OPTIMAL PHASE-SHIFTER DESIGN TO CANCEL RF INTERFERENCE IN MULTI-ANTENNA SYSTEMS

Vijay Venkateswaran and Alle-Jan van der Veen

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

This paper introduces an analog preprocessing network (APN) operating in radio frequency (RF), to cancel signals from interferers in an antenna array system. Interference cancellation facilitates the use of low resolution ADCs. For a given ADC resolution, we will propose the optimal beamformer to minimize the overall mean squared error (MSE) between the desired user and its received estimate, while maximizing the desired signal to quantization noise ratio (SQNR). Subsequently, propose a matching pursuit design technique to represent the analytical APN transform as a linear combination of implementable phase shifters.
Keywords: *Analog beamforming, RF phase shifters, ADC power consumption, Matching pursuit*.

1. INTRODUCTION

In existing wireless receivers, the power consumed in an ADC operation is equivalent to that of hundreds and thousands of logic gates [1]. The ADC power can be approximated as $P_{ADC} \propto f_s 2^{2res}$, where f_s is the sampling frequency and res is the ADC bits to quantize desired user and interferers. In a multi-user scenario with strong interferers, the ADCs sample unwanted interferers, leading to increased power consumption. RF interference cancellation allows use of low resolution ADCs, reducing the power consumption.

One well known sub-optimal technique to reduce the RF and ADC power is to select the antennas with the highest energy [2, 3]. Current hardware offers possibilities using phase shift networks [4]. However these techniques only select the strongest signal(s). They do not take multi-user interference into account, do not consider minimizing the overall MSE, do not factor the ADC quantization, and do not account for the hardware limitations.

Setup: These factors motivate us to introduce an analog preprocessing network (APN), operating on the RF signals, to cancel the interfering users as shown in Fig.1. This APN can be designed as passive phase shifts similar to [4] and maps N_r antennas to N_D ADCs. Typically $N_r = 4 \rightarrow N_D = 2$. In practice, the RF preprocessors are coarsely quantized [5] and further corrupted by $5-7\%$ phase errors at $2-3$ GHz.

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

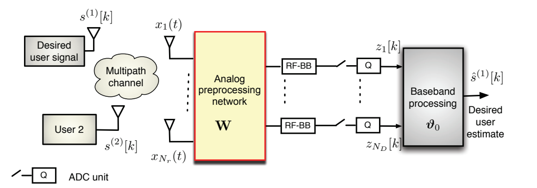

Fig. 1: Proposed setup with APN transform $N_r = 4 \rightarrow N_D = 2$

Contributions: In this paper, we propose a $N_r \times N_D$ preprocessing matrix with focus on interference cancellation. Somewhat similar work in the context of beamspace array processing was done in [6], where the authors have proposed a digital preprocessor and a heuristic/iterative design technique. Though well presented, the approach fails to express the preprocessor in closed form. In this paper, we derive an analytical APN to minimize the overall MSE while maximizing SQNR at the receiver. To implement such architectures in practice, we propose a quantized version of matching pursuit approximation [7]. In summary, the APN can be seen as a *sparse* representation of the desired user and digitally combined to reconstruct the high resolution desired user signals.

Notation: $*$, $(.)^T$, $(.)^H$, $(.)^\dagger$ and $\|.\|^2$ represent convolution, transpose, Hermitian, pseudo inverse and Frobenius norm. Continuous time and sampled signals have time indices represented respectively by $(.)$ and $[.]$.

2. SYSTEM SETUP

2.1. RF data model

Consider an RF signal $x(t)$ received at the antenna and containing multiple rays from N_t users transmitting over a narrow-band (NB) channel. For simplicity, let N_t =2 with user 1 being the desired user transmitting $s^{(1)}(t)$ and user 2 the interferer $s^{(2)}(t)$, both over a carrier frequency f_c. The received signal can be represented as a function of the antenna gain pattern $a^{(l)}(\theta)$ and the channel impulse response $g^{(l)}(t-\tau_l)$ as

$$x(t)= \text{Re}\,\{\sum_{l=1}^{N_t=2} a^{(l)}(\theta) g^{(l)}(t-\tau_l) * \Big[s^{(l)}(t)e^{j2\pi f_c t}\Big] + n(t)\}$$

where $n(t)$ is the noise term independent of $s^{(l)}(t)$. For NB signals, where the channel bandwidth is much less than the operating frequencies, we can approximate $g^{(l)}(t-\tau_l)$ as a constant. The time delayed signal $s^{(l)}(t-\tau_l)e^{j2\pi f_c t}$ can then be replaced by a phase shift $s^{(l)}(t-\tau_l)e^{j2\pi f_c t} \rightarrow e^{j2\pi f_c (t-\tau_l)} s^{(l)}(t)$ and

$$x(t) = \text{Re}\,\{\sum_{l=1}^{N_t=2} a^{(l)}(\theta) e^{j2\pi f_c (t-\tau_l)}\, s^{(l)}(t) + n(t)\}.$$

Extending this for an array of N_r antennas, the $N_r \times 1$ vector $\mathbf{x}(t) = [x_1(t), \cdots, x_{N_r}(t)]^T$ is

$$\mathbf{x}(t)= \text{Re}\,\{e^{j2\pi f_c t}\Big[\mathbf{a}^{(1)}(\theta)\ \mathbf{a}^{(2)}(\theta)\Big]\mathbf{s}(t) + \mathbf{n}(t)\}, \quad (1)$$

$\mathbf{a}^{(l)}(\theta) = e^{-j2\pi f_c \tau_l}[a_1^{(l)}(\theta), \cdots, a_{N_r}^{(l)}(\theta)]^T$, $\mathbf{s}(t) = [s^{(1)}(t),\ s^{(2)}(t)]^T$ and $\mathbf{n}(t)$ is a $N_r \times 1$ noise vector.

From (1), we see that the received signal is a phase shift modulation of the transmitted user signals. A $N_r \times 1$ vector $\mathbf{w} = [e^{j\phi_1}, \cdots e^{j\phi_{N_r}}]^T$ can be designed with its elements of $\phi_i \in [-\pi, \pi]$, to combine the antenna array signals and steer towards the desired user. E.g. $\mathbf{w} = \frac{\mathbf{a}^{(1)}(\theta)}{\|\mathbf{a}^{(1)}(\theta)\|^2}$ such that $\|\mathbf{w}^H \mathbf{a}^{(1)}(\theta)\|^2 = 1$.

2.2. APN setup

Consider a $N_r \times N_D$ APN $\mathbf{W} = [\mathbf{w}_1, \cdots \mathbf{w}_{N_D}]$ operating on $\mathbf{x}(t)$ as in Fig. 1, with $\mathbf{W}$ having fewer outputs than inputs ($N_r > N_D$). The APN transforms $\mathbf{x}(t)$ to N_D signals $\tilde{z}_i(t) = \mathbf{w}_i^H \mathbf{x}(t)\ \forall\, i \in \{1, \cdots, N_D\}$ or $\tilde{\mathbf{z}}(t) = \mathbf{W}^H \mathbf{x}(t)$. The RF signal $\tilde{\mathbf{z}}(t)$ is down-converted to baseband as $\mathbf{z}(t) = e^{-j2\pi f_c t}\tilde{\mathbf{z}}(t)$. For details of APN implementation using RC circuits, refer [4, 5].

For simplicity, we restrict the representations to equivalent baseband (BB). The APN output is digitized into a $N_D \times 1$ vector $\mathbf{z}[k] = Q\{\mathbf{W}^H \mathbf{x}(t)\}_{t=kT}$ using N_D ADCs with resolution R_i, $i \in \{1, \cdots N_D\}$. Here $Q\{.\}$ denotes the sample and quantize operation. These N_D signals outputs are combined digitally a fitting vector $\boldsymbol{\vartheta} = [\vartheta_1, \cdots, \vartheta_{N_D}]^T$ to estimate the desired user $\hat{s}^{(1)}[k] = \boldsymbol{\vartheta}_0^H \mathbf{z}[k]$ and to minimize overall MSE

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

Problem formulation: The main question is to design $\mathbf{z}[k] = Q\{\mathbf{W}^H \mathbf{x}(t)\}$ using the APN $\mathbf{W}$. Following that, the design of $\boldsymbol{\vartheta}$ to minimize the overall MSE is specified by the Wiener-Hopf solution of (2): $\boldsymbol{\vartheta}_0 = \mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{zs}}$, where $\mathbf{R}_{\mathbf{z}} = E\{\mathbf{z}[k]\mathbf{z}^H[k]\}$ and $\mathbf{r}_{\mathbf{zs}} = E\{\mathbf{z}[k]\bar{s}^{(1)}[k]\}$. The APN design to minimize D involves the following constraints:

- A1 The APN circuits consist of a limited number of phase shift combinations (the elements of $\mathbf{W}$ can be selected from a dictionary $\mathcal{D}$ of size $4 \sim 16$).
- A2 Each ADC performs coarse quantization $Q\{.\}$ with a fixed R_i: $z_i[k] = Q\{\mathbf{w}_i^H \mathbf{x}(t)\}$.

We approach the APN design in the following order:

- P1 We initially relax [A1], and assume a continuous APN. Can we represent the overall MSE D as a function of R_i and $\mathbf{W}$: $D_0 = D(\mathbf{W}, R_i)$?
- P2 Given R_i and D_0 obtained from [P1], select $\mathbf{w}_i \in \mathcal{D}$, to minimize (2).

The design techniques P1 and P2 are the core of this paper and are covered respectively in Sec. 3 and 4.

3. PREPROCESSOR DESIGN

This section assumes a continuous APN [P1] and we will

- Proposition 1 - Express D in terms of $\mathbf{W}$ and propose the *range* of $\mathbf{W}$ minimizing D.
- Theorem 1 - Incorporate additional design criterion to maximize the SQNR to compute a unique $\mathbf{W}$.

The design techniques assume knowledge of a $N_r \times N_r$ covariance matrix $\mathbf{R}_{\mathbf{x}} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\}$ and a $N_r \times 1$ cross covariance vector $\mathbf{r}_{\mathbf{xs}} = E\{\mathbf{x}(t)\bar{s}^{(1)}(t)\}$. In [8], we explain an *online* technique to estimate $\mathbf{R}_{\mathbf{x}}$ and $\mathbf{r}_{\mathbf{xs}}$ from a set of 1-bit low resolution beamformers.

3.1. APN design to minimize the MSE

Define the "whitened" correlation matrices

$$\underline{\mathbf{r}}_{\mathbf{xs}} = \mathbf{R}_{\mathbf{x}}^{-1/2}\mathbf{r}_{\mathbf{x}s}\,, \quad \underline{\mathbf{W}} = \mathbf{R}_{\mathbf{x}}^{1/2}\mathbf{W}\,. \quad (3)$$

Proposition 1: Define the orthogonal projection matrix

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

Assume for the time being, that the quantization noise is negligible. For any $\mathbf{W}$, and the corresponding optimal $\boldsymbol{\vartheta}_0$, the MMSE solution $\mathbf{W}_0$ satisfies

$$\underline{\mathbf{W}}_0 = \arg\max_{\underline{\mathbf{W}}}\ \underline{\mathbf{r}}_{\mathbf{xs}}^H \mathbf{P}_{\underline{\mathbf{W}}} \underline{\mathbf{r}}_{\mathbf{xs}} \quad (4)$$

Proof. Given that $\mathbf{z}[k] = \mathbf{W}^H\mathbf{x}(t)|_{t=kT} + \mathbf{e}[k]$ and $\mathbf{e}[k] = 0$. We also know that $\boldsymbol{\vartheta}_0 = \mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}_{\mathbf{zs}}$, where

$\mathbf{R_z} = \mathbf{W}^H\mathbf{R_x}\mathbf{W}$ and $\mathbf{r_{zs}} = \mathbf{W}^H\mathbf{r_{xs}}$.

$$\begin{aligned} D &= E\|s^{(1)}[k] - \mathbf{r}_{\mathbf{z}s}^H\mathbf{R_z}^{-1}\mathbf{W}^H\mathbf{x}(t)\|^2 \\ &= 1 - \mathbf{r}_{\mathbf{x}s}^H\mathbf{W}(\mathbf{W}^H\mathbf{R_x}\mathbf{W})^{-1}\mathbf{W}^H\mathbf{r}_{\mathbf{x}s} \\ &= 1 - \mathbf{r}_{\mathbf{xs}}^H\mathbf{W}(\mathbf{W}^H\mathbf{R_x}\mathbf{W})^{-1}\mathbf{W}^H\mathbf{r}_{\mathbf{x}s} \\ &= 1 - \underline{\mathbf{r}}_{\mathbf{xs}}^H\mathbf{P}_{\underline{\mathbf{W}}}\underline{\mathbf{r}}_{\mathbf{xs}} \end{aligned} \quad (5)$$

such that $\{\mathbf{W}_0\} = \arg\min_{\mathbf{W}} D = \arg\max_{\underline{\mathbf{W}}}\ \underline{\mathbf{r}}_{\mathbf{xs}}^H\mathbf{P}_{\underline{\mathbf{W}}}\underline{\mathbf{r}}_{\mathbf{xs}}$

□

Clearly, this only specifies that $\underline{\mathbf{r}}_{\mathbf{xs}} \in \text{colspan}\{\underline{\mathbf{W}}\}$, and there can be many solutions. To narrow down the set of the available $\underline{\mathbf{W}}$, we incorporate the ADC resolution in (4) and compute $\underline{\mathbf{W}}$ to maximize the received SQNR.

3.2. APN design to maximize the desired user SQNR

For simplicity, assume that $N_D = 2$, i.e. $\mathbf{z}[k] = [z_1[k],\, z_2[k]]^T$ and $\mathbf{e}[k] = [e_1[k],\, e_2[k]]^T$. The quantization noise variance at ADC 1, $\sigma_{e_1}^2$, depends on $\sigma_{z_1}^2 = E\{z_1[k]\bar{z}_1[k]\}$ and resolution the R_1, expressed using the well known Lloyd-Max equation [9]

$$\sigma_{e_i}^2 = \frac{\sigma_{z_i}^2 2^{-2R_i}}{12} \quad i \in \{1,\, 2\}$$

Let $R_2 = R_1$ and $e_1[k]$ be uncorrelated with $z_1[k]$ as well as $e_2[k]$. The noise covariance matrix can be expressed as

$$\mathbf{R_e} = \begin{bmatrix} \sigma_{e_1}^2 & \\ & \sigma_{e_2}^2 \end{bmatrix} = \mathbf{D_z}\frac{2^{-2R_1}}{12}\,,\ \mathbf{D_z} = \text{diag}\{\mathbf{R_z}\}.$$

Given that $\boldsymbol{\vartheta}$ operates on $\mathbf{z}[k]$ and $\mathbf{e}[k]$, the energy due to contributions of the desired user and quantization noise is respectively

- $E\|\boldsymbol{\vartheta}^H\mathbf{z}[k]\|^2 = \mathbf{r_{zs}}^H\mathbf{R_z}^{-1}\mathbf{r_{zs}}$ and
- $E\|\boldsymbol{\vartheta}^H\mathbf{e}[k]\|^2 = \mathbf{r_{zs}}^H\mathbf{R_z}^{-1}\mathbf{R_e}\mathbf{R_z}^{-1}\mathbf{r_{zs}}$.

The SQNR at digital baseband then results as

$$\frac{E\|\boldsymbol{\vartheta}^H\mathbf{z}[k]\|^2}{E\|\boldsymbol{\vartheta}^H\mathbf{e}[k]\|^2} = \frac{\mathbf{r_{zs}}^H\mathbf{R_z}^{-1}\mathbf{r_{zs}}}{\mathbf{r_{zs}}^H\mathbf{R_z}^{-1}\mathbf{D_z}\mathbf{R_z}^{-1}\mathbf{r_{zs}}\frac{2^{-2R_1}}{12}} \quad (6)$$

$E\|\boldsymbol{\vartheta}^H\mathbf{z}[k]\|^2$ is obtained by projection over $\mathbf{P}_{\underline{\mathbf{W}}}$, and is independent of $\underline{\mathbf{W}}$. Without loss of generality, maximizing SQNR in (6) is equivalent to minimizing $E\|\boldsymbol{\vartheta}^H\mathbf{e}[k]\|^2$

$$\begin{aligned} \underline{\mathbf{W}}_0 = &\ \min_{\boldsymbol{\vartheta}} \boldsymbol{\vartheta}^H\mathbf{D_z}\boldsymbol{\vartheta} \\ = &\ \min_{\underline{\mathbf{W}}} \mathbf{r_{zs}}^H\mathbf{R_z}^{-1}\mathbf{D_z}\mathbf{R_z}^{-1}\mathbf{r_{zs}}\frac{2^{-2R_1}}{12} \\ = &\ \min_{\underline{\mathbf{W}}} \underline{\mathbf{r}}_{\mathbf{xs}}^H\underline{\mathbf{W}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\mathbf{D_z}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1}\underline{\mathbf{W}}^H\underline{\mathbf{r}}_{\mathbf{xs}} \end{aligned} \quad (7)$$

Note that $\mathbf{D_z} = \text{diag}\{\mathbf{R_z}\} = \text{diag}\{\underline{\mathbf{W}}^H\underline{\mathbf{W}}\}$. For R_1 constant, omit the exponential term in the quantization noise. To solve (7), we first parametrize $\underline{\mathbf{W}}$ such that $\underline{\mathbf{W}} = \mathbf{U}\mathbf{V}^H = \mathbf{u}_1\mathbf{v}_1^H + \mathbf{u}_2\mathbf{v}_2^H$, where $\mathbf{V}$ is a 2×2 unitary matrix and $\mathbf{u}_1 = \underline{\mathbf{r}}_{\mathbf{xs}}/\|\underline{\mathbf{r}}_{\mathbf{xs}}\|$.

<u>Theorem 1</u>: 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 is obtained if all columns of $\mathbf{W}$ are equal to the MMSE beamformer, specified by the Wiener-Hopf solution as $\mathbf{R_x}^{-1}\mathbf{r}_{\mathbf{x}s}$, up to scaling and certain linear transformations.

Proof. Define $p = \|\mathbf{u}_2\|\,,\ \ \alpha = \mathbf{u}_1^H\mathbf{u}_2/p$ (Note that $|\alpha| \le 1$.) Then

$$(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1} = \mathbf{V}\frac{1}{p^2(1-|\alpha|^2)}\begin{bmatrix} p^2 & -\alpha p \\ -\bar{\alpha}p & 1 \end{bmatrix}\mathbf{V}^H$$

$$\begin{aligned} \frac{\underline{\mathbf{r}}_{\mathbf{xs}}^H}{\|\underline{\mathbf{r}}_{\mathbf{xs}}\|}\underline{\mathbf{W}}(\underline{\mathbf{W}}^H\underline{\mathbf{W}})^{-1} &= [1,\quad \alpha p]\frac{1}{p^2(1-|\alpha|^2)} \\ &\quad \cdot \begin{bmatrix} p^2 & -\alpha p \\ -\bar{\alpha}p & 1 \end{bmatrix}\mathbf{V}^H = \mathbf{v}_1^H \end{aligned}$$

$$\begin{aligned} \mathbf{D_z} &= \text{diag}\{\underline{\mathbf{W}}^H\underline{\mathbf{W}}\} \\ &= \text{diag}\{\mathbf{v}_1\mathbf{v}_1^H + p^2\mathbf{v}_2\mathbf{v}_2^H + \bar{\alpha}p\mathbf{v}_2\mathbf{v}_1^H + \alpha p\mathbf{v}_1\mathbf{v}_2^H\}. \end{aligned}$$

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

$$\mathbf{V} = \begin{bmatrix} \phi_1 & \\ & \phi_2 \end{bmatrix}\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \quad (8)$$

where $\phi_1 = e^{j\Phi_1}, \phi_2 = e^{j\Phi_1}$, $c = \cos(\theta)$ and $s = \sin(\theta)$. Here $\theta, \Phi_1,\ \Phi_2$ are in the range $(-\pi,\ \pi]$. Then

$$\mathbf{D_z} = \begin{bmatrix} c^2 + p^2s^2 - 2\beta psc & 0 \\ 0 & s^2 + p^2c^2 + 2\beta psc \end{bmatrix}$$

where $\beta = \text{Re}(\alpha)$; note that $-1 \le \beta \le 1$. The cost function (7) reduces to

$$J(\underline{\mathbf{W}}) \ \ = \ \ (c^4 + 2p^2s^2c^2 + s^4) + 2\beta p(s^3c - sc^3)$$

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

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

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. □

The above proof drives the following facts:

- In the absence of phase errors and RF imperfections, the two APN outputs compute the same estimate, upto scaling. These estimates are sampled by low resolution ADCs, followed by digital aveaging to reduce the quantization noise.

- From the quantization noise perspective, the APN can be seen as a (spatially) oversampled quantizer of MMSE estimates.
- From the ADC power consumption perspective, a multi-channel APN with ADC resolution R_1 leads to same quantization noise as a single channel APN with ADC resolution R_0 as long as $N_D = 2^{\frac{R_0}{R_1}}$.
- The multi-channel APN is robust to multi-access interference, allows a more efficient way to predict and quantize $\mathbf{z}[k]$, especially in the presence of RF phase errors and discrete phase shifts.

4. DISCRETE PHASE SHIFT IMPLEMENTATION

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. The APN design is now reduced to a problem of choosing N_D vectors from $\{\boldsymbol{\phi}_m\}_{m=1}^{M} \in \mathcal{D}$ as $\underline{\mathbf{W}} = [\boldsymbol{\phi}_{M_1}, \cdots, \boldsymbol{\phi}_{M_{N_D}}]$, such that the MSE distortion is minimized

$$\min_{\boldsymbol{\vartheta}, M\, \boldsymbol{\phi}_m \in \mathcal{D}} \quad E\|s^{(1)}[k] - \boldsymbol{\vartheta}^H Q\{\underline{\mathbf{W}}^H \mathbf{x}(t)\}\|^2 \qquad (9)$$

Here $\mathcal{D}$ is a fixed phase shifter containing the set of all possible $N_r \times 1$ phase shift vectors that the columns in $\underline{\mathbf{W}}$ can take. Typically $\mathcal{D} = \{\boldsymbol{\phi}_m\}_{m=1}^{M}$ is overdetermined i.e. $M \gg N_r$, where M is the size of the dictionary $\mathcal{D}$. For a $N_r = 2$ setup and APN taps indexed by say $R_{\mathbf{W}} =$ 2-bits; $M = 2^{N_D \mathrm{R}_{\mathbf{W}}} = 16$.

One technique, robust to phase errors and discrete APN, with reduced computation complexity is through matching pursuit (MP) [7] to successively choose $\boldsymbol{\phi}_m, m \in \{1, \cdots, M\}$. We start from the cost function (9), replace $\underline{\mathbf{W}}$ with $\boldsymbol{\phi}_{M_1}$ and continue as proposition 1.

$$\begin{aligned}\boldsymbol{\phi}_{M_1} &= \min_{\boldsymbol{\phi}_m \in \mathcal{D}} \; E\|s^{(1)}[k] - \bar{\vartheta}_1 Q\{\boldsymbol{\phi}_m^H \mathbf{x}(t)\}\|^2 \\ &= \max_{\boldsymbol{\phi}_m \in \mathcal{D}} \; \underline{\mathbf{r}}_{\mathbf{xs}}^H \boldsymbol{\phi}_m \boldsymbol{\phi}_m^H \underline{\mathbf{r}}_{\mathbf{xs}}. \qquad (10)\end{aligned}$$

Due to space constraints, we briefly explain the iterative phase shift approximation of the APN using Quantized MP using the following table. For detailed proofs, refer the more extended version [8].

We also show in [8] using detailed simulation results, that the APN setup with $M = 16$, $N_D = 2$and$R = 6$ lead to equivalent MSE and BER performance when compared to a reference setup of N_r ADCs operating at float precision.

Objective: Select the discrete phase vectors of APN
Step 1: Given: Input signal $\mathbf{x}(t)$ and the training $s^{(1)}[k]$

- Select $\boldsymbol{\phi}_{M_1}$ and $\boldsymbol{\phi}_m$ unitary, such that $\|\boldsymbol{\phi}_{M_1}^H \underline{\mathbf{r}}_{\mathbf{xs}}\|^2 \geq \|\boldsymbol{\phi}_m^H \underline{\mathbf{r}}_{\mathbf{xs}}\|^2 \; \forall m \in \{1, \cdots, M\}$.
- Compute $\Delta_1 \mathbf{x}(t) = (\mathbf{I} - \beta_{M_1} \mathbf{P}_{M_1})\mathbf{x}(t)$, where $\mathbf{P}_{M_1} = \boldsymbol{\phi}_{M_1} \boldsymbol{\phi}_{M_1}^H$ and β_{M_1} is scaling factor.
  - $\mathbf{R}_{\Delta \mathbf{x}} = E\{\Delta_1 \mathbf{x}(t) \Delta_1 \mathbf{x}^H(t)\}$ and $\mathbf{r}_{\Delta \mathbf{s}} = E\{\Delta_1 \mathbf{x}(t) \bar{s}^{(1)}[k]\}$
  - Select $\|\boldsymbol{\phi}_{M_2}^H \underline{\mathbf{r}}_{\Delta \mathbf{s}}\|^2 \geq \|\boldsymbol{\phi}_m^H \underline{\mathbf{r}}_{\Delta \mathbf{s}}\|^2$.
- Continue till $\underline{\mathbf{W}} = \left[\boldsymbol{\phi}_{M_1}, \cdots, \boldsymbol{\phi}_{M_{N_D}}\right]$.

5. REFERENCES

[1] B. Murmann, "Digitally assisted analog circuits," *IEEE Micro*, vol. 26, no. 2, pp. 38–47, Feb 2006.

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

[3] X. Zhang, A. F. Molisch, and S. Y. Kung, "Variable phase shift based RF baseband codesign for MIMO antenna selection," *IEEE Transactions Signal Processing*, vol. 53, no. 11, pp. 4091–4103, Nov 2005.

[4] A. Hajimiri, H. Hashemi, A. Natarajan, X Guang, and A. Babakhani, "Integrated phased arrays in silicon," *Proceedings of IEEE*, vol. 93, no. 9, pp. 1637–1655, Sep 2005.

[5] P. Baltus and R. Dekker, "Optimizing RF front ends for low power," *Proceedings of IEEE*, vol. 88, no. 10, pp. 1546–1559, Jan 2000.

[6] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Transactions Information Theory*, vol. 44, no. 7, pp. 2943–2959, Nov 1998.

[7] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec 1993.

[8] V. Venkateswaran and A. J. van der Veen, "Analog beamforming in MIMO communications – with phase shift networks and online channel estimation," *Submitted Nov 2009 to IEEE Transactions on Signal Processing*.

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