Distortion-Free Delta-Sigma Beamforming

Carl-Inge Colombo Nilsen, *Student Member, IEEE*, and Sverre Holm, *Senior Member, IEEE*

***Abstract*—We have analyzed a compact, digital delay-and-sum beamformer using 1-bit delta-sigma A/D converters, where dynamic delays are applied directly to the delta-sigma modulated bitstream. It has previously been shown that the introduction of this time-variant operation into the time-invariant delta-sigma demodulator distorts the output signal, and several methods for minimizing this effect have been suggested. We have described the effect using standard system terminology. Based on this description, a new method for removing the distortion without any significant increase in system complexity is introduced. Its advantages, performance-wise, are supported by theoretical assessments and simulations.**

I. Introduction

The beamformer is a central component in most systems where arrays of transducers are used to acquire signals. It performs electronic steering or focusing of beams during transmission and reception of signals and operates simultaneously on the signals from each active transducer element. If there is a large number of active elements, as is often the case in high-resolution imaging applications such as medical ultrasound and sonar, the beamformer becomes both physically large and expensive to manufacture. These characteristics put restraints on availability (due to high cost) and applications (due to a low degree of portability) of such systems. To counteract such unwanted effects, a lot of research has been devoted to the field of compact beamformers.

Compactness in a beamformer has traditionally been attained in one of 2 ways:

- Reducing the number of active channels and possibly compensating for decreased resolution using more sophisticated beamforming algorithms (this is the case in, for instance, synthetic aperture schemes [1] or sparse arrays [2])
- Reducing the per channel complexity of the beamformer by using simpler hardware (e.g., [3]) or optimizing existing algorithms (e.g., [4])

One particular instance in which the latter line of thought has been applied is delta-sigma beamforming ([5]–[7], among others). By using simple, low-cost delta-sigma A/D converters, size, complexity, and cost can be reduced. The oversampled nature of delta-sigma bitstreams has the added advantage of making post-sampling interpolation unnecessary, further simplifying the beamformer. (Interpolation is usually required for attaining sufficient delay accuracy in a conventional beamformer; see [8] e.g.) However, the modulated/coded, 1-bit nature of delta-sigma signals has opened for some rather interesting challenges pertaining to processing in the beamformer. Performing dynamic receive focusing on such signals has been known to distort the resulting image, due to what has been described as loss of modulator-demodulator synchronization [6].

The 1-bit delta-sigma beamforming principle is shown for a single channel in Fig. 1 together with the classical pulse coded modulation (PCM), i.e., multi-bit word, approach to provide an intuitive feeling for how the time-variant operation of dynamic delays has unwanted effects on a bitstream, as one in reality operates on a subsample level with respect to the desired PCM stream. The reason for this is that the PCM stream is the output of an M-tap reconstruction finite impulse response (FIR) filter ($H_{\mathrm{LPF}}(z)$), so each 1-bit input sample will actually contribute to M/R output samples, and each output sample will have contributions from M 1-bit samples. In other words, by dynamically delaying the delta-sigma modulated bitstream, we affect the corresponding PCM samples on a subsample level, an effect not present in a beamformer operating directly on PCM samples.

Several methods have been suggested to compensate for this by minimizing the error that is introduced when applying dynamic delays. In [6] and [9], Freeman *et al.* showed that the distortion following the introduction of dynamic focusing manifested itself as an increase of the noise floor with a magnitude of almost 20 dB in the resulting image. (It is important to note that this number is a specific example, only included here to indicate that the order of magnitude is large and not to be understood as being generally true.) They suggested several corrections, all designed to make the delta-sigma modulation adjust itself to the new bit(s) introduced as a result of delay increments. One of their suggestions is the insert-zero/split-sample technique (from here on referred to as *recoding*). The technique is based on the fact that no 1-bit sample should ever be repeated due to delay adjustments. Insert-zero recoding will, as the name suggests, insert a zero in the stream where the repeated sample would be placed. Split-sample recoding will split the value of the sample to be repeated (which is ± 1 for a 1-bit delta-sigma modulator) between its original position in the stream and the position where the duplicate would be placed. This latter approach calls for more than 2 levels (i.e., more bits) to represent the recoded signal. An additional technique, based on nonuniform sampling, has been suggested [7]. The method will not be discussed further in this article, be-

Manuscript received March 27, 2007; accepted January 28, 2008.
The authors are with the University of Oslo, Department of Informatics, Oslo, Norway (e-mail: carlingn@ifl.uio.no).
Digital Object Identifier 10.1109/TUFFC.2008.857

0885–3010/$25.00 © 2008 IEEE

Luxottica et al. v. Solos
IPR2026-00412
Petitioners' Ex. 1020

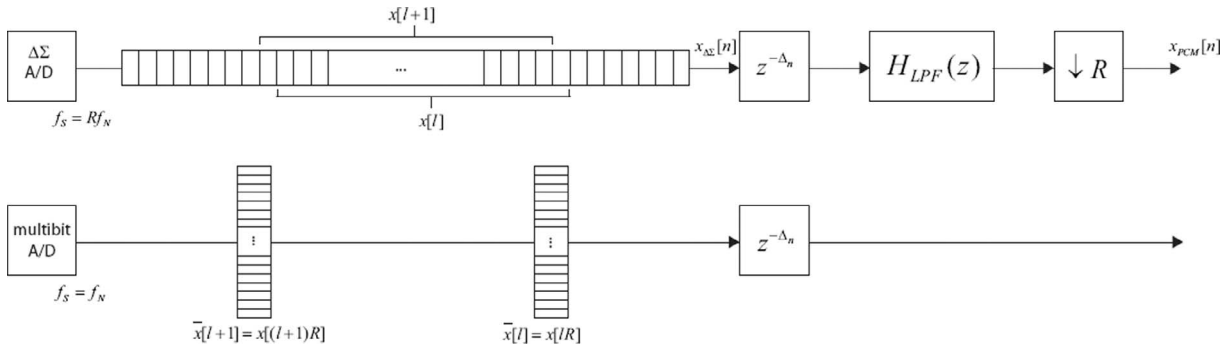

Fig. 1. The delta-sigma beamforming principle (upper branch), shown with the standard PCM approach (lower branch). The figure shows how the information pertaining to one sample of measurement is "spread out" in time in the delta-sigma beamformer, whereas it is located at one specific time instance with higher amplitude accuracy in the standard PCM beamformer. The signal $x_{\mathrm{PCM}}[n]$ is the (demodulated) PCM signal resulting from reconstruction filtering of the delta-sigma 1-bit signal.

cause it approaches the problem in a fundamentally different way by adjusting the sampling process itself instead of manipulating the digital samples after delta-sigma modulation.

A fundamental problem is that the underlying nature of the distortion has never been fully investigated or explained, making the quality of the above-mentioned corrections hard to evaluate beyond the fact that they have been shown to be effective through simulations. As a result of this, comparing different corrections or improving on them is not a simple task.

In this paper, we explain the distortion following dynamic delays in a new way, and the associated effects on the signal are interpreted and approximated using the theory of linear and time-invariant (LTI) systems. Simulations are performed to verify these approximations. This analysis makes it possible to identify the strengths and limitations of existing corrections such as insert-zero recoding. Specifically, we have found that it is not only the introduction of new bits by repetition, but also the perturbation of samples in the modulated stream, that causes the distortion. This has made it possible to specify an optimal correction with a compact implementation. The meaning of *optimal* in this context is that no additional distortion is introduced by the beamformer, and the meaning of *compact implementation* is that the number of operations performed is reduced from that of a trivial application of delta-sigma A/D converters in a digital beamformer. The idea behind this correction, which is called *subfiltering*, has previously been introduced in [10]. This paper presents an extension of these earlier results by taking the complete system description into consideration and comparing it to existing techniques.

Delta-sigma modulation and the nature of delay increments following time-variant dynamic receive focusing will be briefly explained in Section II, because these are central topics in the following analysis. The analysis will be based on the notion of an ideal and noncompact delta-sigma beamformer, which is introduced in Section III. Section IV contains the actual analysis and an example on how it can be applied to an existing technique, namely insert-zero recoding. Finally, an optimal and compact beamformer is introduced and explained in Section V. An example showing how this beamformer eliminates the remaining distortion is shown in Section VI, along with an overview of the computational complexity of the different solutions.

II. Delta-Sigma Modulation and Digital Beamforming with Dynamic Focusing

A. Delta-Sigma A/D Conversion

The primary virtue of the delta-sigma A/D converter is that it requires very little analog circuitry. Because it additionally shows a relatively large degree of tolerance for inaccuracies in such circuitry, there is little need for high-precision (i.e., expensive) analog components. This is due to the fact that, contrary to conventional A/D converters, the delta-sigma converter generates digital sequences of few bits per sample, often only one single bit. A satisfactory resolution (i.e., low sampling period) and accuracy (i.e., small quantization effects) in the resulting digital sequence is attained due to a combination of oversampling and quantization-noise shaping. The theory behind first- and second-order delta-sigma modulators and their discrete time equivalent systems are described in detail in [11]. For modulators of orders 3 and above, stability becomes an issue, and modulator design is no longer trivial.

The effect of using finite-word-length samples in the digital representation of a signal can be viewed as an additive noise process. This process can usually be approximated as white noise with a magnitude depending on the word-length, i.e., the number b of bits/sample. This so-called quantization noise floor can be reduced by 6 dB by increasing b by 1 bit/sample or vice versa, depending on how this model is viewed. For a first-order delta-sigma modulator, the system function with output signal $y[n]$, input signal $x[n]$, and quantization noise component $e[n]$ can be written as

$$Y(z) = z^{-1}X(z) + (1 - z^{-1})E(z) \tag{1}$$

meaning that the quantization noise is high-pass filtered (hence the term *noise shaping*) by an FIR filter with frequency response $H(e^{j\omega}) = 1 - e^{-j\omega}$. The corresponding system function for a second-order delta-sigma modulator is

$$Y(z) = z^{-1}X(z) + (1 - 2z^{-1} + z^{-2})E(z), \qquad (2)$$

from which we can observe that the high-pass filtering of the quantization noise becomes more efficient when the modulator order is increased. When oversampling is performed, a relatively low amount of quantization noise energy is contained in the signal band. This means that the amount of quantization noise decreases as the oversampling ratio (OSR) increases, causing better amplitude resolution because a lower amount of quantization noise can be interpreted as more bits per sample. To retrieve the desired Nyquist-rate PCM-valued signal, the delta-sigma modulated sequence must be low-pass filtered to remove the high-pass quantization noise. This process is often referred to as reconstruction filtering or demodulation and requires a very sharp low-pass filter (i.e., of a high filter order) to attenuate the high-pass filtered quantization noise. This requirement is also due to the fact that oversampling relaxes the analog anti-aliasing filter applied before sampling, resulting in a certain amount of unfiltered out-of-band signal components remaining in the digital signal. Note that the analysis in this article is based on the assumption that the reconstruction low-pass filter is an FIR filter.

B. Digital Beamforming in Medical Ultrasound

The main task of a beamformer is to steer or focus a beam. In medical ultrasound, the sources are always located in the near-field, making focusing necessary. The actual focusing can be performed in the time domain by delaying the signal from each active element in the array by the propagation time from the focal point relative to a reference point called the phase center, which is often the center element of the array. A digital receiving beamformer can focus on all ranges $r_n = nT_s c/2$ [m] associated with the nth sample acquired at a sampling rate $f_s = 1/T_s$ [Hz] when the acoustic waves propagate at a speed of c [m/s]. The (discrete-time) delay for an angle θ away from broadside, for element i at the time of sample n (assuming $n = 0$ represents the time of transmission) is

$$\Delta_n^i = \left\lfloor \frac{r_n^i f_s}{c} \right\rfloor = \left\lfloor \frac{\left(\frac{n}{2f_s c} - \sqrt{d_i^2 - \frac{d_i n \sin(\theta)}{f_s c} + \left(\frac{n}{2f_s c}\right)^2} \right) f_s}{c} \right\rfloor \text{ [samples]} \qquad (3)$$

where d_i is the signed distance between the ith element and the phase center, which for a symmetric, uniformly spaced linear array (ULA) with N elements positioned along the x-axis is

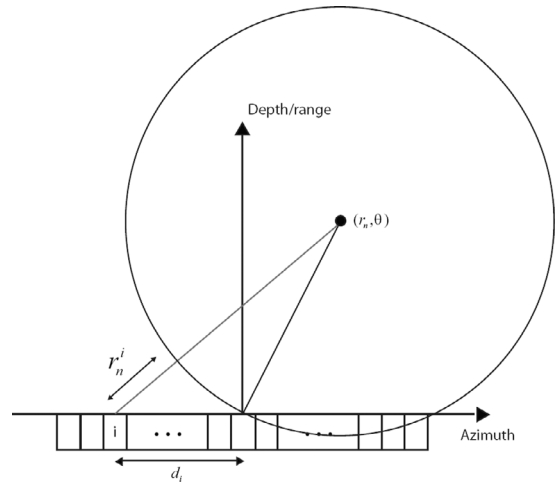

Fig. 2. The geometry of focusing in ultrasound imaging: element i is located a distance d_i from the center of the array (which coincides with the origin of the coordinate system). When focusing at a point with polar coordinates (r_n, θ), r_n^i is the additional distance the wave/pulse must travel when compared with the center element. This distance divided by the speed of propagation represents the delay that is applied to the signal transmitted from or received at element i.

$$d_i = \left(i - \frac{N-1}{2} \right) d \text{ [m]}. \qquad (4)$$

The geometry of the problem is illustrated in Fig. 2. The delayed contributions, $w_i[n] = x_i[n - \Delta_n^i]$, from all channels $i = 0, \cdots, N-1$ are combined to form the beamformer output:

$$y[n] = \sum_{i=0}^{N-1} w_i[n] = \sum_{i=0}^{N-1} x_i[n - \Delta_n^i]. \qquad (5)$$

The expression in (3) depends on n, meaning that the system is time-variant; Δ_n needs only be adjusted for $n = kD$, where k is an integer and D is the down-sampling factor, to be applied after dynamic focusing at the oversampled rate. The effect of dynamic focusing on the signal can be interpreted as sample repetition. To see this, let K be the difference between 2 subsequent dynamic delays—$\Delta_{n+1} = \Delta_n + K$, $K > 0$—and find the output at time $n + 1$:

$$\begin{aligned} w_i[n+1] &= x_i[n + 1 - \Delta_{n+1}] = x_i[n + 1 - (\Delta_n + K)] \\ &= x_i[n - \Delta_n - (K-1)] = w_i[n - (K-1)] \end{aligned} \qquad (6)$$

where the last equality assumes time invariance over at least K samples. This is equivalent to $\Delta_{n-(K-1)} = \Delta_{n-(K-2)} = \cdots = \Delta_n$, which is guaranteed for $K < D$ when dynamic delays are only adjusted at the downsampled rate. This is a reasonable assumption, because the downsampled rate samples are the only ones that will be used in the resulting image (or whatever alternative output the system provides).

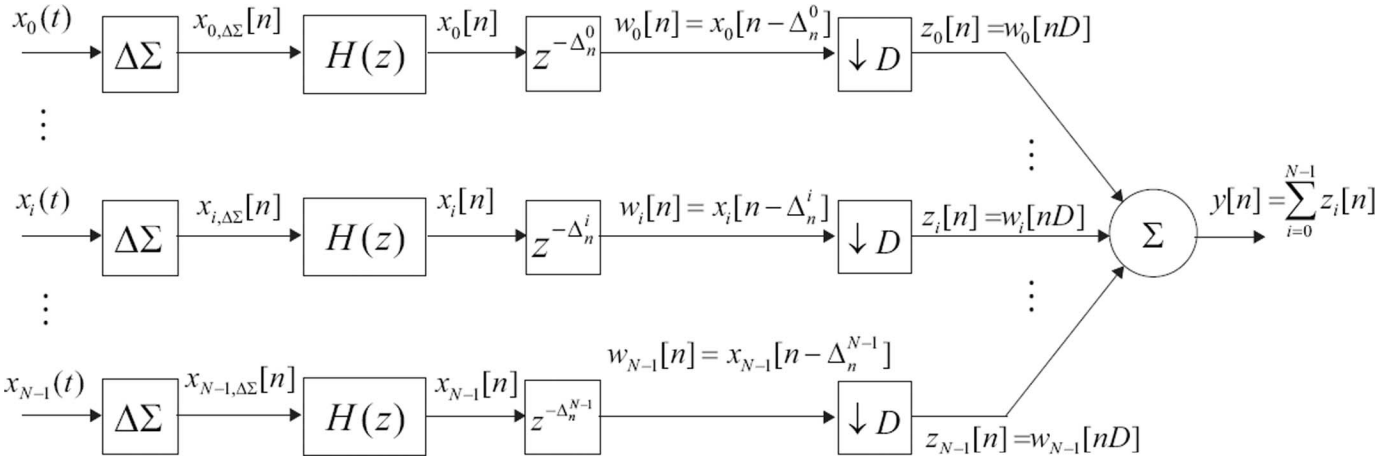

Fig. 3. An ideal delta-sigma-based beamformer, applying steering delays on PCM data at an oversampled rate and subsequently decimating the steered sequences to the Nyquist rate. Note that the signal will only take on 1-bit modulated values immediately after sampling is performed and before any actions associated with the actual beamformer take place, meaning that the choice of a delta-sigma A/D converter does not affect beamformer performance. The M-tap FIR filter $H(z)$ can therefore be included in a "black-box" A/D converter, creating a non-delta-sigma digital beamformer. The signals before the decimator are sampled at a rate f_s and the signals following the decimator at a rate f_s/D.

For the case of $K = 1$ in (6), which will be shown to be common for some ranges, this means that $w_i[n+1] = w_i[n]$. In other words, the same input sample $w_i[n] = x_i[n - \Delta_n]$ contributes to the beamforming twice. In general, a block of K samples is repeated. When this time-variant operation of dynamic focusing is performed in the middle of the LTI-system of delta-sigma modulation and PCM reconstruction, the resulting system is no longer LTI. In general, this may or may not alter the output to any significant degree. It has been shown through simulations that the alteration is significant in this case [9].

III. Ideal Delta-Sigma Beamforming

The time-invariance of the delta-sigma modulator-demodulator system will not be compromised if we apply dynamic delays after reconstruction filtering has been performed. Down-sampling to the Nyquist rate should not take place until after the delays have been applied, to achieve maximum resolution in time and a suitably low degree of delay quantization. This system is shown in Fig. 3 and represents a trivial application of delta-sigma modulator technology to a digital beamformer. It is deemed trivial because the organization and structure of the remaining components in the beamformer are not affected by introducing the delta-sigma A/D converter and demodulator into the system, i.e., no attempt to minimize the number of digital components (i.e., the reconstruction filters) is carried out. Having such a large number of parallel and identical M-tap FIR filters is a significant and seemingly unnecessary overhead in the hardware. However, because they are part of LTI systems (the delta-sigma demodulators) separated by a time-variant operation (the dynamic delays), we cannot use standard LTI-system theory to combine them into one single operation without creating the distortions that were discussed earlier.

The output of the ideal delta-sigma beamformer can serve as a gold standard to which the output of any other kind of delta-sigma based beamformer is compared. This is plausible because the only degradation of beamformer output quality as a result of applying delta-sigma A/D converters will be due to the choice of modulator structure and OSR, i.e., due to the introduction of the delta-sigma A/D converter itself and not as a result of how the beamforming is performed. The contributions to the beamformer output $y[n]$ from the delta-sigma signals on each channel are the terms in the first sum in the following expression:

$$\begin{aligned} y[n] &= \sum_{i=0}^{N-1} z_i[n] = \sum_{i=0}^{N-1} w_i[nD] = \sum_{i=0}^{N-1} x_i[nD - \Delta_{nD}^i] \\ &= \sum_{i=0}^{N-1} \sum_{k=0}^{M-1} h[k] x_{\Delta\Sigma,i}[nD - \Delta_{nD}^i - k]. \end{aligned} \tag{7}$$

IV. Analysis of the Effect of Dynamic Focusing on Delta-Sigma Modulated Sequences

To analyze the effects of dynamic focusing on a delta-sigma modulated signal, we should first quantify to which degree delay increments will occur in the individual channels.

The mean delay increments associated with focusing at a certain f-number (i.e., aperture-normalized distance)

$$f_\# = \frac{r}{L} = \frac{r}{|d_{N-1} - d_0|} \tag{8}$$

have been calculated for an example system with a 3.5 MHz center frequency, a 10 MHz Nyquist frequency, a speed of sound of 1540 m/s, and an OSR of 32, and they are shown in Fig. 4 as a function of relative element position on the array and focal f-number. Contours for $K = 0.5$,

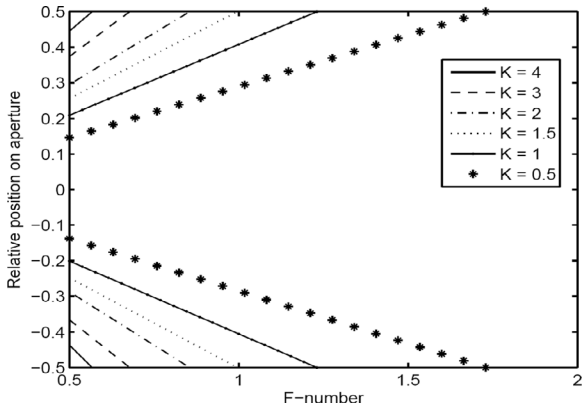

Fig. 4. Contour plot of the surface describes mean delay increments across steering angles between $\pm\pi$, which has been calculated for a system with center frequency 3.5 MHz, a Nyquist rate of 10 MHz, and an OSR of 32.

1, 1.5, 2, 3, and 4 are shown for clarity. This figure can be used to estimate the degree of delay incrementation for many different array configurations.

We can now compare the existing solution of insert-zero recoding to the ideal beamformer. Assuming that $K = 1$, an uncompensated, dynamically delayed delta-sigma sequence will be given as

$$w_{i,\text{uncomp}}[n] = \begin{cases} x_i[n(1-1/R)] & \text{for } n = mR, m \in \mathbb{Z} \\ x_i[\lfloor n/R \rfloor (R-1) + (m \bmod R)] & \text{otherwise} \end{cases} \tag{9}$$

where R is the oversampling ratio; $\lfloor \cdot \rfloor$ denotes rounding up or down to the closest integer; and n mod R is the remainder of n when divided by R. By defining the sequences

$$x_{i,\text{rep}}[n] = \begin{cases} x_i[n(1-1/R)] & \text{for } n = mR, m \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

and

$$x_{i,\text{del}}[n] = \begin{cases} 0 & \text{for } n = mR, m \in \mathbb{Z} \\ x_i[\lfloor n/R \rfloor (R-1) + (n \bmod R)] & \text{otherwise} \end{cases} \tag{11}$$

we can express the output as $w_{i,\text{uncomp}}[n] = x_{i,\text{del}}[n] + x_{i,\text{rep}}[n]$. From this, we see that insert-zero recoding can be expressed as $w_{i,\text{rec}}[n] = w_{i,\text{uncomp}}[n] - x_{i,\text{rep}}[n] = x_{i,\text{del}}[n]$. We can now comment on the performance of recoding by comparing $w_{i,\text{rec}}[n] = x_{\text{del}}[n]$ with $w_i[n]$ from the ideal beamformer in Fig. 3, an expression for which is given in (7). Comparing this to (7), we can determine that recoding is not an optimal correction in the sense that its output is not identical to that of the ideal beamformer. The recoding output sequence $x_{\text{del}}[n]$ is equal to the input sequence $x[n]$ with a zero inserted between every $(R-1)$ samples. This is approximately equivalent to the upsampling step in a fractional rate changer of order $I = (1 + 1/R)$, which can be shown as

$$\begin{aligned} X_{\text{del}}(\omega) &= \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega\left(n + \lfloor \frac{n}{R} \rfloor\right)} \approx \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega\left(n + \frac{n}{R}\right)} \\ &= X\left(\omega\left[1 + \frac{1}{R}\right]\right). \end{aligned} \tag{12}$$

From multirate theory, we know that the effect is a frequency domain scaling of order $R/(R+1)$. For $R = 32$, this means that approximately 3% of out-of-band frequencies will leak into the signal band, which is equivalent to a 150 kHz frequency displacement for an example system sampling at a rate of 10 MHz. This contraction/displacement takes place before reconstruction filtering, with the immediate result that the transition band is no longer where it was designed to be, with respect to the intended pass- and stop-bands of the sampled signal. There are 2 noise components in the stop-band: quantization noise from the delta-sigma converter and noise acquired in the sampling process (regular noise which is to be removed by the anti-aliasing filter). The magnitude of the former can be approximated using classical quantization noise approximation techniques, such as (1.21) in [12]:

$$|N(f)| = \sqrt{2T}\left[2\sin\left(\frac{2\pi f T}{2}\right)\right]^L \tag{13}$$

where T is the sampling period and L is the modulator order. The noise increase in the pass-band due to quantization noise can be found using the expression

$$20\log_{10}\left(\frac{\int_0^{f_S/2R}\left|N\left(f\left[1+\frac{K}{R}\right]\right)\right| df}{\int_0^{f_S/2R}|N(f)|\,df}\right), \tag{14}$$

where R is the OSR and k is the size of the delay increment. For a second-order modulator and $K = 2$, this only corresponds to a noise increase of 1 dB. The increase from noise not caused by quantization will vary with the situation, depending on the amount of noise located in the transition band and beginning of the stop band.

V. Subfiltering

An existing method for delta-sigma beamforming, with output that does not differ from that of the ideal beamformer, is the sparse sample-processing concept that is used in a delta-sigma beamforming context by Tomov and Jensen to implement a compact FPGA-based beamformer in [13]. It has previously been proposed in a general digital beamforming context in [14]. A similar technique has been suggested by Han *et al.* in [15], but this implementation is optimized for sinc reconstruction filters (accumulators). In this section, we will suggest a general method with low computational complexity that works for any type of reconstruction filter, along with a complete mathematical description of the method.

If we study the input-output equations associated with the ideal beamformer of Fig. 3, we see that each output sample depends on M input samples from each of the N channels. Samples are selected from the oversampled-rate, reconstructed signal $x_{\mathrm{PCM},i}[n]$ by the dynamic delay module and the down-sampler, and added together to generate the focused signal. As in any other decimator, the output $x_{\mathrm{PCM},i}[n]$ only needs to be calculated for each Dth sample with respect to the oversampled rate:

$$z_i[n] = \sum_{m=0}^{M-1} h[m] x_{\Delta\Sigma,i}[nD - \Delta^i_{nD} - m]. \tag{15}$$

By considering the corresponding equations for the time-variant system of Fig. 1 (upper branch), i.e., the system considered by previous articles, we get the following expression:

$$z_i[n] = \sum_{m=0}^{M-1} h[m] x_{\Delta\Sigma,i}[nD - \Delta^i_{nD-m} - m]. \tag{16}$$

When comparing (15) and (16), we need to make the following assumptions:

1. The first assumption has two possibilities.
   (a) There is a time index n such that $\Delta_{n+D} = \Delta_n + K$ (where K is some positive integer).
   (b) For all time indices nD, we have $\Delta_{nD+m} = \Delta_{nD}$ whenever $m < D$.

   The first possibility means that a delay increment of size K will occur at some point in time, which is a reasonable assumption according to the calculations presented in Section IV for f-numbers between 1 and 2. The second possibility means that delay increments only occur at time indices that are integer multiples of D. This is accomplished by only adjusting delays at a rate D, which makes sense if D is interpreted as the down-sampling factor.
2. $M \gg D$, i.e., the filter length is much greater than the down-sampling factor, which covers the general case of reconstruction filtering for significantly oversampled signals (but not the specific case of $sinc^P$-type filters, where $M = P(D-1)+1$ and P is relatively low).

Under these conditions, the output in (16) will always be different from that in (15). This is a more formal verification of the existence of the original flaw.

An alternative technique can be found by examining the output of some general decimator of degree D with a M-tap anti-aliasing filter with impulse response $\overline{h}[n]$, working on an input signal denoted $\overline{x}[n]$ to create the output $\overline{y}[n]$. Assuming $M = LD$, $L \in \mathbb{N}$ (which can be satisfied for any FIR filter by using zero padding), we can divide the input signal into L sub-signals filtered by L subfilters:

$$\begin{aligned}
\overline{y}[n] &= \sum_{m=0}^{M-1} \overline{h}[m]\overline{x}[nD-m] \\
&= \sum_{d=0}^{D-1} \overline{h}[d]\overline{x}[nD-d] \\
&\quad + \sum_{d=0}^{D-1} h[D+d]\overline{x}[nD-(D+d)] + \cdots \\
&\quad + \sum_{d=0}^{D-1} \overline{h}[(L-1)D+d]\overline{x}[nD-((L-1)D+d)] \\
&= \sum_{l=0}^{L-1}\sum_{d=0}^{D-1} \overline{h}[lD+d]\overline{x}[nD-(lD+d)].
\end{aligned} \tag{17}$$

Defining L new subfilters of length D as

$$\overline{h}_l[n] \triangleq \begin{cases} \overline{h}[lD+n] & \text{for } 0 \le n < D \\ 0 & \text{otherwise} \end{cases} \quad \text{for } 0 \le l < L, \tag{18}$$

with associated outputs

$$\overline{y}_l[n] = \sum_{d=0}^{D-1} \overline{h}_l[d]\overline{x}[nD-d] \tag{19}$$

we can write the output in (17) as

$$\overline{y}[n] = \sum_{l=0}^{L-1} \overline{y}_l[n-l]. \tag{20}$$

This means that each sample in the output sequence can be constructed from L samples from the different subfilter output sequences. This technique will now be applied to the delta-sigma beamformer. Because of this, we have the ability to control the dynamic delay for the input to each subfilter individually. The delayed, delta-sigma modulated contribution from the ith and lth channel to the input to the lth subfilter at time n is $x_{\Delta\Sigma,i}[n - \Delta^i_{n+lD}]$. This can be written as

$$\begin{aligned}
z_i[n] &= \sum_{l=0}^{L-1} y_l[n-l] \\
&= \sum_{l=0}^{L-1}\sum_{d=0}^{D-1} h_l[d] x_{\Delta\Sigma,i}[(n-l)D - d - \Delta^i_{nD}] \\
&= \sum_{m=0}^{M-1} h[m] x_{\Delta\Sigma,i}[nD - \Delta^i_{nD} - m]
\end{aligned} \tag{21}$$

which is identical to the output of (15). In other words, the contribution from one channel to the subfiltering beamformer is identical to the contribution from the same channel in the ideal beamformer, so the performances of the ideal and subfiltering beamformers must be equal. Note that one can use any decimation factor less than or equal to D (the length of the subfilters and the rate at which the dynamic delays are updated), as long as the contribution from i channel to subfilter l at time n is specified

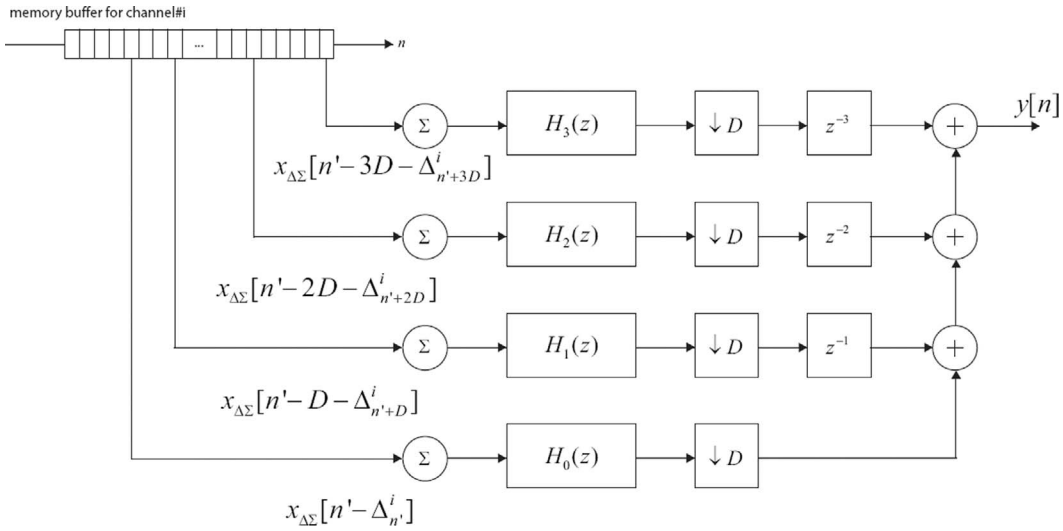

Fig. 5. The schematic diagram of the subfiltering algorithm for the case $L = 4$. The summation blocks show where the remaining channel sequences would be added.

as $x_{\Delta\Sigma,i}[n - l(D-1) - \Delta^i_{n+l}]$. The different subfilters can work in parallel, generating contributions to the D output samples $z_i[n]$ through $z_i[n - D + 1]$ simultaneously. This means that we do not need one large low-pass filter per contributing channel as in Fig. 3, but rather one set of small subfilters working in parallel. The final beamformer output can be written as

$$\begin{aligned} y[n] &= \sum_{i=0}^{N-1} z_i[n] \\ &= \sum_{l=0}^{L-1} \sum_{d=0}^{D-1} h_l[d] \sum_{i=0}^{N-1} x_{\Delta\Sigma,i}[(n-l)D - d - \Delta^i_{(n)D}] \end{aligned} \tag{22}$$

which means that each of the L subfilters will have an input signal that is a sum of one sample from each of the N channels. This equation reflects the sequence of operations in the subfiltering beamformer: First, the delta-sigma sequences from all channels are delayed and added together, then they are filtered by all subfilters, and finally the outputs of the subfilters are added.

The structure for the subfiltering delta-sigma beamformer is shown in Fig. 5, where L has been chosen as low as 4 for illustration purposes. The primary increase in complexity is due to the control system routing signals to the different subfilters, which could be implemented as a simple last-in-first-out (LIFO) buffer acting as an index to the internal memory module of the beamformer. Modules summing the N 1-bit contributions from all channels are required for each subfilter. Additional delays and sums following the individual subfilters are also necessary. This means that the system complexity grows with the number of subfilters, which in turn is proportional to the number of reconstruction filter coefficients and inversely proportional to the downsampling factor.

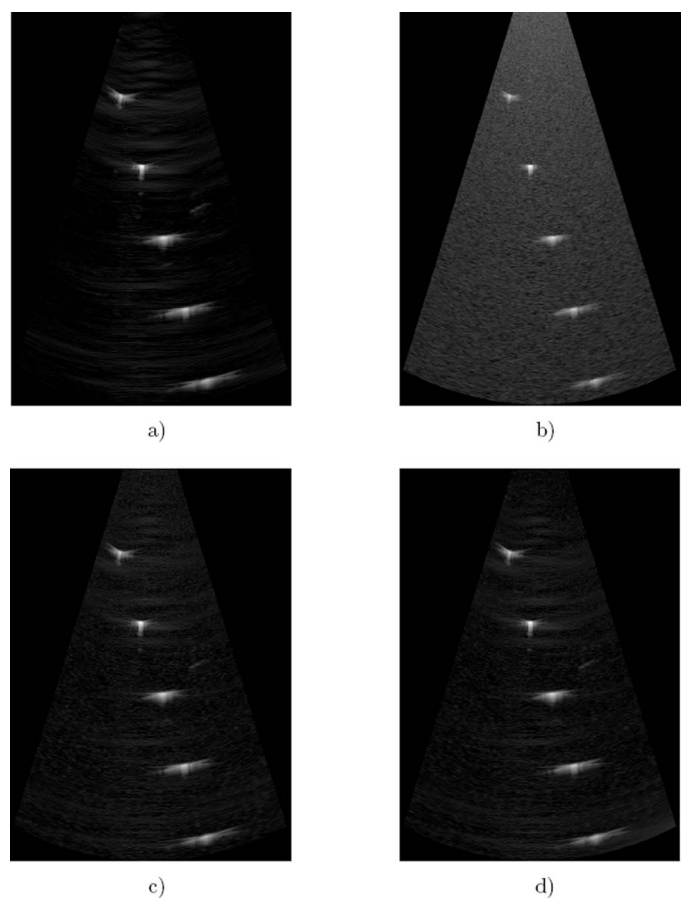

Fig. 6. B-mode images of wire phantoms formed using a) 10-bit A/D converter; b) uncompensated delta-sigma A/D converter; c) delta-sigma A/D converter and insert-zero recoding; and d) delta-sigma A/D converter and subfiltering.

VI. Performance and Simulations

In the previous sections, we have shown mathematically that subfiltering and all other solutions that generate an output equal to that of the ideal beamformer (like sparse sample processing) will remove an additional noise component present in delta-sigma beamformers. We have also shown that the magnitude of this noise due to quantization only is very small for a second-order delta-sigma A/D converter. The presence of this noise component, and its dependence on out-of-band noise, will be verified in this section through simulations. The different delta-sigma beamforming techniques have been simulated using Matlab (MathWorks, Natick, MA). We have used a wire phantom data set from Michigan Biomedical Ultrasonics Laboratory, which is also used in [9] with similar simulation parameters. The following are the most important simulation/data set parameters:

- Speed of sound: 1480 m/s
- Number of elements in array: 128
- Element spacing: $0.22 \cdot 10^{-3}$ m
- Delta-sigma modulator order: 2
- Original sampling rate: 13.8889 MHz
- OSR: 16
- Coefficients in reconstruction filter: 1500
- Effective number of bits in delta-sigma A/D converter: approximately 9
- Scan lines: 160 lines between $\pm\pi/8$.

All the images in Fig. 6(a)–(d) are shown with 70 dB dynamic range. Fig. 6(a) shows the resulting image using 10 bit A/D conversion, static transmit focusing, and dynamic receive focusing. Fig. 6(b) shows the resulting image using delta-sigma A/D conversion, static transmit focusing, and dynamic receive focusing with no attempt to correct the induced errors. The increase in noise is easily observed.

Fig. 6(c) shows the resulting image using delta-sigma A/D conversion and applying insert-zero recoding to reduce the noise. Fig. 6(d) shows the resulting image using delta-sigma A/D conversion and applying subfiltering to reduce the noise. The difference between these 2 B-mode images is not easy to observe, but more information can be obtained by inspecting plots of the cross section for a given depth. The cross sections across all angles for the 2 correcting methods are shown in the 3 lower curves of Fig. 7(a) and (b) together with those of the multibit beamformer. It is clear from the figures that recoding compensates for most of the distortion because it results in a noise level almost as low as that of the multibit A/D beamformer. The noise level for subfiltering is somewhat lower than this, and it yields the best performance that can be achieved with this combination of delta-sigma modulator order and OSR, which has been demonstrated in the previous section. The difference between the performance of the subfiltering beamformer and the PCM beamformer is due to the fact that the subfiltering beamformer has an approximate sample resolution of 9 bits, while the PCM samples are 10 bit.

The 2 upper lines in Figs. 7 and 8 show what happens when artificial noise is added to the signal before beamforming. The noise has been created by high-pass filtering white noise so that it only has frequency components for frequencies higher than half the sampling rate divided by the OSR, in this case, approximately MHz, giving it a bandwidth of approximately 6.9 MHz. The noise on different channels are independent of each other. The magnitude of the received noise, after reconstruction filtering, is approximately 10 dB higher than that of the noise in the original image. Prior to reconstruction filtering, the magnitude of the artificially added noise is approximately equal to that of the reflected signal. The degree of frequency contraction induced by the delta-sigma beamformer for the worst-case scenario of a delay increment of one sample occurring for every 32 samples in the stream is approximately equal to 200 kHz. The mean noise level is raised by about 10 dB for both recoding and subfiltering (indicated by the solid and dotted lines, respectively), but it is apparent that the difference between the 2 upper noise levels in Fig. 7 is larger than the difference between the 2 lower. The same can be seen from Fig. 8, in which the average intensity has been calculated across all depths of the image. This means that the distorting effect of the delay increments is increased as more out-of-band noise is introduced into the system, which confirms our earlier observations that the potential for improvement by using a delta-sigma beamformer with ideal performance depends on the amount of noise in the vicinity of the transition band of the reconstruction filter. We observe that, as the distance from the array increases from Fig. 7(a) to Fig. 7(b), the effect of the delay increments decreases rapidly. Note that the type of noise we have added to the signal is not meant to illustrate a realistic noise scenario for an ultrasound system, but to verify that the presence of out-of-band noise will have an effect on the beamformer output when delay increments are present. It is also important to note that all the simulations presented in this section only deal with B-mode image formation, which is just one of the many functions of an ultrasound beamformer. The impact of subfiltering on other such functions, like Doppler processing and flow imaging, is outside the scope of this article. Details on the application of recoding to these functions can be found in [16].

We will now compare the computational complexities of the different methods. The main results are summarized in Table I. Given an array of N elements and a delta-sigma modulator sampling with an OSR of R using a reconstruction filter with M taps and downsampling by a factor D, then recoding consists of

- $N-1$ additions (the sum of the beamformer)
- M multiplications (in the reconstruction filter)
- $M-1$ additions (in the reconstruction filter)

for each Rf_N/Dth time instant, giving a total of $N+M-2$ additions and M multiplications. Considering the same scenario with $L = M/R$, subfiltering consists of

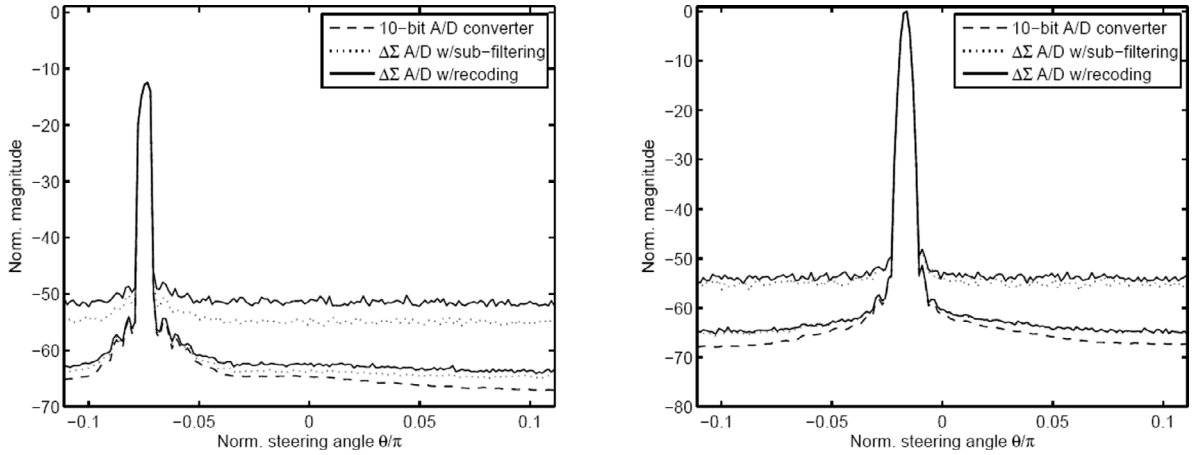

Fig. 7. Cross sections of the B-mode images in Fig. 6 for a) the first wire target and b) the second wire target. The upper curves show the cross sections when artificial noise is added to the signal. The lower curves show the cross section without any additional noise. It is apparent that the effect of the delay increments diminishes as the focal range increases. It is also clear that the difference between recoding and subfiltering is greater in the presence of out-of-band noise.

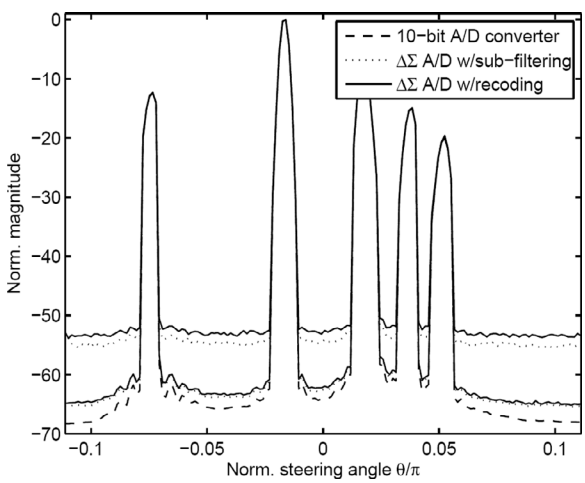

Fig. 8. The mean background noise level plotted against the peaks of the wire targets, taken from the B-mode images in Fig. 6. The 2 upper curves labeled subfiltering and recoding represent the signal after artificial noise has been added. The 2 lower curves labeled subfiltering and recoding represent the signal without any additional noise. The 10-bit signal is also without any additional noise. The difference between recoding and subfiltering is greater when the artificial noise is added to the signal.

TABLE I
Computational Complexity.

Method	Computational Complexity	Other
Recoding	$O(M)$	Increased word length into filter.
Subfiltering	$O(M)$	Fetch more samples from memory.
Ideal	$O(MN)$	None

The computational complexity is measured according to how the numbers of operations per output sample for all methods grow together with the number of elements in the array (N) and the number of coefficients in the reconstruction filter (M).

- $N-1$ additions before processing with each of the L subfilters
- $R-1$ additions in each of the L subfilters
- R multiplications in each of the Lsubfilters
- $L-1$ additions to combine the outputs of all the subfilters into one output sample for each R/Dth time instant, giving a total of $(N-1)L+M-1$ additions and M multiplications. We see that there is an increase in additions of $(N-1)(L-1)$ from recoding to subfiltering. Note that the number of multiplications stays the same.

For the ideal system, we have

- $N-1$ additions in the beamformer
- M multiplications in each of the N filters
- $M-1$ additions in each of the N filters

which yields a total of $NM-1$ additions and NM multiplications. This implementation is clearly more computationally complex than both subfiltering and recoding.

VII. Conclusion

We have analyzed the noise caused by applying dynamic delays in a delta-sigma beamformer. This analysis can be used for getting a more detailed understanding of the previously suggested methods for overcoming the problem, by addressing both their virtues and limitations. It also motivated the specification of an alternative solution for noise removal. Our solution was shown to remove completely all effects caused by inserting a time-variant operation in the delta-sigma demodulator, while retaining much of the compactness that is achieved by applying delays before demodulation. An existing solution, recoding, was shown to be less computationally complex and still able to remove most of the noise.

Future research should address more specifically how implementations of the present and previously suggested

solutions differ with respect to chip area, power consumption, and other different measures of compactness. This will give a more accurate view of where further optimization is necessary. Simulations of Doppler and flow imaging should also be performed to examine whatever effect the remaining error has in these cases. It would also be interesting to implement an actual data acquisition system using different kinds of delta-sigma A/D converters to capture the signals, because all results known to date are obtained through simulations performed on PCM data. As mentioned, the previously unresolved distortion caused by the interpolation-like effect has a magnitude that depends on the out-of-band noise to be removed by the anti-aliasing filter; a component that is not present in pre-captured data sets with a Nyquist sampling rate.

References

[1] M. Karaman, P.-C. Li, and M. O'Donnell, "Synthetic aperture imaging for small scale systems," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 42, no. 3, pp. 429–442, 1995.

[2] A. Austeng and S. Holm, "Sparse 2d arrays for 3d phased array imaging—design methods," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 49, no. 8, pp. 1073–1086, 2002.

[3] K. Ranganathan, M. K. Santy, T. N. Blalock, J. A. Hossack, and W. F. Walker, "Direct sampled i/q beamforming for compact and very low-cost ultrasound imaging," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 51, no. 9, pp. 1082–1094, 2004.

[4] H. T. Feldkamper, R. Schwann, V. Gierenz, and T. G. Noll, "Low power delay calculation for digital beamforming in handheld ultrasound systems," in *Proc. IEEE Ultrason. Symp.*, 2002, vol. 2, pp. 1763–1766.

[5] S. E. Noujaim, S. L. Garverick, and M. O'Donnell, "Phased array ultrasonic beam forming using oversampled a/d converters," U.S. Patent 5461389, 1993.

[6] S. R. Freeman, M. K. Quick, M. A. Morin, R. C. Anderson, C. S. Desilets, T. E. Linnenbrink, and M. O'Donnell, "An ultrasound beamformer using oversampling," in *Proc. IEEE Ultrason. Symp.*, 1997, pp. 1687–1690.

[7] M. Kozak and M. Karaman, "Digital phased array beamforming using single-bit delta-sigma conversion with non-uniform oversampling," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 48, no. 4, pp. 922–931, 2001.

[8] S. Holm and K. Kristoffersen, "Analysis of worst-case phase quantization sidelobes in focused beamforming," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 39, no. 5, pp. 593–599, 1992.

[9] S. R. Freeman, M. K. Quick, M. A. Morin, R. C. Anderson, C. S. Desilets, T. E. Linnenbrink, and M. O'Donnell, "Delta-sigma oversampled ultrasound beamformer with dynamic delays," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46, no. 2, pp. 320–332, 1999.

[10] C. I. C. Nilsen and S. Holm, "Distortion-free delta-sigma beamforming with dynamic delays," in *Proc. IEEE Ultrason. Symp.*, 2006, pp. 2136–2139.

[11] S. R. Norsworthy, R. Schreier, and G. C. Themes, *Delta-Sigma Data Converters: Theory, Design, and Simulation*. New York: IEEE Press, 1992.

[12] R. M. Gray, "Quantization noise in delta-sigma a/d converters," in *Delta-Sigma Data Converters*. S. R. Norsworthy, R. Schreier, and G. C. Themes, Eds. New York: IEEE Press, 1992, pp. 44–74.

[13] B. G. Tomov and J. A. Jensen, "Compact FPGA-based beamformer using oversampled 1-bit a/d converters," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 52, no. 5, pp. 870–880, 2005.

[14] M. Karaman, A. Atalar, and H. Koymen, "VLSI circuits for adaptive digital beamforming in ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 711–720, 1993.

[15] H.-S. Han, H.-J. Park, and T.-K. Song, "A new architecture for ultrasound sigma-delta modulation beamformer," in *Proc. IEEE Ultrason. Symp.*, 2002, vol. 2, pp. 1631–1634.

[16] S. R. Freeman, M. K. Quick, M. A. Morin, R. C. Anderson, C. S. Desilets, T. E. Linnenbrink, and M. O'Donnell, "Heterodyning technique to improve performance og delta-sigma-based beamformers," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46, no. 4, pp. 771–790, 1999.

Carl-Inge C. Nilsen (S'06) received his M.Sc. degree in computer science from the University of Oslo, Oslo, Norway, in 2005. He is currently pursuing his Ph.D. degree. His research interests include signal and array processing for ultrasound imaging and other acoustic applications.

Sverre Holm (M'82–SM'02) was born in Oslo, Norway, in 1954. He received the M.S. and Ph.D. degrees in electrical engineering from the Norwegian Institute of Technology, Trondheim, in 1978 and 1982.

He has done research and development in speech coding and frequency estimation at SINTEF, Trondheim, Norway, and worked with synthetic aperture radar processing and spectral estimation at Informasjonskontroll AS, Asker, Norway. From 1990 to 1994 he was at GE Vingmed Ultrasound, Horten, Norway, doing research and development work in digital beamforming, ultrasound probes, and ultrasound contrast agents. He spent the fall of 1998 on a sabbatical at GE Global Research, Schenectady, NY.

He has been an assistant professor in the Electrical Engineering Department of Yarmouk University in Jordan (1984–1986) and from 1989 he was an adjunct professor, until 1992 at the Norwegian Institute of Technology, then at the University of Oslo, Oslo, Norway. Since 1995 he has been a full professor of signal processing at the University of Oslo working in ultrasound and sonar imaging, acoustic field simulation, and indoor positioning systems.

Dr. Holm was an associate editor of the *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* from 1997–2002. He was elected as a member of the Norwegian Academy of Technological Sciences in 2002.