

Reducing The Effect of Channel Time Variations in MIMO Broadcast Systems

Soroush Akhlaghi and Amir K. Khandani

Coding & Signal Transmission Laboratory (www.cst.uwaterloo.ca)

E&CE Department, Univ. of Waterloo, Waterloo, ON, Canada, N2L 3G1
 {soroush,khandani}@cst.uwaterloo.ca, Tel: (519) 885-1211

Abolfazl Falahati

Department of Electrical Engineering,

Iran University of Science & Technology (IUST), Tehran, IRAN

afalahati@iust.ac.ir

Abstract—A simple signaling method for Multiple Input Multiple Output (MIMO) broadcast channels with Zero-Forcing Beamforming (ZFBB) in the presence of outdated Channel State Information (CSI) is proposed. We address the problem of channel variation for the case of fixed-rate users using Block Coded Modulation (BCM). It is assumed that the perfect CSI is known at the beginning of each block, while the channel varies over time. The impact of the outdated CSI is evaluated both analytically, and experimentally through Monte-Carlo simulation. The effect of the channel aging is modeled as an additive noise with time-varying power. We introduce a new method based on Hadamard transformation in order to equalize the noise power across the coding block. Simulation results show that the proposed method results in a significant improvement in the error performance with a negligible increase in the complexity.

Index Items— MIMO broadcast channels, channel aging, Hadamard transformation.

I. INTRODUCTION

Multiple Input Multiple Output (MIMO) channels have received considerable attention due to their potential to dramatically increase the throughput of wireless systems [1], [2]. Recent information theoretic results [3]–[5] show that multiple transmit antennas at a Base Station (BS) can be used to simultaneously transmit data to several users by converting the channel into data pipes that overlap in both time and frequency [3]–[6]. It has been demonstrated that in a MIMO broadcast channel with K users (with single antenna) and a transmitter with M antennas, the sum-capacity grows linearly with the minimum of M and K [7].

It has been shown that the sum-capacity of MIMO broadcast channels can be achieved by applying Dirty-Paper Coding (DPC) at the transmitter [3]–[6]. However, existence

results do not easily translate to practical implementation. Beamforming is a suboptimal method with low complexity that achieves a large fraction of DPC performance [3]. Finding the optimal beamforming weight vectors, however, is a difficult task.

To maximize the sum throughput in a network with a large number of users, the BS should select the best set of users to communicate with. This form of selection diversity is called multiuser diversity [8] [9]. It has been shown that to achieve the optimum sum-capacity in a MIMO broadcast channel with M transmit antennas, the number of active users is between M and M^2 [10]. However, experimental results show that M users are adequate to achieve a significant portion of this capacity [11]. In this case, the optimum user selection requires the BS to perform an exhaustive search over the subsets of users with cardinality M which is computationally infeasible for large networks [12].

In [13], the authors propose a suboptimum algorithm, called random-beamforming, for user selection and show that this scheme asymptotically achieves the optimum sum-capacity. However, for practical number of users, random-beamforming does not perform well [14]. References [15], [16] use a very simple precoding scheme of Zero-Forcing Beamforming (ZFBB), where the weight vectors are chosen to avoid interference among users. This method selects a group of semi-orthogonal users in a greedy fashion, and then the beamforming weights are computed by inverting the corresponding channel matrix. As a result, the downlink is decomposed into M orthogonal sub-channels between the BS and the uncoordinated receivers. This method has two shortcomings: (i) It wastes the transmit power as the channel matrix of the selected users may be far from orthogonal, and (ii) It suffers from channel aging, resulting in interference among users. The current article addresses the above shortcomings by: (i) Selecting an appropriate set of users that are nearly orthogonal, and (ii) Applying a Hadamard transformation to the coding block to reduce the effect of

This work is financially supported by Communications and Information Technology Ontario (CITO), Nortel Networks, and Natural Sciences and Engineering Research Council of Canada (NSERC).

the channel aging. The use of the Hadamard transformation averages the impact of the channel aging over all symbols in a given block. Relying on pair-wise error probability, we show that the use of the Hadamard transformation is nearly optimal in reducing the effect of channel aging.

This article is organized as follows: Section II contains a short overview and some background information on MIMO broadcast channels. In Section III, we investigate the effect of the channel aging in MIMO broadcast channels. Section IV discusses BCM and its performance evaluation for a time-varying channel. In Section V, the proposed precoding strategy is introduced. Numerical results are provided in Section VI. Finally, Section VII concludes the paper. Throughout the paper, vectors are boldfaced lowercase and matrices are boldfaced uppercase, respectively. $x^{(i)}$ is the i 'th element of vector \mathbf{x} and $[\mathbf{X}]_{i,j}$ is the (i,j) 'th element of matrix \mathbf{X} . $\mathcal{I}(\mathbf{x})$ and $\mathcal{R}(\mathbf{x})$ are the real and imaginary parts of vector \mathbf{x} , respectively, and $(\cdot)^*$ denotes the conjugate transpose.

II. SYSTEM MODEL

We consider a MIMO broadcast channel with M transmit antennas at the BS and K ($K \geq M$) users, each with a single receive antenna¹. Block Coded Modulation (BCM) and ZFBF is employed at the BS. It is assumed that perfect CSI is available at the start of the coding block, while the CSI varies over time. The BS selects M users (among the total K users), indexed by $\pi(j)$, $j = 1, \dots, M$, in a greedy fashion such that the corresponding vectors are nearly orthogonal and have a large norm. The BS transmits M simultaneous coded data streams (with equal rate) to the selected users by using a precoding matrix which is the inverse of the corresponding channel matrix. The mathematical model of the system at the t 'th channel use is given by (the i 'th row in this matrix equation corresponds to the i 'th user),

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{W} \mathbf{x}_t + \mathbf{n}_t, \quad 1 \leq t \leq T \quad (1)$$

where $\mathbf{H}_t \in \mathcal{C}^{M \times M}$ and $\mathbf{W} \in \mathcal{C}^{M \times M}$ denote the channel matrix from the BS to the selected users and the beamforming matrix, respectively, $\mathbf{x}_t \in \mathcal{C}^{M \times 1}$ contains the transmitted signals, $\mathbf{y}_t \in \mathcal{C}^{M \times 1}$ contains the received signals, $\mathbf{n}_t \in \mathcal{C}^{M \times 1}$ is circularly symmetric complex Gaussian noise with a zero-mean and unit-variance, i.e., $\mathbf{n}_t \sim \mathcal{N}_c(\mathbf{0}, \mathbf{I}_M)$, where \mathbf{I}_M denotes the M dimensional identity matrix. We assume flat, time correlated Rayleigh fading, where the Jakes model is used for numerical simulations.

The precoding matrix for ZFBF is selected according to the initial CSI, i.e.,

$$\mathbf{W} = \mathbf{H}_0^* (\mathbf{H}_0 \mathbf{H}_0^*)^{-1}, \quad (2)$$

where $(\cdot)^*$ denotes the conjugate transpose, and \mathbf{H}_0 is the channel matrix at time instant zero. For rectangular matrices, we have $\mathbf{W} = \mathbf{H}_0^{-1}$. We assume that $E[\mathbf{x}\mathbf{x}^*] = P\mathbf{I}_M$.

¹The extension result of this paper for users with multiple-receive antennas is given in [17].

The matrix \mathbf{H}_0 corresponds the initial CSI of the selected users, and can be uniquely decomposed as,

$$\mathbf{H}_0 = \mathbf{\Lambda} \mathbf{R} \mathbf{Q}, \quad (3)$$

where \mathbf{Q} defines an orthonormal basis composed of vectors $\left\{ \frac{\mathbf{g}_{\pi(1)}}{\|\mathbf{g}_{\pi(1)}\|}, \frac{\mathbf{g}_{\pi(2)}}{\|\mathbf{g}_{\pi(2)}\|}, \dots, \frac{\mathbf{g}_{\pi(M)}}{\|\mathbf{g}_{\pi(M)}\|} \right\}$, and \mathbf{R} , $\mathbf{\Lambda}$ are lower triangular and diagonal matrices, respectively [16]. Matrix \mathbf{R} is composed of elements,

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \epsilon_{2,1} & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \dots & 0 \\ \epsilon_{M,1} & \epsilon_{M,2} & \dots & \epsilon_{M,M-1} & 1 \end{pmatrix}, \quad (4)$$

where the off diagonal elements are equal to,

$$\epsilon_{i,j} = \frac{\mathbf{h}_{\pi(i)}^* \mathbf{g}_{\pi(j)}}{\|\mathbf{h}_{\pi(i)}\| \|\mathbf{g}_{\pi(j)}\|} \quad 1 \leq i \leq M, 1 \leq j \leq i-1, \quad (5)$$

where $\mathbf{h}_{\pi(i)}$ denotes the CSI of the i 'th user. Matrix $\mathbf{\Lambda}$ is composed of elements,

$$\mathbf{\Lambda} = \text{diag}\{\|\mathbf{g}_{\pi(1)}\|, \|\mathbf{g}_{\pi(2)}\|, \dots, \|\mathbf{g}_{\pi(M)}\|\}, \quad (6)$$

where $\|\mathbf{g}_{\pi(i)}\|$ denotes the gain of the i 'th user in the null-space of the previously selected users [16].

Using the above matrix decomposition, the ZFBF precoding matrix is equal to,

$$\mathbf{W} = \mathbf{H}_0^{-1} = \mathbf{Q}^* \mathbf{R}^{-1} \mathbf{\Lambda}^{-1}, \quad (7)$$

and the received signal is,

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{W} \mathbf{x}_t + \mathbf{n}_t, \quad 1 \leq t \leq T. \quad (8)$$

III. EFFECT OF CHANNEL AGING ON THE RECEIVED SIGNAL

We use the following expression as statistical model for the time variations of the channel:

$$\mathbf{H}_t = \rho_t \mathbf{H}_0 + \sqrt{1 - \rho_t^2} \mathbf{E}_t, \quad 1 \leq t \leq T, \quad (9)$$

where ρ_t denotes the correlation coefficient between the time instants 0 and t , \mathbf{E}_t is a circularly symmetric complex Gaussian matrix with i.i.d entries, i.e. $[\mathbf{E}_t]_{i,j} \sim \mathcal{N}_c(0, \frac{1}{M})$. It is assumed that the matrices \mathbf{E}_t , $t = 1, \dots, T$, are uncorrelated (independent). The received signal is equal to,

$$\mathbf{y}_t = \rho_t \mathbf{x}_t + \sqrt{1 - \rho_t^2} \mathbf{E}_t \mathbf{H}_0^{-1} \mathbf{x}_t + \mathbf{n}_t. \quad (10)$$

Assuming that the correlation coefficient ρ_t is known at the receivers ($\rho_t \neq 0$), it follows:

$$\begin{aligned} \tilde{\mathbf{y}}_t &= \mathbf{x}_t + \frac{\sqrt{1 - \rho_t^2}}{\rho_t} \mathbf{E}_t \mathbf{H}_0^{-1} \mathbf{x}_t + \frac{1}{\rho_t} \mathbf{n}_t \\ &= \mathbf{x}_t + \mathbf{i}_t + \frac{1}{\rho_t} \mathbf{n}_t. \end{aligned} \quad (11)$$

Since each active user knows only the corresponding beamforming direction, the matrix \mathbf{H}_0 is unknown at the receivers. As a result, the channel tracking techniques [18] are not applicable here and the second term in (11) is treated as an additive interference term.

Since the transmitted vectors \mathbf{x}_t are uncorrelated across time, i.e., $E[\mathbf{x}_l \mathbf{x}_m^*] = \mathbf{0}$, for $l \neq m$, it follows:

$$\begin{aligned} E[\mathbf{i}_l \mathbf{i}_m^*] &= E_{\mathbf{E}_l, \mathbf{E}_m} \left[E_{\mathbf{x}_l, \mathbf{x}_m} [\mathbf{i}_l \mathbf{i}_m^* | \mathbf{E}_l, \mathbf{E}_m] \right] \\ &= E_{\mathbf{E}_l, \mathbf{E}_m} [\mathbf{E}_l \mathbf{H}_0^{-1} E[\mathbf{x}_l, \mathbf{x}_m^*] \mathbf{H}_0^{*-1} \mathbf{E}_m^*] \\ &= 0 \quad \text{for } l \neq m. \end{aligned} \quad (12)$$

Using (7) and (11), the interference power for the i 'th receiver at time instant t can be decomposed as,

$$\begin{aligned} [E[\mathbf{i}_t \mathbf{i}_t^*]]_{i,i} &= \frac{P}{M} \left\{ \frac{1 - \rho_t^2}{\rho_t^2} \sum_{i=1}^M \frac{1}{\|\mathbf{g}_{\pi(i)}\|^2} \right. \\ &\quad \left. + \frac{1 - \rho_t^2}{\rho_t^2} \sum_{i=1}^M \frac{1}{\|\mathbf{g}_{\pi(i)}\|^2} \sum_{j=i+1}^M \epsilon_{j,i}^2 \right\} \\ &= \frac{P}{M} (I_{t,1} + I_{t,2}), \end{aligned} \quad (13)$$

The performance loss due to the interference is caused by two factors : (i) The channel aging due to the time varying fading (term $I_{t,1}$ in (13)), and (ii) The lack of orthogonality between beamforming directions (term $I_{t,2}$ in (13)).

Using (11) and (13), it follows:

$$\begin{aligned} SINR_t &= \frac{P}{\frac{P}{M} (I_{t,1} + I_{t,2}) + \frac{\sigma^2}{\rho_t^2}} \\ &= \frac{1}{\frac{I_{t,1} + I_{t,2}}{M} + \frac{1}{\rho_t^2 SINR}}. \end{aligned} \quad (14)$$

IV. ERROR PERFORMANCE

As discussed earlier, the effect of aging can be modeled as a time-varying channel in which each active user experiences an extra additive noise term with unknown and varying power at each time instant. The baseband model of this system over one coding block of length T is given by,

$$\mathbf{r} = \sqrt{P} \mathbf{s} + \mathbf{z}, \quad (15)$$

where \mathbf{s} , \mathbf{r} and \mathbf{z} are T dimensional vectors represented the transmitted, the received and the noise, respectively, and $E[\mathbf{z} \mathbf{z}^*] = \text{diag}[\sigma_1^2, \sigma_2^2, \dots, \sigma_T^2]$. Let \mathbf{c} and $\hat{\mathbf{c}}$ denote two distinct codewords, mapped to the modulation vectors \mathbf{s}_c and $\mathbf{s}_{\hat{c}}$, respectively. The Pair-wise Error Probability (PEP) [19], $\Pr(\mathbf{s}_c \rightarrow \mathbf{s}_{\hat{c}} | \mathbf{s}_c)$, is the probability that the transmitted modulation vector \mathbf{s}_c is received incorrectly as $\mathbf{s}_{\hat{c}}$. Assume that \mathbf{c} and $\hat{\mathbf{c}}$ have a Hamming distance $d(\mathbf{c}, \hat{\mathbf{c}}) = k$ and let $\dot{\mathbf{s}} = [s_c^{(i_1)}, s_c^{(i_2)}, \dots, s_c^{(i_k)}]$ and $\ddot{\mathbf{s}} = [s_{\hat{c}}^{(i_1)}, s_{\hat{c}}^{(i_2)}, \dots, s_{\hat{c}}^{(i_k)}]$ denote the symbols of corresponding block in which the two modulation vectors are different. The receiver performs maximum likelihood detection by assuming a Gaussian noise with

a constant power throughout the block (minimum distance decoding). Hence, the PEP ($P_E = \Pr(\mathbf{s}_c \rightarrow \mathbf{s}_{\hat{c}} | \mathbf{s}_c)$) can be evaluated as (Appendix A):

$$P_E = Q\left(\frac{\sqrt{P} \|\mathbf{d}\|^2}{2\sqrt{\mathcal{R}(\mathbf{d}^T) \mathbf{M}_d \mathcal{R}(\mathbf{d}) + \mathcal{I}(\mathbf{d}^T) \mathbf{M}_d \mathcal{I}(\mathbf{d})}}\right) \quad (16)$$

where \mathbf{M}_d is the correlation matrix of real and imaginary parts of vector \mathbf{z} in the positions that the two modulation vectors are different, and $\mathbf{d} = \dot{\mathbf{s}} - \ddot{\mathbf{s}}$ is the corresponding difference.

For BPSK and QPSK modulation, the squared Euclidean distance between a pair of coded modulation vectors is proportional to the product of the constellation minimum Euclidean distance (\mathcal{X}_{min}^2) and the Hamming distance between the two codewords (k), i.e., $\|\mathbf{d}\|^2 = k \mathcal{X}_{min}^2$. In what follows, it is shown that (assuming BPSK/QPSK modulation) by employing Hadamard transformation at both transmitters and receivers, the resulting PEP is lower than the PEP of the conventional method.

Consider two channels, one with constant noise power across time and correlation matrix $\sigma^2 \mathbf{I}_T$, and the second with an increasing noise power across time with correlation matrix $\text{diag}[\sigma_1^2, \sigma_2^2, \dots, \sigma_T^2]$. Also, $\sigma^2 = \frac{\sum_{i=1}^T \sigma_i^2}{T}$ which means that the average energy of the noise is the same in both cases. The average PEP for two codewords with Hamming distance $d(\mathbf{c}, \hat{\mathbf{c}}) = k$ in the first channel is independent from the positions in which the two modulation vectors are different. This depends on the Hamming distance as follows:

$$P_{E_1} = Q\left(\sqrt{\frac{k \mathcal{X}_{min}^2 P}{4 \sigma^2}}\right). \quad (17)$$

For the channel with varying noise power, the average PEP in the presence of random interleaving would be,

$$P_{E_2} = \frac{1}{\binom{T}{k}} \sum_{m=1}^{\binom{T}{k}} Q\left(\sqrt{\frac{k \mathcal{X}_{min}^2 P}{4 \mu_m^2}}\right), \quad (18)$$

where, $\mu_m^2 = \frac{\sum_{j=1}^k \sigma_{m_j}^2}{k}$ is the average of noise power in the positions in which the two modulation vectors are different (these positions are indexing by m_j for $j = 1, \dots, k$).

Denoting $\overline{\mu^2} = \frac{1}{\binom{T}{k}} \sum_{m=1}^{\binom{T}{k}} \mu_m^2$, one can see

$$\begin{aligned}
\overline{\mu^2} &= \frac{1}{\binom{T}{k}} \sum_{m=1}^{\binom{T}{k}} \mu_m^2 \\
&= \sum_{m=1}^{\binom{T}{k}} \sum_{j=1}^k \frac{\sigma_{m_j}^2}{\binom{T}{k} k} \\
&= \sum_{l=1}^T \frac{\binom{T-1}{k-1} \sigma_l^2}{\binom{T}{k} k} \\
&= \sum_{l=1}^T \frac{\sigma_l^2}{T} \\
&= \sigma^2,
\end{aligned} \tag{19}$$

where (19) is computed relying on the observation that each value of the noise variance participate in exact $\binom{T-1}{k-1}$ equations. Using (18), (20) and after some manipulations showing that the Q-function is a convex function with respect to the noise power, if the average SINR exceeds a certain value (Appendix B), we obtain:

$$\begin{aligned}
P_{E_2} &\geq Q\left(\sqrt{\frac{k\mathcal{X}_{min}^2 P}{4\mu^2}}\right) \\
&= Q\left(\sqrt{\frac{k\mathcal{X}_{min}^2 P}{4\sigma^2}}\right) \\
&= P_{E_1}.
\end{aligned} \tag{21}$$

Equation (21) implies that in the case of the same average noise power, the performance of coded modulation techniques over an AWGN channel with fixed noise power across time results in a better performance as compared to the channel with different noise power.

V. USING HADAMARD MATRIX TO DECREASE THE EFFECT OF THE CHANNEL AGING

In this section, it is shown that applying Hadamard transformation to the vector of transmitted signals results in two benefits: (i) All the transmitted symbols will experience the same noise power, which is equal to the average noise power of the original channel. (ii) Relying on the central limit theorem, the resulting interference plus noise terms becomes nearly Gaussian, and hence, the common minimum distance decoding can be used for maximum likelihood detection. In this case, the transmitted vector is computed as,

$$\begin{aligned}
\mathbf{X} &= [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T] \\
\mathbf{U} &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_T] \\
\mathbf{X} &= \frac{1}{\sqrt{T}} \mathbf{U} \mathbf{\Xi},
\end{aligned} \tag{22}$$

where \mathbf{U} and $\mathbf{\Xi}$ are M by T matrices of the information vectors and the Hadamard transformation, respectively, and $\frac{1}{\sqrt{T}}$ is a scaling factor to normalize power.

Let \mathbf{r} represent the received vector corresponding to the k 'th user and $\hat{\mathbf{r}}$ represent the vector in which the t 'th entry corresponds to the $\frac{r^{(t)}}{\rho_t}$. After applying the Hadamard transformation at the receiver, we obtain:

$$\begin{aligned}
\tilde{\mathbf{r}} &= \frac{1}{\sqrt{T}} \hat{\mathbf{r}} \mathbf{\Xi} \\
\tilde{\mathbf{r}} &= \mathbf{u} + \tilde{\mathbf{i}} + \tilde{\mathbf{n}},
\end{aligned} \tag{23}$$

where $\tilde{\mathbf{i}} = \frac{1}{\sqrt{T}} \mathbf{i} \mathbf{\Xi}$ and $\tilde{\mathbf{n}} = \frac{1}{\sqrt{T}} \mathbf{n} \mathbf{\Xi}$ are the resulting interference and noise by applying the Hadamard transformation at the receiver, respectively. As a result, the correlation matrix of the resulting interference term would be:

$$E[\tilde{\mathbf{i}} \tilde{\mathbf{i}}^*] = \frac{1}{T} \mathbf{\Xi}^t E[\mathbf{i} \mathbf{i}^*] \mathbf{\Xi}, \tag{24}$$

where \mathbf{i} and $\tilde{\mathbf{i}}$ are the vectors (in time) of the interference prior and after applying the Hadamard transformation, respectively. Equations (12) and (13) show that the correlation matrix $E[\mathbf{i} \mathbf{i}^*]$ is diagonal with diagonal entries $[E[\mathbf{i} \mathbf{i}^*]]_{t,t} = \frac{P}{M} (I_{t,1} + I_{t,2})$. As a result, the diagonal elements of the resulting correlation matrix are equal to each other, i.e.,

$$[E[\tilde{\mathbf{i}} \tilde{\mathbf{i}}^*]]_{i,i} = \frac{P}{M} \frac{\sum_{t=1}^T I_{t,1} + I_{t,2}}{T}, \tag{25}$$

and by following the same approach, the resulting noise power would be,

$$[E[\tilde{\mathbf{n}} \tilde{\mathbf{n}}^*]]_{i,i} = \frac{\sigma^2}{T} \sum_{t=1}^T \frac{1}{\rho_t^2}, \tag{26}$$

In summary, the Hadamard transformation converts the channel into a new channel for which the total noise power across time remains constant. The resulting noise is no longer independent. However, simulation results show that this correlation is negligible in most situations.

VI. NUMERICAL RESULTS

We consider a broadcast channel with $M = 4$ transmit antennas at the BS and $K = 32$ users, each equipped with one receive antenna. The Jakes fading model is being used to simulate the time correlation of the underlying fading process. $M = 4$ users are serviced at a given time and the information bits of each active user are first encoded by a Turbo code with block length 2048 and rate $\frac{1}{2}$, and then mapped to the QPSK constellation points using an interleaver. The resulting stream after applying the Zero-Forcing precoding matrix is transmitted through the channel. Average Bit Error Rate (BER) and Frame Error Rate (FER) for various normalized Doppler frequencies are plotted in Fig.1 and Fig.2, respectively. Simulation results show that for the case of Quasi-Static fading channel, i.e. $f_d T = 0$, employing Hadamard transformation does not result in any improvement. However, as the normalized Doppler frequency increases, the proposed scheme results in a remarkable improvement in the performance.

APPENDIX

A. The Pair-wise Error Probability

Assume that the received signal is given by (15). Consider two codewords \mathbf{c} and $\hat{\mathbf{c}}$ with Hamming distance k which are mapped to the modulation vectors \mathbf{s}_c and $\mathbf{s}_{\hat{c}}$, respectively. Let

$$\begin{aligned} \dot{\mathbf{s}} &= [s_c^{(i_1)}, s_c^{(i_2)}, \dots, s_c^{(i_k)}] \\ \ddot{\mathbf{s}} &= [s_{\hat{c}}^{(i_1)}, s_{\hat{c}}^{(i_2)}, \dots, s_{\hat{c}}^{(i_k)}] \end{aligned} \quad (27)$$

denote the positions in which the two modulation vectors are different. Hence, the PEP can be evaluated as

$$\begin{aligned} P_E &= \Pr(\mathbf{s}_c \rightarrow \mathbf{s}_{\hat{c}} | \mathbf{s}_c) \\ &= \Pr\left(\sum_{j=1}^k |r^{(i_j)} - \sqrt{P}s_c^{(i_j)}|^2 \geq \sum_{j=1}^k |r^{(i_j)} - \sqrt{P}s_{\hat{c}}^{(i_j)}|^2\right) \\ &= \Pr\left(\sum_{j=1}^k \mathcal{R}\{z^{(i_j)*}(s_{\hat{c}}^{(i_j)} - s_c^{(i_j)})\}\right) \\ &\geq \frac{\sqrt{P}}{2} \sum_{j=1}^k |s_{\hat{c}}^{(i_j)} - s_c^{(i_j)}|^2 \\ &= \Pr\left(\sum_{j=1}^k \mathcal{R}(z^{(i_j)})\mathcal{R}(d^{(j)}) + \mathcal{I}(z^{(i_j)})\mathcal{I}(d^{(j)})\right) \\ &\geq \frac{\sqrt{P}}{2} \sum_{j=1}^k |d^{(j)}|^2 \\ &= \Pr(u \geq \frac{\sqrt{P}}{2} \|\mathbf{d}\|^2) \\ &= Q\left(\frac{\sqrt{P} \|\mathbf{d}\|^2}{2\sigma_u}\right) \end{aligned} \quad (28)$$

where the following definitions are being used for the sake of simplicity

$$\begin{aligned} d^{(j)} &= s_{\hat{c}}^{(i_j)} - s_c^{(i_j)} \\ u &= \sum_{j=1}^k \mathcal{R}(z^{(i_j)})\mathcal{R}(d^{(j)}) + \mathcal{I}(z^{(i_j)})\mathcal{I}(d^{(j)}) \\ &= \mathcal{R}(\mathbf{d}^T)\mathcal{R}(\mathbf{z}) + \mathcal{I}(\mathbf{d}^T)\mathcal{I}(\mathbf{z}), \end{aligned} \quad (29)$$

where the random variable u has the following variance,

$$\begin{aligned} \sigma_u^2 &= E[(\mathcal{R}(\mathbf{d}^T)\mathcal{R}(\mathbf{z}) + \mathcal{I}(\mathbf{d}^T)\mathcal{I}(\mathbf{z})) \\ &\quad \cdot (\mathcal{R}(\mathbf{z}^T)\mathcal{R}(\mathbf{d}) + \mathcal{I}(\mathbf{z}^T)\mathcal{I}(\mathbf{d}))] \\ &= \mathcal{R}(\mathbf{d}^T)E[\mathcal{R}(\mathbf{z})\mathcal{R}(\mathbf{z}^T)]\mathcal{R}(\mathbf{d}) \\ &\quad + \mathcal{I}(\mathbf{d}^T)E[\mathcal{I}(\mathbf{z})\mathcal{I}(\mathbf{z}^T)]\mathcal{I}(\mathbf{d}) \\ &= \mathcal{R}(\mathbf{d}^T)\mathbf{M}_d\mathcal{R}(\mathbf{d}) + \mathcal{I}(\mathbf{d}^T)\mathbf{M}_d\mathcal{I}(\mathbf{d}) \end{aligned} \quad (30)$$

where, \mathbf{M}_d is the correlation matrix of real and imaginary parts of vector \mathbf{z} in the positions that the two modulation vectors are different. As a result, the PEP would be

$$P_E = Q\left(\frac{\sqrt{P} \|\mathbf{d}\|^2}{2\sqrt{\mathcal{R}(\mathbf{d}^T)\mathbf{M}_d\mathcal{R}(\mathbf{d}) + \mathcal{I}(\mathbf{d}^T)\mathbf{M}_d\mathcal{I}(\mathbf{d})}}\right) \quad (31)$$

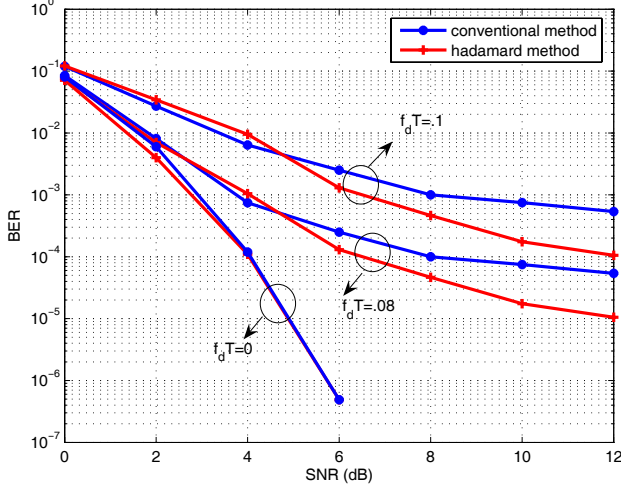


Fig. 1. BER performance of the proposed method as compared to the conventional method

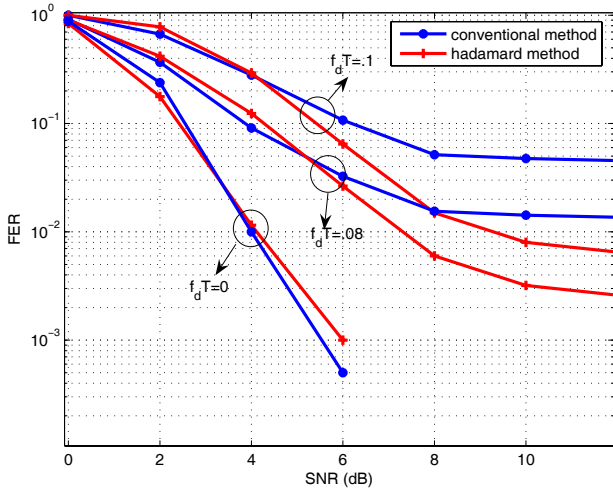


Fig. 2. FER performance of the proposed method as compared to the conventional method

VII. CONCLUSIONS

The impact of the CSI aging in MIMO broadcast systems with ZFBF is modeled as a time varying channel in which the noise power varies across time. Then, the effect of coded modulation is studied for both the underlying channel and a channel with fixed noise power across time obtained by applying a Hadamard transformation. Analytical and simulation results show that the proposed algorithm results in a significant improvement in the performance as compared to the conventional methods, with a negligible increase in the complexity.

B. On the Convexity of Q-Function with respect to the noise power

In this appendix, we will show that for the cases that the SNR exceeds a certain value which depends on the type of modulation method, the Q-function with respect to the noise power is a convex function. One can easily show that the Q-function is continuous and differentiable with respect to the noise power. Taking Hessian of this function with respect to the σ^2 , yields:

$$Q\left(\sqrt{\frac{c^2}{\sigma^2}}\right) = \int_{\frac{c}{\sigma}}^{\infty} e^{-\frac{x^2}{2}} dx$$

$$\nabla_{\sigma^2} Q\left(\sqrt{\frac{c^2}{\sigma^2}}\right) = \frac{c}{4\sigma^5} \left(\frac{c^2}{\sigma^2} - 3\right) e^{-\frac{c^2}{2\sigma^2}} \quad (32)$$

As a result, for $\sigma \leq \frac{c}{\sqrt{3}}$ the Q-function is convex. For coded modulation systems,

$$c_{min}^2 = Pd_{min}\mathcal{X}_{min}^2 \quad (33)$$

where, d_{min} and \mathcal{X}_{min} denote the minimum hamming distance of corresponding coding scheme and the normalized minimum Euclidean distance of the underlying modulation, respectively. The acceptable SNR region is defined as the region in which the Q-function is convex. As a result, the acceptable SNR region is lower bounded as

$$SNR \geq \frac{3}{d_{min}\mathcal{X}_{min}^2} \quad (34)$$

One can easily observe that for good coding schemes (with acceptable minimum Hamming distance), the practical SNR regions reside in this area.

REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.
- [2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, pp. 585–595, Nov. 1999.
- [3] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1691–1706, July 2003.
- [4] W. Yu and J. Cioffi, "Sum capacity of vector Gaussian broadcast channels," *IEEE Trans. Inform. Theory*, 2004.
- [5] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2658–2668, Oct. 2003.
- [6] P. Viswanath and D.N.C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1912 – 1921, Aug. 2003.
- [7] M. Sharif and B. Hassibi, "A comparison of time-sharing, dpc, and beamforming for mimo broadcast channels with many users," *IEEE Trans. Commun.*, submitted for publication.
- [8] H. Hochwald and S. Vishwanath, "Space-time multiple access: Linear growth in the sum rate," *Proceedings of 42nd Allerton Conference, Allerton, IL*, Oct 2002.
- [9] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. on Inform. theory*, vol. 48, pp. 1277–1294, 2002.
- [10] M. A. Maddah-Ali, M. Ansari, and Amir K. Khandani, "An efficient signaling method over MIMO broadcast channels," *Proceedings of 42nd Allerton Conference, Monicello, IL*, Sept 2004.
- [11] Mohammad Ali Maddah-Ali, Mehdi Ansari Sadrabadi, and Amir K. Khandani, "An efficient signaling scheme for mimo broadcast systems: Design and performance evaluation," *IEEE Trans. on Info. Theory*, Submitted for publication, 2005.
- [12] S. Venkatesan and H. Huang, "System capacity evaluation of multiple antenna systems using beamforming and dirty paper coding," *Bell Labs, unpublished*.
- [13] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Trans. on Info. Theory*, 2003.
- [14] T. Yoo and A. Goldsmith, "Optimality of zero-forcing beamforming with multiuser diversity," *ICC 2005, submitted for publication*.
- [15] Q.H. Spencer, A.L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 52, pp. 461–471, Feb. 2004.
- [16] Taesang Yoo and Andrea Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *To appear: IEEE JSAC Special Issue on 4G Wireless Systems*.
- [17] Soroush Akhlaghi, Amir K. Khandani, and A. Falahati, "Reducing the effect of channel time variations in mimo broadcast systems," *Technical Report is available at www.cst.uwaterloo.ca*, 2006.
- [18] j. R. Zeidler P. Yee S. Haykin, A. H. Sayed and P. C. Wei, "Adaptive tracking of linear time-varient systems by extended rls algorithms," *IEEE Trans. Signal Processing*, 1997.
- [19] C. Tellambura, "Evaluation of exact union bound for trellis-coded modulations over fading channels," *IEEE Trans. on Communications*, vol. 44, Dec 1999.