AN ADAPTIVE BEAMFORMING SCHEME FOR ENHANCED COCHANNEL INTERFERENCE MITIGATION ON SHORT ARRAY SIGNAL INTERVALS

Thomas Hunziker and Makoto Taromaru

ATR Wave Engineering Laboratories
2-2-2 Hikaridai, Keihanna Science City, Kyoto, 619-0288 Japan

ABSTRACT

The problem of mitigating Gaussian cochannel interference with unknown spatial correlation within short array signal observation intervals is addressed in this paper. We show that the classic sample matrix inversion schemes can be represented as the projection of the ideally combined signal onto a subspace that is uniformly distributed on the Grassmann manifold. This brings us to a novel beamforming scheme with enhanced mean-square error performance.

1. INTRODUCTION

Cochannel interference (CCI) mitigation is a crucial feature in wireless networks without stringent channel access policy. In emerging decentralized multi-hop networks, both data and control frames sent from peer to peer need to be very short in order to limit latency over multiple hops, since the relaying peers normally cannot receive and transmit simultaneously. This leads to CCI with rapidly fluctuating characteristics in unsynchronized networks, requiring an adaptation of the spatial filters within very limited observation intervals.

Many of the popular adaptive beamforming techniques for array antenna-enhanced receivers, discussed in textbooks like [1, 2], rely on an inversion of a sample covariance matrix (SCM). Their performance depends on the sample support, i.e., the number of observed vector samples for the SCM computation, relative to the number of antennas. Besides, the presence of the desired signal in the sample basis sometimes dramatically impairs the accuracy of the SCM [3]. Diagonal loading of the SCM has been shown to improve the beamforming performance [4, 5], however, the methods are only effective for certain CCI constellations. With N antennas, for instance, and $N-1$ similarly strong interferers constituting the spatial covariance matrix, any diagonal loading increases the sample error variance in the spatially filtered signal.

In this paper we present a novel beamforming scheme whose performance compared to an ideal signal combining is independent of the number of interferers and the spatial correlation of the CCI. The enhanced beamforming scheme is derived in Sect. 4 following the system model and a discussion on the classic SCM-based methods, and thereafter the complexity issue is addressed.

2. SYSTEM MODEL

Throughout the paper we use boldfaced lowercase characters for row and column vectors and boldfaced uppercase characters for matrices. The Hermitian transpose of $\mathbf{X}$ is written as $\mathbf{X}^{\mathrm{H}}$, and $[\mathbf{X}\ \mathbf{Y}]$ represents the horizontal concatenation of $\mathbf{X}$ and $\mathbf{Y}$.

This work is part of "Research and development of ultra-high speed giga-bit rate wireless LAN systems" granted by National Institute of Information and Communication Technology (NICT).

Assume a staggered, block-wise transmission of an information-bearing signal over a narrow-band single-input/N-output channel. A block comprises K data symbols. The channel gain is constant over many blocks and perfectly known, however, the characteristics of the CCI vary arbitrarily from block to block, and they are unknown. Following a proper sampling of the array signal at the symbol rate, the baseband receiver observes a block as the sequence $\mathbf{y}_1, \ldots, \mathbf{y}_K$ of complex $N \times 1$-vectors, where $\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_K]$ is given as

$$\mathbf{Y} = \mathbf{h}\mathbf{s} + \mathbf{W}. \tag{1}$$

The column vector $\mathbf{h} \in \mathbb{C}^N$ defines the signal attenuation at the N receiver antennas, the $1 \times K$-row vector $\mathbf{s}$ comprises the data symbols, and $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_K]$ includes the CCI and thermal noise (CCIN).

The interference may stem from an arbitrary number of diverse, unsynchronized sources. It seems reasonable to model the CCIN as temporally white, spatially correlated Gaussian: The independent random vectors $\mathbf{w}_1, \ldots, \mathbf{w}_K$ are $\mathcal{CN}(\mathbf{0}, \mathbf{R})$, i.e., zero-mean circularly symmetric complex Gaussian with the covariance matrix $\mathbf{R}$, and $\mathbf{s}$ and $\mathbf{W}$ are independent. The covariance matrix $\mathbf{R}$ depends on the radio channels between the CCI sources and the receiver. In environments with multipath signal propagation $\mathbf{R}$ is solely known to be Hermitian positive definite.

Preambles in the form of a number of leading zeros may facilitate the beamforming at the receiver side. Including a preamble of length M, the vector $\mathbf{s}$ comprises M zeros and $K-M$ data symbols, i.e., $\mathbf{s} = [\mathbf{0}\ \mathbf{s}_{\mathrm{D}}]$. Likewise, $\mathbf{Y} = [\mathbf{Y}_{\mathrm{P}}\ \mathbf{Y}_{\mathrm{D}}]$, where $\mathbf{Y}_{\mathrm{P}}$ and $\mathbf{Y}_{\mathrm{D}}$ relate to the preamble and the data sections, respectively, of the block.

The signals from the N antennas are linearly combined prior to the demodulation/decoding for the sake of limiting complexity. The adaptive beamforming aims at choosing appropriate weights in order to attain a favorable signal-to-interference-and-noise ratio (SINR) in the combined signal. Varying SINR from block to block, due to fluctuating CCI, can be dealt with by a forward error control (FEC) coding and interleaving over a large number of blocks.

3. BEAMFORMING VIA SAMPLE MATRIX INVERSION

In the ideal case where the covariance matrix $\mathbf{R}$ of the CCIN is perfectly known, $\mathbf{f}_{\mathrm{MVDR}} = (\mathbf{h}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{h})^{-1}\mathbf{h}^{\mathrm{H}}\mathbf{R}^{-1}$ maximizes the SINR of the linearly filtered signal $\mathbf{f}_{\mathrm{MVDR}}\mathbf{Y}$ subject to the constraint $\mathbf{f}_{\mathrm{MVDR}}\mathbf{h} = 1$. The spatial filter $\mathbf{f}_{\mathrm{MVDR}}$ is usually referred to as the MVDR (*minimum variance distortionless response*) beamformer. The error in the filter output equals $\mathbf{f}_{\mathrm{MVDR}}\mathbf{W}$. Hence, the sample errors are complex normally distributed with the variance

$$\vartheta_{\mathrm{MVDR}} = \frac{1}{K} E\left[\|\mathbf{f}_{\mathrm{MVDR}}\mathbf{W}\|^2\right] = \left(\mathbf{h}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{h}\right)^{-1}, \tag{2}$$

where $E[\cdot]$ and $\|\cdot\|$ denote the expectation and the 2-norm of a row/column vector, respectively.

3.1. Preamble-Based SMI Beamforming

An estimate $\widehat{\mathbf{R}}_{\mathrm{P}} = M^{-1}\mathbf{Y}_{\mathrm{P}}\mathbf{Y}_{\mathrm{P}}^{\mathrm{H}}$ of the covariance matrix is termed *SCM*, and the spatial filtering by

$$\mathbf{f}_{\mathrm{PSMI}} = \left(\mathbf{h}^{\mathrm{H}}\widehat{\mathbf{R}}_{\mathrm{P}}^{-1}\mathbf{h}\right)^{-1}\mathbf{h}^{\mathrm{H}}\widehat{\mathbf{R}}_{\mathrm{P}}^{-1} \tag{3}$$

$$= \left(\mathbf{h}^{\mathrm{H}}(\mathbf{Y}_{\mathrm{P}}\mathbf{Y}_{\mathrm{P}}^{\mathrm{H}})^{-1}\mathbf{h}\right)^{-1}\mathbf{h}^{\mathrm{H}}(\mathbf{Y}_{\mathrm{P}}\mathbf{Y}_{\mathrm{P}}^{\mathrm{H}})^{-1} \tag{4}$$

is known as the *sample matrix inversion* (SMI) technique. Employing $\mathbf{f}_{\mathrm{PSMI}}$, the conditional sample error variance of the combined signal $\mathbf{f}_{\mathrm{PSMI}}\mathbf{Y}_{\mathrm{D}}$ is given as $\vartheta_{\mathrm{PSMI}} = \mathbf{f}_{\mathrm{PSMI}}\mathbf{R}\mathbf{f}_{\mathrm{PSMI}}^{\mathrm{H}}$. Both the beamformer $\mathbf{f}_{\mathrm{PSMI}}$ and $\vartheta_{\mathrm{PSMI}}$ are random quantities that depend on $\mathbf{Y}_{\mathrm{P}}$. If $M \geq N$, the ratio $\vartheta_{\mathrm{MVDR}}/\vartheta_{\mathrm{PSMI}}$ is known to be beta distributed with the parameters $M-N+2$ and $N-1$ [6]. It follows that $E[\vartheta_{\mathrm{PSMI}}/\vartheta_{\mathrm{MVDR}}] = M/(M-N+1)$ and, consequently, the unconditional mean-square sample error (MSSE) equals

$$\bar{\vartheta}_{\mathrm{PSMI}} = \frac{1}{K-M}E\left[\|\mathbf{f}_{\mathrm{PSMI}}\mathbf{Y}_{\mathrm{D}} - \mathbf{s}_{\mathrm{D}}\|^2\right] = \frac{M}{M-N+1}\vartheta_{\mathrm{MVDR}} \tag{5}$$

provided that $K > M \geq N$. A comprehensive study on the statistical properties of the output of this beamforming scheme is found in [7].

3.2. Data-Based SMI Beamforming

In the absence of a preamble, the SCM $\widehat{\mathbf{R}} = K^{-1}\mathbf{Y}\mathbf{Y}^{\mathrm{H}}$ may be used as an estimate of $\mathbf{R}$, resulting in the spatial filter

$$\mathbf{f}_{\mathrm{DSMI}} = (\mathbf{h}^{\mathrm{H}}\widehat{\mathbf{R}}^{-1}\mathbf{h})^{-1}\mathbf{h}^{\mathrm{H}}\widehat{\mathbf{R}}^{-1}. \tag{6}$$

In this approach, the sample basis for the SCM computation and the target of the beamformer are *identical*. Choosing $M = 0$ saves bandwidth, however, the presence of the desired signal in the sample basis degrades the accuracy of the SCM.

If $K \geq N$, the MSSE after this beamforming variant equals

$$\bar{\vartheta}_{\mathrm{DSMI}} = \frac{1}{K}E\left[\|\mathbf{f}_{\mathrm{DSMI}}\mathbf{Y} - \mathbf{s}\|^2\right] = \frac{K-N+1}{K}\vartheta_{\mathrm{MVDR}} + \frac{N-1}{K}\varepsilon_{\mathrm{s}} \tag{7}$$

with $\varepsilon_{\mathrm{s}} = E\left[\|\mathbf{s}_{\mathrm{D}}\|^2\right]/(K-M)$ the 2^{nd} moment of a data symbol. The two terms on the right hand side in (7) may be regarded as due to the CCIN and due to the desired signal in the sample basis, respectively. Both MSSE terms were derived by [8], and an alternative way for their calculation is outlined in the Sect. 4. We note that if K is not significantly larger than N, the presence of the desired signal leads to a large MSSE even in situations with little interference.

4. ENHANCED BEAMFORMING

In applications with a ratio K/N not much larger than one, the above data-based beamforming may not achieve satisfactory interference mitigation, whereas the preamble-based variant may meet the SINR requirements of the demodulator/decoder only at the cost of an unacceptable degradation of spectral efficiency. To find alternatives, let us first make the following definitions:

- The matrix $\mathbf{I}_K$ represents the $K \times K$-identity matrix.
- The so-called *signal blocking matrix* $\mathbf{C} = [\mathbf{c}_1 \cdots \mathbf{c}_{N-1}]$ is an $N \times (N-1)$-matrix with orthonormal column vectors, i.e., $\mathbf{C}^{\mathrm{H}}\mathbf{C} = \mathbf{I}_{N-1}$, such that $\mathbf{h}^{\mathrm{H}}\mathbf{C} = \mathbf{0}$.
- The rows of $\mathbf{C}^{\mathrm{H}}\mathbf{Y}$ span the $(N-1)$-dimensional subspace $\mathcal{B}$ of $\mathbb{C}^K$, i.e., $\mathcal{B} = \mathrm{span}(\mathbf{c}_1^{\mathrm{H}}\mathbf{Y}, \ldots, \mathbf{c}_{N-1}^{\mathrm{H}}\mathbf{Y})$, and $\mathcal{B}^\perp$ represents the $(K-N+1)$-dimensional orthogonal subspace in $\mathbb{C}^K$.

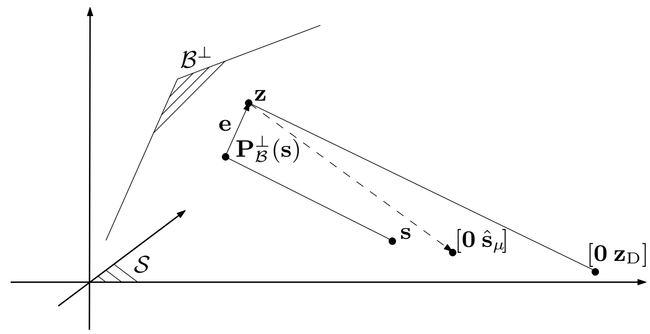

Figure 1: Illustration of the vector signal space $\mathbb{C}^K$ with the subspaces $\mathcal{S}$ and $\mathcal{B}^\perp$ for the case $K=3$, $N=2$, $M=1$.

- $\mathbf{P}_{\mathcal{B}}(\mathbf{x})$ stands for the orthogonal projection of $\mathbf{x}$ onto the subspace $\mathcal{B}$, and $\mathbf{P}_{\mathcal{B}}^\perp(\mathbf{x}) = \mathbf{x} - \mathbf{P}_{\mathcal{B}}(\mathbf{x})$ for the projection onto $\mathcal{B}^\perp$.

Using these notations, we can express the signal $\mathbf{z} = \mathbf{f}_{\mathrm{DSMI}}\mathbf{Y}$ obtained by the *data-based* SMI beamforming in the following form.

Proposition 1: The signal $\mathbf{z} = \mathbf{f}_{\mathrm{DSMI}}\mathbf{Y}$ may be written as

$$\mathbf{z} = \mathbf{P}_{\mathcal{B}}^\perp(\mathbf{f}_{\mathrm{MVDR}}\mathbf{Y}). \tag{8}$$

That is, $\mathbf{z}$ may be seen as the result of an ideal combining, however, projected onto the subspace $\mathcal{B}^\perp$. As a direct consequence of (8),

$$\mathbf{z} = \mathbf{P}_{\mathcal{B}}^\perp(\mathbf{s}) + \mathbf{e} \tag{9}$$

$$= \mathbf{s} + \mathbf{e} - \mathbf{P}_{\mathcal{B}}(\mathbf{s}), \tag{10}$$

where $\mathbf{e} = \mathbf{P}_{\mathcal{B}}^\perp(\mathbf{f}_{\mathrm{MVDR}}\mathbf{W})$.

Proposition 2: The subspace $\mathcal{B}^\perp$ is uniformly distributed on the Grassmann manifold $G(K, K-N+1)$ of all $(K-N+1)$-dimensional subspaces of $\mathbb{C}^K$. Furthermore, the subspace defining matrix $\mathbf{C}^{\mathrm{H}}\mathbf{Y}$, the CCIN vector $\mathbf{f}_{\mathrm{MVDR}}\mathbf{W}$, and the data vector $\mathbf{s}$ are independent.

The MSSE (7) can easily be derived from (10) using proposition 2. The two MSSE terms that make up $\bar{\vartheta}_{\mathrm{DSMI}}$ correspond to $K^{-1}E[\|\mathbf{e}\|^2]$ and $K^{-1}E[\|\mathbf{P}_{\mathcal{B}}(\mathbf{s})\|^2]$, respectively. The former expectation results from (2) and the dimension of the subspace $\mathcal{B}^\perp$, while the latter equals ε_{s} scaled by the ratio of the dimension of $\mathcal{B}$ over K. Adding the two expectations is permissible since the two error terms following $\mathbf{s}$ in (10) reside in orthogonal subspaces.

From (9) we note that $\mathbf{z}$ contains no *a posteriori* information about the desired signal component in the subspace $\mathcal{B}$. The signal part $\mathbf{P}_{\mathcal{B}}(\mathbf{s})$ is completely lost. This loss could be avoided if *a priori* information about the signal $\mathbf{s}$ was available, for instance, if $\mathbf{s}$ was known to be an element of some subspace of $\mathbb{C}^K$. Let us once again assume the inclusion of a preamble of the above described form, thereby constraining $\mathbf{s}$ to a $(K-M)$-dimensional subspace $\mathcal{S}$.

An obvious way for reconstructing $\mathbf{s}$ would be to choose the element in $\mathcal{S}$ for which the projection onto $\mathcal{B}^\perp$ is nearest to $\mathbf{z}$, i.e.,

$$\mathbf{z}_{\mathrm{D}} = \arg\min_{\mathbf{x}\in\mathbb{C}^{K-M}} \left\|\mathbf{P}_{\mathcal{B}}^\perp([\mathbf{0}\ \mathbf{x}]) - \mathbf{z}\right\|. \tag{11}$$

If $M = N-1$, the dimensions of $\mathcal{B}^\perp$ and $\mathcal{S}$ equal and $\mathbf{P}_{\mathcal{B}}^\perp([\mathbf{0}\ \mathbf{z}_{\mathrm{D}}]) = \mathbf{z}$ with probability one (see Fig. 1).

Proposition 3: The signal $\mathbf{z}_{\mathrm{D}}$ defined in (11) corresponds to the result of the *preamble-based* SMI beamforming: $\mathbf{z}_{\mathrm{D}} = \mathbf{f}_{\mathrm{PSMI}}\mathbf{Y}_{\mathrm{D}}$.

When the largest canonical angle between the subspaces $\mathcal{B}^\perp$ and $\mathcal{S}$ is close to 90°, the least-squares estimation (11) of the data vector suffers from severe noise enhancement, as can be seen in Fig. 1. To avoid this effect, we may rather adopt a minimum mean-square error (MMSE) approach. To this end, let $\mathbf{B}$ represent a $(K{-}N{+}1){\times}K$-matrix whose orthonormal rows constitute a basis of $\mathcal{B}^{\perp}$. It follows from (9) that the signal $\mathbf{z}$, written w. r. t. this basis, is given as

$$\mathbf{z}\mathbf{B}^{\mathrm{H}} = \mathbf{s}_{\mathrm{D}}\mathbf{B}_{\mathrm{D}}^{\mathrm{H}} + \mathbf{e}\mathbf{B}^{\mathrm{H}}, \tag{12}$$

where the $(K{-}N{+}1){\times}(K{-}M)$-matrix $\mathbf{B}_{\mathrm{D}}$ comprises the rightmost $K-M$ columns of $\mathbf{B}$. The row vector $\mathbf{e}\mathbf{B}^{\mathrm{H}} = \mathbf{f}_{\mathrm{MVDR}}\mathbf{W}\mathbf{B}^{\mathrm{H}}$ in (12) is $\mathcal{CN}(\mathbf{0}, \vartheta_{\mathrm{MVDR}}\mathbf{I}_{K-N+1})$ as a consequence of proposition 2 and independent of $\mathbf{s}_{\mathrm{D}}$. The MMSE estimator of $\mathbf{s}_{\mathrm{D}}$ on the basis of the linear model (12) reads [9]

$$\hat{\mathbf{s}}_{\mathrm{MMSE}} = \mathbf{z}\mathbf{B}^{\mathrm{H}}\left(\varepsilon_{\mathrm{s}}\mathbf{B}_{\mathrm{D}}\mathbf{B}_{\mathrm{D}}^{\mathrm{H}} + \vartheta_{\mathrm{MVDR}}\,\mathbf{I}_{K-N+1}\right)^{-1}\varepsilon_{\mathrm{s}}\mathbf{B}_{\mathrm{D}}. \tag{13}$$

The matrices $\mathbf{B}$ and $\mathbf{B}_{\mathrm{D}}$ as well as $\mathbf{z} = \mathbf{f}_{\mathrm{DSMI}}\mathbf{Y}$ can be computed from the observation $\mathbf{Y}$, however, $\vartheta_{\mathrm{MVDR}}$ is unknown. Resorting to some estimate $\mu > 0$ of $\vartheta_{\mathrm{MVDR}}$, we obtain

$$\hat{\mathbf{s}}_{\mu} = \mathbf{z}\mathbf{B}^{\mathrm{H}}\left(\varepsilon_{\mathrm{s}}\mathbf{B}_{\mathrm{D}}\mathbf{B}_{\mathrm{D}}^{\mathrm{H}} + \mu\,\mathbf{I}_{K-N+1}\right)^{-1}\varepsilon_{\mathrm{s}}\mathbf{B}_{\mathrm{D}} \tag{14}$$

as our *enhanced beamforming* scheme.

To investigate the MSSE in $\hat{\mathbf{s}}_{\mu}$ when $M \geq N{-}1$, we first express the conditional sample error variance given $\mathbf{B}_{\mathrm{D}}$. Making use of the singular value decomposition of $\mathbf{B}_{\mathrm{D}}$ and some algebra leads to

$$\frac{1}{K-M}E\left[\|\hat{\mathbf{s}}_{\mu} - \mathbf{s}_{\mathrm{D}}\|^2 \;\middle|\; \mathbf{B}_{\mathrm{D}}\right] = \frac{\varepsilon_{\mathrm{s}}}{K-M}\sum_{i=1}^{K-M}\frac{\mu^2 + \varepsilon_{\mathrm{s}}v_i^2\vartheta_{\mathrm{MVDR}}}{(\varepsilon_{\mathrm{s}}v_i^2 + \mu)^2}, \tag{15}$$

where $v_1, \ldots, v_{K-M}$ denote the singular values of $\mathbf{B}_{\mathrm{D}}$. Integrating over the density $p(\mathbf{q})$ of the random vector $\mathbf{q} = [v_1 \cdots v_{K-M}]$ with the (zero or) positive singular values yields the unconditional MSSE

$$\begin{aligned}\bar{\vartheta}_{\mu} \;&=\; (K-M)^{-1}E\left[\|\hat{\mathbf{s}}_{\mu} - \mathbf{s}_{\mathrm{D}}\|^2\right] \\ &=\; \frac{\varepsilon_{\mathrm{s}}}{K-M}\int\limits_{\mathbf{q}\in\mathbb{R}_{+}^{K-M}}\sum_{i=1}^{K-M}\frac{\mu^2 + \varepsilon_{\mathrm{s}}v_i^2\vartheta_{\mathrm{MVDR}}}{(\varepsilon_{\mathrm{s}}v_i^2 + \mu)^2}p(\mathbf{q})\mathrm{d}\mathbf{q}. \end{aligned} \tag{16}$$

A closed-form expression for $\bar{\vartheta}_{\mu}$ is not available[1], however, we can make the following statements:

1. As a consequence of the uniform distribution of $\mathcal{B}^{\perp}$ on $G(K, K-N+1)$, the distribution of the singular values depends only on K, N, and M. Similar to $\bar{\vartheta}_{\mathrm{PSMI}}$ and $\bar{\vartheta}_{\mathrm{DSMI}}$, the MSSE performance $\bar{\vartheta}_{\mu}$ of the enhanced beamforming thus depends on $(\mathbf{h}, \mathbf{R})$ only via $\vartheta_{\mathrm{MVDR}}$.
2. For any given $(\mathbf{s}_{\mathrm{D}}, \mathbf{h}, \mathbf{Y})$, the MMSE estimate $\hat{\mathbf{s}}_{\mu}$ tends to $\mathbf{z}_{\mathrm{D}}$ as μ approaches zero. In the same time, the MSSE $\bar{\vartheta}_{\mu}$ tends to $\bar{\vartheta}_{\mathrm{PSMI}}$.
3. Via partial derivation of (15) w. r. t. μ we find that $\bar{\vartheta}_{\mu}$ has only one minimum within $\mu \in (0, \infty)$, namely at $\mu = \vartheta_{\mathrm{MVDR}}$. As a consequence of statement 2, $\bar{\vartheta}_{\vartheta_{\mathrm{MVDR}}} \leq \bar{\vartheta}_{\mathrm{PSMI}}$.
4. The above result can be extended as follows:
   Proposition 4: If $\mu \in (0, 2\vartheta_{\mathrm{MVDR}}]$ then $\bar{\vartheta}_{\mu} \leq \bar{\vartheta}_{\mathrm{PSMI}}$.

[1] The singular values actually relate to the cosines of the canonical angles between the subspaces $\mathcal{S}$ and $\mathcal{B}^{\perp}$. The joint density of the singular values is formulated in [10], theorem 3. 3. 4., for the corresponding real case.

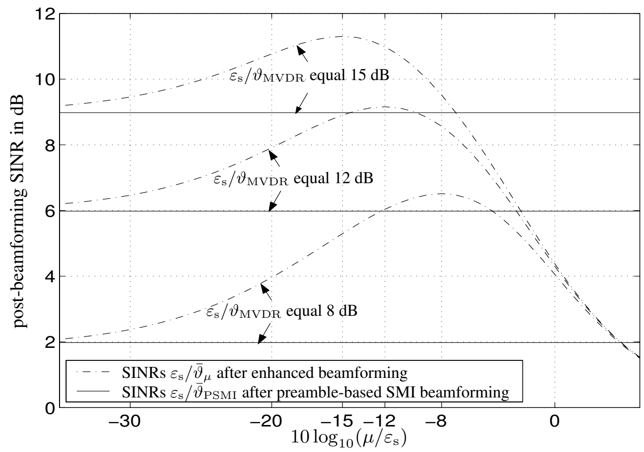

Figure 2: Performance comparison of the preamble-based SMI beamforming and the enhanced beamforming scheme versus the parameter μ.

We shall now discuss some exemplary transmission scenarios. First, assume that the receiver employs $N = 4$ receive antennas, a block comprises $K = 16$ samples, of which the first $M = 4$ symbols constitute the preamble, and that an ideal filtering would achieve an MSSE of $8, 12,$ or 15 dB below the desired signal power. Fig. 2 compares the respective SINRs after a preamble-based SMI beamforming (according to (5)) and after the enhanced beamforming as a function of μ. The numerically evaluated $\varepsilon_{\mathrm{s}}/\bar{\vartheta}_{\mu}$ are in line with the above statements 2–4. Obviously, the enhanced beamforming with the optimal choice $\mu = \vartheta_{\mathrm{MVDR}}$ achieves gains of more than 4 dB over the classic beamforming when $\varepsilon_{\mathrm{s}}/\vartheta_{\mathrm{MVDR}}$ equals 8 dB.

It arises the question of how μ should be chosen without prior knowledge about the amount of interference. With the CCI varying from block to block and an FEC coding over a large number of blocks, the decoder requires a certain *average* SINR for reliable information recovery. Choosing μ around this known average SINR seems reasonable, since this lets the blocks with above-average CCI (i.e., large $\vartheta_{\mathrm{MVDR}}$) always benefit from the enhanced beamforming. The classic preamble-based SMI beamforming becomes only superior in the blocks with little interference (small $\vartheta_{\mathrm{MVDR}}$) due to statement 4, however, these have minor impact on the average SINR.

This can be seen in Fig. 3, showing the achieved SINRs in a scenario with $N = 4$, $K = 16$, and either $M = 4$ or $M = 5$ preamble symbols, versus the SINR $\varepsilon_{\mathrm{s}}/\vartheta_{\mathrm{MVDR}}$ by an ideal combining. The parameter μ is set to 12 dB below the desired signal power. Obviously, the enhanced scheme is superior unless the SINR $\varepsilon_{\mathrm{s}}/\vartheta_{\mathrm{MVDR}}$ exceeds $16\frac{3}{4}$ and $18\frac{1}{2}$ dB, respectively. The figure further shows the performance of the data-based SMI beamforming as given in (7), outperforming the other adaptive schemes at very low SINRs.

5. IMPLEMENTATION

The classic SMI schemes involve an inversion of the $N \times N$-SCM, whereas (14) contains the inverse of a matrix of dimension $K{-}N{+}1$. The gain in MSSE performance thus comes at the cost of a considerable increase in computational complexity if $\hat{\mathbf{s}}_{\mu}$ was computed according to (14). There are, however, more favorable methods for obtaining $\hat{\mathbf{s}}_{\mu}$. For any $\mathcal{B}^{\perp}$, the matrix $\mathbf{B}$ can assume the form

$$\mathbf{B} = \begin{bmatrix} \mathbf{u} & \mathbf{v} \\ \mathbf{U} & \mathbf{V} \\ \mathbf{0} & \mathbf{X} \end{bmatrix} \text{ with } \begin{array}{l} [\mathbf{u}\;\mathbf{v}] = \mathbf{P}_{\mathcal{B}}^{\perp}(\mathbf{h}^{\mathrm{H}}\mathbf{Y})/\|\mathbf{P}_{\mathcal{B}}^{\perp}(\mathbf{h}^{\mathrm{H}}\mathbf{Y})\|, \\ \mathbf{U}: \text{an } M \times M\text{-matrix}, \\ \mathbf{V}: \text{an } M \times (K-M)\text{-matrix}, \\ \mathbf{X}: \text{a } (K-N-M) \times (K-M)\text{-matrix}. \end{array}$$

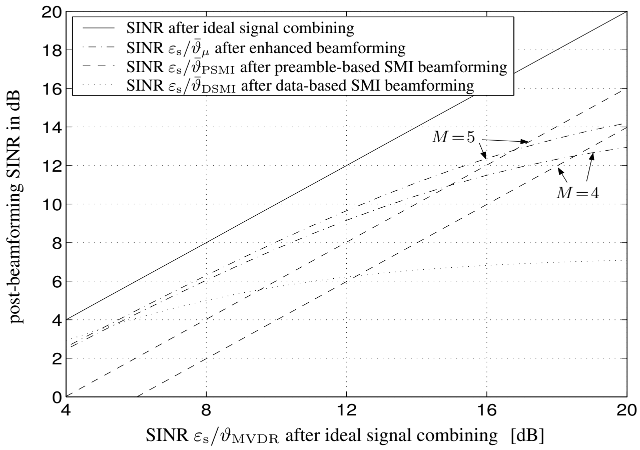

Figure 3: Performance of the various beamforming schemes versus the SINR by an ideal combining, choosing ε_s/μ equal 12 dB.

Now, $\mathbf{zB}^{\rm H}$ in (14) simplifies to $\left[\|\mathbf{P}_{\mathcal{B}}^{\perp}(\mathbf{h}^{\rm H}\mathbf{Y})\|/\|\mathbf{h}\|^2 \;\; 0 \;\cdots\; 0\right]$ and

$$\left(\varepsilon_s \mathbf{B}_{\rm D}\mathbf{B}_{\rm D}^{\rm H} + \mu\,\mathbf{I}_{K-N+1}\right)^{-1} = \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & (\varepsilon_s+\mu)^{-1}\mathbf{I}_{K-N-M} \end{bmatrix} \tag{17}$$

with

$$\mathbf{G} = \left((\varepsilon_s+\mu)\mathbf{I}_{M+1} - \varepsilon_s \begin{bmatrix} \mathbf{u} \\ \mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{U} \end{bmatrix}^{\rm H}\right)^{-1}. \tag{18}$$

Finally, the result of the enhanced beamforming can be expressed as

$$\hat{\mathbf{s}}_\mu = \left(\left\|\mathbf{P}_{\mathcal{B}}^{\perp}(\mathbf{h}^{\rm H}\mathbf{Y})\right\| \Big/ \|\mathbf{h}\|^2\right) \mathbf{g} \begin{bmatrix} \mathbf{v} \\ \mathbf{V} \end{bmatrix} \tag{19}$$

with $\mathbf{g}$ denoting the first row of $\mathbf{G}$.

Hence, only the upper $M+1$ rows of $\mathbf{B}$ need to be computed. These actually correspond to the rows N to $N+M$ of the matrix $\mathbf{Q}$ resulting from the $\mathbf{LQ}$-decomposition

$$\mathbf{LQ} = \begin{bmatrix} \mathbf{C}^{\rm H}\mathbf{Y} \\ \mathbf{h}^{\rm H}\mathbf{Y} \\ [\mathbf{I}_M \; \mathbf{0}] \end{bmatrix}$$

(with $\mathbf{L}$ left-triangular and $\mathbf{QQ}^{\rm H} = \mathbf{I}_{N+M}$). The scalar preceding $\mathbf{g}$ in (19) corresponds to the Nth diagonal element in $\mathbf{L}$, and $\mathbf{g}$ is the first row of the inverse of the matrix (18) of dimension $M+1$.

6. CONCLUSION AND OUTLOOK

A novel beamforming scheme has been proposed, featuring better adaptive CCI suppression on short signal intervals than the classic preamble-based SMI beamforming methods. The enhanced scheme has turned out to reduce the gap to the SINR $\varepsilon_s/\vartheta_{\rm MVDR}$ attained by an ideal signal combining based on perfectly known spatial covariance, over a large range of $\vartheta_{\rm MVDR}$. Unlike diagonal loading techniques, the performance gain is independent of the particular structure of the CCI.

The enhanced beamforming turns out similarly superior to the classic SMI methods with/without diagonal loading in scenarios with look direction mismatches, i.e., mismatches in $\mathbf{h}$. This issue will be discussed in an upcoming paper.

APPENDIX

Proof of proposition 1:
Using first (6) and $(\|\mathbf{h}\|^{-2}\mathbf{hh}^{\rm H} + \mathbf{CC}^{\rm H}) = \mathbf{I}_N$, and second the identity $(\mathbf{h}^{\rm H}\mathbf{X}^{-1}\mathbf{h})^{-1}\mathbf{h}^{\rm H}\mathbf{X}^{-1}\mathbf{C} = -\|\mathbf{h}\|^{-2}\mathbf{h}^{\rm H}\mathbf{XC}(\mathbf{C}^{\rm H}\mathbf{XC})^{-1}$, which holds for any non-singular $N\times N$-matrix $\mathbf{X}$,

$$\mathbf{f}_{\rm DSMI}\mathbf{Y} = (\mathbf{h}^{\rm H}\widehat{\mathbf{R}}^{-1}\mathbf{h})^{-1}\mathbf{h}^{\rm H}\widehat{\mathbf{R}}^{-1}(\|\mathbf{h}\|^{-2}\mathbf{hh}^{\rm H} + \mathbf{CC}^{\rm H})\mathbf{Y} \tag{20}$$
$$= \|\mathbf{h}\|^{-2}\mathbf{h}^{\rm H}\mathbf{Y} - \|\mathbf{h}\|^{-2}\mathbf{h}^{\rm H}\widehat{\mathbf{R}}\mathbf{C}(\mathbf{C}^{\rm H}\widehat{\mathbf{R}}\mathbf{C})^{-1}\mathbf{C}^{\rm H}\mathbf{Y}. \tag{21}$$

Expanding $\widehat{\mathbf{R}}$ as $\widehat{\mathbf{R}} = K^{-1}\mathbf{YY}^{\rm H}$ and using that $\mathbf{C}^{\rm H}\mathbf{h} = \mathbf{0}$ leads to $\mathbf{f}_{\rm DSMI}\mathbf{Y} = \mathbf{P}_{\mathcal{B}}^{\perp}(\|\mathbf{h}\|^{-2}\mathbf{h}^{\rm H}\mathbf{Y})$, and furthermore

$$\mathbf{f}_{\rm DSMI}\mathbf{Y} = \mathbf{P}_{\mathcal{B}}^{\perp}(\mathbf{f}_{\rm MVDR}\mathbf{Y} + (\|\mathbf{h}\|^{-2}\mathbf{h}^{\rm H} - \mathbf{f}_{\rm MVDR})\mathbf{Y}) \tag{22}$$
$$= \mathbf{P}_{\mathcal{B}}^{\perp}(\mathbf{f}_{\rm MVDR}\mathbf{Y}) + \tag{23}$$
$$\mathbf{P}_{\mathcal{B}}^{\perp}((\|\mathbf{h}\|^{-2}\mathbf{h}^{\rm H} - \mathbf{f}_{\rm MVDR})(\|\mathbf{h}\|^{-2}\mathbf{hh}^{\rm H} + \mathbf{CC}^{\rm H})\mathbf{Y}).$$

The second line of (23) equals zero as $(\|\mathbf{h}\|^{-2}\mathbf{h}^{\rm H} - \mathbf{f}_{\rm MVDR})\mathbf{h} = \mathbf{0}$ and since $(\cdots)\mathbf{CC}^{\rm H}\mathbf{Y}$ is in the subspace $\mathcal{B}$.

Proof of proposition 2:
The subspace $\mathcal{B}^{\perp}$ is uniformly distributed on $G(K, K-N+1)$ since $\mathrm{span}\{\mathbf{c}_1^{\rm H}\mathbf{Y}, \ldots, \mathbf{c}_{N-1}^{\rm H}\mathbf{Y}\} = \mathrm{span}\{\mathbf{c}_1^{\rm H}\mathbf{W}, \ldots, \mathbf{c}_{N-1}^{\rm H}\mathbf{W}\}$ and $\mathbf{C}^{\rm H}\mathbf{W}$ does not change its distribution if post-multiplied by some unitary $K\times K$-matrix. The subspace defining matrix $\mathbf{C}^{\rm H}\mathbf{W}$ and the CCIN vector $\mathbf{f}_{\rm MVDR}\mathbf{W}$ are independent because the columns of $\mathbf{W}$ are independent Gaussian and $\mathbf{f}_{\rm MVDR}\mathbf{RC} = \mathbf{0}$. And the independence of $\mathbf{C}^{\rm H}\mathbf{W}$, $\mathbf{f}_{\rm MVDR}\mathbf{W}$ and $\mathbf{s}$ follows from the independence of $\mathbf{W}$ and $\mathbf{s}$.

Proofs of propositions 3 and 4: Left to the interested reader.

REFERENCES

[1] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York, NY: Wiley, 1980.

[2] H. L. Van Trees, *Optimum Array Processing*. New York, NY: Wiley, 2002.

[3] D. D. Feldman and L. J. Griffiths, "A projection approach for robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 42, pp. 867–876, Apr. 1994.

[4] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, pp. 397–401, July 1988.

[5] S. Shahbazpanahi, A. B. Gershman, Z.-Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Processing*, vol. 51, pp. 2257–2269, Sept. 2003.

[6] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 10, pp. 853–863, Nov. 1974.

[7] C. D. Richmond, "Derived PDF of maximum likelihood signal estimator which employs an estimated noise covariance," *IEEE Trans. Signal Processing*, vol. 44, pp. 305–315, Feb. 1996.

[8] B. D. Van Veen, "Adaptive convergence of linearly constrained beamformers based on the sample covariance matrix," *IEEE Trans. Signal Processing*, vol. 39, pp. 1470–1473, June 1991.

[9] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[10] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York, NY: Wiley, 1982.