# Multichannel $\Sigma\Delta$ ADCs With Integrated Feedback Beamformers to Cancel Interfering Communication Signals

Vijay Venkateswaran, *Member, IEEE*, and Alle-Jan van der Veen, *Fellow, IEEE*

*Abstract*—In multiuser multiantenna communication receivers, the use of high resolution ADCs is costly. In the presence of (strong) interference, more bits are used than would be necessary for quantizing only the signal of interest. Thus, if it is possible to cancel the interference in the analog domain, considerable savings can be realized. In this paper, we exploit the fact that a multiantenna receiver consists of a bank of ADCs and propose a new architecture, wherein a feedback beamformer (FBB) takes a linear combination of the ADC outputs and feeds back the result to be subtracted at the input. This ADC architecture is especially compatible with existing $\Sigma\Delta$ ADCs, that already consist of a digital-to-analog converter (DAC) in the feedback loop and enables sophisticated source separation algorithms designed in the digital baseband to cancel the interfering users in the analog domain. Subsequently, the $\Sigma\Delta$ ADCs digitize only the desired user signals and achieve considerable savings in power consumption. Using a mean squared error criterion and assuming that a training sequence is available, we present an algorithm to design the weights of the FBB. The interference suppression and power savings of the proposed approach are demonstrated via simulation results.

*Index Terms*—Linear prediction, multiantenna/MIMO systems, oversampling $\Sigma\Delta$ ADCs, predictive quantization, source separation.

## I. INTRODUCTION

### A. Interference Cancellation at the ADC

ANTENNA arrays in multiantenna receivers exploit spatial diversity to achieve reliable communications at low signal energies and in the presence of interferers [3]. However, multiple receive antennas lead to multiple RF and ADC chains, which increases the circuit area and the power consumption. Already, the power consumption per ADC operation is comparable

to that of hundred thousands of digital logic gates [4] and the improvements in ADC technology evolve at a much slower pace when compared to Moore's law [5].

Consider a multiuser cellular/WLAN scenario, where the ADC input signals contain contributions from the desired user, noise and the interfering users. Although advanced beamforming techniques mentioned in [3] can achieve interference cancellation in the digital baseband, the presence of strong interferers forces the ADCs to spend a significant part of their dynamic range and resolution on digitizing the unwanted components. If we are able to cancel most of the interference before it reaches the ADC, we can use lower resolution ADCs, which directly translates into reduced power consumption [5].

One well known suboptimal approach to reduce the number of ADC chains, thereby reducing the ADC power consumption, is to select the antennas with the largest signal energies and only quantize these [6]. However, such techniques do not fully exploit the advantages of multiantenna systems, cannot track for variations in the wireless channel and fail in the presence of interferers. Instead of antenna selection, if we can cancel the interference a priori in the analog domain, then considerable power savings can be achieved. For example, the paper [7] shows a single channel feedforward cancellation architecture that attempts to reconstruct the interfering signal and subtract it in the analog domain prior to the ADC. It is also possible to integrate analog phase shifters within the RF architecture [8], so that we obtain a set of analog beamformers to cancel interferers. In [9], we detailed one such approach. We discussed algorithms for estimating the channel and designing the beamforming coefficients to minimize the mean squared error (MSE) between the desired user and its estimate at receiver.

The present paper introduces an alternative approach. Here we consider a bank of ADCs connected to the antenna array. A space-time digital beamformer operates on the ADC outputs to identify the interferer signals. These interferer estimates are fed back via a digital to analog converter (DAC) to cancel the interfering user signals before the quantization operation. For a given ADC resolution, interference cancellation would allow a more faithful digital representation of the desired user signals; alternatively, we can quantize with fewer bits to achieve the same performance[1]. The space-time beamformer coefficients are designed using a training sequence from the desired user signal.

One subclass of ADCs, especially compatible with the above DAC feedback architecture, is the family of $\Sigma\Delta$ ADCs. These have been extensively studied in the literature [11], [12]. Typ-

---

[1]This is a generally known idea also employed e.g., in differential quantization or predictive coding techniques [10].

ically, $\Sigma\Delta$ ADCs sample at a frequency much higher than the Nyquist frequency and obtain coarsely quantized signals (say 1-bit), which are subsequently integrated to obtain a high resolution signal in the digital baseband. Further improvements in signal prediction and reconstruction, exploiting the bandlimited (BL) nature of the input, have been studied previously in [13]–[15].

The DAC feedback included in the $\Sigma\Delta$ architecture offers possibilities to suppress BL interferers and to improve dynamic range. Philips *et al.* [16] perform interference cancellation with a high pass filter in the DAC feedback loop of a second order $\Sigma\Delta$ modulator. This is a single-channel and nonadaptive solution. More generally, ADCs with feedback DAC to improve dynamic range in the context of cognitive radios have been proposed recently in [17]. These papers focus on the implementation and not on the design of beamformer weights to cancel interferers. Interference cancellation with feedback ADCs utilizing the spatial diversity offered by multiantenna receivers has not been addressed yet. However [16], [17] offer a starting point for the realization of such architectures in silicon.

### B. Setup and Objectives

We consider a narrowband (NB) multiuser setup, where the desired user (band limited by frequency $f_0$) and interfering users transmit over a shared wireless channel as in Fig. 1(a). These signals are received by an array of $N_r$ antennas, downconverted to baseband, sampled (at sampling frequency $f_s$) and coarsely quantized using an oversampled ADC. In Fig. 1(a), the digital postprocessor does beamforming to suppress the interfering signal. Alternatively, in Fig. 1(b) the multiple ADC outputs are fed back via a digital beamformer to cancel the interferers. We will refer to this architecture as a multichannel (MC) $\Sigma\Delta$ ADC with a feedback beamformer (FBB). Note that the $\Sigma\Delta$ ADC is just one relevant architecture to cancel the interferers; this could be generalized to other types of oversampled ADCs.

With the objective to cancel the interferer before the ADC operation, our aim in this paper is to design the FBB coefficients. If done successfully, this will lead to reduced power consumption. Even a partial cancellation of the interfering user energy would lower the requirements on the dynamic range of the ADC and lead to an overall power reduction in the ADC units. As will be shown using simulation results in Section VI, introducing the FBB may lead to a reduction in power consumption by a factor of four.

In the signal processing literature, several types of equalizers have been designed where the equalizer output is fed back to cancel the incoming signals [18]. One related context is the classical least mean squares (LMS) technique [19]. Feedback equalization techniques in the context of blind channel estimation have been proposed in [20], [21]. In these methods, a feedback beamformer operating on the equalizer outputs estimates and cancels the redundancy in the incoming signals. Our aim is to integrate such techniques within the ADC architecture. Somewhat similar work on designing a quantizer with a single channel FBB in the context of subband coding is given by [13]. In this work the authors specify the ADC output as an interpolation of bandlimited signals and design the quantizer (as an *oversampled frame*) to cancel the redundancy in the incoming signals. More recently, in [22] FBBs are designed in the context of a compressive sensing framework.
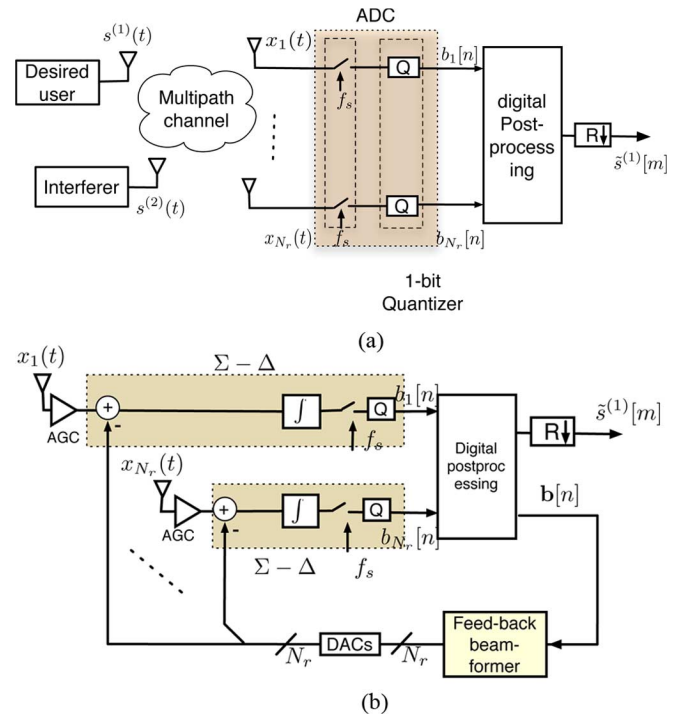


Fig. 1. (a) Antenna array configuration with desired and interfering user signals quantized by oversampled ADCs (operating at $R$ times the Nyquist rate of $2f_0$ : $f_s = 2Rf_0$) followed by baseband combining to estimate the desired user (b) Proposed multichannel (MC) ADC architecture with a feedback beamformer (FBB) to identify and cancel interfering user signals.

The above mentioned approaches aim to perform noise shaping or to cancel redundancy in the incoming signals. In contrast, our aim is to design the quantizer output to cancel the interferers and to reconstruct the signal of interest. This signal is identified via a training sequence.

### C. Outline and Contributions

We progressively address the different design issues of an FBB arrangement used with a $\Sigma\Delta$ ADC. In Section II, we specify the oversampled $\Sigma\Delta$ ADC setup and provide general background on higher-order $\Sigma\Delta$ ADCs that provide a filter in the feedback loop. In Section III, we then consider a single-channel $\Sigma\Delta$ ADC operating on the incoming band limited signals. We formulate a mean squared error (MSE) cost function on the difference between the desired user training signal and its estimate at the output of the ADC. Solving this cost function is not directly feasible, but it can be modified into a prediction error cost function that can be solved in closed form. The original cost function can then be solved iteratively using the previous solution as a starting point. In Section IV, we consider a bank of ADCs receiving contributions from the desired and the interfering users and derive an extension of the single channel FBB design to a multichannel FBB design. Again, the FBB can be designed in a closed form for a prediction error cost function and the minimum MSE solution can be obtained iteratively. All designs depend on the availability of a training signal, i.e., the output of the ADCs in the absence of interferers and noise. In Section V, we consider some aspects of this. Finally, in Section VI, we show simulations to indicate the performance of the proposed algorithms. It is seen that in
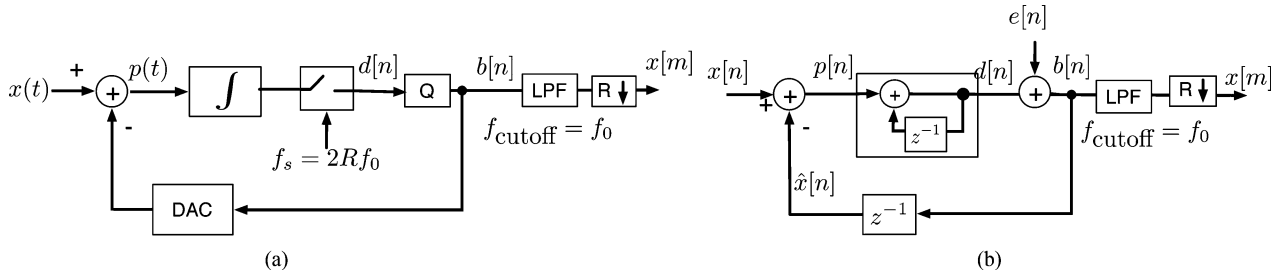
Fig. 2. (a) First-order continuous time $\Sigma\Delta$ modulator with 1-bit output. (b) Discrete time equivalent model where the sampling frequency $f_s = 2Rf_0$.

many cases an interference suppression of over 20–25 dB can be achieved, leading to comparable power savings.

*Notation:* $(\bar{\cdot})$, $(.)^T$, $(.)^H$ and $(.)^\dagger$ denote conjugation, transpose, Hermitian transpose and pseudo-inverse operations. $\otimes$ denotes the Kronecker product, $\text{vec}(\cdot)$ a vectorization operator (stacking all columns of its argument into a vector) and $\|.\|$ denotes the Frobenius norm. Vectors and matrices are represented in lower and upper case bold letters and we use an underscore to denote some multichannel vectors and matrices (for distinction with the single-channel case). $\mathbf{I}_K$ denotes the identity matrix of size $K \times K$, $\mathbf{1}_K$ and $\mathbf{0}_K$ are $K \times 1$ vectors of ones and zeros, respectively.

## II. PREDICTIVE QUANTIZATION WITH $\Sigma\Delta$ ADCS

In this section, we review the $\Sigma\Delta$ ADC architecture and in particular extensions with a higher-order feedback loop. We derive a transfer function model and formulate the design problem to be solved in subsequent sections.

### A. Signal Sampling and Reconstruction

Consider an ADC operating on a continuous time input signal $x(t)$. We will assume that $x(t)$ is band limited by the highest frequency of operation $f_0$ and time-limited to a given observation interval $t \in [0, T)$, where $T$ would correspond to the duration of a transmission packet. Outside this interval, the signal is supposed to be zero.

If $x(t)$ is uniformly sampled $N$ times in the interval, then its samples are denoted as $x[n] = x(n/NT)$, for $n = 0, \ldots, N-1$ and they are stacked in an $N \times 1$ vector $\mathbf{x} = [x[0], \ldots, x[N-1]]^T$.

Let $M$ correspond to the number of samples of $x(t)$, obtained when sampled at the Nyquist rate of $2f_0$ in $[0, T)$, i.e., $M = \lfloor 2f_0 T \rfloor$. If sampled at this rate, the input signal is uniquely represented by its samples and can be reconstructed as

$$x(t) = \sum_{m=0}^{M-1} x[m] \text{sinc}\left(\frac{t - m\left(\frac{T}{M}\right)}{\frac{T}{M}}\right). \tag{1}$$

This reconstruction can be implemented via a lowpass filter. If the input signal has variance $\sigma_x^2$ and the ADC quantizes the signal at $res$ bits, the quantization noise $e[n]$ is modeled to be uniformly distributed and independent of $x[n]$ with a variance $\sigma_e^2 = c2^{-2res}\sigma_x^2$, where $c$ is a constant [23][2].

[2]We assume that the ADC unit includes an automatic gain control (AGC) which scales the input signal to match the range of the ADC without overload. The design and the analysis of the AGC is left out of the considerations in this paper.

In oversampling ADCs, the sampling frequency is much greater than the Nyquist rate. Let $N$ be an integral multiple of $M$ and define the oversampling ratio (OSR)

$$R = \frac{N}{M}.$$

Typically, we could have $R = 32, 64, 128, \ldots$. If sampled at this rate, the input signal is redundantly specified by its samples. Indeed, given the $M \times 1$ Nyquist sample vector $\mathbf{x}_M$, the oversampled sample vector $\mathbf{x}_N$ of size $N \times 1$ is given by inserting $t = nT/N$ in (1), leading to the interpolation formula (ignoring quantization noise)

$$\mathbf{x}_N = \mathcal{I}\mathbf{x}_M$$

where $\mathcal{I}$ is a tall $N \times M$ matrix with entries

$$\mathcal{I}_{nm} = \text{sinc}\left(\frac{n\left(\frac{T}{N}\right) - m\left(\frac{T}{M}\right)}{\frac{T}{M}}\right). \tag{2}$$

Conversely, given $\mathbf{x}_N$, the Nyquist sample vector $\mathbf{x}_M$ can be obtained by premultiplying $\mathbf{x}_N$ by any left inverse of $\mathcal{I}$, leading to an infinite number of possible reconstructions that differ in the way that the quantization noise is processed. One straightforward reconstruction of the Nyquist samples can be implemented via a lowpass filter followed by a downsampler.

In this case, the quantization noise on the samples in $\mathbf{x}_N$ has variance $\sigma_{e,N}^2 = c2^{-2res}\sigma_x^2$, whereas after reconstruction the samples of $\mathbf{x}_M$ have quantization noise variance $\sigma_{e,M}^2 = 1/Rc2^{-2res}\sigma_x^2$. This is a factor $R$ lower.

However, it is possible to do much better. Since the input signal is bandlimited, it is possible to predict the current input sample $x[n]$ from previous samples $x[n-1], x[n-2], \ldots$. By subtracting this prediction from the input signal and quantizing only the prediction error $p[n]$, the required dynamic range at the input of the quantizer is much smaller and so will be the quantization error. This is the principle of differential quantization, or predictive ADCs [13], [23]. A special case is the $\Sigma\Delta$ ADC, reviewed in the next subsection.

### B. Oversampled $\Sigma\Delta$ ADCs

Consider a first-order $\Sigma\Delta$ ADC operating at a sampling frequency $f_s = 2Rf_0$ as in Fig. 2(a). Its discrete-time equivalent is shown in Fig. 2(b). We assume an idealistic model of the ADC and the DAC, i.e.,

- ADC operation is instantaneous (no sample and hold delays);
- DAC linear with infinite precision (although in simulations we ultimately use 1-bit precision).

In this setup, the prediction error $p[n]$ is integrated and the quantizer $Q\{\cdot\}$ digitizes the integrator output $d[n]$ to obtain $b[n] = Q\{d[n]\}$. The quantizer output $b[n]$ is fed back with a one-sample delay and subtracted from the input through a DAC to obtain a prediction of $x[n]$. As before, Nyquist-rate samples are obtained after low-pass filtering and decimation by a factor $R$. For modeling purposes, we replace the quantizer operator $Q\{\cdot\}$ by an i.i.d. additive noise source $e[n]$ as in Fig. 2(b), such that $b[n] = d[n] + e[n]$. For details, see [12].

As indicated in the figure, $\hat{x}[n] = b[n-1]$ is the prediction of $x[n]$ and $p[n] = x[n] - \hat{x}[n]$ is the prediction error. The discrete-time integrator operating on $p[n]$ has output $d[n]$, satisfying the difference equation

$$
\begin{aligned}
d[n] &= p[n] + d[n-1] \\
&= (x[n] - b[n-1]) + d[n-1]. \quad (3)
\end{aligned}
$$

We assume, for simplicity throughout the paper, that the initial states are $d[-1] = 0$ and $b[-1] = 0$. The ADC output $b[n]$ can be rewritten from (3), utilizing $b[n] = d[n] + e[n]$, as

$$
\sum_{j=0}^{n} b[j] = \sum_{j=0}^{n} x[j] + e[n]. \quad (4)
$$

Stack the $\Sigma\Delta$ modulator output $b[n]$ for $n = 0, \ldots, N-1$ as a $N \times 1$ vector $\mathbf{b} = [b[0], \ldots, b[N-1]]^T$ and likewise for the input vector $\mathbf{x} = [x[0], \ldots, x[N-1]]^T$ and the quantization noise vector $\mathbf{e} = [e[0], \ldots, e[N-1]]^T$. From (4), the relation between $\mathbf{b}$ and $\mathbf{x}$ can be written as

$$
\mathbf{Lb} = \mathbf{Lx} + \mathbf{e} \quad (5)
$$

where $\mathbf{L}$ is an $N \times N$ lower triangular matrix whose nonzero elements are equal to 1; it corresponds to the integration operation.

It is thus seen that

$$
\mathbf{b} = \mathbf{x} + \mathbf{L}^{-1}\mathbf{e},
$$

where $\mathbf{L}^{-1}$ is a lower bi-diagonal matrix with $+1$ and $-1$ on the main diagonal and off diagonal positions, respectively. One can interpret $\mathbf{L}^{-1}$ as a high pass filter, suppressing the low frequency quantization noise terms. For this reason, $\Sigma\Delta$ ADCs are commonly referred to as noise shaping ADCs. The remaining high frequency quantization noise terms will be canceled by a LPF with cut-off frequency $f_0$ prior to the decimation of the output $\mathbf{b}$. It is well known that this makes the quantization noise power drop off with a factor $1/R^3$, rather than $1/R$ as we had with a straightforward oversampling ADC [24]. Since the power consumption of an ADC can be approximated as $P_{adc} \propto f_s 2^{res}$, we see that it is advantageous to maximize $R$ and use a 1-bit converter.

### C. Generalized Higher Order $\Sigma\Delta$ ADC

The discrete time equivalent model specified by Fig. 2 can be generalized to a higher order $\Sigma\Delta$ ADC [12, Ch.6]. Consider Fig. 3, explaining a $K$th order $\Sigma\Delta$ ADC. In this case, the predictor of $x[n]$ is formed by the output of a $K$th order FIR filter with coefficients $w_k$, $k = 0, \ldots, K-1$, which we will stack in a vector $\mathbf{w} = [w_0, \ldots, w_{K-1}]^T$. The prediction of $x[n]$ is $\hat{x}[n] = \mathbf{w}^H \mathbf{b}_K[n-1]$, where $\mathbf{b}_K[n-1] = [b[n-1], \ldots, b[n-K]]^T$.

Similar to the first order case, the prediction error $p[n] = x[n] - \mathbf{w}^H \mathbf{b}_K[n-1]$ is integrated and subsequently quantized,
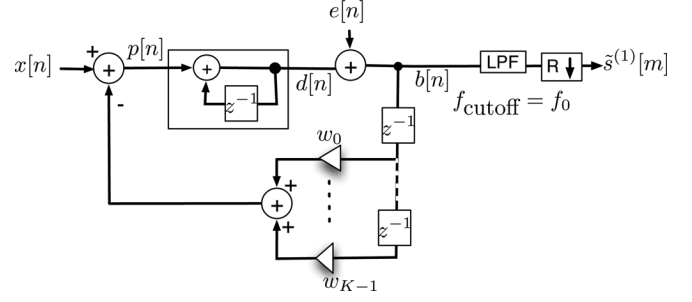


Fig. 3. Discrete time equivalent model of a $K$th order $\Sigma\Delta$ ADC with weighted feedback, represented by $\mathbf{w} = [w_0, \ldots, w_{K-1}]^T$.

represented by the additive quantization noise $e[n]$. The output thus becomes

$$
b[n] = \sum_{j=0}^{n} p[j] + e[n].
$$

Similar to (4), the ADC output can be written as

$$
b[n] + \sum_{j=0}^{n} \mathbf{w}^H \mathbf{b}_K[j-1] = \sum_{j=0}^{n} x[j] + e[n]. \quad (6)
$$

Stack the left-hand side (LHS) and right-hand side (RHS) terms of (6) for $n = 0, \ldots, N-1$ and assume that $b[j] = 0$ for $j < 0$. It follows that the ADC output vector $\mathbf{b}$ satisfies

$$
\mathbf{Mb} = \mathbf{Lx} + \mathbf{e} \quad (7)
$$

where $\mathbf{M}$ is an $N \times N$ matrix whose elements are made of $\mathbf{w}$ (as specified in detail in the next section) and $\mathbf{L}$ is an $N \times N$ matrix as defined in (5). As an example, a first-order ($K = 1$) $\Sigma\Delta$ ADC with $\mathbf{w} = w_0$ leads to $\mathbf{M}$ described by

$$
\mathbf{M} = \begin{bmatrix} 1 & & & \\ \bar{w}_0 & 1 & & \\ \vdots & \ddots & \ddots & \\ \bar{w}_0 & \cdots & \bar{w}_0 & 1 \end{bmatrix}.
$$

For the special case where $w_0 = 1$, $\mathbf{M}$ is equal to $\mathbf{L}$ and (7) reduces to (5).

Following (7), we see that $\mathbf{b}$ is given by

$$
\mathbf{b} = (\mathbf{M}^{-1}\mathbf{L})\mathbf{x} + \mathbf{M}^{-1}\mathbf{e}.
$$

The main question with this generalized architecture is how to design $\mathbf{w}$. E.g., we can select the objective to minimize the low-frequency components of the quantization noise terms $\mathbf{M}^{-1}\mathbf{e}$, i.e., noise shaping. Alternatively, we can design $\mathbf{w}$ such that $b[n]$ is an estimate of a target signal $b_0[n]$.

### D. Problem Formulation

The recurring theme in this paper is to design the feedback weights specified by $\mathbf{w}$, focusing on the cancellation of unwanted interfering signals. This reduces the required dynamic range at the quantizer, leading to lower requirements on the oversampling ratio or ADC resolution and consequently minimizes the power consumption.

The measured signal $\mathbf{x}$ contains contributions from the desired user (including channel impairments), interferers and

noise terms. The measured signal vector $\mathbf{x}$ can be expressed as a sum of $N \times 1$ vectors $\mathbf{x}^{(j)}$, $j \in \{1, \ldots, N_t\}$ in noise

$$\mathbf{x} = \mathbf{x}^{(1)} + \cdots + \mathbf{x}^{(N_t)} + \mathbf{v}$$

where $\mathbf{v}$ is a $N \times 1$ vector denoting the thermal noise terms. $\mathbf{x}^{(1)}$ corresponds to the contributions of the desired user and $\mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(N_t)}$ correspond to that of the interfering users.

For the design of $\mathbf{w}$, it is clear that the receiver needs a way to distinguish interfering signals from the desired signal. To simplify the design procedure, we make the following assumptions throughout the paper:

- There is a training phase, namely one entire transmission packet during which the receiver has knowledge of a training signal corresponding to the desired user and specified by a $N \times 1$ vector $\mathbf{b}_0 = [b_0[0], \ldots, b_0[N-1]]^T$.
- The desired user signal is narrowband, limited by the frequency of operation $f_0$, whereas the interferers can be wide-band signals with operating frequencies up to $Rf_0$.
- To enable interference cancellation, the samples of $\mathbf{x}^{(1)}$ are assumed to be uncorrelated with the samples of $\mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(N_t)}$.

This enables us to design $\mathbf{w}$ that minimizes the mean squared error (MSE)[3] between the desired and observed signal at the output of the quantizer, i.e., such that the ADC quantizes mostly the desired signal. After this, the FBB $\mathbf{w}$ is kept fixed and is used for subsequent transmission packets assuming stationarity of the interference.

We approach the design problem in the following order:

- We initially consider a single channel ADC. Can we design $\mathbf{w}$ to minimize the MSE?
- Subsequently, we extend the single channel setup to a multichannel scenario. Each channel has an ADC and can use the feedback from the other ADCs as well. How is the feedback beamformer $\mathbf{W}$ designed?
- Finally, we consider the application of this multichannel ADC in a multiuser communication scenario and discuss how a suitable training sequence $\mathbf{b}_0$ is obtained.

## III. SINGLE-CHANNEL FEEDBACK BEAMFORMER DESIGN

We consider the single channel feedback setup in Fig. 3. Our aim will be to derive an estimation algorithm for the feedback weights $\mathbf{w}$. To start, we will first rederive the data model (7) in more details.

### A. Data Model

Let $x(z)$ and $b(z)$ denote the $z$-transformations of the sequence $\mathbf{x}$ and $\mathbf{b}$ (denoted by $N \times 1$ vectors) respectively, i.e., $x(z) = \sum_{n=0}^{N-1} x[n] z^{-n}$ and $b(z) = \sum_{n=0}^{N-1} b[n] z^{-n}$. Similarly, let $w(z)$ represent the $z$-transform of $\mathbf{w}$: $w(z) = \sum_{n=0}^{K-1} w_n z^{-n}$. The transfer function in Fig. 3 of the setup is represented in the $z$-domain as

$$\left(1 + \frac{z^{-1}\bar{w}(z)}{1 - z^{-1}}\right) b(z) = \frac{1}{1 - z^{-1}} x(z) + e(z)$$
$$\Leftrightarrow \quad B(z)b(z) = L(z)x(z) + e(z) \tag{8}$$

[3]Similarly, we could design $\mathbf{w}$ to maximize SINR or to minimize BER, but minimizing MSE seems more straightforward.

where

$$L(z) = \frac{1}{1 - z^{-1}}$$

is the transfer function of the integrator and

$$B(z) = 1 + \frac{z^{-1}\bar{w}(z)}{1 - z^{-1}}$$

is the transfer function of the feedback arrangement. Define (with some abuse of notation, since as written here the matrix $\mathbf{Z}$ is not invertible)

$$\mathbf{Z}^{-1} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}.$$

We can then make the correspondence

$$L(z)x(z) \leftrightarrow \mathbf{Lx} = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ \cdots \\ x[N-1] \end{bmatrix}$$

where $\mathbf{L} = (\mathbf{I} - \mathbf{Z}^{-1})^{-1} = \mathbf{I} + \mathbf{Z}^{-1} + \cdots + \mathbf{Z}^{-(N-1)}$. Similarly, $\bar{w}(z)b(z) \leftrightarrow$

$$\bar{\mathcal{W}}\mathbf{b} = \begin{bmatrix} \bar{w}_0 & & & \mathbf{0} \\ \vdots & \bar{w}_0 & & \\ \bar{w}_{K-1} & \ddots & \ddots & \\ & \ddots & \ddots & \ddots \\ \mathbf{0} & & \bar{w}_{K-1} & \cdots & \bar{w}_0 \end{bmatrix} \begin{bmatrix} b[0] \\ b[1] \\ \cdots \\ b[N-1] \end{bmatrix}$$

where $\mathcal{W} = w_0\mathbf{I} + w_1\mathbf{Z}^{-1} + \cdots + w_{K-1}\mathbf{Z}^{-(K-1)}$. The transfer function (8) can thus be written in matrix form as

$$(\mathbf{I} + \mathbf{Z}^{-1}\mathbf{L}\bar{\mathcal{W}})\mathbf{b} = \mathbf{Lx} + \mathbf{e} \tag{9}$$

which corresponds to our earlier model (7), setting $\mathbf{M} = \mathbf{I} + \mathbf{Z}^{-1}\mathbf{L}\bar{\mathcal{W}}$.

Since $\mathbf{Mb}$ is a linear function of the entries of $\mathbf{w}$, we can write this alternatively as a matrix multiplied by $\mathbf{w}$. Specifically

$$\bar{\mathcal{W}}\mathbf{b} = \begin{bmatrix} b[0] & & & \mathbf{0} \\ b[1] & b[0] & & \\ \vdots & & \ddots & \\ b[N-1] & \cdots & b[1] & b[0] \end{bmatrix} \begin{bmatrix} \bar{w}_0 \\ \vdots \\ \bar{w}_{K-1} \\ 0 \\ \vdots \end{bmatrix}$$

$$= \begin{bmatrix} b[0] & & \mathbf{0} \\ b[1] & \ddots & \\ \vdots & \ddots & b[0] \\ \vdots & \vdots & \vdots \\ b[N-1] & \cdots & b[N-K] \end{bmatrix} \begin{bmatrix} \bar{w}_0 \\ \vdots \\ \bar{w}_{K-1} \end{bmatrix} =: \mathcal{B}\bar{\mathbf{w}}.$$

It follows that

$$\mathbf{Mb} = \mathbf{b} + \mathbf{Z}^{-1}\mathbf{L}\bar{\mathcal{W}}\mathbf{b} = \mathbf{b} + \mathbf{Z}^{-1}\mathbf{L}\mathcal{B}\bar{\mathbf{w}}.$$

Starting from (7), an alternative model formulation is thus

$$(\mathbf{Z}^{-1}\mathbf{L}\mathcal{B})\bar{\mathbf{w}} = (\mathbf{Lx} - \mathbf{b}) + \mathbf{e}. \tag{10}$$

*B. Estimation of* $\mathbf{w}$

To estimate $\mathbf{w}$, we assume that a desired output (training) sequence $\mathbf{b}_0$ is given, along with measured input data $\mathbf{x}$. Ideally, we would aim to minimize the expected mean squared error at the output, i.e., minimize

$$J(\mathbf{w}) = E\|\mathbf{b} - \mathbf{b}_0\|^2 \tag{11}$$

where $\mathbf{b}$ is given by the model (9), written as

$$\mathbf{b} = \mathbf{M}^{-1}\mathbf{L}\mathbf{x} + \mathbf{M}^{-1}\mathbf{e},$$

and the expectation is with respect to the quantization noise $\mathbf{e}$ ($\mathbf{x}$ is considered deterministic here). A complication with this formulation is that $\mathbf{w}$ enters the model via the denominator, $\mathbf{M}^{-1}$, i.e., in a nonlinear way. The same complication occurs in autoregressive moving-average (ARMA) model identification [25], where the direct Least Squares problem is considered not attractive. More practical is to formulate, as in the Prony method [25], a prediction error problem, i.e., to premultiply (10) by $\mathbf{M}$ and move the denominator to the other side and minimize

$$\begin{aligned} J'(\mathbf{w}) &= E\|\mathbf{M}\mathbf{b} - \mathbf{M}\mathbf{b}_0\|^2 \\ &= E\|\mathbf{L}\mathbf{x} + \mathbf{e} - \mathbf{M}\mathbf{b}_0\|^2. \end{aligned} \tag{12}$$

Write $\mathbf{M}\mathbf{b}_0$ as a function of $\mathbf{w}$, i.e., $\mathbf{M}\mathbf{b}_0 = \mathbf{b}_0 + (\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0)\bar{\mathbf{w}}$, where $\mathcal{B}_0$ is obtained using the delayed samples of $\mathbf{b}_0$. We obtain

$$J'(\mathbf{w}) = E\|(\mathbf{L}\mathbf{x} - \mathbf{b}_0) - (\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0)\bar{\mathbf{w}} + \mathbf{e}\|^2.$$

This is a Least Squares cost function. Minimization of $J'(\mathbf{w})$ to $\mathbf{w}$ leads to the closed-form solution

$$\bar{\mathbf{w}} = (\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0)^{\dagger}(\mathbf{L}\mathbf{x} - \mathbf{b}_0). \tag{13}$$

Note that the quantization noise $\mathbf{e}$ is a zero mean random process independent of the signal terms and the expectation is computed with respect to $\mathbf{e}$.

A unique solution exists if $\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0$ has a left inverse, i.e., if this matrix has full column rank. Since $\mathbf{Z}^{-1}$ and $\mathbf{L}$ commute, we can write $\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0 = \mathbf{L}\mathbf{Z}^{-1}\mathcal{B}_0$. The matrix $\mathbf{L}$ is square and invertible and does not change the column rank of $\mathbf{Z}^{-1}\mathcal{B}_0$. The matrix $\mathbf{Z}^{-1}$ introduces a zero row on top of $\mathcal{B}_0$ and drops the last row. We thus require that the $(N-1) \times K$ Toeplitz matrix

$$\begin{bmatrix} b_0[0] & & \mathbf{0} \\ b_0[1] & \ddots & \\ \vdots & \ddots & b_0[0] \\ \vdots & \vdots & \vdots \\ b_0[N-2] & \cdots & b_0[N-K-1] \end{bmatrix}$$

is full column rank. This requires at least that $N > K$. Under this condition, it is sufficient that, e.g., $b_0[0] \neq 0$. More generally, it is sufficient that one of $b_0[0], b_0[1], \ldots, b_0[N-K-1]$ is not zero.

*C. Iterative Refinement*

As in the Prony method, we can go back to solve the original cost function (13) in an iterative fashion. This relates to the it-
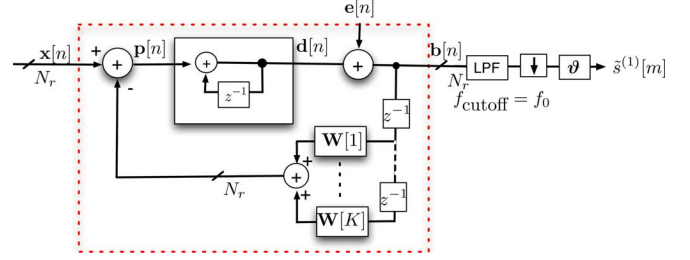


Fig. 4. Interference cancellation with a FBB $\mathbf{W}$ operating on a first-order multichannel $\Sigma\Delta$ ADC.

erative prefiltering technique by Steiglitz and McBride [25] and is also known as Iterative Weighted Least Squares. The convergence of such techniques is shown in [26].

Thus, given the solution $\mathbf{w}$ in (11), form $\mathcal{W}$ and $\mathbf{M} = \mathbf{I} + \mathbf{Z}^{-1}\mathbf{L}\mathcal{W}$. Due to its structure (lower triangular with main diagonal equal to $\mathbf{I}$), the matrix $\mathbf{M}$ is always invertible. Define $\mathbf{\Lambda} = \mathbf{M}^{-1}$. Then

$$\begin{aligned} J(\mathbf{w}) &= E\|\mathbf{b} - \mathbf{b}_0\|^2 \\ &= E\|\mathbf{M}^{-1}(\mathbf{M}\mathbf{b} - \mathbf{M}\mathbf{b}_0)\|^2 \\ &= E\|\mathbf{\Lambda}(\mathbf{L}\mathbf{x} + \mathbf{e} - \mathbf{M}\mathbf{b}_0)\|^2. \end{aligned}$$

We now freeze $\mathbf{\Lambda}$ and solve the resulting Weighted Least Squares problem for $\mathbf{w}$ present in $\mathbf{M}$

$$\bar{\mathbf{w}} = (\mathbf{\Lambda}\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0)^{\dagger}\mathbf{\Lambda}(\mathbf{L}\mathbf{x} - \mathbf{b}_0).$$

If necessary, this process is repeated until convergence (From simulations, it appears that one iteration is sufficient). In this manner, we obtain a good approximation to the solution of the original cost function (11).

IV. MULTICHANNEL FEEDBACK BEAMFORMER DESIGN

*A. Data Model*

We now consider an extension of the FBB design to a multichannel (MC) setup, where a bank of $N_r$ $\Sigma\Delta$ ADCs is used to quantize an $N_r \times 1$ vector $\mathbf{x}[n] = [x_1[n], \ldots, x_{N_r}[n]]^T$, as shown in Fig. 4. The $K$th order feedback filter $\mathbf{w}$ is replaced by a feedback beamformer (FBB) or space-time filter $\mathbf{W}$ of size $KN_r \times N_r$.

As before, $N_r$ low-resolution quantizers digitize the integrator output $\mathbf{d}[n]$ as $\mathbf{b}[n] = Q\{\mathbf{d}[n]\} = \mathbf{d}[n] + \mathbf{e}[n]$, where $\mathbf{e}[n]$ is the quantization noise and these are now $N_r \times 1$ vectors. The ADC outputs are fed back via the FBB matrix $\mathbf{W}$. The prediction of $\mathbf{x}[n]$ is a $N_r \times 1$ vector $\hat{\mathbf{x}}[n] = \mathbf{W}^H\underline{\mathbf{b}}_K[n-1]$, where $\underline{\mathbf{b}}_K[n-1] = [\mathbf{b}^T[n-1], \ldots, \mathbf{b}^T[n-K]]^T$ is a $KN_r \times 1$ vector and $\mathbf{W}$ is a $KN_r \times N_r$ filtering matrix $\mathbf{W} = [\mathbf{W}_0^T, \ldots, \mathbf{W}_{K-1}^T]^T$ with $\mathbf{W}_j$ the $N_r \times N_r$ beamforming matrix for the $j$th lag, $j = 0, \ldots, K-1$. Similar to Section III, the $N_r \times 1$ prediction error vector is

$$\mathbf{p}[n] = \mathbf{x}[n] - \mathbf{W}^H\underline{\mathbf{b}}_K[n-1] \tag{14}$$

and the ADC output satisfies

$$\mathbf{b}[n] + \sum_{j=0}^{n} \mathbf{W}^H\underline{\mathbf{b}}_K[j-1] = \sum_{j=0}^{n} \mathbf{x}[j] + \mathbf{e}[n]. \tag{15}$$

Let $\underline{\mathbf{b}}$ be a $NN_r \times 1$ vector stacking the $\mathbf{b}[n]$ as $\underline{\mathbf{b}} = [\mathbf{b}^T[0], \dots, \mathbf{b}^T[N-1]]^T$ and likewise for $\underline{\mathbf{x}}$ and $\underline{\mathbf{e}}$. As before, the measured signal $\underline{\mathbf{x}}$ contains contributions from the desired user (including channel imperfections), the interferer and the noise terms. Stacking the LHS and RHS terms of the ADC output in (15) for $n = 0, \dots, N-1$ and assuming that $\mathbf{b}[j] = 0$ for $j < 0$, leads to a MC relation

$$\underline{\mathbf{M}}\,\underline{\mathbf{b}} = \underline{\mathbf{L}}\,\underline{\mathbf{x}} + \underline{\mathbf{e}} \qquad (16)$$

where $\underline{\mathbf{L}} = \mathbf{L} \otimes \mathbf{I}_{N_r}$ is a $NN_r \times NN_r$ lower triangular matrix and $\underline{\mathbf{M}}$ is a $NN_r \times NN_r$ lower triangular matrix constructed from the elements of $\mathbf{W}$ (see below). $\underline{\mathbf{L}}$ and $\underline{\mathbf{M}}$ denote the MC equivalents of $\mathbf{L}$ and $\mathbf{M}$, respectively. For example, when $K = 1$,

$$\underline{\mathbf{M}} = \begin{bmatrix} \mathbf{I}_{N_r} & & & \\ \mathbf{W}^H & \mathbf{I}_{N_r} & & \\ \mathbf{W}^H & \ddots & \mathbf{I}_{N_r} & \\ \mathbf{W}^H & \cdots & \mathbf{W}^H & \mathbf{I}_{N_r} \end{bmatrix}.$$

More specifically, we can generalize (9) to the MC context. Indeed, using multichannel $z$-transforms, we can write the equation as follows:

$$\begin{aligned} \mathbf{d}(z) &= (\mathbf{I} - z^{-1}\mathbf{I})^{-1}\mathbf{p}(z) \\ \mathbf{b}(z) &= (\mathbf{I} - z^{-1}\mathbf{I})^{-1}\mathbf{x}(z) \\ &\quad - (\mathbf{I} - z^{-1}\mathbf{I})^{-1}\mathbf{W}^H(z)z^{-1}\mathbf{b}(z) + \mathbf{e}(z) \\ \Leftrightarrow \quad & [\mathbf{I} + z^{-1}(\mathbf{I} - z^{-1}\mathbf{I})^{-1}\mathbf{W}^H(z)]\mathbf{b}(z) \\ &= (\mathbf{I} - z^{-1}\mathbf{I})^{-1}\mathbf{x}(z) + \mathbf{e}(z). \end{aligned}$$

Let $\underline{\mathbf{Z}}^{-1} = \mathbf{Z}^{-1} \otimes \mathbf{I}$, then we can rewrite this as

$$\underline{\mathbf{M}}\,\underline{\mathbf{b}} = \underline{\mathbf{L}}\,\underline{\mathbf{x}} + \underline{\mathbf{e}} \qquad (17)$$

where

$$\underline{\mathbf{M}} = \mathbf{I} + \underline{\mathbf{Z}}^{-1}\underline{\mathbf{L}}\,\mathcal{W}^H,$$

and

$$\mathcal{W}^H = \begin{bmatrix} \mathbf{W}_0^H & & & & \mathbf{0} \\ \vdots & \mathbf{W}_0^H & & & \\ \mathbf{W}_{K-1}^H & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & \mathbf{W}_{K-1}^H & \cdots & \mathbf{W}_0^H \end{bmatrix}.$$

To generalize (10) to the MC context, we first write $\mathcal{W}^H\underline{\mathbf{b}}$ in terms of entries of $\mathbf{W}$, as follows:

$$\begin{aligned} \mathcal{W}^H\underline{\mathbf{b}} &= \begin{bmatrix} \mathbf{W}_0^H\mathbf{b}[0] \\ \mathbf{W}_0^H\mathbf{b}[1] + \mathbf{W}_1^H\mathbf{b}[0] \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{b}[0]^T \otimes \mathbf{I})\mathrm{vec}(\mathbf{W}_0^H) \\ (\mathbf{b}[1]^T \otimes \mathbf{I})\mathrm{vec}(\mathbf{W}_0^H) + (\mathbf{b}[0]^T \otimes \mathbf{I})\mathrm{vec}(\mathbf{W}_1^H) \\ \vdots \end{bmatrix} \\ &= (\underline{\mathcal{B}} \otimes \mathbf{I})\underline{\mathbf{w}} \end{aligned}$$

where

$$\underline{\mathcal{B}} = \begin{bmatrix} \mathbf{b}[0]^T & & & \mathbf{0} \\ \mathbf{b}[1]^T & \ddots & & \\ \vdots & & \ddots & \mathbf{b}[0]^T \\ \vdots & & \vdots & \vdots \\ \mathbf{b}[N-1]^T & \cdots & \mathbf{b}[N-K]^T \end{bmatrix},$$

$$\underline{\mathbf{w}} = \mathrm{vec}(\mathbf{W}^T) = \begin{bmatrix} \mathrm{vec}(\mathbf{W}_0^T) \\ \mathrm{vec}(\mathbf{W}_1^T) \\ \vdots \\ \mathrm{vec}(\mathbf{W}_{K-1}^T) \end{bmatrix}.$$

($\underline{\mathcal{B}}$ has size $N \times KN_r$, whereas $\underline{\mathbf{w}}$ has size $KN_r^2 \times 1$.) Furthermore

$$\begin{aligned} \underline{\mathbf{Z}}^{-1}\underline{\mathbf{L}}\,\mathcal{W}^H\underline{\mathbf{b}} &= (\mathbf{Z}^{-1} \otimes \mathbf{I})(\mathbf{L} \otimes \mathbf{I})(\underline{\mathcal{B}} \otimes \mathbf{I})\underline{\mathbf{w}} \\ &= (\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}} \otimes \mathbf{I})\underline{\mathbf{w}} \end{aligned}$$

so that

$$\underline{\mathbf{M}}\,\underline{\mathbf{b}} = \underline{\mathbf{b}} + (\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}} \otimes \mathbf{I})\underline{\mathbf{w}}. \qquad (18)$$

We can thus write the MC data model (17) as

$$\underline{\mathbf{M}}\,\underline{\mathbf{b}} = \underline{\mathbf{L}}\,\underline{\mathbf{x}} + \underline{\mathbf{e}} \Leftrightarrow (\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}} \otimes \mathbf{I})\underline{\mathbf{w}} = (\underline{\mathbf{L}}\,\underline{\mathbf{x}} - \underline{\mathbf{b}}) + \underline{\mathbf{e}}. \quad (19)$$

Due to the Kronecker structure, we can also rearrange this in a more compact form as

$$(\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}})\bar{\mathbf{W}} = (\mathbf{LX} - \mathbf{B}) + \mathbf{E} \qquad (20)$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}[0]^T \\ \mathbf{x}[1]^T \\ \vdots \\ \vdots \\ \mathbf{x}[N-1]^T \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{b}[0]^T \\ \mathbf{b}[1]^T \\ \vdots \\ \vdots \\ \mathbf{b}[N-1]^T \end{bmatrix}$$

and with a similar definition for $\mathbf{E}$.

### B. Estimation of $\mathbf{W}$

Similar to the single-channel case, we aim to design the FBB $\mathbf{W}$ such that $\mathbf{b}[n]$ approximates a known training sequence $\mathbf{b}_0[n]$, i.e., we aim to minimize

$$J(\mathbf{W}) = E\|\underline{\mathbf{b}} - \underline{\mathbf{b}}_0\|^2 \qquad (21)$$

where $\underline{\mathbf{b}}$ is given by the model (17), i.e.

$$\underline{\mathbf{b}} = \underline{\mathbf{M}}^{-1}\underline{\mathbf{L}}\,\underline{\mathbf{x}} + \underline{\mathbf{M}}^{-1}\underline{\mathbf{e}}.$$

Similar to before, this minimization is complicated and we replace it by the minimization of a modified cost function

$$\begin{aligned} J'(\mathbf{W}) &= E\|\underline{\mathbf{M}}\,\underline{\mathbf{b}} - \underline{\mathbf{M}}\,\underline{\mathbf{b}}_0\|^2 \\ &= E\|\underline{\mathbf{L}}\,\underline{\mathbf{x}} + \underline{\mathbf{e}} - \underline{\mathbf{M}}\,\underline{\mathbf{b}}_0\|^2 \\ &= E\|(\mathbf{LX} - \mathbf{B}_0) - (\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}}_0)\bar{\mathbf{W}} + \mathbf{E}\|_F^2. \quad (22) \end{aligned}$$

The closed form solution is given by

$$\bar{\mathbf{W}} = (\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}}_0)^\dagger(\mathbf{LX} - \mathbf{B}_0). \qquad (23)$$

Using this as a starting point, we can solve (21) iteratively by setting $\underline{\mathbf{M}} = \mathbf{I} + \underline{\mathbf{Z}}^{-1}\mathbf{L}\,\overline{\underline{\mathcal{W}}}$, defining $\underline{\boldsymbol{\Lambda}} = \underline{\mathbf{M}}^{-1}$ and writing

$$J(\mathbf{W}) = E\|\underline{\mathbf{b}} - \underline{\mathbf{b}}_0\|^2$$
$$= E\|\underline{\boldsymbol{\Lambda}}(\mathbf{L}\,\underline{\mathbf{x}} + \underline{\mathbf{e}} - \underline{\mathbf{M}}\,\underline{\mathbf{b}}_0)\|^2. \qquad (24)$$

For a fixed $\underline{\boldsymbol{\Lambda}}$, we can rewrite $\underline{\mathbf{M}}\,\underline{\mathbf{b}}_0$ using (18) and solve for $\mathbf{W}$ present in $\underline{\mathbf{w}}$

$$\bar{\underline{\mathbf{w}}} = (\underline{\boldsymbol{\Lambda}}[\underline{\mathbf{Z}}^{-1}\mathbf{L}\underline{\mathcal{B}}_0 \otimes \mathbf{I}])^{\dagger}\underline{\boldsymbol{\Lambda}}(\mathbf{L}\,\underline{\mathbf{x}} - \underline{\mathbf{b}}_0). \qquad (25)$$

If necessary, this can be repeated iteratively a number of times.

A few remarks are in order. First of all, (23) is seen to treat each channel "independently": each column of $\mathbf{W}$ depends only on a corresponding column of the input data $\mathbf{X}$ and training data $\mathbf{B}_0$. This is different for the solution of the weighted problem (25), because the weighting $\underline{\boldsymbol{\Lambda}}$ does not have a Kronecker structure (i.e., it cannot be factored into a Kronecker product of two matrices). The weighting relates the channels to each other.

Secondly, we can verify the conditions for the existence of a unique left inverse. For the first solution (23), we require $\underline{\mathcal{B}}_0$ at least to be tall, i.e., $N > KN_r$. However, this is a necessary but not sufficient condition. If each channel has the same training sequence, then $\mathbf{b}_0[n] = \mathbf{1}b_0[n]$ and it is seen that each column $\underline{\mathcal{B}}_0$ is $N_r$ times repeated: we can write $\underline{\mathcal{B}}_0 = \mathcal{B}_0 \otimes \mathbf{1}^T$. Thus, for such training sequences, the first solution is not unique. However, the Kronecker structure may also be an advantage, as it will lead to simpler calculations: we can take as the (nonunique) left inverse

$$(\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}}_0)^{\dagger} = [(\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0) \otimes \mathbf{1}^T]^{\dagger}$$
$$= (\mathbf{Z}^{-1}\mathbf{L}\mathcal{B}_0)^{\dagger} \otimes \frac{1}{N}.$$

For the second solution (25), the matrix to be left-inverted has size $NN_r \times KN_r^2$ and it is tall if $N > KN_r$. If $\underline{\mathcal{B}}_0$ has repeated columns, then the same will hold for $\underline{\boldsymbol{\Lambda}}[\mathbf{Z}^{-1}\mathbf{L}\underline{\mathcal{B}}_0 \otimes \mathbf{I}]$. Thus, in general also this solution will not be unique and again, the structure can be exploited to facilitate the computations.

## V. CONSTRUCTION OF THE TRAINING SEQUENCE

The FBB designs as discussed in the previous sections all depend on the critical assumption that a training sequence $\mathbf{b}_0$ (or $\underline{\mathbf{b}}_0$) is available. In this section, we discuss a few cases under which such a training sequence may be obtained.

### A. Single-Channel ADC

We consider first the single-channel case. Ideally, we should have $\mathbf{b}_0 = \mathbf{x}_0$: the training sequence is equal to an interference- and noise-free version of the incoming signal. For the case where the channel is instantaneous and not convolutive, this measured signal is (up to a complex scaling $\alpha$) equal to the transmitted signal, presumably an oversampled version of the Nyquist-rate symbol sequence $\mathbf{s}^{(1)}$ of $M$ entries (the superscript $(1)$ denotes the desired user index). In this case, we can set

$$\mathbf{b}_0 = \alpha \mathcal{I}\mathbf{s}^{(1)} \qquad (26)$$

where $\mathcal{I}$ is the interpolation function as defined in (2). If the channel is convolutive, then this generalizes to

$$\mathbf{b}_0 = \mathbf{G}^{(1)}\mathbf{s}^{(1)} \qquad (27)$$

where the $N \times M$-matrix $\mathbf{G}^{(1)}$ is the oversampled channel response of the desired user. The matrix $\mathbf{G}^{(1)}$ matrix can be estimated using standard channel estimation techniques [27], although the presence of the interference may complicate this. Alternatively, the estimation of $\alpha$ or $\mathbf{G}^{(1)}$ can be integrated in the MSE estimation of $\mathbf{w}$, changing the problem into a more general ARMA prediction estimation problem. The details of this estimation are omitted here and the reader is referred to [25].

Note that the FBB ADC structure consists of a feedback filter and is capable of doing some equalization. If the channel is convolutive but we still use (26) and set $\mathbf{b}_0 = \mathcal{I}\mathbf{s}^{(1)}$, then the FBB will attempt to do interference cancellation and equalization. The quality of the result depends on how well the equivalent $K$th order AR filter can equalize the convolutive channel. Presumably, this would work best for a channel that is FIR of order less than $K$ after oversampling.

### B. Multichannel ADC

In principle, the multichannel case is a straightforward generalization of the single-channel case, although we have somewhat more design freedom. We will assume that all channel outputs are to reconstruct the same desired user sequence $\mathbf{s}^{(1)}$, up to complex scalings denoted by $\alpha_i$ for the $i$th channel. In that case, we obtain for the instantaneous channel case as suitable training sequence the generalization of (26) as

$$\underline{\mathbf{b}}_0 = \mathcal{I}\mathbf{s}^{(1)} \otimes \mathbf{a} = (\mathcal{I} \otimes \mathbf{a})\mathbf{s}^{(1)} \qquad (28)$$

where $\mathbf{a} = [\alpha_1, \ldots, \alpha_{N_r}]^T$. We can recognize that $\mathbf{a}$ is the array response vector (or direction vector) of the desired user.

For a convolutive channel, we obtain as generalization

$$\underline{\mathbf{b}}_0 = \underline{\mathbf{G}}^{(1)}\mathbf{s}^{(1)} \qquad (29)$$

where $\underline{\mathbf{G}}^{(1)}$ is an $NN_r \times M$ matrix, with little structure, that contains the oversampled channel responses of the desired user to each of the $N_r$ antennas.

As in the single-channel case, if $\underline{\mathbf{G}}^{(1)}$ (or $\mathbf{a}$) is unknown, it will have to be estimated using a prior channel estimation phase, or the estimation has to be integrated in the estimation of $\mathbf{W}$, making this an ARMA estimation problem.

### C. Digital Postprocessing

After the ADC outputs $\underline{\mathbf{b}}$ have been obtained, typically two postprocessing steps are applied:
1) The outputs are low-pass filtered and downsampled. This operation can be represented by a wide rectangular matrix $\mathbf{F}$, e.g., $\mathbf{F} = \mathcal{I}^{\dagger}$. The downsampled ADC output as shown in Fig. 4 is a $MN_r \times 1$ vector $\underline{\mathbf{y}}_M : \underline{\mathbf{y}}_M = (\mathbf{F} \otimes \mathbf{I})\underline{\mathbf{b}}$.
2) These downsampled signals $\underline{\mathbf{y}}_M = [\mathbf{y}^T[0], \ldots, \mathbf{y}^T[M-1]]^T$ are used with a $N_r \times 1$ digital beamforming vector $\boldsymbol{\vartheta}$ to obtain an estimate of the desired user symbol sequence $\mathbf{s}^{(1)} = [s^{(1)}[0], \ldots s^{(1)}[M-1]]^T$ as $\tilde{s}^{(1)}[m] = \boldsymbol{\vartheta}^H \mathbf{y}[m]$. A reasonable approach to estimate $\boldsymbol{\vartheta}$ is to minimize the output MSE

$$\boldsymbol{\vartheta}_1 = \arg\min_{\boldsymbol{\vartheta}} \|(\mathbf{s}^{(1)})^T - \boldsymbol{\vartheta}^H\underline{\mathbf{Y}}\|^2$$
$$\Rightarrow \boldsymbol{\vartheta}_1 = (\underline{\mathbf{Y}}\,\underline{\mathbf{Y}}^H)^{-1}\underline{\mathbf{Y}}\,\bar{\mathbf{s}}^{(1)} \qquad (30)$$

where $\underline{\mathbf{Y}} = [\mathbf{y}[0], \ldots, \mathbf{y}[M-1]]$ and $\boldsymbol{\vartheta}$ is the well known Wiener beamformer. In this case $\mathbf{s}^{(1)}$ is known from training.

Ideally, we should have designed the FBB using a MSE criterion based on the resulting single output $\mathbf{s}^{(1)}$ and jointly design $\mathbf{W}$ and $\boldsymbol{\vartheta}$. It can be seen that this leads to a design problem with (too) many degrees of freedom and no unique solution, unless additional constraints are taken into account, such as the quantization errors.

For a related design problem, we were able to derive that the output MSE is minimized if all $N_r$ ADC outputs are nominally equal to the same signal, up to complex scaling, but with independent (quantization) noise [9]. This then motivates the design of the training as in the previous subsection. In the simplest case, all ADCs reconstruct the same output signal and the optimal (Wiener) beamformer $\boldsymbol{\vartheta}$ in that case becomes a simple average of the ADC outputs, after correcting for any phase differences. For convolutive channels, we can model the output as (29) and base the design of the beamformer on this model.

The effect of the lowpass filtering should be integrated in the MSE cost functions (11) and (21), i.e., we do not minimize $\|\mathbf{b} - \mathbf{b}_0\|$ but rather filtered versions $\|\mathbf{F}(\mathbf{b} - \mathbf{b}_0)\|^2$. The generalization is straightforward and will lead to a projection matrix based on $\mathbf{F}$. It will enter in the equations in a similar way as $\boldsymbol{\Lambda}$.

## VI. SIMULATION RESULTS

To assess the performance of the proposed algorithms, we have applied them to a multiantenna setup receiving multiple user signals and computer generated data. We present results that incorporate a first order FBB in a bank of first order $\Sigma\Delta$ ADCs as presented in Section IV.

In the simulations, the input signal-to-noise ratio (SNR) is the signal to noise power ratio between the desired user signal and the thermal noise as received at antenna 1; it is the same for all antennas. The input signal-to-interference ratio (SIR) is defined as the ratio of the power of the desired user signal to the sum of powers of all interfering user signals as received at antenna 1; it is the same for all antennas. All users (desired and interfering) transmit QPSK signals with zero mean and unit variance and the interfering users have equal powers. The performance indicators are

1) The average signal to interference and noise ratio (SINR) of the prediction error signal $\mathbf{p}[n]$ in the MC $\Sigma\Delta$ ADC. A high SINR indicates that less power is spent in quantizing the interferers for a given ADC resolution.
2) The MSE of $\tilde{s}^{(1)}[m]$, observed after digital postprocessing as in Section V-C.

All results are obtained by averaging 1000 Monte Carlo runs, each with instantaneous independent Rayleigh fading channel realizations and independently generated data signals. Each run transmits data packages of length 8192 symbols as in a WLAN transmission, where the first 256 symbols are used for training. The beamformer design techniques proposed in Section IV are used to design the FBB weights from the training sequence, followed by the digital beamformer $\boldsymbol{\vartheta}$. Unless specified otherwise, we used $N_t = 4$ transmitters and $N_r = 4$ receive antennas. The SIR at the antennas is $-5$ dB and the oversampling $\Sigma\Delta$ ADCs each have 1-bit resolution. After downsampling the effective resolution of each ADC at the Nyquist rate would be
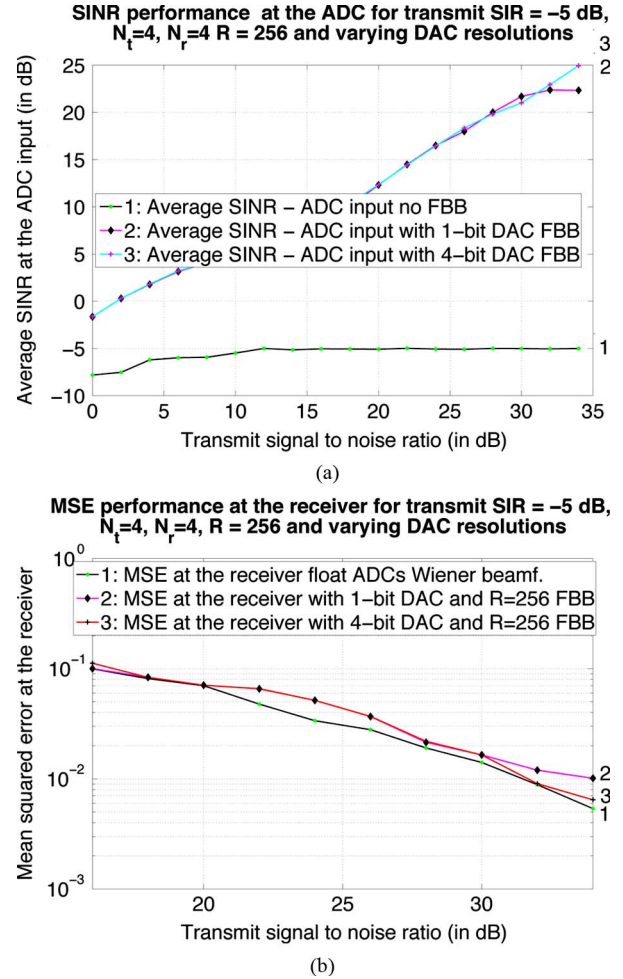


Fig. 5. Performance comparison as a function of the resolution of $\hat{\mathbf{x}}[n] = \mathrm{DAC}(\mathbf{W}^H \underline{\mathbf{b}}_K[n])$. (a) Average SINR at the ADC input. (b) MSE at the receiver.

$\log_2 R$, which corresponds to 6 bits for $R = 64$ (not taking into account the increase in resolution due to noise shaping).

### A. Effect of Fixed Precision DAC Feedback

Fig. 5 shows the basic performance of the MC FBB ADC receiver, where we set $R = 256$. We show the SINR at the input of the ADC, i.e., prediction error $\mathbf{p}[n]$ and the output MSE as a function of the input SNR. We further consider that in practice the feedback loop uses poorly quantized signals: the DAC in the feedback loop often has only 1 bit. Therefore, we plot several curves, for varying resolution of $\hat{\mathbf{x}}[n] = \mathrm{DAC}(\mathbf{W}^H \underline{\mathbf{b}}_K[n-1])$. A 1 bit DAC output has two possible levels i.e., the DAC output signal takes the possible values $\pm 1.0$ depending on the sign of the elements of $\hat{\mathbf{x}}[n]$. For a 4-bit DAC (assuming uniform quantization), the DAC output is rounded to one of the 16 uniformly spaced values between $+1.0$ and $-1.0$. For a precision greater than 1 bit, most DAC implementations are not linear, which is why they are less popular.

For reference, curve 1 in Fig. 5(a) plots the SINR performance for the case without FBB, i.e., $\mathbf{W} = \mathbf{I}$ when $K = 1$. Similarly, curve 1 in Fig. 5(b) plots the MSE performance for the optimal (Wiener) beamformer acting in the digital domain with full precision ADCs.
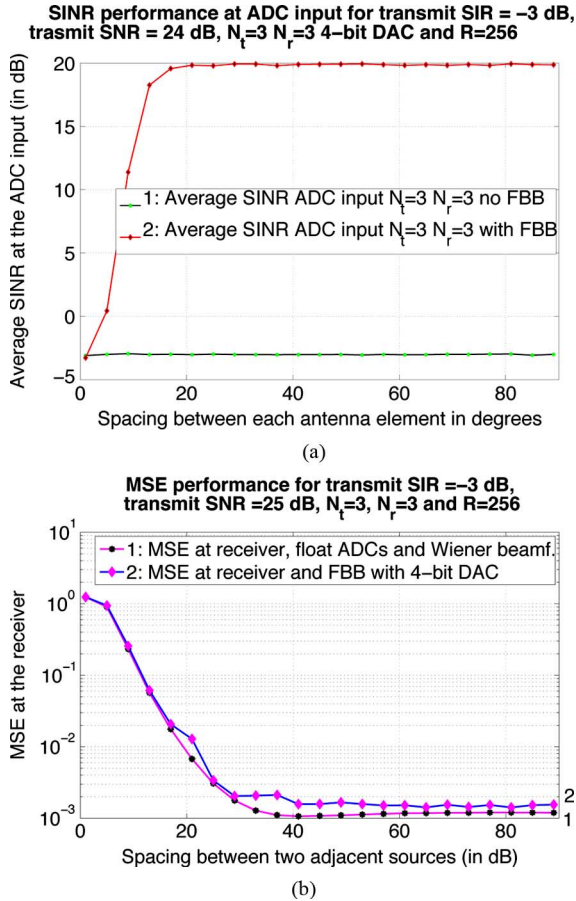
Fig. 6. Performance comparison as a function of the angular spacing between the desired user and 2 interferers. (a) Average SINR at the ADC input. (b) MSE at the receiver output.
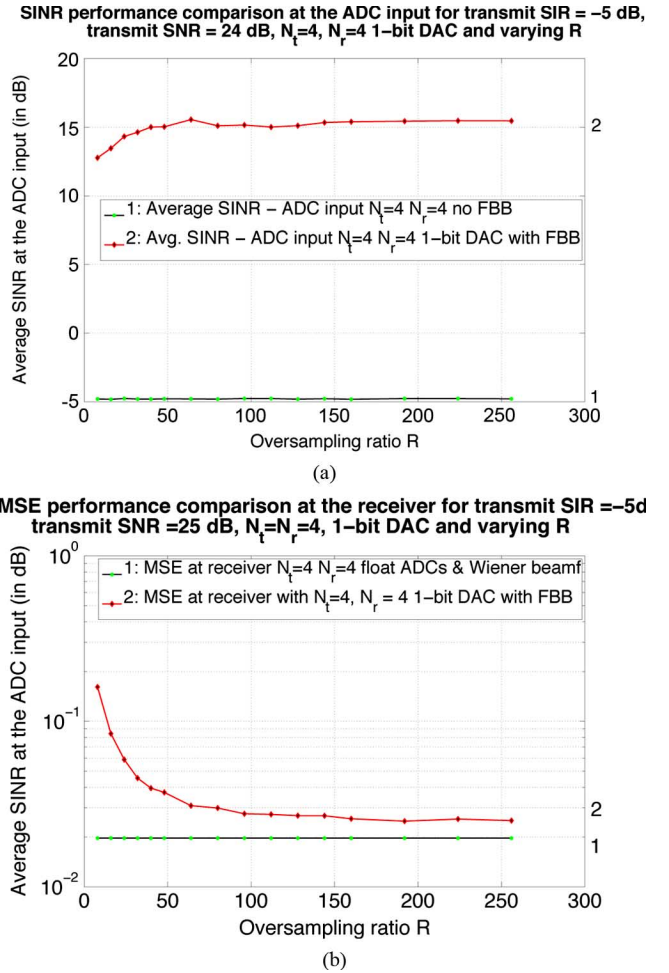


Fig. 7. Performance comparison as a function of varying oversampling ratios with 1-bit DAC (a) Average SINR at the ADC input (b) MSE at the receiver output.

Fig. 5(a) compares the average SINR at the input of the ADCs as a function of transmit SNR for different resolutions of the DAC. We see that the introduction of the FBB (curves 2, 3) improves the SINR by a factor of 20 to 25 dB, when compared to a setup without the FBB (curve 1). The introduction of the fixed precision DAC in the feedback loop introduces a small performance loss at the receiver as seen in Fig. 5(b). In this case, curve 1 acts as the reference and curves 2–3 show the MSE performance of the MC $\Sigma\Delta$ ADC setup for different DAC precisions. We conclude that a 1-bit DAC is probably sufficient because the performance degradation is small and the reduction in complexity is significant.

### B. Effect of Source Spacing

Fig. 6(a) and (b) show the SINR and MSE performance as a function of the angular spacing between two adjacent sources. The simulations consider a line of sight scenario without multipath and the results are observed for $N_t = 3$ sources, with the desired user transmitting from an angle of $\theta_0 = 0°$ and received by an $N_r = 3$ uniform linear array whose elements are spaced half-wavelengths $\lambda/2$ apart. The DAC output $\hat{x}[n] = \text{DAC}(\mathbf{W}^H\hat{\mathbf{b}}_K[n-1])$ is of 4-bit precision and $R = 256$. Two interferers are located at varying equidistant angles $\theta_1 = \theta$ and $\theta_2 = -\theta$. The plots show the SINR and MSE as function of $\theta$, where the transmit SNR is 24 dB and the transmit SIR is

$-3$ dB. For reference, curve 1 plots the SINR performance for the case without FBB and the MSE performance for the optimal (Wiener) beamformer, for $N_r = 3$ antennas.

Looking at Fig. 6(a) and comparing curves 1 and 2, we see that for $\theta > 20°$ the introduction of the FBB improves the SINR at the first ADC by a factor 25 dB. The FBB setup cannot suppress the interferers for angular spacing $\theta < 10°$. Fig. 6(b) shows the MSE performance after digital postprocessing. It is sen that the performance of the FBB follows that of the optimal Wiener beamformer closely, both cannot suppress interferers for small angular deviations.

### C. Effect of the ADC Oversampling Factor

We now keep the DAC precision at 1 bit and vary the oversampling ratio $R$ for transmit $\text{SNR} = 24 \text{ dB}$ and $N_t = N_r = 4$. Fig. 7(a) and (b) shows the SINR at the ADC input and the MSE performance at the receiver for varying $R$ (curve 2). For reference, curve 1 plots the SINR performance for the case without feedback and the MSE performance for the optimal (Wiener) float precision beamformer.

In Fig. 7(a), we observe that the introduction of the FBB improves the SINR by a factor of 20 dB (when $R \geq 8$). In Fig. 7(b), we see that MSE saturates for a FBB with $R \geq 64$ due to the
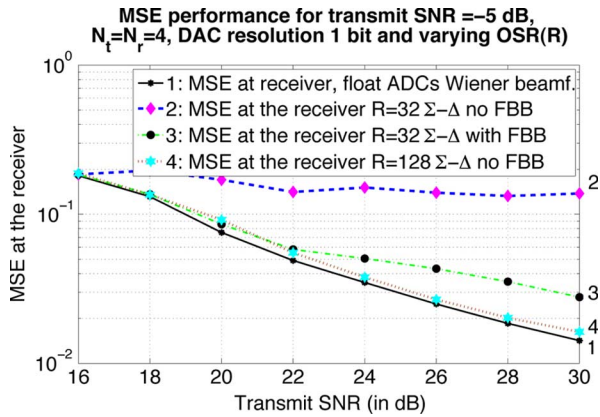
Fig. 8. Performance comparison illustrating oversampling ratios and ADC resolution.

limited resolution of the DAC. Note that although the MSE performance with the 1-bit DAC is considerably worse when compared with the optimal antenna array setup and float precision, the energy consumption in the proposed setup is considerably less than that of the reference setup.

### D. Extent of ADC Power Savings

We now keep the DAC precision fixed at 1-bit and compare the MSE performance of our $\Sigma\Delta$ ADC setup with a fixed precision first order $\Sigma\Delta$ ADC without FBB. The ADC resolution is kept fixed at 1-bit. In Fig. 8, curves 1, 2 and 4 respectively show the MSE performance at the receiver for float precision ADCs, $R = 32$ and $R = 128$ with no FBB i.e., $\mathbf{W} = \mathbf{I}$. Curve 3 corresponds to a MC $\Sigma\Delta$ ADC setup with $R = 32$ and a FBB $\mathbf{W}$ with 1-bit DAC arrangement as in Section IV.

From Fig. 8, we observe that the introduction of the FBB leads to a 5-bit ($R = 32$) ADC with MSE performance close (up to MSE 0.05) to that of an equivalent 7-bit ($R = 128$) ADC without FBB. From the power consumption relation $P_{adc} \propto N_r R(f_0 2^{res})$, we can conclude that interference cancellation with fixed precision ADCs in a dense multiuser setup leads to quadruple improvement in power consumption at a similar performance.

### VII. CONCLUDING REMARKS

In this paper, we have proposed a multichannel ADC setup, employing a space-time feedback beamformer that complements the usual DAC feedback. The prime advantage of this architecture is that it reduces the interference at the input of the ADCs, so that less dynamic range and fewer bits are required to reconstruct the desired user signals. These requirements can lead to significant power savings in the ADCs.

For this architecture, we designed the optimal feedback beamformer coefficients that minimize the output MSE compared to a training sequence. Simulations showed that the SINR at the input of the quantizer can be improved by more than 10–25 dB (depending on the SNR), thus indicating the potential power savings.

Further research is required along the following directions to implement the proposed class of ADCs in practice:

- We did not analyze the effect of the coarse quantization of the ADC. In practice, this setup operates on 1 bit signals

of the $\Sigma\Delta$ ADC output and the simulations indicated that this is adequate.

- We also did not analyze the effect of the quantization of the feedback DAC. When the DAC resolution is greater than 1 bit, the conversion may not be linear in practice, thus showing a preferred resolution of 1 bit. Simulations showed that this is sufficient for improving the SINR at the input of the quantizer, with a small loss in MSE.
- We did not consider the operating limits on the signal amplitude, slew rate, AGC, stability and other practical aspects. We also assumed an ideal ADC/DAC operation. In practice, there is a sample and hold circuit in the ADC, leading to latency in the ADC operation.
- The beamformer estimation employs training signals corresponding to the desired user. This is not practical in all cases. Furthermore, we need to do a forward channel estimation (or equivalently, replace the design problem of the AR feedback channel by an ARMA prediction error problem). This extension is a topic for future research.
- Initially, we are not synchronized to the desired user and in a dense setup, interference may overwhelm the ADCs. This requires a good initialization strategy.

### REFERENCES

[1] V. Venkateswaran, A. J. van der Veen, and D. Slock, "Sigma-delta interference canceling ADC's for antenna arrays," in *Proc. SPAWC*, 2009, pp. 459–463.

[2] V. Venkateswaran and A. J. van der Veen, "Feedback beamformer design with oversampling ADCs in multiantenna systems," in *Proc. ICASSP*, 2010, pp. 3426–3429.

[3] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 281–302, Mar. 2003.

[4] B. Murmann, "Digitally assisted analog circuits," *IEEE Micro*, vol. 26, no. 2, pp. 38–47, Mar./Apr. 2006.

[5] R. Walden, "Analog to digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, pp. 539–550, Apr. 1999.

[6] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[7] J. Yang, R. W. Brodersen, and D. Tse, "Addressing the dynamic range problem in cognitive radios," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2007, pp. 5183–5188.

[8] A. Hajimiri, H. Hashemi, A. Natarajan, X. Guang, and A. Babakhani, "Integrated phased arrays systems in silicon," *Proc. IEEE*, vol. 93, pp. 1637–1655, Sep. 2005.

[9] V. Venkateswaran and A. J. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, pp. 4131–4143, Aug. 2010.

[10] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer Academic, 1994.

[11] P. A. Aziz, H. V. Sorensen, and J. van der Spiegel, "An overview of sigma-delta convertors," *IEEE Signal Process. Mag.*, vol. 1, no. 1, pp. 61–84, Jan. 1996.

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

[13] H. Boelcskei and F. Hlawatsch, "Noise reduction in oversampled filter bank using predictive quantization," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 155–172, Jan. 2001.

[14] N. Thao and M. Vetterli, "Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates," *IEEE Trans. Signal Process.*, vol. 42, no. 3, pp. 519–531, Mar. 1994.

[15] S. Hein and A. Zakhor, "Reconstruction of oversampled bandlimited signals from sigma-delta encoded single bit binary sequences," *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 799–811, Apr. 1994.

[16] K. Philips, P. Nuijten, R. Roovers, A. v. Roermund, F. Chavero, M. Pallares, and A. Torralba, "A continuous time sigma delta ADC with increased immunity to interferers," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2170–2178, Dec. 2004.

[17] R. Winoto and B. Nikolic, "A downconverting sigma-delta modulator for a highly reconfigurable radio receiver," *J. Solid State Circuits*, vol. 45, no. 4, Apr. 2010.

[18] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2000.

[19] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.

[20] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error method for second order blind identification," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 694–705, Mar. 1997.

[21] D. T. M. Slock, "Blind fractionally spaced equalization, perfect reconstruction filter-banks and multichannel linear prediction," in *Proc. IEEE ICASSP*, 1994, vol. 4, pp. 585–588.

[22] P. Boufounos and R. G. Baraniuk, "Sigma delta quantization for compressive sensing," *SPIE Wavelets XII*, vol. 6701, Aug. 2007.

[23] N. Jayanth and P. Noll, *Digital Coding of Waveforms*. Upper Saddle River, NJ: Prentice-Hall, 1964.

[24] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[25] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 2002.

[26] J. McClellan and D. Lee, "Exact equivalence of the Steiglitz-Mcbride iteration and IQML," *IEEE Trans. Signal Process.*, vol. 39, no. 2, pp. 509–512, Feb. 1991.

[27] G. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advances in Wireless and Mobile Communications, Volume 1: Trends in Channel Estimation and Equalization*. Upper Saddle River, NJ: Prentice Hall/PTR, 2000.

**Vijay Venkateswaran** (M'09) received the M.S. degree in electrical engineering from the University of Arizona, Tucson, in 2003 and the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands in 2010.

He is currently a Member of Technical Staff at the RF & Antenna Technologies group, Bell Labs, Alcatel-Lucent Ireland, Dublin. From 2003 to 2005, he worked as an engineer at the Sony Corporation, Tokyo, Japan. He has held visiting/intern positions in Seagate Tech., Institute Eurecom and Nanyang Technological University. His research interests are in the general area of signal processing advancements in wireless communications and, in particular, joint RF-digital communications design.

**Alle-Jan van der Veen** (F'05) was born in The Netherlands in 1966. He received the Ph.D. degree (*cum laude*) from TU Delft, The Netherlands, in 1993.

Throughout 1994, he was a postdoctoral scholar with Stanford University, Stanford, CA. At present, he is a Full Professor with the Circuits and Systems Department, TU Delft.

Dr. van der Veen is the recipient of a 1994 and a 1997 IEEE Signal Processing Society (SPS) Young Author Paper award and was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1998–2001), Chairman of IEEE SPS Signal Processing for Communications Technical Committee (2002–2004), Editor-in-Chief of the IEEE SIGNAL PROCESSING LETTERS (2002–2005), Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2006–2008), and Member-at-Large of the Board of Governors of IEEE SPS (2006–2008). He is currently member of the IEEE SPS Awards Board, member of the IEEE Periodicals Review and Advice Committee, Technical Program Chair of IEEE ICASSP 2011 (Prague), and Chairman of the IEEE SPS Fellow Reference Committee.