Analog Beamforming in MIMO Communications With Phase Shift Networks and Online Channel Estimation

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

***Abstract*—In multiple-input multiple-output (MIMO) systems, the use of many radio frequency (RF) and analog-to-digital converter (ADC) chains at the receiver is costly. Analog beamformers operating in the RF domain can reduce the number of antenna signals to a feasible number of baseband channels. Subsequently, digital beamforming is used to capture the desired user signal. In this paper, we consider the design of the analog and digital beamforming coefficients, for the case of narrowband signals. We aim to cancel interfering signals in the analog domain, thus minimizing the required ADC resolution. For a given resolution, we will propose the optimal analog beamformer to minimize the mean squared error between the desired user and its receiver estimate. Practical analog beamformers employ only a quantized number of phase shifts. For this case, we propose a design technique to successively approximate the desired overall beamformer by a linear combination of implementable analog beamformers. Finally, an online channel estimation technique is introduced to estimate the required statistics of the wireless channel on which the optimal beamformers are based.**

***Index Terms*—ADC power consumption, analog beamforming, matching pursuit, passive RF phase shifters.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) and multi-sensor communication systems employ multiple receive antennas to exploit selection diversity and improve multiplexing gains. The aim is to achieve reliable communication close to theoretical limits [2]. However, the introduction of multiple antennas at the receiver leads to separate radio frequency (RF) front ends and analog to digital converter (ADC) units, i.e., increased circuit size and power consumption.

The implementation of digital baseband algorithms follows Moore's law, resulting in a power reduction by a factor of 32 for every ten years. In contrast, ADC power was reduced only by a factor of 10 in the past decade [3]. In existing multiantenna receivers, an ADC operation requires the same power as that of hundred thousands of logic gates [4]. To enable the full potential promised by MIMO capacity theory to become reality, there is a need for novel RF architectures with digital assistance. One particular option is to consider a multiuser cellular/WLAN scenario, where the bandpass RF signals contain contributions from the desired user, noise and interfering users. In the presence of strong interferers, the ADCs are forced to spend a significant part of their dynamic range on digitizing the unwanted interferers and noise. 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.

If the number of receiver antennas is given, one well known suboptimal technique to reduce the number of RF and ADC chains is to use antenna/diversity selection. Basically, we select the antenna with the highest signal energy [5]. This technique does not enable interference cancellation before the ADC.

An advancement over antenna selection is the use of analog preprocessing networks (APNs) for linearly combining the antennas, i.e., beamforming. Current hardware developments offer many possibilities. In [6]–[8], a phase shift preprocessor is implemented, which uses active and passive weighting elements to combine signals from the antenna array in the RF domain. Hajimiri *et al.* [9] propose a design where the required phase delays are implemented in the RF to baseband demodulation step, by using a bank of several phase-shifted local oscillators. Typically these designs provide about 16 possible phases (4-bit phase resolution) and no variable amplitude, thus implement a poorly quantized set of possible beamformers. These papers focus on the hardware design and only briefly touch upon the question how these beamformer coefficients should be selected. E.g., in [9], a set of beamforming vectors is precomputed to steer beams in predefined directions, with a resolution of about 22°. This only allows to select the direction with highest energy, which does not necessarily result in the desired signal in the presence of multipath and interference.

Improvements are possible by considering multiple output streams. Shown in Fig. 1 is an architecture where an analog preprocessing network (APN) directly operates on the RF signals, mapping N_r antenna array signals to $N_D < N_r$ receiver chains (i.e., ADCs). A digital beamformer $\boldsymbol{\vartheta}$ subsequently combines the ADC outputs to generate the desired user estimate. Zhang *et al.* [10] considered such an architecture, and proposed several MIMO transmitter/receiver beam steering techniques.

Problem Statement: Our aim in this paper is to design an optimal APN beamforming matrix. Our focus is to minimize the interference at the input of the ADCs, so that reduced resolution is possible, leading to reduced power consumption. Some design issues are (1) to choose N_D, (2) to select the beamforming

Manuscript received November 19, 2009; accepted April 07, 2010. Date of publication April 15, 2010; date of current version July 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Martin Schubert. This research was supported by Senter-Novem under the IOP-Gencom program (IGC-0502B). Parts of this paper was presented at the International Conference on Acoustics, Speech and Signal Processing (ICASSP), Dallas, TX, March 15–19, 2010 [1].

The authors are with the Delft University of Technology, Faculty of the Electrical Engineering, Mathematics and Computer Science, 2628 CD Delft, The Netherlands (email: a.j.vanderveen@tudelft.nl; v.venkateswaran@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2048321

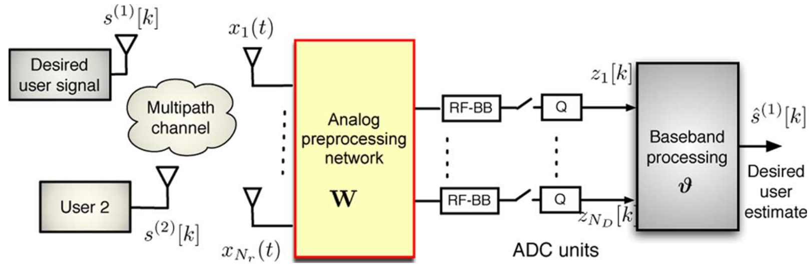


Fig. 1. Proposed receiver architecture: the analog preprocessing network (RF beamformer) cancels interference and reduces the number of antenna signals to a smaller number of ADC chains.

coefficients, and (3) to determine how many bits are needed in each of the ADCs. A design constraint is the poor resolution of the APN coefficients. Design criteria are the mean-square error (MSE) at the output of the digital beamformer and the ADC power consumption.

As an example, consider a wireless channel with one interfering user transmitting signals with the same energy as that of the desired user, with $N_r = 4$ antennas and $N_D = 2$ receiver chains. It will be seen from the simulation results in Section VI, that RF interference cancellation with an APN can reduce the ADC power consumption by half for the same MSE at the receiver output.

A. Connections

In the array signal processing literature, several types of preprocessing matrices have been designed to reduce the number of receiver chains. One related context is "beamspace array processing," where a preprocessing is done on the receive antennas to reduce dimensionality; see, e.g., [11]. In earlier work, this is called "partially adaptive beamforming," where the preprocessor is fixed and the digital beamformer is adaptive; see, e.g., [12]. In this literature, the design of the beamspace transformation matrix is based on prior knowledge of the location of the signal of interest and/or on the interference scenario. Reduced dimension transformations, to cancel interferers using statistics from the desired user, have also been proposed [13]. In the present paper, we aim to design the APN using feedback from the baseband processor, so that it can be optimized for the actual situation on a block-by-block basis.

Transformation preprocessors have also been pursued in direction-of-arrival (DOA) estimation problems, e.g., [14]. In that paper, a set of preprocessors is applied over time, and the results are combined to estimate the DOA. In contrast, our aim is to minimize interference and reconstruct the signal of interest; however, we will apply techniques from [14] to estimate the required channel parameters.

In practice, the APN coefficients are quantized. There exists a significant amount of literature on (adaptive) beamforming using variable phase only (cf. equal gain combining). Most literature considers a single weight vector (with variable phases only) that should be designed to match certain performance criteria, e.g., [6], [15], and [16]. The paper [10] considers an APN with multiple outputs, and it is shown that any desired weight vector can always be obtained by linearly combining two phase-only beamformers, thus $N_D = 2$ is sufficient. The work of [10] has generated some follow-up work, focusing on phase shift beamforming at the transmitter and receiver to linearly combine the signals. However, these approaches do not emphasize on interference cancellation nor on implementable APN weights. We will consider a more restricted case where the APN weights and ADC taps are severely quantized.

B. Contributions and Outline

In the paper, we progressively study various aspects of the APN beamformer design. In Section II, the system setup and the data model is specified. In Section III we consider the case where the APN and ADC have sufficiently high resolution. We will aim to design an $N_r \times N_D$ preprocessing matrix that minimizes the MSE. This leads to a non-unique design. To make it unique, we also take the ADC quantization error into account and we design the APN to maximize the signal to quantization noise ratio (SQNR). For an APN with infinite precision, we will derive that it is sufficient to consider only $N_D = 1$ output chain.

However, as mentioned earlier, in practice the APN is a programmable discrete phase shifter with a coarse quantization. In this case, using $N_D > 1$ allows us to extract different low resolution streams, some representing the user of interest and some the interferers, followed by digital combining. Thus, in Section IV, we study the case where the APN quantization is the limiting factor in the design. To minimize the MSE, we try to match an optimal beamformer to a linear combination of N_D vectors from the set of available quantized beamformers; for this we propose a quantized version of the matching pursuit (MP) algorithm [17], [18]. All the above mentioned designs depend on knowledge of the channel statistics: the antenna covariance matrix and the antenna cross-correlation vector with the desired user signal. Note that the digital receiver does not have direct access to the antenna signals. In Section V, we provide an algorithm to deduce this information from various observations of beamformed outputs (the algorithm is related to that of Sheinvald *et al.* [14]). The results are combined in the digital beamformer to obtain a high resolution estimate and illustrated using simulation results.

Notation: Vectors and matrices are represented in lower and upper case bold letters. $(.)^T$, $(.)^H$, and $(.)^\dagger$ represent transpose, complex conjugate transpose and pseudo inverse, respectively. $\otimes$ and $||.||^2$ represent Kronecker product and Frobenius norm. The operation $\mathbf{f} = \mathrm{vec}(\mathbf{F})$ transforms a matrix $\mathbf{F}$ to a vector $\mathbf{f}$ by stacking its columns, while $\mathbf{F} = \mathrm{vec}^{-1}(\mathbf{f})$ does the opposite. Continuous time signals are represented with round braces as in $(\cdot)$ and sampled signals as $[\cdot]$.

II. System Setup and Data Model

A. RF Data Processing

Consider an RF signal $\tilde{x}(t)$ received at an antenna. Assuming suitable bandpass prefiltering, only a narrow frequency band around a carrier frequency f_c is of interest, and we can write

$$\tilde{x}(t) = \text{real}\{x(t)e^{j2\pi f_c t}\}$$

where $x(t)$ is the complex envelope or baseband signal. In the receiver, the "RF to baseband" processing block recovers $x(t)$ using quadrature demodulation. This signal subsequently enters the ADC unit. Here it is sampled at time instants $t = kT$ (where T is the sampling period) leading to $x[k]$ and quantized using R bits, leading to $Q(x[k])$. We will always assume that the Nyquist condition holds. The ADC unit includes an automatic gain control (AGC) that scales the input signal such that its amplitude matches the range of the ADC without overload.

If the RF signal $\tilde{x}(t)$ is delayed by τ, we obtain

$$\begin{aligned}\tilde{x}(t-\tau) &= \text{real}\{x(t-\tau)e^{j2\pi f_c(t-\tau)}\}\\ &\approx \text{real}\{x(t)e^{-j2\pi f_c\tau}e^{j2\pi f_c t}\}.\end{aligned}$$

The approximation $x(t-\tau) \approx x(t)$ is valid if $f\tau \ll 1$ for all frequencies f in the bandwidth of $x(t)$, i.e., the "narrowband condition". After RF to baseband conversion, the delayed baseband signal is $x(t)e^{-j2\pi f_c\tau}$ and the sampled signal is $x[k]e^{-j2\pi f_c\tau}$.

If we have an array with N_r receive antennas, it will be convenient to stack all signals into $N_r \times 1$ vectors $\tilde{\mathbf{x}}(t)$, $\mathbf{x}(t)$, and $\mathbf{x}[k]$, respectively.

B. Received Data Model

Consider now a communication setup, where N_t user signals $s^{(j)}(t),\ j \in \{1,\ldots,N_t\}$ are transmitted over the same carrier f_c, propagate over a multipath channel, and are received by the array with N_r antennas. Without loss of generality, let $s^{(1)}(t)$ is the desired user signal, and the other signals are considered interferers. Assume that the narrowband condition holds for all propagation delays (except for a bulk delay that we will ignore here) so that they can be represented by phase shifts. We can write the equivalent discrete time data model as

$$\mathbf{x}[k] = \mathbf{H}\mathbf{s}[k] + \mathbf{n}[k]$$

where $\mathbf{s}[k] = [s^{(1)}[k],\ldots,s^{(N_t)}[k]]^T$ is an $N_t \times 1$ vector of user signals, and $\mathbf{n}[k]$ is an $N_r \times 1$ vector of noise signals. $\mathbf{H}$ is a $N_r \times N_t$ matrix denoting the MIMO channel response with complex entries h_{ij}, representing the channel coefficients for the propagation of the jth user signal to the ith receive antenna, which includes the transmit/receive filters, array response, amplitude scalings, and phase delays.

Throughout the paper, we will make the following standard assumptions on this model.

- The user signals $s^{(j)}[k]$ are modeled as random processes that are zero mean, independent, wide-sense stationary, with equal powers normalized to 1.
- The noise signal vector $\mathbf{n}[k]$ is sampled from an i.i.d. Gaussian process, zero mean, with unknown covariance matrix $\mathbf{R}_n$.

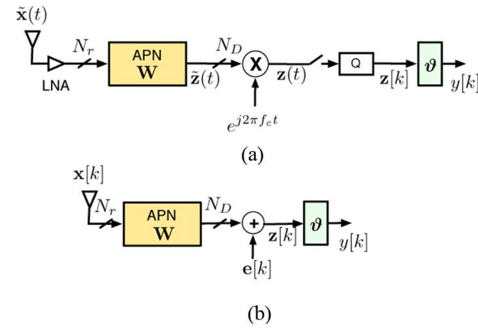


Fig. 2. (a) Proposed receiver architecture with RF beamformer. (b) Discrete time equivalent.

C. High-Resolution Digital Beamforming

Given observations $\mathbf{x}[k]$, our goal is to obtain an estimate $\hat{s}^{(1)}[k]$ of the desired user signal $s^{(1)}[k]$. We will consider linear beamforming and a minimum mean-square error (MMSE) criterion. Thus, let $\boldsymbol{\theta}$ be an $N_r \times 1$ weight vector, then the digital beamforming output is

$$y[k] = \boldsymbol{\theta}^H\mathbf{x}[k]$$

and the MMSE beamformer is obtained as the solution of

$$\boldsymbol{\theta}_0 = \arg\min_{\boldsymbol{\theta}} \quad E\|s^{(1)}[k] - \boldsymbol{\theta}^H\mathbf{x}[k]\|^2. \tag{1}$$

As is well known, the solution is given by the Wiener beamformer [19]

$$\boldsymbol{\theta}_0 = \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{r}_{\mathbf{x}s} \tag{2}$$

where $\mathbf{R}_{\mathbf{x}} = E\{\mathbf{x}[k]\mathbf{x}^H[k]\}$ and $\mathbf{r}_{\mathbf{x}s} = E\{\mathbf{x}[k]\bar{s}^{(1)}[k]\}$. Estimates of $\mathbf{R}_{\mathbf{x}}$ and $\mathbf{r}_{\mathbf{x}s}$ are obtained from the sample covariance matrix and sample cross-correlation vector; this requires access to all antenna signals and the availability of a reference signal (training sequence) for the desired user.

The above solution (2) will serve as our reference design. In its derivation, the effect of the quantizer operator $Q(\cdot)$ was ignored.

D. Analog Preprocessing Network

As mentioned in the Introduction, it is expensive to insert a complete RF receiver chain and high-rate high-resolution ADC unit for each antenna. We thus consider an APN inserted in the RF domain immediately after the low-noise amplifiers and bandpass filter [see Fig. 2(a) and its discrete time equivalent Fig. 2(b)]. Although there are various implementations, we will model the APN as an analog beamformer that constructs linear combinations of slightly delayed antenna signals $\tilde{x}_i(t)$. This results in output signals $\tilde{z}_j(t)$, $j = 1,\ldots,N_D$, where the number of outputs $N_D < N_r$. As before, the output signals are stacked into $N_D \times 1$ vectors $\tilde{\mathbf{z}}(t)$, downconverted (leading to baseband signals $\mathbf{z}(t)$), sampled and quantized (leading to $\mathbf{z}[k]$). The jth ADC has resolution R_j bits.

The effect of the APN on the baseband signal is modeled using a discrete time equivalent matrix operation

$$\mathbf{z}[k] = Q(\mathbf{W}^H\mathbf{x}[k])$$

where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_{N_D}]$ is a matrix of size $N_r \times N_D$ and $\mathbf{w}_j = [w_{1j}, \ldots, w_{N_r j}]^T$ is a $N_r \times 1$ vector. Each entry w_{ij} corresponds to the phase delay introduced by the APN for the ith receive antenna and the jth output signal. Practical implementations limit w_{ij} to a small set of possible phase shifts, perhaps 8 to 16 choices (3 to 4 bits). Amplitude changes are usually not possible. The weights w_{ij} can be controlled by the baseband processor but note that $\mathbf{x}[k]$ is not directly available at the processor, making the design of w_{ij} a challenge.

The digital baseband signals $\mathbf{z}[k]$ are subsequently combined using a digital beamformer $\boldsymbol{\vartheta} = [\vartheta_1, \ldots, \vartheta_{N_D}]^T$, resulting in the output signal

$$y[k] = \boldsymbol{\vartheta}^H \mathbf{z}[k].$$

To obtain an estimate of $s^{(1)}[k]$, $\boldsymbol{\vartheta}$ can be designed as the MMSE beamformer solving

$$\boldsymbol{\vartheta}_0 = \arg\min_{\boldsymbol{\vartheta}} \quad E\|s^{(1)}[k] - \boldsymbol{\vartheta}^H \mathbf{z}[k]\|^2 \tag{3}$$

leading to a Wiener beamformer $\boldsymbol{\vartheta}_0 = \mathbf{R}_{\mathbf{z}}^{-1} \mathbf{r}_{\mathbf{z}s}$ specified in terms of correlations of $\mathbf{z}[k]$ where $\mathbf{R}_{\mathbf{z}} = E\{\mathbf{z}[k]\mathbf{z}^H[k]\}$ and $\mathbf{r}_{\mathbf{z}s} = E\{\mathbf{z}[k]\bar{s}^{(1)}[k]\}$. For given $\mathbf{W}$, $\mathbf{z}[k]$ is known and this problem can be solved, hence $\boldsymbol{\vartheta}_0$ is a function of $\mathbf{W}$.

E. Problem Formulation

Our aim in this paper is to design the APN $\mathbf{W}$. After that, $\mathbf{z}[k]$ is fixed and the design of $\boldsymbol{\vartheta}$ is relatively straightforward. For the design, there are a number of side conditions or assumptions:

[A1] The APN circuits consist of a limited number of phase shift combinations, hence the elements of $\mathbf{W}$ are selected from a finite set, denoted as a dictionary $\mathcal{D}$.

[A2] Each ADC performs uniform quantization with a resolution of R_j bits, hence $Q\{z_j(t)\} \in [-2^{-(R_j-1)}, 2^{-(R_j-1)}]$.

The ADC power consumption can be approximated as $P_{\mathrm{adc}} \propto f_s 2^{2R}$, where f_s is the sampling frequency (in our case the Nyquist rate) and R is the ADC resolution in bits. We wish to minimize the number of bits in the ADC, since this is directly related to the power consumption in the analog section of the receiver. The number of bits is determined by the required dynamic range, which is partially controlled by $\mathbf{W}$. Indeed, if more interference cancellation is performed, fewer bits are needed for the same MSE performance.

Design objectives are 1) minimizing the MSE at the output of the digital beamformer, including the quantization noise, and 2) minimize the energy consumption in the ADCs, represented by $\sum_{j=1}^{N_D} 2^{2R_j}$. Several problems can be formulated using these objectives, but they do not all have feasible solutions.

We therefore approach the APN design in the following order:

[P0] We initially relax [A1] and [A2] and assume a perfect and continuous APN, and high resolution ADCs. What are then the constraints on the design of $\mathbf{W}$? (It will follow that $\mathbf{W}$ is not unique.)

[P1] Assuming low-resolution ADCs, each with $R_j = R$ bits, how does the design change? Can we compute a unique $\mathbf{W}$?

[P2] Now considering the discrete nature of the APN, select $\mathbf{W}$ from a fixed set of discrete phase shifts present in $\mathcal{D}$, such that the overall MSE is minimized. Here it is assumed that the ADC resolution is not limiting.

The above design problems [P1] and [P2] form the core of this paper and are covered respectively in Sections III and IV.

The design techniques will assume knowledge of the $N_r \times N_r$ antenna array covariance matrix $\mathbf{R}_{\mathbf{x}} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\}$ and a $N_r \times 1$ cross covariance vector $\mathbf{r}_{\mathbf{x}s} = E\{\mathbf{x}(t)\bar{s}^{(1)}(t)\}$. Note that the introduction of the APN implies that this covariance matrix is not available in the digital part of the receiver. In Section V, we explain a technique to estimate $\mathbf{R}_{\mathbf{x}}$ and $\mathbf{r}_{\mathbf{x}s}$ from a set of low-rate beamformers in digital baseband, assuming that a training sequence of the desired user is available.

III. Preprocessor Design-APN Not Quantized

In this section, we consider problem [P1]: design the APN considering only the quantization by the ADCs, and design the number of bits which each ADC should use. We do not consider the limited choice of phase shifts for the APN—this case is deferred to Section IV.

A. Conditions on $\mathbf{W}$ to Minimize the MSE

Let us also ignore the quantization operation by the ADCs for the moment, i.e., consider problem [P0]. The problem is to design

$$\boldsymbol{\theta} = \mathbf{W}\boldsymbol{\vartheta}, \quad \mathbf{W} : N_r \times N_D.$$

We have $\mathbf{z}[k] = \mathbf{W}^H \mathbf{x}[k]$, and the output $y[k] = \boldsymbol{\vartheta}^H \mathbf{z}[k]$ should be close to the desired signal $s^{(1)}[k]$ in MMSE sense, i.e., we require

$$\boldsymbol{\theta} = \boldsymbol{\theta}_0 = \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{r}_{\mathbf{x}s}.$$

Thus, $\mathbf{W}$ implements a rank reduction on the space spanned by $\mathbf{x}[k]$. Some common but suboptimal designs based on rank reduction are listed in Appendix I.

Instead, we will now consider which conditions $\mathbf{W}$ has to satisfy such that we can optimize the MSE. Note that, once $\mathbf{W}$ is specified, we know $\mathbf{z}[k] = \mathbf{W}^H \mathbf{x}[k]$, and if we want to minimize the MSE of the output $y[k] = \boldsymbol{\vartheta}^H \mathbf{z}[k]$, we know we will select

$$\boldsymbol{\vartheta}_0 = \mathbf{R}_{\mathbf{z}}{}^{-1} \mathbf{r}_{\mathbf{z}s} = (\mathbf{W}^H \mathbf{R}_{\mathbf{x}} \mathbf{W})^{-1} \mathbf{W}^H \mathbf{r}_{\mathbf{x}s}$$

where

$$\mathbf{R}_{\mathbf{z}} = \mathbf{W}^H \mathbf{R}_{\mathbf{x}} \mathbf{W}, \quad \mathbf{r}_{\mathbf{z}s} = \mathbf{W}^H \mathbf{r}_{\mathbf{x}s}. \tag{4}$$

Thus, the MSE is a function of $\mathbf{W}$ only. We can see immediately that $\mathbf{W}$ will not be unique: e.g., we could choose $\mathbf{W} = \boldsymbol{\theta}_0$, $\boldsymbol{\vartheta} = 1$ (for $N_D = 1$), or $\mathbf{W} = [\boldsymbol{\theta}_0,\ \boldsymbol{\theta}_0]$, $\boldsymbol{\vartheta} = [1/2,\ 1/2]^T$ (for $N_D = 2$), among many other possibilities.

Define the "whitened" correlation matrices

$$\underline{\mathbf{r}}_{\mathbf{x}s} = \mathbf{R}_{\mathbf{x}}^{-1/2} \mathbf{r}_{\mathbf{x}s}, \quad \underline{\mathbf{W}} = \mathbf{R}_{\mathbf{x}}^{1/2} \mathbf{W}. \tag{5}$$

The following lemma characterizes all solutions $\mathbf{W}$ that lead to MMSE-optimal beamformers $\boldsymbol{\theta} = \mathbf{W}\boldsymbol{\vartheta}$.

Lemma 1: Consider the scenario [P0]: the APN is not quantized, and the quantization error of the ADCs is ignored. Then all beamformers $\mathbf{W}$ that lead to the MMSE-optimal solution $\boldsymbol{\theta}_0 = \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{r}_{\mathbf{x}s}$ are characterized by the condition

$$\underline{\mathbf{r}}_{\mathbf{x}s} \in \text{colspan}\{\underline{\mathbf{W}}\} \Leftrightarrow \mathbf{r}_{\mathbf{x}s} \in \text{colspan}\{\mathbf{R}_{\mathbf{x}}\mathbf{W}\}$$

Proof: $\boldsymbol{\theta} = \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{r}_{\mathbf{x}s} = \mathbf{W}\boldsymbol{\vartheta}$ implies that $\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{r}_{\mathbf{x}s}$ is in the column span of $\mathbf{W}$. Hence, a necessary and also sufficient condition on $\mathbf{W}$ is $\underline{\mathbf{r}}_{\mathbf{x}s} \in \text{colspan}\{\underline{\mathbf{W}}\}$.

An alternative proof is as follows. Define the orthogonal projection matrix

$$\mathbf{P}_{\underline{\mathbf{W}}} = \underline{\mathbf{W}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\underline{\mathbf{W}}^H.$$

For any $\mathbf{W}$, and corresponding optimal $\boldsymbol{\vartheta}_0 = (\mathbf{W}^H\mathbf{R}_{\mathbf{x}}\mathbf{W})^{-1}\mathbf{W}^H\mathbf{r}_{\mathbf{x}s}$, the MSE is given by

$$\begin{aligned} &E\|s^{(1)}[k] - (\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{z}s})^H\mathbf{z}[k]\|^2 \\ &=E\|s^{(1)}[k] - \mathbf{r}_{\mathbf{z}s}^H\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{W}^H\mathbf{x}(t)\|^2 \\ &=1 - \mathbf{r}_{\mathbf{x}s}^H\mathbf{W}(\mathbf{W}^H\mathbf{R}_{\mathbf{x}}\mathbf{W})^{-1}\mathbf{W}^H\mathbf{r}_{\mathbf{x}s} \\ &=1 - \mathbf{r}_{\mathbf{x}s}^H\mathbf{W}(\mathbf{W}^H\mathbf{R}_{\mathbf{x}}\mathbf{W})^{-1}\mathbf{W}^H\mathbf{r}_{\mathbf{x}s} \\ &=1 - \underline{\mathbf{r}}_{\mathbf{x}s}^H\mathbf{P}_{\underline{\mathbf{W}}}\underline{\mathbf{r}}_{\mathbf{x}s}. \end{aligned} \tag{6}$$

Thus, the MMSE solution $\mathbf{W}_0$ satisfies

$$\underline{\mathbf{W}}_0 = \arg\max_{\underline{\mathbf{W}}} \ \underline{\mathbf{r}}_{\mathbf{x}s}^H\mathbf{P}_{\underline{\mathbf{W}}}\underline{\mathbf{r}}_{\mathbf{x}s}. \tag{7}$$

Clearly, this only specifies that $\underline{\mathbf{r}}_{\mathbf{x}s} \in \text{colspan}\{\underline{\mathbf{W}}\}$, and there can be many solutions.

B. Maximizing the SQNR

The next question is whether we can narrow down the set of available solutions for $\mathbf{W}$, by satisfying additional design objectives. Our approach is to incorporate the power consumption or ADC resolution in (7), and minimize the MSE while maximizing the signal to quantization noise ratio (SQNR) of the output estimate. This leads to a design for $\mathbf{W}$, and also to criteria on the number of bits which each ADC should use.

We thus consider problem [P1]. Incorporating the effect of the ADC on the signal at the output of the analog beamformer, we have

$$\mathbf{z}[k] = Q(\mathbf{W}^H\mathbf{x}[k]).$$

The effect of the quantizer will be modeled by an $N_D \times 1$ additive noise vector $\mathbf{e}[k] = [e_1[k], \ldots, e_{N_D}]^T$, i.e.,

$$\mathbf{z}[k] = \mathbf{W}^H\mathbf{x}[k] + \mathbf{e}[k].$$

As usual, $\mathbf{e}[k]$ is modeled as uniformly distributed noise, entry-wise independent, and uncorrelated to $\mathbf{x}[k]$. The corresponding covariance matrix of $\mathbf{z}[k]$ is [replacing (4)]

$$\mathbf{R}_{\mathbf{z}} = \mathbf{W}^H(\mathbf{R}_{\mathbf{x}} + \mathbf{R}_{\mathbf{e}})\mathbf{W}, \quad \mathbf{r}_{\mathbf{z}s} = \mathbf{W}^H\mathbf{r}_{\mathbf{x}s}, \tag{8}$$

where $\mathbf{R}_{\mathbf{e}}$ is a diagonal matrix whose diagonal entries represent the quantization noise variance. Suppose that the ith ADC has a resolution of R_i number of bits. We assume that an automatic gain control (AGC) is used such that the dynamic range of the ADC is optimally used. The quantization noise variance $\sigma_{e_i}^2 = E\{e_i[k]\bar{e}_i[k]\}$ will then depend on the signal variance at the input of the ADC, with some abuse of notation[1] denoted as $\sigma_{z_i}^2$. Using the well-known Lloyd–Max equation [20], we have

$$\sigma_{e_i}^2 = \sigma_{z_i}^2\gamma\frac{2^{-2R_i}}{12}$$

where γ is an AGC scaling factor that models the difference between the average input power and the peak input power.

To reach a feasible optimization problem, we will limit ourselves to the case where all ADCs use an equal number of bits, $R_i = R$. The noise covariance matrix can then be expressed as

$$\mathbf{R}_{\mathbf{e}} = \mathbf{D}_{\mathbf{z}}\frac{\gamma 2^{-2R}}{12} \quad \text{where} \quad \mathbf{D}_{\mathbf{z}} = \text{diag}\{\mathbf{W}^H\mathbf{R}_{\mathbf{x}}\mathbf{W}\}.$$

Given $\mathbf{W}$, the optimal digital beamformer, acting on $\mathbf{z}[k]$, is still $\boldsymbol{\vartheta} = \mathbf{R}_{\mathbf{z}}{}^{-1}\mathbf{r}_{\mathbf{z}s}$. At the output of the beamformer, the average energies of the desired user signal and quantization noise are, respectively

$$\begin{aligned} E\|\hat{s}^{(1)}[k]\|^2 &= E\|\boldsymbol{\vartheta}^H\mathbf{z}[k]\|^2 = \mathbf{r}_{\mathbf{z}s}^H\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{z}s} \\ E\|\boldsymbol{\vartheta}^H\mathbf{e}[k]\|^2 &= \mathbf{r}_{\mathbf{z}s}^H\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{R}_{\mathbf{e}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{z}s} \end{aligned}$$

and the corresponding SQNR of the output signal is

$$\text{SQNR} = \frac{E\|\boldsymbol{\vartheta}^H\mathbf{z}[k]\|^2}{E\|\boldsymbol{\vartheta}^H\mathbf{e}[k]\|^2} = \frac{\mathbf{r}_{\mathbf{z}s}^H\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{z}s}}{\mathbf{r}_{\mathbf{z}s}^H\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{D}_{\mathbf{z}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{z}s}\frac{\gamma 2^{-2R}}{12}} \tag{9}$$

where $\mathbf{R}_{\mathbf{z}}$ and $\mathbf{r}_{\mathbf{z}s}$ are functions of $\mathbf{W}$ as specified in (8). The objective is to design $\mathbf{W}$, first of all, to minimize the MSE as discussed in the previous subsection, and this leads to design freedom, which is used to maximize the SQNR.

Regarding the minimization of the MSE, we can follow the derivation that led to (6), however, $\mathbf{R}_{\mathbf{z}}$ is now slightly different since it also involves the quantization noise $\mathbf{R}_{\mathbf{e}}$. For a reasonable number of bits, we will have that $\|\mathbf{R}_{\mathbf{e}}\| = \|\mathbf{D}_{\mathbf{z}}\|\gamma 2^{-2R}/12 \ll \|\mathbf{W}^H\mathbf{R}_{\mathbf{x}}\mathbf{W}\|$. In that case, we can ignore the influence of $\mathbf{R}_{\mathbf{e}}$ on $\mathbf{R}_{\mathbf{z}}$, (6) still holds, and the MMSE is obtained for any $\mathbf{W}$ satisfying (7),

$$\underline{\mathbf{r}}_{\mathbf{x}s} \in \text{colspan}\{\underline{\mathbf{W}}\}. \tag{10}$$

Using the whitened quantities (5), the numerator of the SQNR expression can be written as

$$\begin{aligned} \mathbf{r}_{\mathbf{z}s}^H\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{z}s} &= \mathbf{r}_{\mathbf{x}s}^H\mathbf{W}^H(\mathbf{W}\mathbf{R}_{\mathbf{x}}\mathbf{W})^{-1}\mathbf{W}^H\mathbf{r}_{\mathbf{x}s} \\ &= \underline{\mathbf{r}}_{\mathbf{x}s}\underline{\mathbf{W}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\underline{\mathbf{W}}^H\underline{\mathbf{r}}_{\mathbf{x}s} \end{aligned}$$

and the denominator (up to scaling by $\gamma 2^{-2R}/12$) as

$$\begin{aligned} &\mathbf{r}_{\mathbf{z}s}^H\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{D}_{\mathbf{z}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{z}s} \\ &\quad = \underline{\mathbf{r}}_{\mathbf{x}s}^H\underline{\mathbf{W}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\mathbf{D}_{\mathbf{z}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\underline{\mathbf{W}}^H\underline{\mathbf{r}}_{\mathbf{x}s} \end{aligned}$$

[1]since $z_i[k]$ denotes the signal at the output of the ADC in this subsection.

where $\mathbf{D}_\mathbf{z} = \mathrm{diag}\{\underline{\mathbf{W}}^H\underline{\mathbf{W}}\}$. Subject to (10), the numerator is equal to $\|\underline{\mathbf{r}}_{\mathbf{x}s}\|^2$, which is a constant independent of $\underline{\mathbf{W}}$. It suffices to minimize the denominator. It further follows from the expression of the SQNR that the scaling of $\underline{\mathbf{W}}$ is not important.

The SQNR optimization problem (subject to optimal output MSE) becomes

$$\underline{\mathbf{W}}_0 = \arg\min_{\underline{\mathbf{W}}} \; \frac{\underline{\mathbf{r}}_{\mathbf{x}s}^H}{\|\underline{\mathbf{r}}_{\mathbf{x}s}\|}\underline{\mathbf{W}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\mathbf{D}_\mathbf{z}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\underline{\mathbf{W}}^H\frac{\underline{\mathbf{r}}_{\mathbf{x}s}}{\|\underline{\mathbf{r}}_{\mathbf{x}s}\|}$$
$$\text{subject to} \quad \underline{\mathbf{r}}_{\mathbf{x}s} \in \mathrm{colspan}\{\underline{\mathbf{W}}\}. \tag{11}$$

We will solve this problem in closed form for the case $N_D = 2$.

Theorem 1: Consider the scenario [P1]: the APN is not quantized, the ADCs are quantized at R bits. Assume $N_D = 2$. Then the optimal APN that minimizes the MSE and maximizes the SQNR (subject to optimal MSE) is obtained if all columns of $\mathbf{W}$ are equal to the MMSE beamformer, $\mathbf{R}_\mathbf{x}^{-1}\mathbf{r}_{\mathbf{x}s}$, up to scaling and certain linear transformations.

Proof: To solve (11), we first parametrize $\underline{\mathbf{W}}$ such that the constraint is satisfied. Thus, let

$$\underline{\mathbf{W}} = \mathbf{U}\mathbf{V}^H = \mathbf{u}_1\mathbf{v}_1^H + \mathbf{u}_2\mathbf{v}_2^H$$

where $\mathbf{V}$ is a 2×2 unitary matrix since $N_D = 2$, and $\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2]$ is an $N_r \times 2$ matrix such that

$$\mathbf{u}_1 = \frac{\underline{\mathbf{r}}_{\mathbf{x}s}}{\|\underline{\mathbf{r}}_{\mathbf{x}s}\|}.$$

Further define

$$p = \|\mathbf{u}_2\|, \quad \alpha = \frac{\mathbf{u}_1^H\mathbf{u}_2}{p}.$$

(Note that $|\alpha| \le 1$.) Then

$$\begin{aligned}
\underline{\mathbf{W}}^H\underline{\mathbf{W}} &= \mathbf{V}\begin{bmatrix} 1 & \alpha p \\ \bar{\alpha}p & p^2 \end{bmatrix}\mathbf{V}^H \\
(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1} &= \mathbf{V}\frac{1}{p^2(1-|\alpha|^2)}\begin{bmatrix} p^2 & -\alpha p \\ -\bar{\alpha}p & 1 \end{bmatrix}\mathbf{V}^H \\
\frac{\underline{\mathbf{r}}_{\mathbf{x}s}^H}{\|\underline{\mathbf{r}}_{\mathbf{x}s}\|}\underline{\mathbf{W}} &= \mathbf{u}_1^H\mathbf{U}\mathbf{V}^H = [1, \quad \alpha p]\mathbf{V}^H \\
\frac{\underline{\mathbf{r}}_{\mathbf{x}s}^H}{\|\underline{\mathbf{r}}_{\mathbf{x}s}\|}\underline{\mathbf{W}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1} &= [1, \quad \alpha p]\frac{1}{p^2(1-|\alpha|^2)} \\
&\quad \times \begin{bmatrix} p^2 & -\alpha p \\ -\bar{\alpha}p & 1 \end{bmatrix}\mathbf{V}^H = \mathbf{v}_1^H \\
\mathbf{D}_\mathbf{z} = \mathrm{diag}\{\underline{\mathbf{W}}^H\underline{\mathbf{W}}\} &= \mathrm{diag}\{\mathbf{v}_1\mathbf{v}_1^H + p^2\mathbf{v}_2\mathbf{v}_2^H \\
&\quad + \bar{\alpha}p\mathbf{v}_2\mathbf{v}_1^H + \alpha p\mathbf{v}_1\mathbf{v}_2^H\}.
\end{aligned}$$

Introduce a sufficiently general parametrization (θ, Φ_1, Φ_2) for $\mathbf{V}$ as

$$\begin{aligned}
\mathbf{V} &= \begin{bmatrix} e^{j\Phi_1} & \\ & e^{j\Phi_2} \end{bmatrix}\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\
&=: \begin{bmatrix} \phi_1 & \\ & \phi_2 \end{bmatrix}\begin{bmatrix} c & -s \\ s & c \end{bmatrix}
\end{aligned} \tag{12}$$

where $\Phi_1, \ \Phi_2$, and θ are in the range $(-\pi, \ \pi]$. (A completely general parametrization would also have two complex phase factors at the right, but one phase can be absorbed in $\mathbf{v}_2$, and the other can be extracted to multiply the complete matrix $\mathbf{V}$; the form of the cost function shows that that phase will cancel.) Then

$$\begin{aligned}
\mathbf{D}_\mathbf{z} &= \begin{bmatrix} c^2 + p^2s^2 - \bar{\alpha}psc - \alpha psc & 0 \\ 0 & s^2 + p^2c^2 + \bar{\alpha}psc + \alpha psc \end{bmatrix} \\
&= \begin{bmatrix} c^2 + p^2s^2 - 2\beta psc & 0 \\ 0 & s^2 + p^2c^2 + 2\beta psc \end{bmatrix}
\end{aligned}$$

where $\beta = \mathrm{Re}(\alpha)$; note that $-1 \le \beta \le 1$. The cost function (11) becomes

$$\begin{aligned}
J(\underline{\mathbf{W}}) &= \mathbf{v}_1^H\mathbf{D}_\mathbf{z}\mathbf{v}_1 \\
&= [c \quad s]\begin{bmatrix} c^2 + p^2s^2 - 2\beta psc & \\ & s^2 + p^2c^2 + 2\beta psc \end{bmatrix}\begin{bmatrix} c \\ s \end{bmatrix} \\
&= c^2(c^2 + p^2s^2 - 2\beta psc) + s^2(s^2 + p^2c^2 + 2\beta psc) \\
&= (c^4 + 2p^2s^2c^2 + s^4) + 2\beta p(s^3c - sc^3).
\end{aligned}$$

Since $-1 \le \beta \le 1$, minimizing the cost function will require choosing β at extremes,

$$\beta = -\mathrm{sign}(s^3c - sc^3).$$

Subsequently, optimal values for θ will follow as well, as a function of p. The location of the minima of $J(\underline{\mathbf{W}}) = J(\theta, p)$ is shown in Fig. 3. The value of the minimum is 0.5. Although there are multiple minima, in any case, we will have $|\alpha| = 1$: the correlation coefficient between $\mathbf{u}_1$ and $\mathbf{u}_2$ has absolute value 1, which implies that $\mathbf{u}_1$ and $\mathbf{u}_2$ are equal, up to a scaling and phase rotation. □

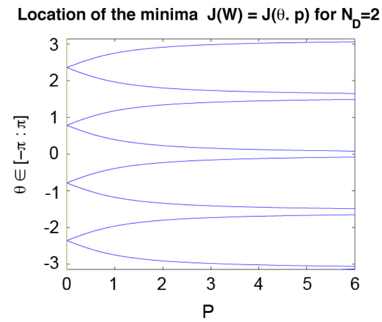


Fig. 3. $J(\underline{\mathbf{W}})$ as a function of $p = \|\mathbf{u}_2\|$ and θ.

In summary, we derived that, given a specific resolution of the ADCs and the number of ADCs, the optimal approach to minimize the MSE and maximize the SQNR is to choose the columns of $\underline{\mathbf{W}}$ all parallel to $\underline{\mathbf{r}}_{\mathbf{x}s}$. Translated to $\mathbf{W}$, it means that each beamformer in the APN is parallel to the Wiener beamformer $\mathbf{R}_\mathbf{x}^{-1}\mathbf{r}_{\mathbf{x}s}$. In actuality, they should differ by at least a phase shift such that the quantization noise on the N_D beamformer outputs becomes uncorrelated. The digital beamformer will simply average the results, i.e., average out the additive noise and quantization noise.

The result was obtained for $N_D = 2$; however, it seems reasonable that it generalizes to larger N_D. Also, the result was obtained for all ADCs having the same resolution $R_i = R$, but the same result will follow also for unequal resolutions. In this case the diagonal terms in $\mathbf{D}_\mathbf{z}$ in $J(\underline{\mathbf{W}})$ will have unequal scaling, but it can be seen that the optimization still leads to $|\beta| = 1$, so that the same conclusion follows.

Finally, let us consider the effect of the APN on the power consumption by the ADCs. It is clear that by inserting an APN, interference cancellation becomes possible, leading to reduced requirements on ADC resolution and hence enabling power reduction. As function of the interference power, the benefits can be arbitrarily large compared to a setup without APN.

Next, we compare $N_D = 1$ to $N_D = 2$, while keeping a constant output MSE after digital beamforming. For the optimal APN, the digital beamformer will simply be averaging the outputs of the ADCs, so that the quantization noise power at the output is halved. Quantization noise of one channel is proportional to 2^{-2R_i}. Thus, for $N_D = 2$ and the same SQNR, each ADC needs half a bit less than for $N_D = 1$. However, power consumption is also proportional to 2^{2R_i}. For two ADCs, each with half a bit less, the total power consumption is constant. Thus, there is no particular advantage to choose $N_D > 1$ from this perspective.

More generally, $N_D > 1$ allows us to use multiple ADCs with lower resolution in situations where high rate high resolution ADCs cease to exist due to fundamental limitations [3].

IV. Preprocessor Design-APN With Discrete Phase Shifts

In the previous section, we did not take the quantization of the APN coefficients into account. In practice, the elements of $\mathbf{W}$ can only be selected from a discrete alphabet, usually only from a set of possible phase shifts. We will now study this case, meanwhile ignoring the quantization of the ADCs, i.e., assuming their resolution is high enough such that it does not dominate the design.

A. Matching the Cross-Correlation Vector

Since the elements of $\mathbf{W}$ are quantized, let $\mathcal{D}$ represent the set of all possible $N_r \times 1$ phase shift vectors that a column in $\mathbf{W}$ can take. The entries of $\mathcal{D}$ are denoted as $\mathcal{D} = \{\boldsymbol{\phi}_m\}_{m=1}^{M}$, where M is the size of the dictionary. If the APN taps are quantized by $R_{\mathbf{W}}$ bits then $M = 2^{N_r R_{\mathbf{W}}}$. Typically, $R_{\mathbf{W}}$ would be 2 to 4 bits. Similarly, call $\mathcal{D}^{N_D}$ the set of all possible APN matrices $\mathbf{W}$.

The APN design is now transformed into a problem of selecting $\mathbf{W} \in \mathcal{D}^{N_D}$ such that the MSE distortion is kept at a minimum,

$$\mathbf{W} = \underset{\boldsymbol{\vartheta}, \mathbf{W} \in \mathcal{D}^{N_D}}{\arg\min} \quad E\|s^{(1)}[k] - \boldsymbol{\vartheta}^H \mathbf{W}^H \mathbf{x}[k]\|^2. \tag{13}$$

Lemma 2: The MMSE beamformer solving (13) is obtained as $\mathbf{W} = \mathbf{R}_{\mathbf{x}}^{-1/2} \underline{\mathbf{W}}$, where

$$\underline{\mathbf{W}} = \underset{\boldsymbol{\vartheta}, \underline{\mathbf{W}} \in \underline{\mathcal{D}}^{N_D}}{\arg\min} \quad \|\underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{W}}\boldsymbol{\vartheta}\|^2 \tag{14}$$

and where $\underline{\mathcal{D}} = \mathbf{R}_{\mathbf{x}}^{1/2}\mathcal{D} = \{\mathbf{R}_{\mathbf{x}}^{1/2}\boldsymbol{\phi}_m\}_{m=1}^{M}$.

Proof: Following (7), solving (13) is equivalent to solving

$$\underline{\mathbf{W}} = \arg \max_{\underline{\mathbf{W}} \in \underline{\mathcal{D}}^{N_D}} \underline{\mathbf{r}}_{\mathbf{x}s}^H \mathbf{P}_{\underline{\mathbf{W}}} \underline{\mathbf{r}}_{\mathbf{x}s} = \arg \max_{\underline{\mathbf{W}} \in \underline{\mathcal{D}}^{N_D}} \|\mathbf{P}_{\underline{\mathbf{W}}} \underline{\mathbf{r}}_{\mathbf{x}s}\|^2.$$

This is equivalent to

$$\begin{aligned} \underline{\mathbf{W}} &= \arg \min_{\underline{\mathbf{W}} \in \underline{\mathcal{D}}^{N_D}} \|(\mathbf{I} - \mathbf{P}_{\underline{\mathbf{W}}})\underline{\mathbf{r}}_{\mathbf{x}s}\|^2 \\ &= \arg \min_{\underline{\mathbf{W}} \in \underline{\mathcal{D}}^{N_D}} \|\underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{W}}(\underline{\mathbf{W}}^H \underline{\mathbf{W}})^{-1} \underline{\mathbf{W}}^H \underline{\mathbf{r}}_{\mathbf{x}s}\|^2 \end{aligned}$$

which is equivalent to (14), since for given $\underline{\mathbf{W}}$ the optimal choice for $\boldsymbol{\vartheta}$ is $\boldsymbol{\vartheta} = \underline{\mathbf{W}}^\dagger \underline{\mathbf{r}}_{\mathbf{x}s} = (\underline{\mathbf{W}}^H \underline{\mathbf{W}})^{-1} \underline{\mathbf{W}}^H \underline{\mathbf{r}}_{\mathbf{x}s}$. □

Thus, the problem becomes to match $\underline{\mathbf{r}}_{\mathbf{x}s}$ in least squares sense to linear combinations of columns of $\underline{\mathbf{W}}$, each of which can assume only values in a discrete set. Equivalently, the columns of $\underline{\mathbf{W}}$ should span a subspace to which $\underline{\mathbf{r}}_{\mathbf{x}s}$ is close. The selection complexity is exponential in N_r, N_D, and $R_{\mathbf{W}}$.

B. Quantized Matching Pursuit

To reduce the complexity, the columns of $\underline{\mathbf{W}}$ are selected one-by-one. The matching pursuit (MP) technique [17] is a *greedy* technique that recursively chooses the dictionary elements to obtain the best approximation of an input vector, in this case $\underline{\mathbf{r}}_{\mathbf{x}s}$. Indeed, write

$$\|\underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{W}}\boldsymbol{\vartheta}\|^2 = \|\underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{w}}_1\vartheta_1 - \underline{\mathbf{w}}_2\vartheta_2 - \cdots - \underline{\mathbf{w}}_{N_D}\vartheta_{N_D}\|^2.$$

In the greedy approach, we first solve

$$\{\vartheta_1, \underline{\mathbf{w}}_1\} = \arg \min_{\vartheta_1, \underline{\mathbf{w}}_1 \in \underline{\mathcal{D}}} \|\underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{w}}_1\vartheta_1\|^2. \tag{15}$$

Given $\underline{\mathbf{w}}_1$, the optimal solution for ϑ_1 is $\vartheta_1 = \underline{\mathbf{w}}_1^\dagger \underline{\mathbf{r}}_{\mathbf{x}s} = (\underline{\mathbf{w}}_1^H \underline{\mathbf{w}}_1)^{-1} \underline{\mathbf{w}}_1^H \underline{\mathbf{r}}_{\mathbf{x}s}$, so that the problem reduces to

$$\underline{\mathbf{w}}_1 = \arg \max_{\underline{\mathbf{w}}_1 \in \underline{\mathcal{D}}} \frac{|\underline{\mathbf{w}}_1^H \underline{\mathbf{r}}_{\mathbf{x}s}|}{\|\underline{\mathbf{w}}_1\|}.$$

The solution requires a search in the dictionary, at a complexity exponential in N_r and $R_{\mathbf{W}}$. To facilitate the search, we can first normalize the vectors in $\underline{\mathcal{D}}$ to unit norm, and then search for the vector with maximal correlation to $\underline{\mathbf{r}}_{\mathbf{x}s}$.

After selecting $\underline{\mathbf{w}}_1$ and ϑ_1, we compute the residual vector $\underline{\mathbf{r}}_{\mathbf{x}s}^{(1)} = \underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{w}}_1\vartheta_1$ and proceed similarly as (15),

$$\begin{aligned} \{\vartheta_2, \underline{\mathbf{w}}_2\} &= \arg \min_{\vartheta_2, \underline{\mathbf{w}}_2 \in \underline{\mathcal{D}}} \|(\underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{w}}_1\vartheta_1) - \underline{\mathbf{w}}_2\vartheta_2\|^2 \\ &= \arg \min_{\vartheta_2, \underline{\mathbf{w}}_2 \in \underline{\mathcal{D}}} \|\underline{\mathbf{r}}_{\mathbf{x}s}^{(1)} - \underline{\mathbf{w}}_2\vartheta_2\|^2 \end{aligned}$$

where

$$\underline{\mathbf{r}}_{\mathbf{x}s}^{(1)} = \underline{\mathbf{r}}_{\mathbf{x}s} - \underline{\mathbf{w}}_1\vartheta_1 = \underline{\mathbf{r}}_{\mathbf{x}s} - \left(\frac{\underline{\mathbf{w}}_1^H \underline{\mathbf{r}}_{\mathbf{x}s}}{\|\underline{\mathbf{w}}_1\|^2}\right) \underline{\mathbf{w}}_1.$$

After selecting $\underline{\mathbf{w}}_2$ and ϑ_2, we could continue the process with the residual vector

$$\underline{\mathbf{r}}_{\mathbf{x}s}^{(2)} = \underline{\mathbf{r}}_{\mathbf{x}s}^{(1)} - \underline{\mathbf{w}}_2\vartheta_2.$$

However, $\underline{\mathbf{w}}_1$ and $\underline{\mathbf{w}}_2$ are not orthogonal, and better coefficients ϑ_1 and ϑ_2 can be computed at this point, leading to a smaller residual. This requires to solve

$$\{\vartheta_1, \vartheta_2\} = \arg \min_{\vartheta_1, \vartheta_2} \left\| \underline{\mathbf{r}}_{\mathbf{x}s} - [\underline{\mathbf{w}}_1, \underline{\mathbf{w}}_2] \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} \right\|. \tag{16}$$

TABLE I
QUANTIZED MATCHING PURSUIT (QMP) ALGORITHM

Objective: Select the discrete phase vectors $\mathbf{W}$ of the APN
Given: Input covariance matrix $\mathbf{R_x}$ and cross-correlation matrix $\mathbf{r}_{\mathbf{x}s}$; dictionary $\mathcal{D} = \{\phi_m\}_{m=1}^{M}$.

- Transform the dictionary vectors into the "whitened" domain:

$$\underline{\mathcal{D}} = \mathbf{R_x}^{1/2}\mathcal{D} = \{\mathbf{R_x}^{1/2}\phi_m\}_{m=1}^{M}$$

- $\underline{\mathbf{W}}^{(0)} = [\cdot]$, $\underline{\mathbf{r}}_{\mathbf{x}s}^{(0)} = \mathbf{R_x}^{-1/2}\mathbf{r}_{\mathbf{x}s}$
- Recursion: for $i = 1, \cdots, N_D$,
  - $\underline{\mathbf{w}}_i = \arg\max_{\underline{\mathbf{w}}_i \in \underline{\mathcal{D}}} \dfrac{|\underline{\mathbf{w}}_i^H \underline{\mathbf{r}}_{\mathbf{x}s}^{(i-1)}|}{\|\underline{\mathbf{w}}_i\|}$
  - $\underline{\mathbf{W}}^{(i)} = [\underline{\mathbf{W}}^{(i-1)} \quad \underline{\mathbf{w}}_i]$
  - Update QR factorization: $\mathbf{Q}^{(i)}\mathbf{R}^{(i)} := \underline{\mathbf{W}}^{(i)}$
  - $\underline{\mathbf{r}}_{\mathbf{x}s}^{(i)} = \underline{\mathbf{r}}_{\mathbf{x}s}^{(i-1)} - \mathbf{Q}^{(i)}\mathbf{Q}^{(i)H}\underline{\mathbf{r}}_{\mathbf{x}s}^{(i-1)}$
- $\mathbf{W} = \mathbf{R_x}^{-1/2}\underline{\mathbf{W}}^{(N_D)}$

Define $\underline{\mathbf{W}}^{(2)} = [\underline{\mathbf{w}}_1, \underline{\mathbf{w}}_2]$ and introduce a QR factorization

$$\underline{\mathbf{W}}^{(2)} = \mathbf{Q}^{(2)}\mathbf{R}^{(2)} \tag{17}$$

where $\mathbf{Q}^{(2)}$ is a $N_r \times 2$ orthonormal matrix and $\mathbf{R}^{(2)}$ a 2×2 upper triangular matrix. The solution to (16) is

$$\begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} = (\mathbf{R}^{(2)})^{-1}\mathbf{Q}^{(2)H}\underline{\mathbf{r}}_{\mathbf{x}s}$$

and the corresponding (smaller) residual is

$$\underline{\mathbf{r}}_{\mathbf{x}s}^{(2)} = \underline{\mathbf{r}}_{\mathbf{x}s}^{(1)} - \mathbf{Q}^{(2)}\mathbf{Q}^{(2)H}\underline{\mathbf{r}}_{\mathbf{x}s}^{(1)} = \mathbf{P}^{\perp}_{\underline{\mathbf{W}}^{(2)}}\underline{\mathbf{r}}_{\mathbf{x}s}^{(1)}$$

which is the projection onto the orthogonal complement of the column span of $\underline{\mathbf{W}}^{(2)}$. The recursion follows in an obvious way. Note that the QR factorization (17) is easily updated once new vectors $\underline{\mathbf{w}}_i$ are added, and that, in fact, it is not necessary to explicitly compute $\boldsymbol{\vartheta}$ at intermediate steps. The algorithm is summarized in Table I.

Further refinements of this algorithm are possible. The update of the QR factorization and the computation of the residual $\underline{\mathbf{r}}_{\mathbf{x}s}^{(i)}$ can be integrated into a single update step. In principle, a better selection of $\underline{\mathbf{w}}_i$ could be obtained by computing the residuals for all possible $\underline{\mathbf{w}}_i$ from the dictionary and selecting the one that gives smallest residual; however, the complexity of this is probably too high.

In practical implementations, the beamforming weights are usually quantized phase shifts with unit amplitude. In this case, it may be more accurate to first split a desired weight vector $\mathbf{w}$ into two weight vectors $\mathbf{w}_1$ and $\mathbf{w}_2$, each with entries on the unit circle, such that a linear combination of them gives the desired $\mathbf{w}$. Zhang *et al.* [10] showed that such a partitioning is always possible. Subsequently, the phase vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ are each quantized into discrete phase shifts, resulting in the "greedy" selection, now consisting of a pair of vectors. The process continues as before to N_D recursions.

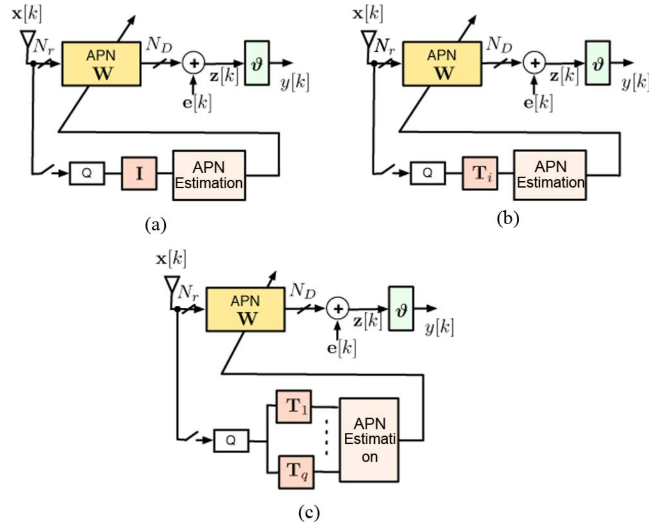


Fig. 4. (a) Architecture 1, where each antenna has its own ADC. (b) Architecture 2, containing time varying beamformers $\mathbf{T}_i$, $i \in \{1, \ldots, p\}$ for different training periods. (c) Architecture 3 with p low resolution beamformers.

In this section, APN has been designed from (13) assuming that the ADC quantization is negligible. It is later shown using simulations in Section VI that for a small $\mathbf{R_W}$ (as is the case in practice) and $R \geq 6$ bits, the MSE is dominated by the quantization of the $\mathbf{W}$.

V. Online Correlation Estimation

The APN phase shift design requires the knowledge of $\mathbf{R_x}$ and $\mathbf{r}_{\mathbf{x}s}$. However, since the digital baseband processor has only access to the beamformed outputs $\mathbf{z}[k]$ and not to the individual antenna signals $\mathbf{x}[k]$, it is not possible to directly compute these correlations from the available observations (and, in the case of $\mathbf{r}_{\mathbf{x}s}$, a training signal $s_1[k]$). In the context of direction of arrival estimation, a similar problem was studied by Sheinvald *et al.* [14] and Tabrikian *et al.* [21].

Regarding system architectures, there are a number of options as enumerated in Fig. 4.

1) Each antenna has its own ADC, operating at a low rate and low resolution. In this way, full (but noisy) information is available and estimates of $\mathbf{R_x}$, $\mathbf{r}_{\mathbf{x}s}$ can be computed straightforwardly.
2) During a training phase, a range of beamformers $\mathbf{W} = \mathbf{T}_1, \ldots, \mathbf{T}_q$ are applied instead of the optimal $\mathbf{W}$, resulting in output sequences $\mathbf{l}_i[k] = Q(\mathbf{T}_i^H\mathbf{x}[k])$, for $i = 1, \ldots, q$. Here $\mathbf{T}_i$ is a $N_r \times p$ matrix with $p > 1$ and $\mathbf{l}_i[k]$ is a $p \times 1$ vector denoting the beamformer output during the training phase. The corresponding correlation statistics of $\mathbf{l}_1[k], \ldots, \mathbf{l}_q[k]$ are computed and the statistics of $\mathbf{x}[k]$ are inferred, as discussed below. Note that the automatic gain controls may have to readjust because the interfering signals will not be suppressed during the estimation phase.
3) A separate APN and bank of low-rate/low-resolution ADCs is used to monitor the inputs, resulting in output sequences $\mathbf{l}_i[k] = Q(\mathbf{T}_i^H\mathbf{x}[k])$, for $i = 1, \ldots, q$. We refer to this setup as low resolution beamformers (LRB's) and the quantization is typically with 1–2 bit ADCs. This is

quite similar to case 2, except that there is more flexibility in the number of APN outputs (can be different than N_D) and ADC resolutions. The APN could simply be a set of switches, making a selection of the antennas towards a low number of ADCs.

The choice of system architecture depends on various criteria, such as avaliable space for additional hardware, the stationarity of the received signals, and the duration and density of the training periods in the desired signal. Regarding the training signal, there are also issues related to acquisition and synchronization to the desired signal; we will not discuss this further.

Without loss of generality, we will consider case 3 and discuss how $\mathbf{R_x}$ and $\mathbf{r_{x}}_s$ are inferred. Let $p > 1$ be the number of APN outputs (dimension of each $\mathbf{l}_i[k]$). From the LRB output sequences $\mathbf{l}_i[k] = Q(\mathbf{T}_i^H \mathbf{x}[k])$, we will be able to form estimates of $\mathbf{R}_\mathbf{l}^{(i)}$, of size $p \times p$, with model

$$\mathbf{R}_\mathbf{l}^{(i)} = \mathbf{T}_i^H \mathbf{R_x} \mathbf{T}_i, \quad i = 1, \ldots, q.$$

As in [14], we subsequently stack the columns of each of these matrices into vectors $\mathrm{vec}(\mathbf{R}_\mathbf{l}^{(i)})$, with model

$$\mathrm{vec}(\mathbf{R}_\mathbf{l}^{(i)}) = (\bar{\mathbf{T}}_i \otimes \mathbf{T}_i)^H \mathrm{vec}(\mathbf{R_x})$$

where the identity $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$ was used. Stacking these vectors results in the model

$$\begin{bmatrix} \mathrm{vec}(\mathbf{R}_\mathbf{l}^{(1)}) \\ \vdots \\ \mathrm{vec}(\mathbf{R}_\mathbf{l}^{(q)}) \end{bmatrix} = \begin{bmatrix} (\bar{\mathbf{T}}_1 \otimes \mathbf{T}_1)^H \\ \vdots \\ (\bar{\mathbf{T}}_q \otimes \mathbf{T}_q)^H \end{bmatrix} \mathrm{vec}(\mathbf{R_x})$$
$$\Leftrightarrow \quad \mathbf{r}_\mathbf{l}^{\mathrm{ef}} = \mathbf{T}_{\mathrm{ef}} \mathbf{r_x}. \tag{18}$$

Assuming $\mathbf{T}_{\mathrm{ef}}$ has a left inverse, we can estimate the data covariance matrix using Least Squares[2] as $\mathbf{R_x} = \mathrm{vec}^{-1}(\mathbf{T}_{\mathrm{ef}}^{\dagger} \mathbf{r}_\mathbf{l}^{\mathrm{ef}})$. The complexity of this step is in the order of $q^2 p^4 N_r^2$.

The Hermitian property of $\mathbf{R_x}$ can be exploited by introducing a vectorization operator "$\mathrm{vech}(\mathbf{R_x})$" that separately stacks the real and imaginary components of the upper triangular (respectively, strictly upper triangular) part of its argument. Similar to [14], [22], this refinement reduces the computational complexity and ensures that the resulting estimate is Hermitian.

A necessary condition for the left invertibility of $\mathbf{T}_{\mathrm{ef}}$ is that this is a "tall" matrix, or $qp^2 \geq N_r^2$. Once this condition is satisfied, simple designs for the $\mathbf{T}_i$ are already sufficient to obtain invertibility. E.g., for $N_r = 4$, $p = 2$, $q = 4$, matrices of the form

$$\mathbf{T}_i = \begin{bmatrix} 1 & 0 \\ w_{i,1} & 0 \\ 0 & 1 \\ 0 & w_{i,2} \end{bmatrix}, \quad w_{i,1} \neq 1, \quad w_{i,2} \neq 1,$$

with distinct $w_{i,1}$ and $w_{i,2}$, lead to a full rank $\mathbf{T}_{\mathrm{ef}}$. Such low complexity selection matrices have also been used for DOA applications in [21].

[2]The paper [14] proposes to use a weighted Least Squares, but it can be shown that, for this unparametrized estimate of $\mathbf{R_x}$, the weight does not change anything.

If only switches are used, then (for $p = 2$) each $\mathbf{T}_i$ gives access to one cross-correlation entry in $\mathbf{R_x}$. For $N_r = 4$ there are 6 such entries, and a minimal design is ($q = 6$)

$$\mathbf{T}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{T}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{T}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},$$
$$\mathbf{T}_4 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{T}_5 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{T}_6 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The vector $\mathbf{r_{x}}_s$ can be estimated in a similar way from estimates of $\mathbf{r}_{\mathbf{l}s}^{(i)} := E(\mathbf{l}^{(i)}[k]\bar{s}^{(1)}[k])$ via the model equations

$$\begin{bmatrix} \mathbf{r}_{\mathbf{l}s}^{(1)} \\ \vdots \\ \mathbf{r}_{\mathbf{l}s}^{(q)} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_1^H \\ \vdots \\ \mathbf{T}_q^H \end{bmatrix} \mathbf{r_{x}}_s.$$

This requires the matrix in the RHS to be tall ($qp \geq N_r$) and full column rank, which is a milder condition than what we had for the estimate of $\mathbf{R_x}$. As mentioned, we need the desired user to be synchronized to the receiver and the receiver must have knowledge of the training sequence transmitted at the start of the packet.

VI. Simulation Results

To assess the performance of the proposed algorithms, we have applied it to a multiuser/antenna setup and computer generated data. We present simulation results that incorporate the impairments, discrete design and channel parameter estimation as covered in Sections III–V.

The input SNR is the signal to noise power ratio for the desired signal and the noise as received at antenna 1; it is the same for all antennas. The input SIR is the signal to interference power ratio for the desired signal and the sum of all interference signals as received at antenna 1; it is the same for all antennas. All users transmit QPSK signals, with zero mean and unit variance as assumed in Section II-B and the interferers have equal powers. The performance indicators are usually as follows:

1) SINR at the first ADC input—a high SINR indicates that less power is spent in quantizing the interferers for a given ADC resolution;
2) MSE—observed at the output of the digital receiver.

All results are obtained by averaging 100 Monte Carlo runs, each with independent Rayleigh fading channel realizations and independently generated data signals. Each run transmits data packages of size 8192 symbols, as in a WLAN transmission packet. The QMP algorithm we proposed in Table I and Section IV is used to design the APN weights from the training sequence. Unless specified otherwise, we used $N_t = 4$ transmitters, $N_r = 4$ receive antennas, and $N_D = 2$ ADC receiver chains. The ADC resolution is $R = 10$ bits and transmit SIR is -5 dB. The APN is represented by a dictionary with $R_\mathbf{W} = 4$ bits. In the cases where the receiver is based on estimated channel coefficients, these are estimated from a training sequence of length 256 symbols incorporated in the data packet.

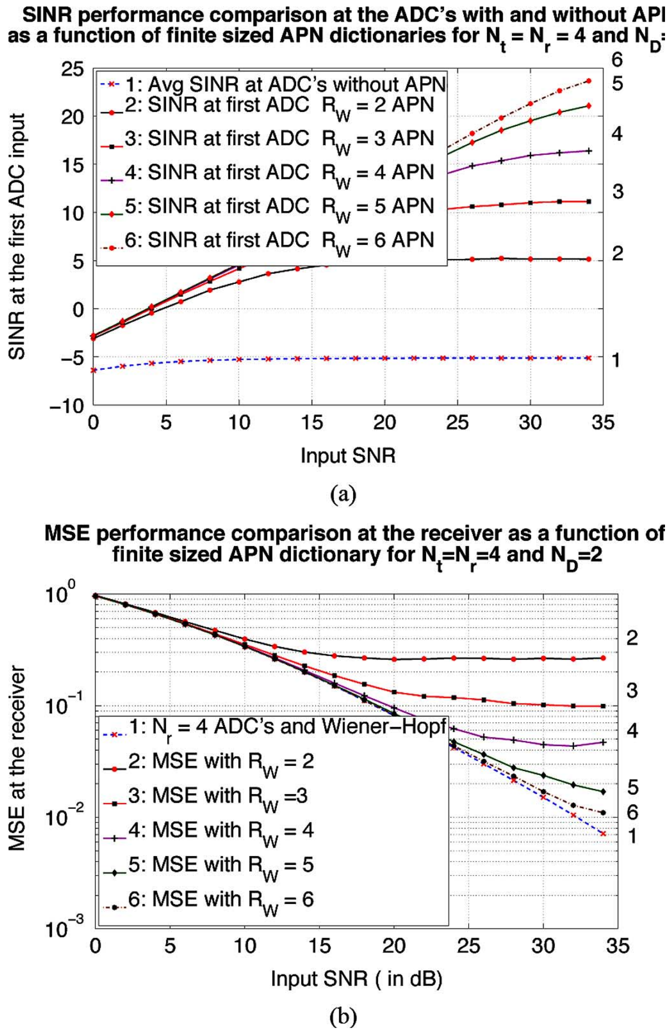


Fig. 5. Performance comparison of the APN setup for different $R_{\mathbf{W}}$ and $R = 10$ as function of transmit SNR: (a) SINR at the input of first ADC; (b) MSE at the output of baseband receiver.

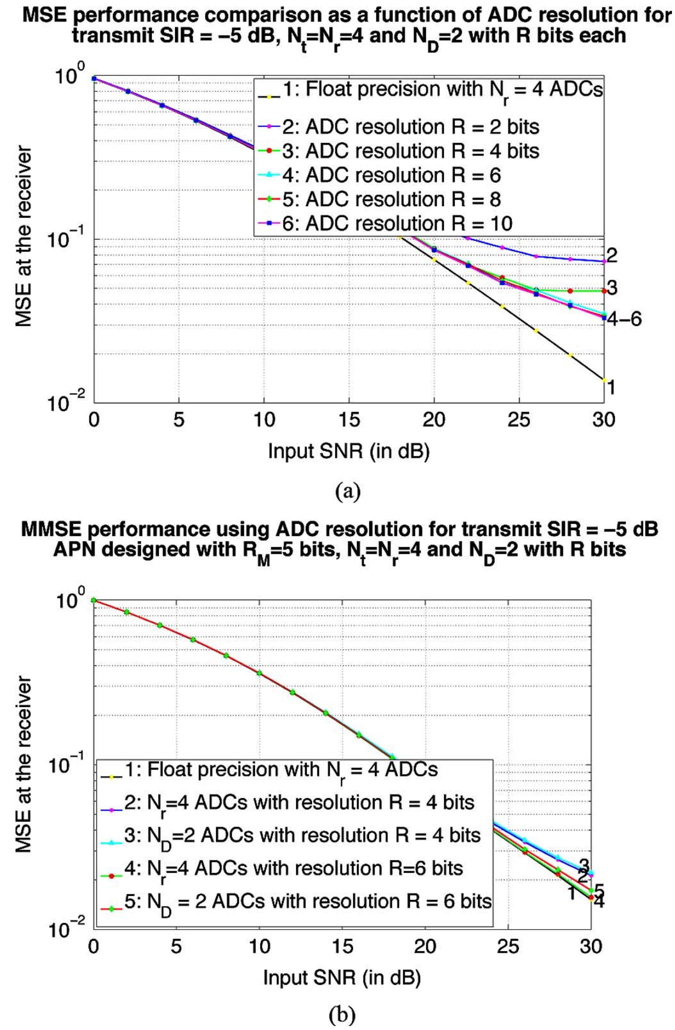


Fig. 6. MSE performance comparison at the output of the baseband receiver as a function of transmit SNR (a) for varied R and (b) for various numbers and resolutions of ADCs.

A. Finite Sized APN Dictionaries

Fig. 5(a) and (b) shows the SINR at the input of the first ADC and the MSE at the receiver output respectively. A training sequence of length 256 is used for the design of the APN. We show curves for varying dictionary size $R_{\mathbf{W}}$, and fixed ADC resolution of $R = 10$ bits. Consider the SINR plot [Fig. 5(a)], curve 1 corresponds to a case with no APN and $N_r = 4$ ADCs operating with float precision and curves 2–6 show the performance of the APN setup for increasing $R_{\mathbf{W}}$. Comparing curves 1–4, the results show that the introduction of the APN with dictionary size $R_{\mathbf{W}} = 4$ bits improves the SINR at the first ADC input up to a factor of 20 dB. For increasing SNR, the performance saturates: it is limited by the residual interference power. The performance can be further improved by increasing $R_{\mathbf{W}}$. Consider Fig. 5(b) comparing the MSE at the receiver. For an MSE 0.05, the setup with $R_{\mathbf{W}} = 4$ bits and $N_D = 2$ ADCs each with $R = 10$ bits (curve 4) performs 2 dB worse than the optimal Wiener beamformer with N_r ADCs and float precision (curve 1).

B. Effect of the ADC Resolution

Fig. 6(a) and (b) shows the MSE performance at the receiver for a similar setup as in Section VI-A, where the APN resolution is $R_{\mathbf{W}} = 4$ bits. In Fig. 6(a), curve 1 corresponds to a case with $N_r = 4$ ADCs, with float precision and curves 2–5 correspond to $N_D = 2$ ADCs and varying ADC resolution $R_i = R$ bits. The transmit SIR is -5 dB. We observe that for $R \geq 6$ bits, MSE curves overlap and the finite precision APN leads to an error floor. In other words, for higher resolution ADCs, the APN resolution is the limiting factor.

Fig. 6(b) gives an idea of ADC power savings with the introduction of an APN, for an APN designed with $R_{\mathbf{W}} = 5$ bits and varying ADC resolution R. Comparing curves 2–3 and 4–5, respectively, we see that the introduction of an APN with $N_D = 2$ ADCs operating with $R = 4$ and $R = 6$ bits results in a similar MSE values as that of a receiver without APN, $N_r = 4$ ADCs with precision $R = 4$ and $R = 6$ bits followed by optimal Wiener beamformer. Since ADC power consumption is related to $\sum_{i=1}^{N_D} 2^{2R_i}$, this suggests that the use of an APN can reduce the ADC power consumption by half.

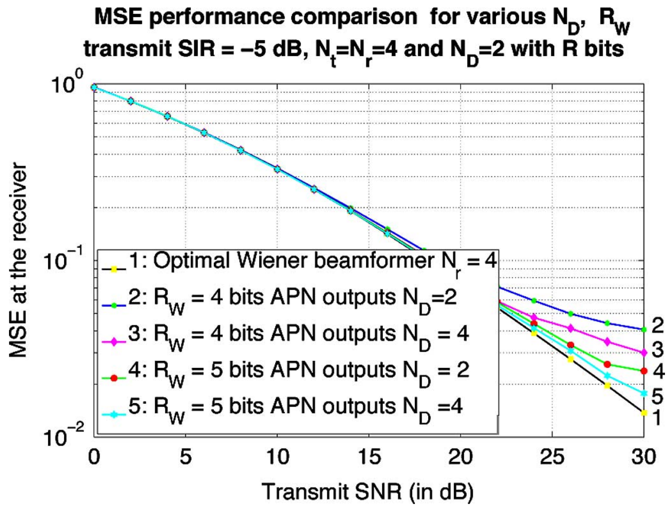


Fig. 7. MSE performance comparison at the output of the baseband receiver for different N_D.

C. Effect of the Number of APN Outputs

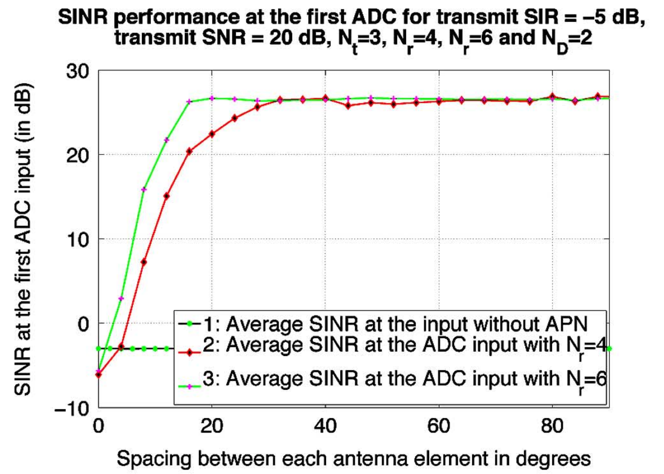
Fig. 7 shows the MSE performance at the output of the baseband receiver for a similar setup as in Section VI-A, where the number N_D of ADCs is varied, $R_{\mathbf{W}} = 4$–5 bits and ADC resolution $R = 10\,\text{bits}$. Typically, for APN with perfect interference cancellation, $\text{MSE} \propto 1/N_D$. However, fixed precision APN and interfering users limit the performance gains. From curve 3, we see that even with $N_D = 4$, the MSE curves lead to an error floor and this suggests that the APN with $R_{\mathbf{W}} = 4$ might be ill-conditioned. Increasing the APN Resolution to $R_{\mathbf{W}} = 5$ bits leads to improved MSE performance as is obvious from curves 4 and 5. However, to limit the APN circuit size it is suggested to choose $R_{\mathbf{W}} = 4$ and $N_D = N_r/2$.

D. Effect of Source Spacing

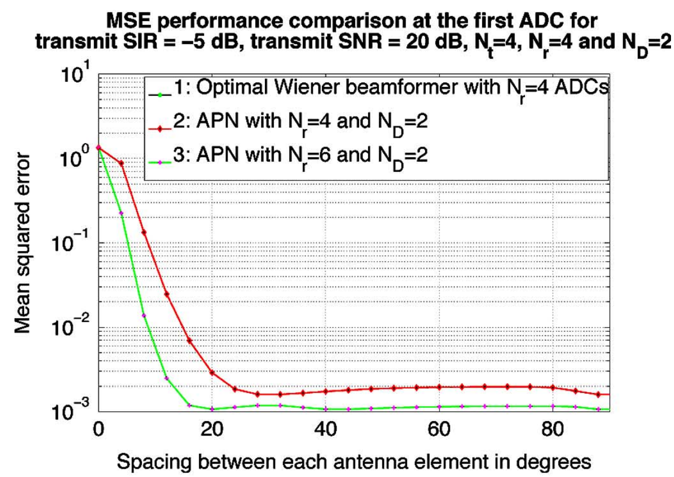
Fig. 8(a) and (b) shows the SINR and MSE performance as a function of the spacing between two adjacent sources. The simulations consider line of sight simulations without multipath, and results are observed for $N_t = N_r = 3$ with the desired user transmitting from an angle say $\theta_1 = 0^\circ$. The MSE is a function of transmit $\text{SNR} = 20$ dB, transmit $\text{SIR} = -5$ dB, and angular spacing θ between the desired user and interferers. We consider two interferers equidistant from the desired user and in opposite directions transmitting from angles $\theta_2 = \theta - \theta_1$ and $\theta_3 = \theta_1 - \theta$. From Fig. 8(a) we see that for $\theta > 20^\circ$, the APN improves the SINR at the first ADC by a factor of 15 dB. In both the cases, we see that the APN performs poorly for angular spacing $< 5^\circ$, and the performance can be improved by increasing the number N_r of the antennas.

E. Communication Setup and Channel Estimation With LRB's

The previous sections have given indications on the improvements in SINR, power consumption and MSE for phase shifter based APN as functions of $R_{\mathbf{W}}$ and R_i. Here we focus on channel estimation using LRBs as specified in Section V. We select the architecture type 3 in Section V and choose $q = 6$ LRBs with $p = 2$ outputs. As specified in that section, switches $\mathbf{T}_1, \ldots, \mathbf{T}_6$ are used to estimate $\mathbf{R}_{\mathbf{x}}$.



(a)



(b)

Fig. 8. Performance comparison for $N_D = 2$ setup as a function of spacing between desired user and two interferers: (a) SINR at the input of the first ADC; (b) MSE at the output of the baseband receiver.

Fig. 9(a) and (b) compares the MSE and BER performance of the fixed precision $R_{\mathbf{W}} = 4$, $N_D = 2$ APN estimated using LRBs as a function of training lengths. The results are compared with the reference $N_r = 4$ ADC case and beamformer designed using true channel parameters. The ADC resolution is kept at $R = 6$ bits.

In Fig. 9(a), the curves 1 and 3 show that going from $N_r = 4$ to $N_D = 2$ ADCs leads to performance degradation of 2 dB at MSE of 0.05. For the sake of completeness, we also show the BER performance between the transmitted and received QPSK signals and see that curves 1 and 3 in Fig. 9(b) show that going from $N_r = 4$ to $N_D = 2$ ADCs APN leads to performance degradation of 2 dB at BER 10^{-3}.

Considering that the above setup reduces the interferer contributions at the ADCs, these results suggest that architectures with reduced RF chains, limited by RF imperfections can perform close to theoretical MIMO while consuming a fraction of the power.

VII. Concluding Remarks

In this paper, we have proposed a MIMO receiver employing an analog preprocessing network (APN) or multichannel

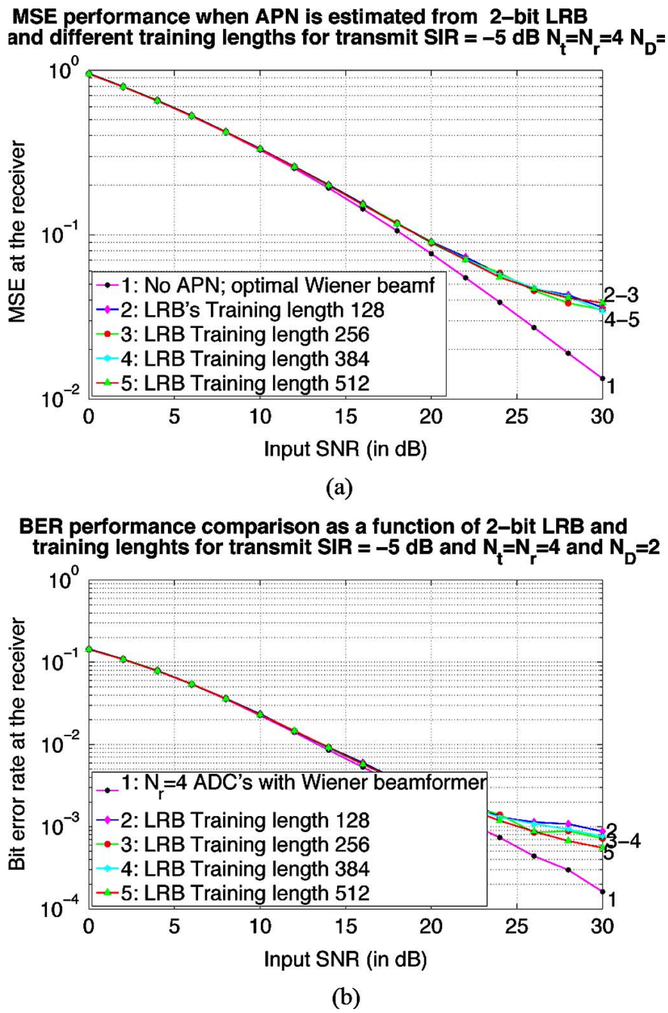


Fig. 9. Performance comparison, when $\mathbf{R_x}$ and $\mathbf{r}_{\mathbf{x}s}$ are estimated for 2-bit LRBs and varied channel lengths: (a) MSE; (b) BER.

beamformer in the RF domain, followed by a digital beamformer in baseband. The prime advantage of this architecture is that it reduces the number of antenna elements to a smaller number of mixers and ADC chains. Further, it can reduce the interference at the inputs of the ADC, so that less dynamic range and fewer bits are required. Overall, significant power savings are possible.

An optimal preprocessor to minimize the MSE and maximize the desired user SQNR at the digital baseband was derived. It was shown that, if the APN quantization is very fine, it is sufficient to consider only $N_D = 1$ analog beamforming output. In practice the quantization is poor, and a larger number of outputs is required such that the cross-correlation vector $\underline{\mathbf{r}}_{\mathbf{x}s}$ is well approximated. Further research is needed in the following directions.

- In practice, the APN coefficients have poor accuracy; implementation errors of up to 7% of phase have been reported [23]. This affects both the channel estimation and the APN design. How can this effect be modeled and incorporated into the design?
- Initially, we are not synchronized to the source of interest. It may then be complicated to estimate $\mathbf{r}_{\mathbf{x}s}$ and design the APN; subsequently, the interference may overwhelm the ADCs and make acquisition impossible. What is a good initialization strategy?

An alternative to using an APN to reduce ADC power, is to exploit spatial and temporal oversampling with a predictive Sigma-Delta ADC as in [24].

Appendix
APN Design Using Cross Spectral Projections

To obtain some intuition on the APN design problem [P0] as specified in Section III-A, we first consider a few suboptimal techniques before deriving a closed form APN in Section III-B. Introduce an eigenvalue decomposition of $\mathbf{R_x}$:

$$\mathbf{R_x} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$$

where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_{N_r}]$ is an $N_r \times N_r$ unitary matrix containing the eigenvectors, and $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{R_x}$, sorted from large to small. $\mathbf{W}$ can be chosen as the N_D dominant eigenvectors, corresponding to the N_D largest eigenvalues, $\mathbf{W} = [\mathbf{u}_1, \ldots, \mathbf{u}_{N_D}]$. This is somewhat similar to one of the approaches proposed in [10]. In this way, computation of $\mathbf{z}[k] = \mathbf{W}^H\mathbf{x}[k]$ retains the components with the largest energy and drops the components with less power. However, this does not distinguish between desired and interfering users. If the interferers are strong, the desired user could be projected out.

This is avoided in the following design, based on "cross spectral projections." Instead of selecting N_D dominant eigenvectors, selecting the eigenvectors that contain a large correlation with the desired user array response given by $\mathbf{r}_{\mathbf{x}s}$ results in a better approximation. This approach is similar to one technique in [25] where the authors sort the basis vectors based on the *cross spectral norm*, defined as

$$\mathbf{c}_{\mathbf{xs}} = \boldsymbol{\Lambda}^{-1}\mathbf{U}^H\mathbf{r}_{\mathbf{x}s}.$$

More precisely, let $\hat{s}^{(1)}[k] = \boldsymbol{\theta}^H\mathbf{x}[k]$ with $\boldsymbol{\theta} = \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{r}_{\mathbf{x}s}$. Then we can write

$$\boldsymbol{\theta} = \mathbf{U}\boldsymbol{\Lambda}^{-1}\mathbf{U}^H\mathbf{r}_{\mathbf{x}s} = \mathbf{U}\mathbf{c}_{\mathbf{x}s}$$

and obtain the energy of the output signal

$$E\{|\hat{s}^{(1)}[k]|^2\} = \boldsymbol{\theta}^H\mathbf{R_x}\boldsymbol{\theta} = \mathbf{c}_{\mathbf{xs}}^H\boldsymbol{\Lambda}\mathbf{c}_{\mathbf{x}s} = \sum_{j=1}^{N_r}|c_j|^2\lambda_j$$

where $\mathbf{c}_{\mathbf{x}s} = [c_1, \ldots, c_{N_r}]^T$ and λ_j is the j-th eigenvalue in $\boldsymbol{\Lambda}$. Selecting the N_D columns of $\mathbf{U}$ corresponding to the largest (weighted) cross spectral norm terms, i.e., the N_D largest terms among

$$|c_j|^2\lambda_j = \frac{|\mathbf{u}_j^H\mathbf{r}_{\mathbf{x}s}|^2}{\lambda_j}, \quad j = 1, \ldots, N_r$$

would lead to the "best" $N_D \times 1$ representation of the desired user signal, in the sense of maximizing the output energy of the desired user.

Collect the selected N_D columns of $\mathbf{U}$ into a matrix $\mathbf{U}'$, and likewise for $\mathbf{c}'_{\mathbf{x}s}$. Although it seems natural to choose $\mathbf{W} = \mathbf{U}'$ and $\boldsymbol{\vartheta} = \mathbf{c}'_{\mathbf{x}s}$, the above technique does in fact not prescribe a partitioning of $\boldsymbol{\theta}$ into $\mathbf{W}$ and $\boldsymbol{\vartheta}$. Moreover, by selecting columns of $\mathbf{U}$ we have also limited our choice: $\mathbf{x}[k]$ is projected on a subspace of eigenvectors. This is not necessarily optimal.

It will be shown later in Lemma 1 that the APN leading to MMSE has to satisfy the condition $\mathbf{r}_{\mathbf{x}s} \in \text{col span}\{\mathbf{R}_{\mathbf{x}}\mathbf{W}\}$. Choosing $\mathbf{W}$ equal to the dominant N_D eigenvectors of $\mathbf{R}_{\mathbf{x}}$ is optimal if $\mathbf{r}_{\mathbf{x}s}$ is in this subspace, which occurs only if there are at most N_D sources in white noise. This also implies that, for the same scenario, the solution from *cross spectral projections*, where $\mathbf{W} = \mathbf{U}'$ consists of eigenvectors maximizing the cross spectral norm would be optimal only if the dominant N_D eigenvectors are selected.

ACKNOWLEDGMENT

The authors would like to thank Prof. J.P. Linnartz, Prof. P. Baltus, and J. H. C. van den Heuvel for stimulating discussions especially for detailing the possibility of using passive weighting elements in RF, and Prof. A. Leshem for his contribution to the problem formulation.

REFERENCES

[1] V. Venkateswaran and A. J. van der Veen, "Optimal phase shifter design to cancel RF interference in multi-antenna systems," in *In Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Dallas, TX, Mar. 2010.
[2] 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.
[3] R. Walden, "Analog to digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, pp. 539–550, Apr. 1999.
[4] B. Murmann, "Digitally assisted analog circuits," *IEEE Micro*, vol. 26, no. 2, pp. 38–47, Feb. 2006.
[5] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.
[6] S. Denno and T. Ohira, "Modified constant modulus algorithm for DSP adaptive antennas with microwave analog beamforming," *IEEE Trans. Antennas Propag.*, vol. 50, no. 6, pp. 850–857, Jun. 2002.
[7] P. Baltus and R. Dekker, "Optimizing RF front ends for low power," *Proc. IEEE*, vol. 88, no. 10, pp. 1546–1559, Jan. 2000.
[8] R. Eickhoff, R. Kraemer, I. Santamaria, and L. Gonzalez, "Developing energy-efficient MIMO radios," *IEEE Veh. Technol. Mag.*, vol. 4, no. 1, pp. 34–41, Mar. 2009.
[9] A. Hajimiri, H. Hashemi, A. Natarajan, X. Guang, and A. Babakhani, "Integrated phased arrays in silicon," *Proc. IEEE*, vol. 93, no. 9, pp. 1637–1655, Sep. 2005.
[10] X. Zhang, A. F. Molisch, and S. Y. Kung, "Variable phase shift based RF baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.
[11] M. D. Zoltowski, S. D. Silverstein, and C. P. Mathews, "Beamspace root MUSIC for minimum redundancy linear arrays," *IEEE Trans. Signal Process.*, vol. 41, no. 1993, pp. 2502–2507, Jul. 1993.
[12] B. D. van Veen and R. Roberts, "Partially adaptive beamformer design via output power minimization," *IEEE Trans. Signal Process.*, vol. 11, pp. 1524–1532, Nov. 1987.
[13] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2943–2959, Nov. 1998.
[14] J. Sheinvald and M. Wax, "Direction finding with fewer receivers via time varying preprocessing," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 2–9, Jan. 1999.
[15] S. T. Smith, "Optimum phase-only adaptive nulling," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1835–1843, Jul. 1999.
[16] G. M. Kautz, "Phase-only shaped beam synthesis via technique of approximated beamaddition," *IEEE Trans. Antennas Propag.*, vol. 47, no. 5, pp. 887–894, May 1999.
[17] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
[18] V. K. Goyal, M. Vetterli, and N. T. Thao, "Quantized overcomplete expansions in $\mathcal{R}^N$: Analysis, synthesis and algorithms," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 16–31, Jan. 1998.
[19] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.
[20] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1994.
[21] D. Tabrikian and A. Faizakov, "Optimal preprocessing for source localization by fewer receivers than sensors," in *Proc. 11th IEEE Statistical Processing Workshop*, Aug. 6–8, 2001, pp. 213–216.
[22] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1136–1155, May 1996.
[23] N. Fourikis, *Phased Array-Based Systems and Applications*. New York: Wiley, 1997.
[24] V. Venkateswaran and A. J. van der Veen, "Multichannel Sigma-Delta ADCs with integrated feedback beamformers to cancel interference in MIMO communications," *IEEE Trans. Signal Process.*, Apr. 2010, submitted for publication.
[25] J. S. Goldstein and I. S. Reed, "Reduced rank adaptive filtering," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 492–496, Feb. 1997.

Vijay Venkateswaran received the M.S. degree in electrical engineering from the University of Arizona, Tempe, in 2003. He is currently working towards the Ph.D. degree from the Delft University of Technology, Delft, The Netherlands.

From 2003 to 2005, he worked as a member of technical staff at Sony Corporation, Tokyo, Japan, and in 2002 as an intern at Seagate LLC. His research interests are in the general area of signal processing advancements in wireless communications and data storage, and, in particular, joint RF-baseband interference cancellation techniques.

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

Throughout 1994, he was a postdoctoral scholar at Stanford University, Stanford, CA. Currently, he is a Full Professor in signal processing at TU Delft. His research interests are in the general area of system theory applied to signal processing, and in particular algebraic methods for array signal processing, with applications to wireless communications and radio astronomy.

Dr. van der Veen is the recipient of a 1994 and a 1997 IEEE Signal Processing Society (SPS) Young Author Paper Award. He was an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1998 to 2001, Chairman of IEEE SPS Signal Processing for the Communications Technical Committee from 2002 to 2004, a Member-at-Large of the Board of Governors of the IEEE SPS, Editor-in-Chief of the *IEEE Signal Processing Letters* from 2002 to 2005, and Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He currently is Technical Co-Chair of the International Conference on the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2011 in Prague, a Member of the IEEE SPS Awards Board and Fellow Reference Committee, and the IEEE TAB Periodicals Review and Advisory Committee.