrm{GCF}_3}(z)$ for $N = 3$ and $D = 4$. The lower subplot shows a recursive implementation of decimation filter $H_{\mathrm{GCF}_3}(z)$.

Let us find a transfer function equivalent to (5). First of all, notice that, assuming $D = 4$, zeros z_2, z_5 and z_8 constitute a classical first-order comb filter with the following transfer function:

$$\prod_{i=1}^{D-1}\left(1 - z^{-1}e^{j\frac{2\pi}{D}i}\right) = \frac{1 - z^{-D}}{1 - z^{-1}}. \tag{6}$$

Using a similar reasoning, it is simple to show that zeros z_1, z_4 and z_7 represent the following transfer function:

$$\prod_{i=1}^{D-1}\left(1 - z^{-1}e^{j\frac{2\pi}{D}i}e^{-j\alpha}\right) = \prod_{i=1}^{D-1}\left(1 - \beta^{-1}e^{j\frac{2\pi}{D}i}\right) \tag{7}$$

whereby we have posed $\beta = z \cdot e^{j\alpha}$. By comparing the right equation in (6) with (1), it is possible to note that (6) is equivalent to the following transfer function:

$$\frac{1 - \beta^{-D}}{1 - \beta^{-1}} = \frac{1 - z^{-D}e^{-j\alpha D}}{1 - z^{-1}e^{-j\alpha}} \tag{8}$$

which is the transfer function of a classic comb filter whose zeros have been rotated anticlockwise by an angle α. A similar reasoning can be applied to zeros z_3, z_6 and z_9, obtaining

$$\frac{1 - z^{-D}e^{+j\alpha D}}{1 - z^{-1}e^{+j\alpha}}. \tag{9}$$

Multiplying together (6), (8), and (9), after some algebra it is possible to obtain

$$H_{\mathrm{GCF}_3}(z) = \frac{1}{H_{o,\mathrm{ev},2}}\frac{1 - cz^{-D} + cz^{-2D} - z^{-3D}}{1 - bz^{-1} + bz^{-2} - z^{-3}} \tag{10}$$

whereby $b = 1 + 2\cos(\alpha)$ and $c = 1 + 2\cos(\alpha D)$.

The frequency response associated to $H_{\mathrm{GCF}_3}(z)$ can be simply evaluated upon noting that the frequency response of (6) is

$$e^{-j2\pi f_d(D-1)}\frac{\sin(\pi f_d D)}{\sin(\pi f_d)}$$

while the frequency responses corresponding to (8) and (9) are, respectively

$$e^{-j\pi\left(f_d - \frac{\alpha}{2\pi}\right)(D-1)}\frac{\sin\left(\pi\left(f_d - \frac{\alpha}{2\pi}\right)D\right)}{\sin\left(\pi\left(f_d - \frac{\alpha}{2\pi}\right)\right)}$$

and

$$e^{-j\pi\left(f_d + \frac{\alpha}{2\pi}\right)(D-1)}\frac{\sin\left(\pi\left(f_d + \frac{\alpha}{2\pi}\right)D\right)}{\sin\left(\pi\left(f_d + \frac{\alpha}{2\pi}\right)\right)}$$

by virtue of the modulation property of the discrete-time Fourier transform [13].

Multiplying together the previous three frequency responses, it is possible to obtain

$$H_{\mathrm{GCF}_3}(e^{j2\pi f_d}) = \frac{1}{H_{o,\mathrm{ev},2}} \cdot \frac{\sin\left[(\pi f_d - \alpha/2)D\right]}{\sin(\pi f_d - \alpha/2)} \cdot \frac{\sin\left[(\pi f_d + \alpha/2)D\right]}{\sin(\pi f_d + \alpha/2)} \frac{\sin(\pi f_d D)}{\sin(\pi f_d)} \cdot \mathrm{e}^{-j3\pi f_d(D-1)}. \quad (11)$$

Beside linear phase, this frequency response is equal to zero at the frequencies $f_k = (k/D) \pm (\alpha/2\pi)$, k/D, $\forall k \in K_k$. Notice that, for $\alpha = 0$, filter $H_{\mathrm{GCF}_3}(e^{j2\pi f_d})$ becomes a classic third-order comb filter $H_{C_3}(f_d)$.

The previous mathematical formulation is valid for any D. In other words, given N, it is possible to collect $D-1$ zeros together in order to obtain a first-order comb filter with rotated zeros, just as discussed above.

A recursive implementation of filter $H_{\mathrm{GCF}_3}(z)$ in (10) is shown in the lower part of Fig. 7. It is obtained in the same way as for a classic cascade integrator-comb (CIC) implementation [4]. The numerator in (10) corresponds to the comb sections at the right of the decimator by D, while the denominator is responsible for the integrator sections at the left of the decimator by D. The scaling factor $1/H_{o,\mathrm{ev},2}$ assures unity gain at baseband.

The drawback of this filter is the presence of the two multipliers b and c. This is the price to be paid to increase the QN rejection around the folding bands with respect to a classic comb filter. These considerations suggest that the recursive implementation deduced above can only be used in digital signal processors operating with floating-point arithmetic, and, as a consequence, it is not amenable to high-speed applications.

The complexity burden of this filter, due to the presence of multipliers, can be greatly reduced by resorting to a nonrecursive implementation. This way, the multipliers are still present but can be approximated as power-of-2 (PO2) coefficients. Indeed, this is the topic addressed in the next sections.

V. Nonrecursive Implementation of GCFs

This section deals with the design of nonrecursive implementations of GCFs, suitable for decimation factors D that can be expressed as the pth power-of-two, i.e., $D = 2^p$ with p an integer greater than zero. For conciseness, we only focus on the implementation of a 3rd-order GCF, but the considerations that follow, can be easily extended to higher order GCFs with D as specified above.

In the previous section, we obtained the following transfer function:

$$H_{\mathrm{GCF}_3}(z) = \frac{1-z^{-D}}{1-z^{-1}} \frac{1-z^{-D}e^{j\alpha D}}{1-z^{-1}e^{j\alpha}} \frac{1-z^{-D}e^{-j\alpha D}}{1-z^{-1}e^{-j\alpha}}. \quad (12)$$

Notice that, for conciseness, in the mathematical formulation that follows, we omit the constant term assuring unity gain at baseband. This is just a scaling factor which can be accounted for after the proposed mathematical derivations.

A nonrecursive implementation of filter $H_{\mathrm{GCF}_3}(z)$ can be obtained by expressing each polynomial ratio in (12) in a non recursive form. By doing so, the first polynomial ratio can be rewritten as follows:

$$\frac{1-z^{-D}}{1-z^{-1}} = \sum_{i=0}^{D-1} z^{-i} = \prod_{i=0}^{\log_2(D)-1} \left(1 + z^{-2^i}\right) \quad (13)$$

whereby last equality holds for $D = 2^p$. Using a similar reasoning, it is simple to observe that the following equality chain, which derives from (13) by imposing $\beta = z \cdot e^{-j\alpha}$, holds as well

$$\frac{1-z^{-D}e^{j\alpha D}}{1-z^{-1}e^{j\alpha}} = \sum_{i=0}^{D-1} z^{-i}e^{+ji\alpha} = \prod_{i=0}^{\log_2(D)-1} \left(1 + z^{-2^i}e^{+j2^i\alpha}\right). \quad (14)$$

A similar reasoning applies to the last polynomial ratio in (12). Upon noting that

$$\left(1 + z^{-2^i}e^{+j2^i\alpha}\right)\left(1 + z^{-2^i}e^{-j2^i\alpha}\right) = 1 + 2\cos(2^i\alpha)z^{-2^i} + z^{-2^{i+1}}$$

after some algebra, (12) can be rewritten as follows:

$$H_{\mathrm{GCF}_3}(z) = \prod_{i=0}^{\log_2(D)-1} \left[\left(1+z^{-2^i}\right)\cdot\left(1+2\cos(2^i\alpha)z^{-2^i}+z^{-2^{i+1}}\right)\right] = \prod_{i=0}^{\log_2(D)-1} \left[1+r_i\cdot\left(z^{-2^i}+z^{-2\cdot 2^i}\right)+z^{-3\cdot 2^i}\right] \quad (15)$$

whereby

$$r_i = 1 + 2\cos(2^i\alpha) = 1 + 2\cos\left(q\frac{2^i\pi}{\rho}\right) = 1 + 2\cos(q2^{i+1}\pi f_c) \qquad \forall i = 0, \ldots, \log_2(D) - 1.$$

Equation (15) is the starting point for obtaining the architecture of a nonrecursive GCF implementation. By applying the commutative property employed in [14], it is possible to obtain the cascaded implementation shown in Fig. 8, whereby the constant scaling factor $H_{o,\mathrm{ev},2}$ is omitted for brevity. Any stage in

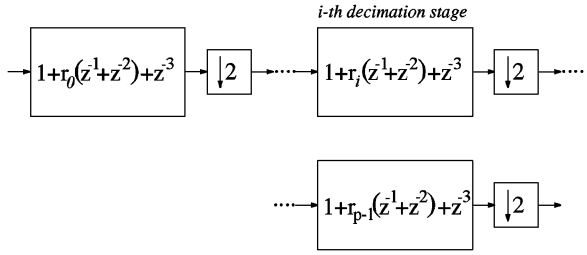

Fig. 8. Architecture of a nonrecursive implementation of decimation filter $H_{\mathrm{GCF}_3}(z)$. Index i, which identifies the decimation stage, belongs to the range $0, 1, \ldots, \log_2 D - 1 = p - 1$.

Fig. 8 is constituted by a simple FIR filter operating at a different rate. Such an example, the ith stage is characterized by the transfer function

$$1 + r_i(z^{-1} + z^{-2}) + z^{-3} \tag{16}$$

and it operates at the sampling rate $f_s/2^i$, whereby f_s is the $\Sigma\Delta$ sampling frequency.

The frequency response related to (15) can be evaluated as follows:

$$H_{\mathrm{GCF}_3}(e^{j\omega}) = 2 \prod_{i=0}^{\log_2(D)-1} e^{-j3\cdot 2^{i-1}\omega} \cdot \left[\cos(3\cdot 2^{i-1}\omega) + r_i \cos(2^{i-1}\omega)\right] \tag{17}$$

whereby $\omega = 2\pi f_d$. Notice that $H_{\mathrm{GCF}_3}(e^{j\omega})$ is a linear-phase filter since it is obtained as the cascade of $\log_2(D)$ linear-phase filter cells.

VI. Sensitivity Analysis of Nonrecursive GCFs

This section provides a sensitivity analysis to derive useful hints for optimizing the nonrecursive architecture in Fig. 8. The goal is to derive a multiplier-less architecture with improved performances in terms of QN rejection with respect to a classical comb filter.

Let us evaluate the sensitivity of the frequency response in (17) with respect to the coefficients r_i. First of all, notice that when the coefficient r_i is approximated, its actual value can be expressed as $\widetilde{r}_i = r_i + \Delta r_i$, whereby Δr_i is the approximation error. Clearly, the approximations of coefficients r_i imply that $H_{\mathrm{GCF}_3}(e^{j\omega})$ be written as $\widetilde{H}_{\mathrm{GCF}_3}(e^{j\omega}) = H_{\mathrm{GCF}_3}(e^{j\omega}) + \Delta H_{\mathrm{GCF}_3}(e^{j\omega})$.

The dependence of the frequency response $H_{\mathrm{GCF}_3}(e^{j\omega})$ on the approximation of multipliers $r_i, \forall i = 0, \ldots, \log_2(D) - 1$, can be evaluated as follows:

$$\Delta H_{\mathrm{GCF}_3}(e^{j\omega}) = \sum_{i=0}^{\log_2(D)-1} \frac{\partial H_{\mathrm{GCF}_3}(e^{j\omega})}{\partial r_i} \Delta r_i. \tag{18}$$

Let us evaluate the derivative of $H_{\mathrm{GCF}_3}(e^{j\omega})$ with respect to r_i

$$\frac{\partial H_{\mathrm{GCF}_3}(e^{j\omega})}{\partial r_i} = 2e^{-j3\cdot 2^{i-1}\omega}\cos(2^{i-1}\omega) \cdot \prod_{m=0, m\neq i}^{\log_2(D)-1} e^{-j3\cdot 2^{m-1}\omega} \times \left[\cos(3\cdot 2^{m-1}\omega) + r_m \cdot \cos(2^{m-1}\omega)\right], \quad \forall i = 0, \ldots, \log_2(D) - 1 \tag{19}$$

Equation (19) can be rewritten as follows:

$$\frac{\partial H_{\mathrm{GCF}_3}(e^{j\omega})}{\partial r_i} = H_{\mathrm{GCF}_3}(e^{j\omega}) \cdot \frac{\cos(2^{i-1}\omega)}{\cos(3\cdot 2^{i-1}\omega) + r_i \cdot \cos(2^{i-1}\omega)}. \tag{20}$$

Upon substituting (20) in (18), it is possible to obtain

$$\Delta H_{\mathrm{GCF}_3}(e^{j\omega}) = H_{\mathrm{GCF}_3}(e^{j\omega}) \cdot \sum_{i=0}^{\log_2(D)-1} \frac{\cos(2^{i-1}\omega)\Delta r_i}{\cos(3\cdot 2^{i-1}\omega) + r_i \cdot \cos(2^{i-1}\omega)}. \tag{21}$$

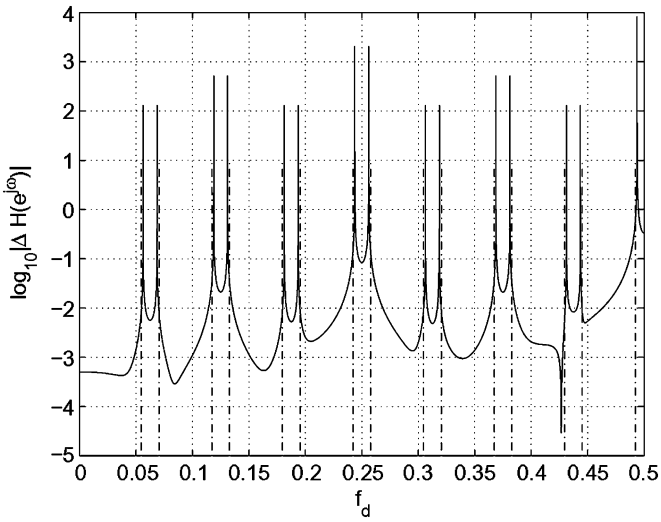

Fig. 9. Frequency behavior (solid line) of the error function $\log_{10}|\Delta H(e^{j\omega})|$ in (23) for $D = 16$, $\nu = 4$, $\alpha = q_o 2\pi f_c$, $q_o = 0.79$, and $\Delta r_i = 5 \cdot 10^{-4}$, $\forall i = 0, \ldots, 3$, along with the folding bands $[(k/D) - f_c, (k/D) + f_c], \forall k \in K_k$ (dash-dot lines).

The actual frequency response $\widetilde{H}_{\mathrm{GCF}_3}(e^{j\omega})$ can be expressed as follows:

$$\begin{aligned}\widetilde{H}_{\mathrm{GCF}_3}(e^{j\omega}) &= H_{\mathrm{GCF}_3}(e^{j\omega}) + \Delta H_{\mathrm{GCF}_3}(e^{j\omega}) \\ &= H_{\mathrm{GCF}_3}(e^{j\omega}) \cdot \left[1 + \sum_{i=0}^{\log_2(D)-1} \frac{\cos(2^{i-1}\omega)\Delta r_i}{\cos(3\cdot 2^{i-1}\omega) + r_i \cdot \cos(2^{i-1}\omega)}\right]\end{aligned} \tag{22}$$

which clearly depends on the approximation errors $\Delta r_i, \forall i = 0, \ldots, \log_2(D) - 1$, between each multiplier r_i and its approximated value $\widetilde{r}_i$.

The effects of the approximation of the multipliers r_i on the actual frequency response, $\widetilde{H}_{\mathrm{GCF}_3}(e^{j\omega})$, can be understood by analyzing the frequency behavior of the following error function:

$$\begin{aligned}\Delta H(e^{j\omega}) &= \sum_{i=0}^{\log_2(D)-1} \frac{\cos(2^{i-1}\omega)\Delta r_i}{\cos(3\cdot 2^{i-1}\omega) + r_i \cdot \cos(2^{i-1}\omega)} \\ &= \sum_{i=0}^{\log_2(D)-1} \Delta H_i(e^{j\omega})\end{aligned} \tag{23}$$

which, to some extent, quantifies the distortion between the desired and the actual frequency response. To this end, Figs. 9 and 10 show, for $D = 16$, $\nu = 4$, $q_o = 0.79$, and $\Delta r_i = 5 \cdot 10^{-4}$, $\forall i = 0, \ldots, 3$, the behaviors of both the error function $\Delta H(e^{j\omega})$ in (23), and its constituent functions $\Delta H_i(e^{j\omega}), \forall i = 0, \ldots, 3$.

Some key observations are in order. Figs. 9 and 10 show that the passband behavior of the filter $H_{\mathrm{GCF}_3}(e^{j\omega})$ is not affected by the approximation of the multipliers r_i. Sensitivity of $H_{\mathrm{GCF}_3}(e^{j\omega})$ is very low for $f_d \in [0, f_c]$, whereby $f_c = 1/2\rho$. This, in turn, suggests that the filter passband drop does not degrade by virtue of multipliers' approximations.

Fig. 9 shows that the frequency response $H_{\mathrm{GCF}_3}(e^{j\omega})$ is very sensitive to coefficients' approximations especially in

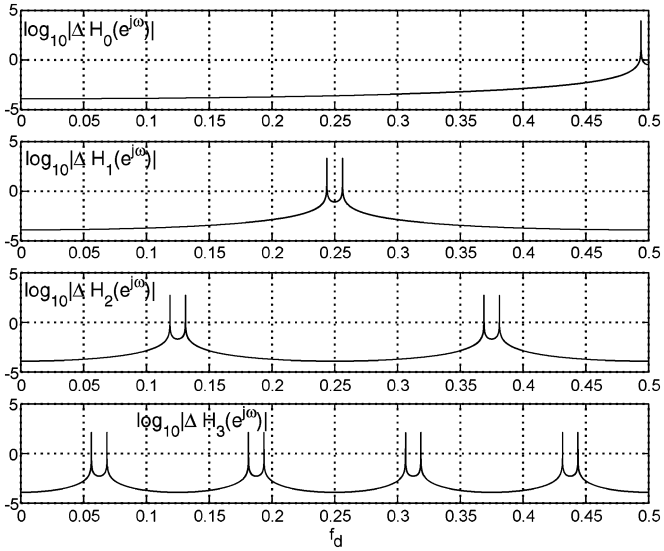

Fig. 10. Frequency behaviors of the functions $|\Delta H_i(e^{j\omega})|$, $\forall i = 0, \ldots, 3$, in (23) for $D = 16$, $\nu = 4$, $\alpha = q_o 2\pi f_c$, $q_o = 0.79$, and $\Delta r_i = 5 \cdot 10^{-4}$, $\forall i = 0, \ldots, 3$.

the folding bands $[(k/D) - f_c, (k/D) + f_c]$, $\forall k \in K_k$. This in turn suggests that particular care must be devoted to the approximation of multipliers r_i as power-of-2 terms in order to save the QN rejection around the folding bands.

Fig. 10 shows that the approximation of each multiplier r_i affects the sensitivity of the frequency response $H_{\mathrm{GCF}_3}(e^{j\omega})$ in disjoint folding bands. This observation paves the way to an effective optimization algorithm whose aim is to approximate each multiplier r_i independently from the others. This is the topic addressed in next section.

VII. Optimization of GCFs

This section presents an optimization framework for approximating each multiplier r_i in (15) as a sum of signed power-of-2 (PO2) terms [15], while reducing the computational complexity of the cascaded architecture shown in Fig. 8.

Let us approximate each real multiplier r_i in (15) as a sum of L_i PO2 coefficients

$$r_i \approx \widetilde{r}_i = \sum_{v=1}^{L_i} b_v^{(i)} 2^{-q_v^{(i)}} \tag{24}$$

whereby

$$b_v^{(i)} \in \{-1, 0, 1\} \qquad \forall i = 0, \ldots, \log_2(D) - 1$$
$$q_v^{(i)} \in \{-1, 0, 1, \ldots, R_m\} \qquad \forall i = 0, \ldots, \log_2(D) - 1$$

and R_m is the maximum word length imposed by the design ($R_m \geq 0$).

The optimization algorithm aims to minimize a cost function related to the computational complexity of each decimation stage in Fig. 8. Let us discuss a cost function which serves our purposes.

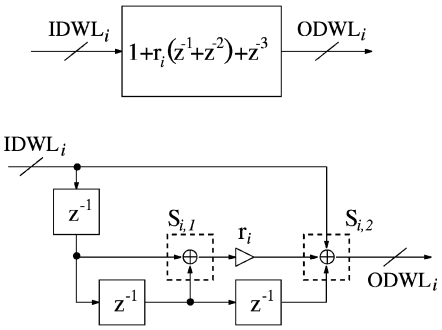

Fig. 11. Architecture of the ith decimation stage $1 + r_i \cdot (z^{-1} + z^{-2}) + z^{-3}$, $\forall i = 0, \ldots, \log_2 D - 1$ in the cascaded architecture shown in Fig. 8.

A suitable cost function accounting for the complexity of the ith decimation stage, can be related to the average power consumption [16] of each stage

$$\frac{f_s}{2^i} \cdot \sum_{h=1}^{N_{\mathrm{add}}} n_{i,S_h} \cdot W_{i,S_h} \tag{25}$$

whereby $f_s/2^i$ is the operating rate of the ith decimation stage, n_{i,S_h} is the number of adders of type h, N_{add} is the number of adders belonging to the ith stage, and W_{i,S_h} is the data word length of the hth adder.

In our setup, we neglect the power dissipation of delay elements. Terms constituting the summation above can be evaluated by analyzing the architecture of each cascaded cell in Fig. 8. Fig. 11 shows a simple implementation of the ith decimation stage, whereby acronyms IDWL_i and ODWL_i stands, respectively, for input and output data word lengths, and the subscript i identifies the decimation stage. Notice that, $\forall i = 0, \ldots, \log_2 D - 1$, $\mathrm{IDWL}_i = \mathrm{ODWL}_{i-1}$, and $\mathrm{IDWL}_0 = \mathrm{ODWL}_{-1}$ is the size of the data output from the ΣΔ modulator.

Given IDWL_i, ODWL_i can be evaluated through a simple, yet rather conservative, calculation. First of all, note that the adder $S_{i,1}$ operates on a data word-length equal to $\mathrm{IDWL}_i + 1$ bits. Second, multiplier r_i is approximated as in (24) by employing

$$T(r_i) = \sum_{v=1}^{L_i} \left| b_v^{(i)} \right|$$

minus one additions/subtractions operating on data with word-length equal to $(\mathrm{IDWL}_i + 1) + (M_i + 2) + (T(r_i) - 1)$, where M_i is defined as

$$M_i = \max_{\forall v \in \{1, \ldots, L_i\}} q_v^{(i)}.$$

Notice that $(M_i + 2)$ accounts for the extra bits required for accomplishing the data-shift $2^{-q_v^{(i)}}$, while $(T(r_i) - 1)$ is the number of extra bits to be allocated for the addition between $T(r_i)$ data. Adder $S_{i,2}$ operates on data word-lengths equal to

$M_i + \mathrm{IDWL}_i + T(r_i) + 4 = \mathrm{ODWL}_i$, since two extra bits should be reserved for accounting for carry bits.

A suitable complexity cost function should minimize the number of additions times the data word-length of the shift-register implementing the approximation of multiplier r_i, i.e., the function $F(r_i) = [T(r_i) - 1] \cdot [M_i + 2 + \mathrm{IDWL}_i + T(r_i)]$. This is tantamount to minimizing summation in (25), since the only degree of freedom is represented by the shift-register implementing multiplier r_i.

Based on the discussion above, the optimization problem can be formulated as follows. For each multiplier r_i, $\forall i = 0, \ldots, \log_2 D - 1$, solve the following optimization algorithm:

$$\begin{aligned} \text{minimize}: \quad & F(r_i) \\ \text{subject to}: \quad & 1)\ |\Delta r_i| = |r_i - \widetilde{r}_i| \leq \Delta r_t \\ & 2)\ \widetilde{P}_{\mathrm{qn}|_{dB}} - P_{\mathrm{qn}|_{dB}} \leq \Delta_P \leq 0 \end{aligned} \tag{26}$$

whereby P_{qn} is defined as in (4) upon using $H_{C_3}(f_d)$ in place of $H_{\mathrm{GCF}_N}(f_d)$, while $\widetilde{P}_{\mathrm{qn}}$ is the $\Sigma\Delta$ QN power using the decimation filter $\widetilde{H}_{\mathrm{GCF}_3}(e^{j\omega})$.[1] Since at step i only multipliers $\widetilde{r}_0, \ldots, \widetilde{r}_{i-1}$ have been already approximated, the frequency response $\widetilde{H}_{\mathrm{GCF}_3}(e^{j\omega})$, needed for evaluating $\widetilde{P}_{\mathrm{qn}}$, considers the approximated multipliers $\widetilde{r}_0, \ldots, \widetilde{r}_i$ plus the real multipliers $r_{i+1}, \ldots, r_{\log_2 D - 1}$.

In brief, the main goal of the optimization algorithm in (26) is to find the best approximation of each multiplier r_i within a predefined error Δr_t, in such a way as to guarantee a QN power noise rejection superior to the one accomplished by a classical comb filter across the folding bands, while minimizing an appropriate complexity metric.

Let us discuss the constraints imposed by the optimization problem in (26). First of all, notice that it has to be solved for each multiplier r_i at a time by starting from the multiplier r_0 belonging to the first decimation stage in the decimation chain in Fig. 8. The reason relies on the fact that the evaluation of the cost function $F(r_i)$ requires the knowledge of the data word-length IDWL_i at the input of the current decimation stage, while the other metrics, M_i and $T(r_i)$, depend on the specific approximation tested during the cost function minimization.

The first constraint in (26), which imposes a maximum approximation error target, Δr_t, on each multiplier r_i, is useful for reducing the complexity burden of the optimization algorithm. The choice of Δr_t can be accomplished by analyzing the effects of such an error on the $\Sigma\Delta$ QN rejection performance. To this aim, we evaluated a set of curves showing the difference between the $\Sigma\Delta$ QN rejection guaranteed by filter $\widetilde{H}_{\mathrm{GCF}_3}(f_d)$ in (22) and the one imposed by a classical third-order comb filter, as a function of the decimation factor D, and assuming various values of the approximation error $\Delta r_i = \Delta r_j$, $\forall i, j = 0, \ldots, \log_2 D - 1$, as parameters. Such curves are shown in Fig. 12 for the approximation errors Δr ranging from 10^{-1} to 10^{-5}. This figure clearly shows that denoising gains with respect to a classical comb filter can be guaranteed upon choosing operating points $(D, \Delta r)$ corresponding to negative values of $\Delta P_{\mathrm{qn}} - [\mathrm{dB}]$. In particular, for any $\Delta r \leq 10^{-4}$, filter $\widetilde{H}_{\mathrm{GCF}_3}(f_d)$ still guarantees denoising gains with respect to $H_{C_3}(f_d)$ in (1). Practically speaking, upon approximating each multiplier r_i within an approximation error less than or equal to $5 \cdot 10^{-5}$, it is possible to obtain $\Delta P_{\mathrm{qn}} \propto 7 \div 8$ dB equal to the one guaranteed by the filter $H_{\mathrm{GCF}_3}(f_d)$ with real multipliers, for any D in the range $2^3 \div 2^{10}$.

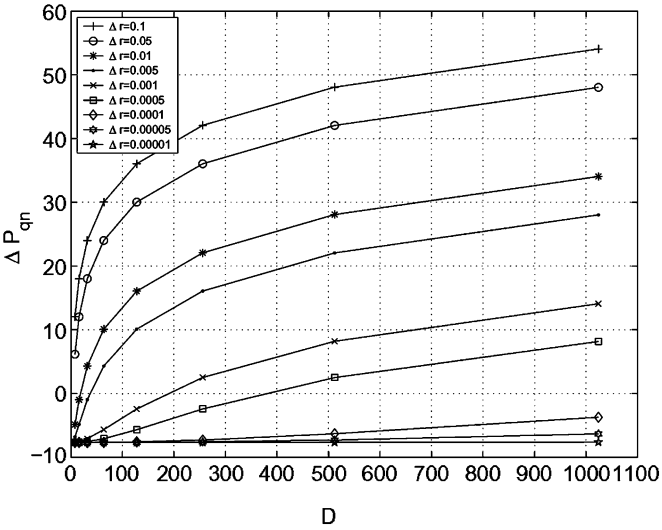

Fig. 12. Behaviour of the difference $\Delta P_{\mathrm{qn}} = \widetilde{P}_{\mathrm{qn}|_{\mathrm{dB}}} - P_{\mathrm{qn}|_{\mathrm{dB}}}$ as a function of the decimation factor D and for various values of the approximation errors Δr shown in the legend.

Finally, notice that we have not considered any constraint on the maximum passband drop error due to multipliers' approximations: this is essentially due to the intrinsic insensitivity of the passband behavior of the frequency response in (17) on the multipliers r_i.

VIII. Design Examples

This section addresses the design of a sample $H_{\mathrm{GCF}_3}(f_d)$ decimation filter considering the following set of parameters: $D = 16$, $\nu = 4$, $\Delta r_t = 10^{-2}, 5 \cdot 10^{-5}$, $R_M = 15$, $L_i = 6$, $\forall i = 0, \ldots, 3$, $\Delta_P = 7$ dB.

Some of these parameters have been identified following the considerations presented in the previous sections. On the other hand, $R_M = 15$ and $L_i = 6$ have been chosen for limiting the complexity burden of the optimization algorithm in (26): both parameters identify the limits of the set of values needed for the approximation in (24). We also assume the use of a $\Sigma\Delta$ modulator employing a 1-bit quantizer.

Results of the optimization algorithm are shown in Table II along with the approximated multipliers $\widetilde{r}_i$s. The values of both M_i and $T(r_i)$ refer to the ones obtained after the optimization for implementing the respective approximated multiplier $\widetilde{r}_i$. Notice that upon reducing the target approximation error Δr_t, it is possible to represent each multiplier $\widetilde{r}_i$ with a smaller number of PO2 coefficients, each one represented on a shorter size.

It is useful to compare the frequency responses of filters $H_{\mathrm{GCF}_3}(f_d)$ and $\widetilde{H}_{\mathrm{GCF}_3}(f_d)$ with the optimization results shown in Table II. To this end, Fig. 13 shows the behaviors

[1] Notice that $\widetilde{H}_{\mathrm{GCF}_3}(f_d)$ is short for $\widetilde{H}_{\mathrm{GCF}_3}(e^{j\omega})$.

TABLE II
RESULTS FOR THE SAMPLE FILTER OPTIMIZATION DISCUSSED IN SECTION VIII FOR $D = 16$

$\Delta r_t = 5 \cdot 10^{-5}$			
IDWL_0	1 bit	IDWL_1	19 bit
$\mathrm{0DWL}_0$	19 bit	$\mathrm{0DWL}_1$	40 bit
M_0	11	M_1	13
$T(r_0)$	4	$T(r_1)$	5
$\widetilde{r_0}$	$3 - 2^{-9} + 2^{-11}$	$\widetilde{r_1}$	$3 - 2^{-7} + 2^{-9} - 2^{-13}$
IDWL_2	40 bit	IDWL_3	59 bit
$\mathrm{0DWL}_2$	59 bit	$\mathrm{0DWL}_3$	80 bit
M_2	13	M_3	12
$T(r_2)$	6	$T(r_3)$	6
$\widetilde{r_2}$	$3 - 2^{-5} + 2^{-7} - 2^{-11} - 2^{-13}$	$\widetilde{r_3}$	$3 - 2^{-3} + 2^{-5} - 2^{-9} + 2^{-12}$
$\Delta r_t = 10^{-2}$			
IDWL_0	1 bit	IDWL_1	17 bit
$\mathrm{0DWL}_0$	17 bit	$\mathrm{0DWL}_1$	32 bit
M_0	9	M_1	7
$T(r_0)$	3	$T(r_1)$	3
$\widetilde{r_0}$	$3 - 2^{-9}$	$\widetilde{r_1}$	$3 - 2^{-7}$
IDWL_2	32 bit	IDWL_3	44 bit
$\mathrm{0DWL}_2$	44 bit	$\mathrm{0DWL}_3$	57 bit
M_2	5	M_3	5
$T(r_2)$	3	$T(r_3)$	4
$\widetilde{r_2}$	$3 - 2^{-5}$	$\widetilde{r_3}$	$3 - 2^{-3} + 2^{-5}$

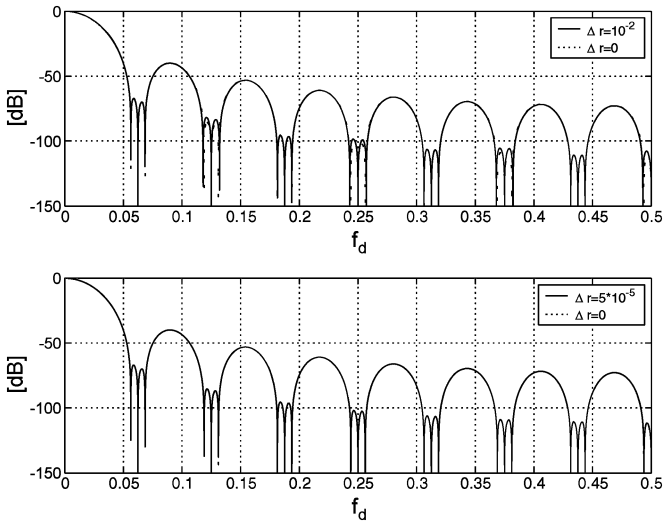

Fig. 13. Modulo of frequency responses of filters $\widetilde{H}_{\mathrm{GCF}_3}(f_d)$ (identified by $\Delta r = 10^{-2}$ and $\Delta r = 5 \cdot 10^{-5}$), embedding the multipliers noted in Table II, and $H_{\mathrm{GCF}_3}(f_d)$ (identified by $\Delta r = 0$ in both subplots) for $f_c = 1/2\rho$, $D = 16, \nu = 4, \alpha = 0.79 \cdot 2\pi f_c$.

of the frequency responses $\widetilde{H}_{\mathrm{GCF}_3}(f_d)$ (identified in the two subplots respectively with the labels $\Delta r = 10^{-2}$ and $\Delta r = 5 \cdot 10^{-5}$) and $H_{\mathrm{GCF}_3}(f_d)$ (identified in both subplots with the label $\Delta r = 0$) as a function of the frequency f_d. Frequency responses $\widetilde{H}_{\mathrm{GCF}_3}(f_d)$ employ the approximated multipliers presented in Table II. The parameters are noted in the respective figure labels. Both subplots clearly show that the frequency responses employing the approximated multipliers shown in Table II, are practically superimposed to the actual frequency responses $H_{\mathrm{GCF}_3}(f_d)$, even for approximation errors as low as 10^{-2}. This in turn suggests that the main performance metrics of the implemented filters, i.e., passband drop and selectivity, are the same as the theoretical ones derived in the previous sections.

The computational complexity of the proposed filters is higher than the one imposed by classical comb filters implemented in according to the architecture shown in Fig. 8. This is the price to be paid for increasing both the ΣΔ QN rejection around the folding bands and the filter selectivity. Let us try to derive a rough estimation of the computational complexity of the proposed filters by making a parallel with the one required by classical comb filters.

A classical third-order comb filter can be obtained upon considering $r_i = 3, \forall i = 0, \ldots, \log_2 D - 1$. By doing so, each decimation stage embeds a decimation filter with transfer function $(1 + z^{-1})^3$. The word length at the input of the ith decimation stage is equal to $\mathrm{IDWL}_i = \mathrm{IDWL}_0 + 3 \cdot i, \forall i = 0, \ldots, \log_2 D - 1$, while the word length at the output of the ith decimation stage is equal to $\mathrm{ODWL}_i = \mathrm{IDWL}_0 + 3 \cdot (i + 1)$, $\forall i = 0, \ldots, \log_2 D - 1$. The number of additions required in each decimation stage is constant and equal to 3.

On the other hand, the number of additions required by the proposed filters is $T(r_i) + 2, \forall i = 0, \ldots, \log_2 D - 1$, since there are two adders, namely $S_{i,1}$ and $S_{i,2}$, in according to the

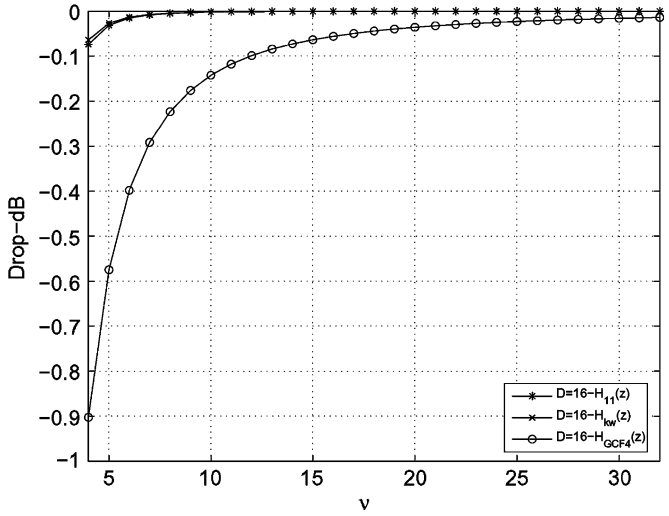

Fig. 14. Passband drops of filters $H_{\mathrm{GCF4}}(f_d)$, $H_{11}(f_d)$ and $H_{kw}(f_d)$, for $D = 16$, as a function of the residual decimation ν.

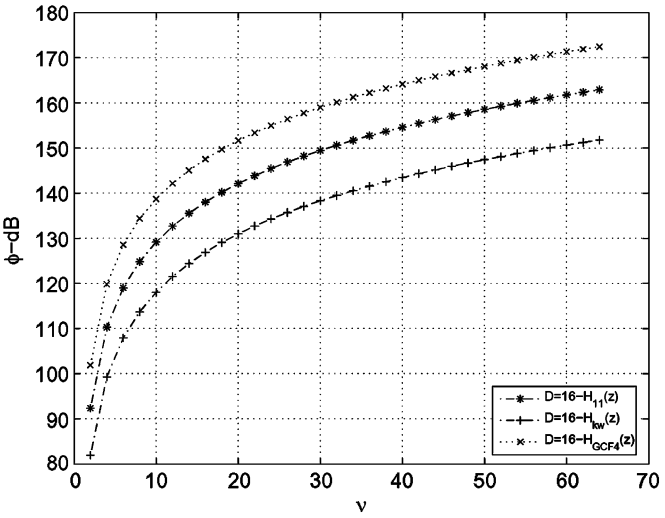

Fig. 15. Selectivities of filters $H_{\mathrm{GCF4}}(f_d)$, $H_{11}(f_d)$ and $H_{kw}(f_d)$, for $D = 16$, as a function of ν.

scheme shown in Fig. 11. In particular, adder $S_{i,2}$ operates on $T(r_i) + 1$ data. As far as the data word lengths are concerned, the results are the one shown in Table II. We notice that appropriately choosing the approximation error Δr, it is possible to reduce both the number of additions and the data word lengths in each decimation stage, as suggested by the design examples shown in Table II.

IX. Comparisons and Simulation Results

In this section, we compare a sample filter from the class of GCFs with fourth-order decimation filters presented in literature, addressing the design of a single decimation stage to be embedded in a multistage implementation. First of all, notice that performance comparisons with respect to classical comb filters have been already dealt with in Figs. 5 and 6 in terms of passband drop and selectivity, respectively.

Fig. 14 compares the passband drops of the proposed filter $H_{\mathrm{GCF4}}(z)$ with both the fourth-order sharpened comb filter proposed in [9] (here identified by $H_{kw}(z)$) and the fourth-order modified sharpened filter proposed in [11] (here identified by $H_{11}(z)$). Filter parameters are noted in the figure label. As expected, sharpened-based comb filters present superior performances in terms of passband band drop with respect to both classical and GCFs, at the cost of an increased computational complexity essentially due to the sharpening cell $[3 - 2H(z)]$, whereby $H(z)$ is the transfer function of a second-order classical comb filter. However, at high frequencies (i.e., for $f_d \gg 0$), the sharpening cell behaves as the constant 3 since $H(f_d) \approx 0$. This multiplicative term increases the values of both frequency responses $H_{kw}(f_d)$ and $H_{11}(f_d)$ in the frequency range above the signal bandwidth $[0, f_c]$. The main problem of such filters is to reduce both the selectivity and the $\Sigma\Delta$ QN rejection around the folding bands. Fig. 15 compares the selectivities of the three decimation filters for $D = 16$.

The effect of the multiplicative term in the stop band of the decimation filter $H_{11}(z)$ can also be seen by resorting to simulation. To this aim, upon employing Matlab we simulated a second-order $\Sigma\Delta$ converter with a 2-level quantizer into the loop, sampling a random signal with digital bandwidth $f_c = 1/256$, whereby $f_x = 100$ Hz, $f_s = 25.6$ kHz, and $\rho = 128$. The oversampled signal is then decimated by $D = 16$ employing filters $H_{11}(z)$ and $H_{\mathrm{GCF_4}}(z)$. The power spectra of the decimated signals at the output of filters $H_{\mathrm{GCF4}}(z)$ and $H_{11}(z)$ are shown in the upper subplot in Fig. 16, while the lower subplot shows the power spectrum of the oversampled signal after decimation by $D = 16$ accomplished through a FIR filter guaranteeing a stopband as low as 100 dB. The latter signal can be interpreted as the desired signal after the first decimation stage by using an ideal FIR filter isolating the useful signal bandwidth and rejecting the $\Sigma\Delta$ QN in the stopband ranging above the digital frequency $D \cdot f_c \approx 0.0625$. The upper subplot shows that sharpened filters do not appropriately reject the $\Sigma\Delta$ QN in the stopband, and so the QN folding down to baseband after decimation is higher than the one guaranteed by GCFs.

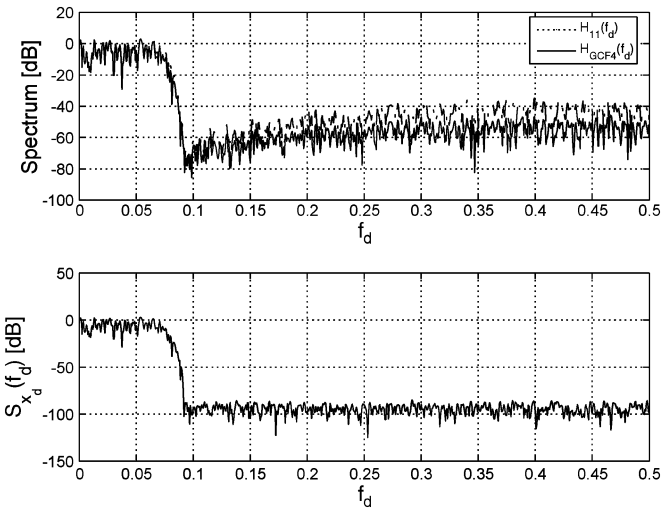

Fig. 16. The upper subplot depicts the power spectra of the signals at the output of the two decimation filters $H_{\mathrm{GCF_4}}(f_d)$ and $H_{11}(f_d)$ for $f_c = 1/256$, and $D = 16$. The lower subplot shows the power spectrum of the desired signal after the first decimation stage, i.e., a signal which has been decimated and filtered by a FIR filter guaranteeing a stopband as low as 100 dB.

X. Conclusion

This paper focused on the design of novel decimation filters, called GCFs, suitable for $\Sigma\Delta$ modulators. GCFs provide better performances in terms of both selectivity and quantization noise rejection with respect to conventional comb decimation filters, at the cost of both a very limited passband drop penalty and an increase in the computational complexity of the decimation filter realizations. Comparisons and tradeoff have been discussed with respect to classic comb filters of orders falling in the range 3 to 6. We also proposed an optimization algorithm for approximating the multipliers embedded in a sample nonrecursive filter architecture, with power-of-2 coefficients.

References

[1] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters, Theory, Design, and Simulation*. New York: IEEE Press, 1997.

[2] J. C. Candy, "Decimation for sigma delta modulation," *IEEE Trans. Comm.*, vol. COM-34, no. 1, pp. 72–76, Jan. 1986.

[3] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[4] E. B. Hogenauer, "An economical class of digital filters for decimation and interpolation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 2, pp. 155–162, Apr. 1981.

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

[6] M. Laddomada, L. Lo Presti, M. Mondin, and C. Ricchiuto, "An efficient decimation sinc-filter design for software radio applications," in *Proc. IEEE SPAWC*, Mar. 20–23, 2001.

[7] Y. Gao, J. Tenhunen, and H. Tenhunen, "A fifth-order comb decimation filter for multi-standard transceiver applications," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'00)*, Geneva, Switzerland, May 28–31, 2000, pp. III-89–III-92.

[8] H. Aboushady, Y. Dumonteix, M. Louërat, and H. Mehrez, "Efficient polyphase decomposition of comb decimation filters in $\Sigma\Delta$ analog-to-digital converters," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 48, no. 10, pp. 898–903, Oct. 2001.

[9] A. Y. Kwentus, Z. Jiang, and A. N. Willson, Jr., "Application of filter sharpening to cascaded integrator-comb decimation filters," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 457–467, Feb. 1997.

[10] G. Jovanovic-Dolecek and S. K. Mitra, "A new two-stage sharpened comb decimator," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 7, pp. 1414–1420, Jul. 2005.

[11] M. Laddomada and M. Mondin, "Decimation schemes for $\Sigma\Delta$ A/D converters based on Kaiser and Hamming sharpened filters," *Proc. IEE Vision, Image Signal Process.*, vol. 151, no. 4, pp. 287–296, Aug. 2004.

[12] M. Laddomada, "Comb-based decimation filters for $\Sigma\Delta$ A/D converters: Novel schemes and comparisons," *IEEE Trans. Signal Process.*, May 2007, to be published.

[13] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*. New York: McGraw-Hill, 1998.

[14] S. Chu and C. S. Burrus, "Multirate filter designs using comb filters," *IEEE Trans. Circuits Syst.*, vol. CAS-31, no. 11, pp. 913–924, Nov. 1984.

[15] 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.

[16] B. A. White and M. I. Elmasry, "Low-power design of decimation filters for a digital IF receiver," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 8, no. 3, pp. 339–345, Jun. 2000.

Massimiliano Laddomada (S'00–M'03) was born in 1973. He received the degree in electronics engineering and the Ph.D. degree in communications engineering from Politecnico di Torino, in 1999 and 2003, respectively.

He is currently a Research Associate at Politecnico di Torino and a part-time faculty at California State University (CSU), Los Angeles. From June 2000 to March 2001, he was a Visiting Researcher at CSU, and a Consultant Engineer at Technoconcepts, Inc., Los Angeles, a start up company specializing in Software Radio. His research is mainly in wireless communications, especially modulation and coding, including turbo codes and, more recently, networks coding.

Dr. Laddomada was awarded a five-year open-ended fellowship by E.D.S.U. in recognition of his university career as an Electronics Engineer. In 2003, he was awarded with the *Premio Zucca per l'Innovazione nell'ICT* from Unione Industriale of Turin. Currently, he is serving as a member of the editorial board of *IEEE Communications Surveys and Tutorials*.