FEEDBACK BEAMFORMER DESIGN WITH OVERSAMPLING ADCS IN MULTI-ANTENNA SYSTEMS

Vijay Venkateswaran and Alle-Jan van der Veen

TU Delft, Fac. EEMCS, Mekelweg 4, 2628 CD Delft, Netherlands

This paper designs a feedback beamformer (FBB) operating in combination with multiple low resolution ADCs of a MIMO communication system, to cancel the interfering users. Interference cancellation before the ADC operation saves the power spent in digitizing the interferers. We specify the necessary conditions required to cancel the interferers and propose a FBB design technique using only the training information from the desired user. We show using simulation results, that the FBB improves the dynamic range by a factor of three.

Keywords: *multi-antenna systems, oversampling ADCs, source separation, linear prediction*.

1. INTRODUCTION

ADCs in multi-antenna systems operating in a dense multi-user environment spend most of their power and dynamic range (DR) in quantizing the interferers. One well known sub-optimal approach to reduce the ADC power in MIMO systems is to select antennas with largest signal energies [1]. However such techniques do not fully exploit the advantages of MIMO, are inherently sub-optimal, fail in the presence of multiple users and cannot pilot for variations in the wireless channel.

An alternative approach is to use ADCs sampling at a rate much higher than the Nyquist rate (*oversampling ADCs*), identify and feedback the interferer using a digital to analog converter (DAC), followed by interference cancellation at the ADC input [2]. Cancelling interference through feedback allows the use of low resolution ADCs and reduces the power consumption.

Set-up: Consider a multi-user setup, with the desired and interfering users transmitting over a common wireless channel and received using antenna array as in Fig. 1. In [2], the authors introduced a feedback beamformer (FBB) operating on a bank of oversampled first order $\Sigma\Delta$ ADCs [3] connected to the antenna array, to cancel the interfering users. They assumed knowledge of the interfering user signals, and designed a first order FBB to partially cancel the interference.

This research was supported in part by the Dutch Min. Econ. Affairs under the IOP-Gencom program (IGC0502B)

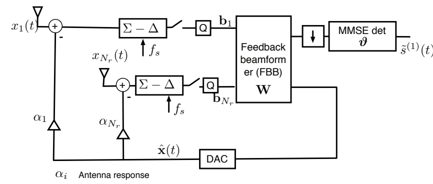

Fig. 1: Proposed ADC architecture with feedback cancellation

Contributions: In this paper, we generalize the FBB design for an arbitrary order $\Sigma\Delta$ ADC. We aim to cancel the interfering user signals, such that the ADC output is an estimate of only the desired user signals. To specify the conditions of interference cancellation, we utilize the regressive structure of the $\Sigma\Delta$ ADC and combine the transfer function of the wireless channel with that of the multi-channel (MC) ADCs. We propose a block adaptive FBB design technique utilizing only the training signals from the desired user. Simulation results show a threefold improvement in the DR along the direction of desired user for a narrowband(NB) scenario.

Notation: $\bar{(.)}$, $(.)^T$, $(.)^H$ and $\|.\|$ represent conjugate, transpose, Hermitian and Frobenius norm. Multi-channel vectors and matrices are represented as *underline bold* letters. Continuous time and sampled signals are indexed as (.) and [.] respectively. $\mathbf{I}$ denotes identity matrix, $\mathbf{0}_K$ represent $K \times 1$ vectors of zeros.

2. $\Sigma\Delta$ MODULATION FOR ANTENNA ARRAYS

2.1. Data model

Assume for simplicity, a $N_t = 2$ user setup as in Fig. 1. Let user 1 be the desired user and user 2 be the interfering user, transmitting $s^{(1)}(t)$ and $s^{(2)}(t)$ respectively, with $\mathcal{N}(0, \sigma_s^2)$. The transmit signals are assumed to be band limited (BL) by frequency f_0 in the observation interval $t \in (0 : T]$. Usually T corresponds to the duration of a transmission packet. Let $M = \lfloor 2f_0T \rfloor$ corresponds to the number of samples in $(0 : T]$, when sampled at the Nyquist rate. $s^{(j)}(t)$ can be described by a $M \times 1$ vector $\mathbf{s}^{(j)} = [s^{(j)}[1], \cdots, s^{(j)}[M]]^T$ and in this case $\mathbf{s}^{(1)}$ has M degrees of freedom, such that $\mathbf{s} = [\mathbf{s}^{(1)T}, \mathbf{s}^{(2)T}]^T$ is a $2M \times 1$ vector.

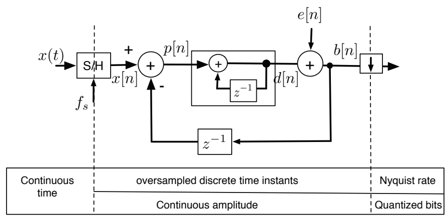

Fig. 2: Classical first order $\Sigma\Delta$ ADC. For details refer [2, 3].

<u>*Multi-antenna data model*</u>: Consider an array of N_r antennas receiving the BL transmitted signals. The $N_r \times 1$ antenna array vector $\mathbf{x}[n] = [x_1[n], \cdots x_{N_r}[n]]^T$ for oversampling instant (OSI) $n \in \{1, \cdots, N\}$ can be modeled as

$$\begin{aligned}\mathbf{x}[n] &= \begin{bmatrix} \mathbf{g}_1^{(1)}[n]^T & \mathbf{g}_1^{(2)}[n]^T \\ \vdots & \vdots \\ \mathbf{g}_{N_r}^{(1)}[n]^T & \mathbf{g}_{N_r}^{(2)}[n]^T \end{bmatrix} \mathbf{s} + \begin{bmatrix} v_1[n] \\ \vdots \\ v_{N_r}[n] \end{bmatrix} \\ &= \mathbf{G}[n]\mathbf{s} + \mathbf{v}[n] \end{aligned} \tag{1}$$

where $\mathbf{G}[n]$ is a $N_r \times MN_t$ matrix denoting the MIMO wireless channel at OSI n, $\mathbf{g}_i^{(j)}[n]$ is a $M \times 1$ vector connecting the antenna i with the user j and $\mathbf{v}[n]$ is a $N_r \times 1$ vector denoting the thermal noise at the antenna. Stack (1) for OSI's $n \in \{1, \cdots, N\}$ as

$$\begin{aligned}\begin{bmatrix} \mathbf{x}[1] \\ \vdots \\ \mathbf{x}[N] \end{bmatrix} &= \begin{bmatrix} \mathbf{G}[1] \\ \vdots \\ \mathbf{G}[N] \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{v}[1] \\ \vdots \\ \mathbf{v}[N] \end{bmatrix} \\ \underline{\mathbf{x}} &= \underline{\mathbf{G}}\mathbf{s} + \underline{\mathbf{v}} \end{aligned} \tag{2}$$

Here $\underline{\mathbf{x}}$ is a $NN_r \times 1$ vector containing signals from N_r antennas and N OSI, $\underline{\mathbf{G}}$ is a $NN_r \times MN_t$ tall matrix and $\underline{\mathbf{v}}$ is a $NN_r \times 1$ thermal noise vector.

2.2. Multi-channel $\Sigma\Delta$ model

From the data model relation (2), we see that the sequence $\underline{\mathbf{x}}$ is an interpolation of the BL sequence $\mathbf{s}$. In a standard $\Sigma\Delta$ ADC setup of first order as in Fig. 2, N_r first order $\Sigma\Delta$ ADCs sample and quantize $\mathbf{x}[n]$ to obtain a $N_r \times 1$ vector $\mathbf{b}[n]$ for $n \in \{1, \cdots, N\}$. Consider our setup as in Fig. 3 and [2], where the ADC output of the previous time instants is fed back via a K^{th} order FBB, denoted by a $KN_r \times N_r$ matrix $\mathbf{W} = [\mathbf{W}^T[1], \cdots, \mathbf{W}^T[K]]^T$.

The input signal $\mathbf{x}[n]$ is predicted using the FBB arrangement as a $N_r \times 1$ vector $\hat{\mathbf{x}}[n] = \mathbf{W}^H \underline{\mathbf{b}}_K[n-1]$, where $\underline{\mathbf{b}}_K[n-1] = [\mathbf{b}[n-K], \cdots, \mathbf{b}[n-1]]$ is a $KN_r \times 1$ vector. The ADC output is

$$\mathbf{b}[n] + \sum_{j=1}^{n} \mathbf{W}^H \underline{\mathbf{b}}_K[j-1] = \sum_{j=1}^{n} \mathbf{x}[j] + \mathbf{e}[n] \tag{3}$$

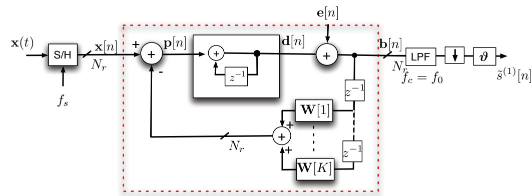

Fig. 3: Interference cancellation using a K^{th} order multi-channel $\Sigma\Delta$ ADC, followed by digital combiner $\boldsymbol{\vartheta}$.

where $\mathbf{e}[n]$ is a $N_r \times 1$ vector denoting the additive quantization noise. The ADC output stacked for $n \in \{1, \cdots, N\}$ as in [2], and for $K = 1$ is

$$\underline{\mathbf{B}}\,\underline{\mathbf{b}} = \underline{\mathbf{L}}_1 \underline{\mathbf{x}} + \underline{\mathbf{e}} \qquad \text{where} \tag{4}$$

$$\underline{\mathbf{B}} = \begin{bmatrix} \mathbf{I}_{N_r} & & & \\ \mathbf{W}^H \mathbf{I}_{N_r} & & & \\ \vdots & \ddots & \mathbf{I}_{N_r} & \\ \mathbf{W}^H & \cdots & \mathbf{W}^H \mathbf{I}_{N_r} & \end{bmatrix}, \quad \underline{\mathbf{L}}_1 = \begin{bmatrix} \mathbf{I}_{N_r} & & & \\ \mathbf{I}_{N_r} \mathbf{I}_{N_r} & & & \\ \vdots & \ddots & \mathbf{I}_{N_r} & \\ \mathbf{I}_{N_r} & \cdots & \mathbf{I}_{N_r} \mathbf{I}_{N_r} & \end{bmatrix}$$

where $\underline{\mathbf{b}}$ is a $NN_r \times 1$ vector denoting the $\Sigma\Delta$ ADC output, and elements of $\underline{\mathbf{b}}$ are ± 1, $\underline{\mathbf{e}}$ is a $NN_r \times 1$ vector denoting the quantization noise [3], uncorrelated with $\underline{\mathbf{x}}$. $\underline{\mathbf{L}}_1$ and $\underline{\mathbf{B}}$ are a $NN_r \times NN_r$ lower triangular (LT) block Toeplitz matrices. We refer to the above setup (4) as multi-channel (MC) $\Sigma\Delta$ ADCs.

Combining the data model relation (2) with the above MC $\Sigma\Delta$ (4) relation leads to

$$\begin{aligned}\underline{\mathbf{B}} \begin{bmatrix} \mathbf{b}[1] \\ \vdots \\ \mathbf{b}[N] \end{bmatrix} &= \begin{bmatrix} \mathbf{I}_{N_r} & & \\ & \ddots & \\ \mathbf{I}_{N_r} & \cdots & \mathbf{I}_{N_r} \end{bmatrix} \begin{bmatrix} \mathbf{G}[1] \\ \vdots \\ \mathbf{G}[N] \end{bmatrix} \mathbf{s} + \begin{bmatrix} \tilde{\mathbf{e}}[1] \\ \vdots \\ \tilde{\mathbf{e}}[N] \end{bmatrix} \\ \underline{\mathbf{B}}\,\underline{\mathbf{b}} &= \underline{\mathbf{L}}_1 \underline{\mathbf{G}}\mathbf{s} + \underline{\tilde{\mathbf{e}}} \end{aligned} \tag{5}$$

where $\underline{\tilde{\mathbf{e}}} = \underline{\mathbf{L}}_1 \underline{\mathbf{v}} + \underline{\mathbf{e}}$, $\underline{\mathbf{G}}$ is a tall matrix with full column rank and (5) can also be seen a block Toeplitz wireless channel $\underline{\mathbf{L}}_1\,\underline{\mathbf{G}}$ operating on BL $\mathbf{s}$. This architecture is somewhat similar to [4], where the authors design the feedback matrix to minimize the quantization noise.

<u>*Digital source separation*</u>: Since the columns of $\underline{\mathbf{L}}_1$ are not orthogonal to the columns of $\underline{\mathbf{G}}$, $\underline{\mathbf{B}}\,\underline{\mathbf{b}}$ can be seen as a linear combination of $\mathbf{s}$. Thus the MC ADC model (5) is a moving average process operating on the transmitted signals. From the array processing literature [5], $\mathbf{s}^{(1)}$ can be estimated if $\underline{\mathbf{L}}_1\,\underline{\mathbf{G}}$ in (5) is invertible.

3. $\Sigma\Delta$ ADC INTERFERENCE CANCELLATION

However our main objective is to exploit the regressive structure and perform interference cancellation before the ADCs. This section will show that the $\Sigma\Delta$ ADC specified by (4) and (5) satisfies the necessary conditions for perfect reconstruction (in a noiseless case) of the desired user signals.

The $\Sigma\Delta$ ADC operation as in Fig. 3 can also be specified by computing the $N_r \times 1$ vectors respectively denoting the prediction error $\mathbf{p}[n]$ and the ADC output $\mathbf{b}[n]$ as

$$\mathbf{p}[n]{=}\mathbf{x}[n]-\mathbf{W}^H\mathbf{b}_K[n-1], \ \& \ \mathbf{b}[n]{=}Q\{\sum_{j=1}^{n}\mathbf{p}[j]\}. \quad (6)$$

The computation of $\mathbf{p}[n]$ in (6) can be related to digital baseband equalization techniques such as *prediction error methods* (PEM) [6]. PEM techniques, however minimize only $\mathbf{p}[n]$, do not perform interference cancellation and are not designed to reconstruct the desired signals.

The main question now is the design of $\mathbf{W}$ to perform interference cancellation. This leads to

- a more faithful representation of the desired user signals for a specified resolution.
- $\mathbf{W}$ designed using the MMSE criteria serves to decorrelate the signal and noise terms.

For simplicity, let $\underline{\mathbf{L}} = \underline{\mathbf{B}}^\dagger \, \underline{\mathbf{L}}_1$. Now $\underline{\mathbf{L}}$ is a function of $\mathbf{W}$ and the MC data model can be rewritten as

$$\underline{\mathbf{b}} = \underline{\mathbf{L}}\,\underline{\mathbf{G}}\mathbf{s} + \underline{\tilde{\varepsilon}} \ \text{ where } \underline{\tilde{\varepsilon}} = \underline{\mathbf{B}}^\dagger\, \underline{\tilde{\mathbf{e}}}.$$

The FBB input $\mathbf{b}_K[n-1]$ in (6) can be modeled as

$$\mathbf{b}_K[n-1] = \mathcal{H}_K[n-1]\mathbf{s} + \tilde{\varepsilon}_K[n-1], \quad (7)$$

where $\tilde{\varepsilon}_K[n-1] = [\tilde{\varepsilon}[n-K], \cdots, \tilde{\varepsilon}[n-1]]^T$ and

$$\mathcal{L}_K[n-1]{=}\begin{bmatrix} \mathbf{L}_{n-K,\,1}\cdots\mathbf{L}_{n-K,\,n-K} & & \\ \vdots \quad \ddots \quad\quad \ddots & & \\ \mathbf{L}_{n-1,\,1}\cdots \quad\quad \cdots & \mathbf{L}_{n-1,\,n-1}\mathbf{0} & \end{bmatrix}\underline{\mathbf{G}}.$$

The ADC output $\mathbf{b}[n]$ from (7) is combined in the digital baseband using a $N_r \times 1$ MSE based fitting vector $\boldsymbol{\vartheta}$, to estimate the desired user signal $\tilde{s}^{(1)}[n] = \boldsymbol{\vartheta}_0^H\mathbf{b}[n]$:

$$\boldsymbol{\vartheta}_0 = \min_{\boldsymbol{\vartheta}} E\|s^{(1)}[n] - \boldsymbol{\vartheta}^H\mathbf{b}[n]\|^2. \quad (8)$$

For simplicity, we ignore the down-sampling operation and represent $s^{(1)}[n]$ for $n \in \{1, \cdots, N\}$.

3.1. Conditions for Interference cancellation

Proposition 1 Consider a MC $\Sigma\Delta$ ADC input $\mathbf{x}[n]$ containing contributions from N_t BL users. There exists $KN_r \times N_r$ filtering matrix $\mathbf{W}$ with $KN_r \geq MN_t$, operating on $\mathbf{b}_K[n-1]$ to compute $\mathbf{b}[n]$ such that:
In the absence of thermal and quantization noise, $s^{(1)}[n]$ can be perfectly reconstructed using a $N_r \times 1$ vector $\boldsymbol{\vartheta}$ operating on $\mathbf{b}[n]$, $\forall n \in \{1, \cdots, N\}$.

Proof. $\underline{\mathbf{L}}$ is a square matrix with linearly independent columns, As long as the columns of $\underline{\mathbf{L}}_1$ are not orthogonal to the columns of $\underline{\mathbf{B}}^\dagger$, $\mathcal{L}_K[n-1]^T$ from (7) is of full column rank KN_r.
$\underline{\mathbf{G}}$ is tall with rank MN_t. If $n > K$ and $N_rK \geq MN_t$ then $\mathcal{H}_K[n-1] = \mathcal{L}_K[n-1]\underline{\mathbf{G}}$ is tall with full column rank.
For $\underline{\tilde{\varepsilon}} = \mathbf{0}$, $\mathbf{x}[n]$ has MN_t degrees of freedom:

$$\mathbf{x}[n] \in \text{span}\{\mathbf{b}[n], \cdots, \mathbf{b}[1]\} \quad \text{From (2)}$$

From (4) and $KN_r \geq MN_t$:

$$\text{span}\{\mathcal{H}_K[n-1]\} \in \{\mathbf{b}_K[n-1]\}$$

Thus we can design $\mathbf{W}$ operating on $\mathbf{b}_K[n-1]$ such that

$$\text{span}\{\mathbf{x}[n] - \mathbf{W}^H\mathbf{b}_K[n-1]\} \in \{\mathbf{b}[n]\} \quad (9)$$

This leads to $\mathbf{p}[n] = \mathbf{x}[n] - \mathbf{W}^H\mathbf{b}_K[n-1]$ as a $N_r \times 1$ innovations vector

$$E\{\mathbf{p}[n]\mathbf{b}^H[n-j]\} = \mathbf{0}$$

and satisfies the *normal equations* or orthogonality conditions for $j \in \{1, \cdots, K\}$ as required for optimal linear prediction [7]. The ADC output $\mathbf{b}[n]$ is combined using a $N_r \times 1$ vector $\boldsymbol{\vartheta}$ to perfectly reconstruct $s^{(1)}[n]$. $\square$

This approach is not limited to $\Sigma\Delta$ ADC, and the signals $\mathbf{p}[n]$ can be quantized directly as $\mathbf{b}[n] = Q\{\mathbf{p}[n]\}$. Since the $\Sigma\Delta$ ADC architecture is especially compatible with our setup, it reduces questions on implementation. The step-wise operations are detailed below.

Objective: Interference cancellation with MC $\Sigma\Delta$ ADC
Step 1: Given: Input signal $\underline{\mathbf{x}}$

- Initialize: $\mathbf{b}[0] = \mathbf{0}$
- For OSI $n = 1$ to N and antennas $i = 1$ to N_r
  - $\mathbf{p}[n] = \mathbf{x}[n] - \mathbf{W}^H\mathbf{b}_K[n-1]$
  - $\Sigma\Delta$ modulator output $\mathbf{b}[n] = Q\{\sum_{j=1}^{n}\mathbf{p}[j]\}$

4. $\Sigma\Delta$ PREDICTIVE BEAMFORMER DESIGN

This section deals with the design of $\mathbf{W}$ to minimize the MSE. We use a first order $\Sigma\Delta$ ADC i.e. $K = 1$ and $\mathbf{W}$ operates on $\mathbf{b}[n-1]$. We assume the following

A1 The transmitted signals $\mathbf{s}$ are uncorrelated with the thermal & quantization noise.

A2 At the start of transmission, a small sequence of data signals known to the receiver, and referred to commonly as training signals are transmitted.

Estimate $\mathbf{W}$ to minimize D, given

$$\mathbf{b}[n] = [\mathbf{I},\ -\mathbf{W}] \begin{bmatrix} \mathbf{x}[n] \\ \mathbf{b}[n-1] \end{bmatrix} + \tilde{\varepsilon}[n] \tag{10}$$

where $\tilde{\varepsilon}[n] = \sum_{j=1}^{n-1} \mathbf{p}[j] + \mathbf{e}[n]$ is a $N_r \times 1$ vector. $\mathbf{e}[n]$ is the $N_r \times 1$ quantization noise vector and from Proposition 1, $\tilde{\varepsilon}[n]$ uncorrelated to $\mathbf{p}[n]$.

Express $\boldsymbol{\vartheta}$ in terms of $\mathbf{W}$: The overall MSE D is minimized, with $\boldsymbol{\vartheta}$ designed using the Wiener-Hopf equation [7] as $\boldsymbol{\vartheta} = \mathbf{R}_{\underline{\mathbf{b}}}^{-1}\mathbf{r}_{\mathbf{bs}}$, where $\mathbf{R}_{\underline{\mathbf{b}}} = E\{\mathbf{b}[n]\mathbf{b}^H[n]\}$ and $\mathbf{r}_{\underline{\mathbf{bs}}} = E\{\mathbf{b}[n]\bar{s}^{(1)}[n]\}$. $\mathbf{r}_{\underline{\mathbf{bs}}}$ is computed from training signals. Representing $\mathbf{R}_{\underline{\mathbf{b}}}$ and $\mathbf{r}_{\underline{\mathbf{bs}}}$ as functions of $\mathbf{W}$, with the notation $\hat{\mathbf{b}}[n] = \mathbf{b}[n-1]$

$$\begin{aligned} \mathbf{R}_{\underline{\mathbf{b}}} &= \begin{bmatrix} \mathbf{I}_{N_r} & -\mathbf{W}^H \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\underline{\mathbf{xx}}} & \mathbf{R}_{\underline{\mathbf{x}\hat{\mathbf{b}}}} \\ \mathbf{R}_{\underline{\hat{\mathbf{b}}\mathbf{x}}} & \mathbf{R}_{\underline{\hat{\mathbf{b}}\hat{\mathbf{b}}}} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{N_r} \\ -\mathbf{W} \end{bmatrix} + \mathbf{R}_{\underline{\tilde{\varepsilon}}} \\ &= \begin{bmatrix} \mathbf{I}_{N_r} & -\mathbf{W}^H \end{bmatrix} \mathbf{R}_{\mathcal{X}} \begin{bmatrix} \mathbf{I}_{N_r} \\ -\mathbf{W} \end{bmatrix} + \mathbf{R}_{\underline{\tilde{\varepsilon}}} \end{aligned} \tag{11}$$

where $\mathbf{R}_{\mathcal{X}}$ is a $(K+1)N_r \times (K+1)N_r$ matrix and $\mathbf{r}_{\mathcal{X}\mathbf{s}} = E\{[\mathbf{x}^T[n],\ \hat{\mathbf{b}}^T[n]]^T\bar{s}^{(1)}[n]\}$ is a $2N_r \times 1$ cross-correlation vector. For $K = 1$ $\mathbf{R}_{\mathcal{X}}$ is a $2N_r \times 2N_r$ square matrix and partitioned as

$$\begin{aligned} \mathbf{R}_{\underline{\mathbf{xx}}} &= E\{\mathbf{x}[n]\mathbf{x}^H[n]\} \;\; \mathbf{R}_{\underline{\mathbf{x}\hat{\mathbf{b}}}} = E\{\mathbf{x}[n]\hat{\mathbf{b}}^H[n]\} \\ \mathbf{R}_{\underline{\tilde{\varepsilon}}} &= E\{\tilde{\varepsilon}[n]\tilde{\varepsilon}^H[n]\} \;\; \mathbf{r}_{\underline{\mathbf{bs}}} = [\mathbf{I}_{N_r} - \mathbf{W}]^H\mathbf{r}_{\mathcal{X}\mathbf{s}} \end{aligned} \tag{12}$$

Minimize the overall D: Inserting the relations (11) and (12) for $\mathbf{R}_{\underline{\mathbf{b}}}$ and $\mathbf{r}_{\underline{\mathbf{bs}}}$ in the overall MSE D, and for simplicity let $\mathbf{W}_{\Delta} = \begin{bmatrix} \mathbf{I}_{N_r} & -\mathbf{W} \end{bmatrix}^T$ be a $2N_r \times N_r$ matrix. The design objective to minimize D is transformed as:

$$\begin{aligned} \mathbf{W}_{\Delta} &= \min_{\mathbf{W}_{\Delta}} E\|s^{(1)}[n] - \mathbf{r}_{\underline{\mathbf{bs}}}^H\mathbf{R}_{\underline{\mathbf{b}}}^{-1}\mathbf{b}[n]\|^2 \\ &= \min_{\mathbf{W}_{\Delta}} \sigma_s^2 - \mathbf{r_{bs}}^H\mathbf{R}_{\underline{\mathbf{b}}}^{-1}\mathbf{r}_{\underline{\mathbf{bs}}} \qquad (13) \\ &= \max_{\mathbf{W}_{\Delta}} \mathbf{r}_{\mathcal{X}\mathbf{s}}^H\mathbf{W}_{\Delta}\Big[\mathbf{W}_{\Delta}^H\mathbf{R}_{\mathcal{X}}\mathbf{W}_{\Delta} + \mathbf{R}_{\underline{\tilde{\varepsilon}}}\Big]^{-1}\mathbf{W}_{\Delta}^H\mathbf{r}_{\mathcal{X}\mathbf{s}} \end{aligned}$$

(14) designs $\mathbf{W}_{\Delta}$ or $\mathbf{W}$ to minimize the MSE between $s^{(1)}[n]$ and its estimate $\tilde{s}^{(1)}[n]$. Assuming that $\mathbf{R}_{\underline{\tilde{\varepsilon}}}$ is independent of $\mathbf{R}_{\mathcal{X}}$, the cost function (8) leads to

$$\begin{aligned} \mathbf{W}_{\Delta} &= \max_{\mathbf{W}_{\Delta}} \mathbf{r}_{\mathcal{X}\mathbf{s}}^H\mathbf{W}_{\Delta}(\mathbf{W}_{\Delta}^H\mathbf{R}_{\mathcal{X}}\mathbf{W}_{\Delta})^{-1}\mathbf{W}_{\Delta}^H\mathbf{r}_{\mathcal{X}\mathbf{s}} \\ &= \max_{\mathbf{W}_{\Delta}} \mathbf{r}_{\mathcal{X}\mathbf{s}}^H\mathbf{P}_{\mathbf{W}_{\Delta}}\mathbf{r}_{\mathcal{X}\mathbf{s}} \end{aligned} \tag{14}$$

where $\mathbf{P}_{\mathbf{W}_{\Delta}} = \mathbf{W}_{\Delta}(\mathbf{W}_{\Delta}^H\mathbf{R}_{\mathcal{X}}\mathbf{W}_{\Delta})^{-1}\mathbf{W}_{\Delta}^H$. $\mathbf{P}_{\mathbf{W}_{\Delta}}$ can also be seen as a projection matrix obtained using the spanning vectors of $\mathbf{W}_{\Delta}$. As a footnote, $\mathbf{R}_{\mathcal{X}}$ in $\mathbf{P}_{\mathbf{W}_{\Delta}}$ can be cancelled with a pre whitener: $\underline{\mathbf{W}}_{\Delta} = \mathbf{R}_{\mathcal{X}}^{1/2}\mathbf{W}_{\Delta}$.

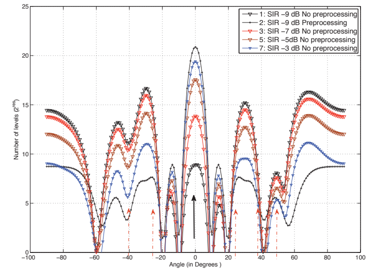

Fig. 4: DR as function of AOA and SNR 20 dB

Due to space constraints, we conclude our design choosing the columns of $\mathbf{W}_{\Delta}$ as $\mathbf{r}_{\mathcal{X}\mathbf{s}}$.

5. SIMULATION RESULTS

To observe the performance of the MC feedback setup and the design algorithm, we consider a $N_r = 6$, $N_t = 6$ NB setup. All users transmit i.i.d sequences over a set of random Rayleigh fading channels. At the start of packet, a training sequence of length $M = 256$ is transmitted. Fig. 4 shows the DR improvement after the introduction of the MC $\Sigma\Delta$ as a function of angle of arrival (AOA). The desired source is at 0° and the interferers at $[30^\circ, -30^\circ, 45^\circ, -50^\circ\ 60^\circ]$ and $\frac{N}{M} = 128$. Curve 2 corresponds to no preprocessing i.e. $\mathbf{W} = \mathbf{I}$ in (6). Comparing the curves 2 with 3, we observe that with the introduction of MC $\Sigma\Delta$ FBB, the ADC gain/DR at 0° improves nearly by a factor of 3.

6. REFERENCES

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

[2] 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.

[3] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data converters: Theory design and simulation*, Wiley Interscience, 1997.

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

[5] H. Krim and M. Viberg, "Two decades of array signal processing research," *Signal Processing Magazine*, vol. 45, no. 4, pp. 69–94, Jan 1996.

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

[7] L. L. Scharf, *Statistical Signal Processing*, Reading, MA: Addison-Wesley, 1991.