Petar SPALEVIĆ [1], Veljko STANKOVIĆ[1], Mihajlo STEFANOVIĆ[2], Stefan PANIĆ[2] Ana SAVIĆ[3]

State University of Novi Pazar(1), Faculty of Electronic Engineering, University of Niš, Serbia(2),
High school of Electrical Engineering and Computing Vocational Study, Belgrade, Serbia (3)

Minimum mean-squared error multi-user MIMO receive filtering

Abstract. *In this paper we use two modifications of the minimum mean-squared-error (MMSE) cost function to derive new multi-user (MU) MIMO receive algorithms. We derive algorithms that have lower complexity than the lattice reduction techniques and are able to extract full diversity inherent in the system. The resulting algorithms work in the overloaded systems, i.e. can process signals at the user terminals (UTs) equipped with the arbitrary number of antennas, and have no bit error floor.*

Streszczenie. *W pracy wykorzystano dwie modyfikacje funkcji MMSE do otrzymania nowego algorytmu otrzymywania danych przez wielu użytkowników. Zaproponowany algorytm jest mniej złożony niż technika redukcji siatki i jest w stanie wydobyć różnorodność właściwą dla systemu. Możliwe jest przetwarzanie sygnału w terminalu użytkownika wyposażonego w dowolna liczbę anten. (**Minimum błędu średnio-kwadratowego w filtrowaniu przesyłu MIMO**)*

Keywords: MIMO systems, Multi-user MIMO, SDMA, receive signal processing
Słowa kluczowe: MIMO, przetwarzanie sygnału, przesyłsygnału.

Introduction

Multi-user (MU) multiple-input, multiple-output (MIMO) systems are a key component of the future wireless communication systems because of their promising improvements in terms of performance and bandwidth efficiency [1]. Such systems have the potential to combine high capacity achievable with MIMO processing with the benefits of space division multiple access (SDMA). It has been shown that linear increase in capacity with the number of antennas in a multiuser scenario can be achieved only by using multi-user MIMO precoding and detection. While maximum likelihood (ML) decoding is optimal, multiple access channel (MAC) sum-rate capacity can also be achieved via a minimum mean-square-error (MMSE) receiver with successive interference cancellation (SIC). However, MMSE receive filtering suffers from a performance loss when it is used to suppress the interference between the signals transmitted from the antennas located at the same user terminal (UT). Since each user can coordinate the processing over all of its antennas, we can combine the signals transmitted from different antennas at the same UT in order to extract higher spatial processing gains. Space-time coding (STC), and spatial multiplexing (SMUX) , provide full diversity and achieve high data rates over MIMO channels, respectively. Group layered space-time architecture has been proposed to achieve a better trade off between multiplexing gain and diversity gain. In this case, the transmit stream is partitioned into different groups, and in each group, STC is applied . It can be regarded as a combination of STC and SMUX. Neither scheme requires channel state information (CSI) at the transmitter and therefore are suboptimal.

It is well known that K co-channel users, each equipped with N antennas, and transmitting uncorrelated signals can be detected with N-order diversity gains if the receiver is equipped with $N(K-1)+1$ antennas. However, the structure of STC can be exploited to reduce the number of receive antennas. It was shown in [2] that only K receive antennas are needed to provide N-order diversity gains and suppress $K-1$ co-channel users. A simple interference cancellation scheme for two co-channel users employing STC scheme was already developed in [2]. It was shown that by using two transmit antennas for each user and two receive antennas at the base station (BS), it is possible to double the system capacity by applying only linear processing at the receiver. However, since the total number of antennas at the UTs is greater than the number of the antennas at the BS, the bit error rate (BER) curve for all of these algorithms has an error floor. One way to separate users on the uplink in a MU MIMO scenario is by using another dimension, like time, frequency or code. A different approach was used in [3], [4] where the authors propose propose two algorithms for detection of MU space-time coded signals. In the first case, an intuitive approach for the detection of MU space-time coded signals is used where the co-channel interference is suppressed first and then the space-time decoding is performed. Another alternative is to do space-time decoding first, and then suppress the co-channel interference. Also, in second case an error floor occurs when the total number of antennas at the UTs is greater than the number of antennas at the BS. In [5] the authors introduce a novel turbo space-time equalizer. In the paper, the authors assume that all users have the same number of antennas, symbol spaced channel model and simple transmission of different data symbols from different antennas at the same UT. The processing is done in time domain and it requires whitening of the spatial interference. The matrix used to generate receive filter is a linear combination of each UT's channel MIMO matrix. Dimensionality of this matrix is proportional to the number of channel taps, the number of UTs and the length of the receive filter. As a consequence, the complexity of this method becomes extremely prohibitive.

Our goal in this paper is to develop an algorithm for MU MIMO detection that will have lower complexity than the lattice reduction techniques, that will be able to extract full diversity inherent in the system, that will work in the overloaded systems, that can process signals at the UTs equipped with the arbitrary number of antennas, and that will have no bit error floor. [6] to derive new MU MIMO receive algorithms that provide higher array and diversity gain than group layered techniques. These cost functions are derived from a modified mean-squared-error (MSE) cost function. Some of the algorithms presented in this paper approach the optimum solution in simulations with significantly lower computational complexity than lattice reduction techniques. The MU MIMO receive matrix is generated in two steps. In the first step we suppress the multi-user interference (MUI) and in the second case we decode each UT signal assuming a set of orthogonal SU channels. This approach was used in [4] to derive a technique called successive MMSE (SMMSE) SIC filtering.

System model

We consider a MU MIMO uplink channel, where M_B antennas are located at the base station and M_{U_i} antennas are located at the i-th UT, $i = 1, 2, \ldots, K$. There are K users (or UTs) in the system. The total number of antennas at the

UTs is

$$M_U = \sum_{i=1}^{K} M_{U_i}.$$

A block diagram of such a system is depicted in Fig. **??**. We use the notation $\{M_{U_1}, \ldots, M_{U_K}\} \times M_B$ to describe the antenna configuration of the system. First, we assume frequency flat slow fading channels. In case of frequency selective channels, we assume transmission using OFDM where the same MIMO processing is performed on each subcarrier. Let the MIMO channel of user i be denoted as $\boldsymbol{H}_i \in \mathbb{C}^{M_{U_i} \times M_B}$. Then, the combined channel matrix is given by

$$\text{(1)} \qquad \boldsymbol{H} = \quad \boldsymbol{H}_1^T \quad \boldsymbol{H}_2^T \quad \cdots \quad \boldsymbol{H}_K^T \quad {}^T \in \mathbb{C}^{M_U \times M_B}.$$

Data vectors $\boldsymbol{x}_k \in \mathbb{C}^{r_k \times 1}$, $k = 1, \ldots, K$, for the K UTs are stacked in the vector $\boldsymbol{x} = \quad \boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_K^T \quad {}^T \in \mathbb{C}^{r \times 1}$. The received vector is given by

$$\text{(2)} \qquad \boldsymbol{y} = \boldsymbol{D} \quad \boldsymbol{H}^T \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{n}$$

where $\boldsymbol{y} = \quad \boldsymbol{y}_1^T \quad \cdots \quad \boldsymbol{y}_K^T \quad {}^T \in \mathbb{C}^{r \times 1}$ is the received data vector, $\boldsymbol{n} \in \mathbb{C}^{M_B \times 1}$ is the vector of the zero mean additive white Gaussian noise at the input of the receive antennas at the BS. The joint detection matrix at the BS and the precoding matrices at the UTs are denoted by $\boldsymbol{D}$ and $\boldsymbol{Q}_i$, respectively.

Let us define the joint detection matrix as

$$\text{(3)} \qquad \boldsymbol{D} = \quad \boldsymbol{D}_1^T \quad \boldsymbol{D}_2^T \quad \cdots \quad \boldsymbol{D}_K^T \quad {}^T \in \mathbb{C}^{r \times M_B}$$

where $\boldsymbol{D}_i \in \mathbb{C}^{r_i \times M_B}$ is the i-th user's detection matrix. Moreover, $r = \sum_{i=1}^{K} r_i \leq \operatorname{rank}(\boldsymbol{H}) \leq \min(M_U, M_B)$ is the total number of the transmitted data streams, whereas r_i is the number of data stream sequences transmitted from the i-th user. The precoding matrix $\boldsymbol{Q}$ can be written as

$$\text{(4)} \qquad \boldsymbol{Q} = \begin{bmatrix} \boldsymbol{Q}_1 & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{Q}_K \end{bmatrix} \in \mathbb{C}^{M_U \times r}$$

where $\boldsymbol{Q}_i \in \mathbb{C}^{M_{U_i} \times r_i}$ is the i-th user's precoding matrix.

Spatial detection of multi-user signals

In order to facilitate generalized design of the detection matrices in a MU MIMO scenario, we separate the multi-user interference (MUI) suppression and the system performance optimization. Therefore, the detection matrix design is performed in two steps. In the first step we balance the MUI suppression which is achieved by reducing the overlap of the row spaces spanned by the effective channel matrices of different users and any MIMO processing gain which requires that the users use as much as possible the available subspaces. In the second step we optimize the system performance assuming parallel single-user (SU) MIMO channels. Thus, the detection matrix in equation (2) is factored as

$$\text{(5)} \qquad \boldsymbol{D} = \boldsymbol{D}_b \boldsymbol{D}_a$$

where

$$\text{(6)} \qquad \boldsymbol{D}_a = \quad \boldsymbol{D}_{a_1}^T \quad \boldsymbol{D}_{a_2}^T \quad \cdots \quad \boldsymbol{D}_{a_K}^T \quad {}^T \in \mathbb{C}^{M_x \times M_B}$$

and

$$\text{(7)} \qquad \boldsymbol{D}_b = \begin{bmatrix} \boldsymbol{D}_{b_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_{b_2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{D}_{b_K} \end{bmatrix} \in \mathbb{C}^{r \times M_x}$$

with $\boldsymbol{D}_{a_i} \in \mathbb{C}^{M_x \times M_B}$ and $\boldsymbol{D}_{b_i} \in \mathbb{C}^{r_i \times M_x}$, $M_x \leq r_i$ and $M_x = \sum_{i=1}^{K} M_{x_i}$ depending on the specific choice of the algorithm to obtain $\boldsymbol{D}_a$. Matrix $\boldsymbol{D}_a$ is used first to suppress the MUI interference, and then the matrix $\boldsymbol{D}_b$ is used to optimize the system performance according to a specific criterion assuming that the MU MIMO channel has been transformed into a set of parallel SU MIMO channels. Note that unless at the UTs we use the STCs, the choice of the matrix $\boldsymbol{D}$ will depend on the matrix $\boldsymbol{Q}$. Moreover, the optimum choice of matrices $\boldsymbol{Q}$ and $\boldsymbol{D}_b$ depends on the matrix $\boldsymbol{D}_a$. In one case, the matrices $\boldsymbol{Q}_i$ could be desiged at the UTs in a distributed way. In this case, we would have to estimate these matrices at the BS in order to perform the optimum decoding. The other option is to design matrices $\boldsymbol{D}$ and $\boldsymbol{Q}$ at the BS and then feed forward matrices $\boldsymbol{Q}_i$ to the UTs. Over-the-air CSI at the BS is used for spatial processing on the downlink and the uplink and for resource allocation. In order to perform optimum detection BS needs to estimate over-the-air CSI and acquire the information about the spatial processing at the UTs. We assume that the BS and the UTs operate in time division duplexing (TDD). The cost of acquiring the CSI at the transmitter is much lower in a TDD system than in a frequency division duplexing (FDD) system, because in a TDD system it is possible to exploit the estimated uplink channel for the downlink transmission due to the reciprocity principle. In the case of FDD we would have to rely on the feedback of the CSI.

We will first obtain matrix $\boldsymbol{D}_a$ assuming matrices $\boldsymbol{D}_{b_i}$ and $\boldsymbol{Q}_i$ are unitary. This assumption corresponds to the assumption that the users use the maximum available channel row subspace for communication. Matrix $\boldsymbol{D}_a$ is used to suppress MUI, and by assuming that the UTs use the maximum available subspace to communicate, the row subspace of each UT's effective channel matrix after the detection, $\boldsymbol{H}_i \boldsymbol{D}_{a_i}^T$, will span the maximum of the available row subspace spanned by the matrix $\boldsymbol{H}_i$ while keeping the MUI to a minimum. After we have defined the matrix $\boldsymbol{D}_a$, matrices $\boldsymbol{D}_{b_i}$ and $\boldsymbol{Q}_i$ are designed according to the specific optimization criterion and UTs' qualty-of-service (QoS) requirements, assuming all effective UTs' channel matrices $\boldsymbol{H}_i \boldsymbol{D}_{a_i}^T$ are orthogonal. We will consider two different cost functions to design matrices $\boldsymbol{D}_{a_i}$. These cost functions represent two modifications of the minimum mean-squared-error (MMSE) cost function. They are introduced in [6] for the generalized design of the downlink MU MIMO precoding matrices. In this paper we will use the same approach to derive in the similar way the MU MIMO detection matrices that will allow easy adaptation to any optimization criterion or the combination of different optimization criteria for different users. For example, if one user requires the transmission under the minimum bit-error-rate (BER), another might transmit under the maximum data rate assumptions, etc. If we consider the "regular" MMSE cost function, assuming that all users are equipped with only one antenna:

$$\text{(8)} \qquad \min_{\boldsymbol{D}} \mathrm{E}\left\{ \| (\boldsymbol{D}\boldsymbol{H}^T - \boldsymbol{I}_r)\boldsymbol{x} + \boldsymbol{D}\boldsymbol{n} \quad {}^2 \right\}$$

the optimum solution for D should be chosen such to minimize the Frobenius norm of the off-diagonal elements of the effective channel DH^T at high signal-to-noise ratios (SNRs), while the elements on the main diagonal should converge to 1.

Regularized block diagonalization (RBD)

Since the matrix D_a is used only to suppress MUI, analogous to the optimization in (8), we can define D_a such to minimize the Frobenius norm of the block matrices off the main diagonal of the effective channel after the detection $D_a H^T$. Thus, the matrices D_{a_i} are derived from the following optimization:

$$(9) \qquad D_a = \min_{D_a} E\left\{ \sum_{i=1}^{K} \quad D_{a_i} \widetilde{H}_i^T \tilde{z}_i \quad ^2 + \|D_{a_i} n\|^2 \right\}$$

where vector $n \in \mathbb{C}^{M_B \times 1}$ contains the samples of a zero mean additive white Gaussian noise (AWGN) at the input of the receive antennas at the BS and the matrix $\widetilde{H}_i$ is defined as

$$(10) \qquad \widetilde{H}_i = \begin{bmatrix} H_1 \\ \vdots \\ H_{i-1} \\ H_{i+1} \\ \vdots \\ H_K \end{bmatrix} \in \mathbb{C}^{(M_U - M_{U_i}) \times M_B}.$$

Vector $\tilde{z}_i \in \mathbb{C}^{(M_U - M_{U_i}) \times 1}$ is given as

$$(11) \qquad \tilde{z}_i = \quad z_1^T \quad \cdots \quad z_{i-1}^T \quad z_{i+1}^T \quad \cdots \quad z_K^T \quad ^T$$

where in general $z_i = Q_i x_i \in \mathbb{C}^{M_{U_i} \times 1}$. However, since we have assumed that each user transmits over all of its modes, i.e. $r_i = M_{U_i}$ and that the matrix Q_i is unitary, we have that the vector z_i has the same statistics as the vector x_i. If we assume that the elements of the vector x_i are independent identically distributed (i.i.d.) zero mean complex uniform random variables with the variance P_t, then the elements of the vector z_i will also be i.i.d., zero mean, and with the variance P_t. We assume throughout this paper that the AWGN samples at the input of different receive antennas are i.i.d., zero mean, with the variance σ_n^2. Each user's detection matrix D_{a_i} can be written as $D_{a_i} = \Phi_{a_i} \cdot M_{a_i}$, where $M_{a_i} \in {}^{M_B \times M_B}$ is a unitary matrix and $\Phi_{a_i} \in \mathbb{R}^{M_B \times M_B}$ is a diagonal power loading matrix with elements on the main diagonal greater than or equal to zero [7]. Let us define the singular value decomposition (SVD) of $\widetilde{H}_i$ as $\widetilde{H}_i = \widetilde{U}_i \, \tilde{}_i \widetilde{V}_i^H$. The solution to the minimization of (9) results in

$$(12) \qquad \begin{aligned} M_{a_i} &= \widetilde{V}_i^T, \\ {}_{a_i} &= \quad \tilde{}_i^T \tilde{}_i + \frac{\sigma_n^2}{P_t} I_{M_B} \quad ^{-1/2}. \end{aligned}$$

From equation (12) we can see that the cost function in equation (9) is minimized if each user transmits over the space spanned by the combined matrix of all other users with the power that is inversely proportional to the singular values of the combined channel matrix of these users $\widetilde{H}_i$. As a result, at high SNRs, and $M_B \geq M_U$, each user transmits only in the null space of all other users and the effective matrix $D_a H^T$ is block diagonal. After we suppress MUI using D_a we optimize the system performance by optimizing a set of SU MIMO channels. One option is that UTs transmit using STCs over the effective channels $D_{a_i} H_i^T$. Other option is that we design the optimum detection matrix D_{b_i} and precoding matrix Q_i at the BS and then feed forward Q_i to the UT. In this case we can use the results for the generalized design of SU MIMO precoding and decoding matrices [7].

RBD successive interference cancellation (RBD SIC)

Successive interference cancellation (SIC) means that users are decoded sequentially, and that the user to be decoded treats all the other users to be decoded as interference, and subtracts out the symbols transmitted by the users already decoded from the received codeword. Since each user is transmitting at an arbitrarily small bit error rate, the users already decoded can be subtracted without introducing additional errors. The RBD SIC decoding filter is derived from the linear RBD receive filter optimization by neglecting the influence of the previously decoded users. Let us assume that the users are ordered in such a way that the first user is decoded first, then the second one, etc. The RBD SIC optimization criterion can be written as:

$$(13) \qquad D_a = \min_{D_a} E\left\{ \sum_{i=1}^{K} \quad D_{a_i} \widehat{H}_i^T \hat{z}_i \quad ^2 + \|D_{a_i} n\|^2 \right\}$$

where the previous $i-1$ users' combined channel matrix $\widehat{H}_i$ is defined as:

$$(14) \qquad \widehat{H}_i = \quad H_{i+1}^T \quad \cdots \quad H_K^T \quad ^T \in \mathbb{C}^{(\sum_{k=i+1}^{K} M_{U_k}) \times M_B},$$

and its corresponding SVD as $\widehat{H}_i = \widehat{U}_i \, \hat{}_i \widehat{V}_i^H$. Elements of the vector $\hat{z}_i = \quad z_{i+1}^T \quad \cdots \quad z_K^T \quad ^T$ are i.i.d zero mean random variables and have variance P_t, assuming that the matrices D_{b_i} and Q_i are unitary and that the elements of the transmit data vector x are also zero mean i.i.d., and with variance P_t.

The optimum solution of the RBD SIC criteria for D_{a_i} is given by:

$$(15) \qquad \begin{aligned} M_{a_i} &= \widehat{V}_i^T, \\ {}_{a_i} &= \quad \hat{}_i^T \hat{}_i + \frac{\sigma_n^2}{P_t} I_{M_B} \quad ^{-1/2}. \end{aligned}$$

Iterative RBD (iRBD)

The performance of each user can be further improved by exploiting the row subspace of H_i that remains unused after the spatial processing. In RBD, the detection matrix for each user is generated under the assumption that the other UTs use all the available modes for transmission. However, after the power loading some modes remain unused and they can be exploited to improve the system performance. We can identify two cases. In the first case, if the number of the transmitted data streams of the i-th user r_i is less than the rank of the i-th user's channel matrix, then the other co-channel users could also transmit in this unused subspace without causing additional interference. In the second case, when $M_B \leq M_U$ and $K \leq M_B$, users must leave a part of their own subspaces unused in order to reduce the overall MUI.

Once we have obtained matrices D_{a_i} using the RBD algorithm, we can determine how much of the channel matrix H_i row subspaces each UT is using for transmission as:

$$H_i^{(1)} = V_{e_i}^{(r_i)\,T} H_i$$

where the matrix $\boldsymbol{V}_{e_i}^{(r_i)}$ contains the first r_i right singular vectors of $\boldsymbol{D}_{a_i}\boldsymbol{H}_i^T$. Iterative RBD (iRBD) is defined as a solution to the following optimization problem:

(16)
$$\boldsymbol{D}_{a_i}^{(l)} = \min_{\boldsymbol{D}_a^{(l)}} \mathrm{E}\left\{\sum_{i=1}^{K} \left\| \boldsymbol{D}_{a_i}^{(l)}\widetilde{\boldsymbol{H}}_i^{(l)\,T}\tilde{\boldsymbol{z}}_i \right\|^2 + \left\| \boldsymbol{D}_{a_i}^{(l)}\boldsymbol{n} \right\|^2 \right\}.$$

where

$$\widetilde{\boldsymbol{H}}_i^{(l)} = \begin{bmatrix} \boldsymbol{H}_1^{(l)} \\ \vdots \\ \boldsymbol{H}_{i-1}^{(l)} \\ \boldsymbol{H}_{i+1}^{(l)} \\ \vdots \\ \boldsymbol{H}_K^{(l)} \end{bmatrix} \in \mathbb{C}^{(r-r_i)\times M_B} \tag{17}$$

is the modified combined channel matrix of co-channel users in the l-th iteration, and $\tilde{\boldsymbol{z}}_i \in \mathbb{C}^{(r-r_i)\times 1}$. Total number of transmitted data streams and the number of data streams transmitted by the i-th user are denoted by r and r_i, respectively. The i-th user's effective channel matrix in the l-th iteration is equal to:

$$\boldsymbol{H}_i^{(l)} = \boldsymbol{V}_{e_i}^{(r_i)\,(l-1)\,T}\boldsymbol{H}_i \tag{18}$$

where $\boldsymbol{V}_{e_i}^{(r_i)\,(l-1)}$ contains the first r_i vectors of $\boldsymbol{V}_{e_i}^{(l-1)}$ which is obtained from the following SVD

$$\boldsymbol{D}_{a_i}^{(l-1)}\boldsymbol{H}_i^T = \boldsymbol{U}_{e_i}^{(l-1)}\boldsymbol{\Sigma}_{e_i}^{(l-1)}\boldsymbol{V}_{e_i}^{(l-1)\,H} \tag{19}$$

The first r_i vectors of $\boldsymbol{V}_{e_i}^{(r_i)\,(l-1)}$ correspond to the r_i strongest singular values of $\boldsymbol{D}_{a_i}^{(l-1)}\boldsymbol{H}_i^T$. Following the analysis for RBD detection, the solution to the optimization problem given in (16) is equal to:

$$\boldsymbol{D}_{a_i}^{(l)} = \left(\widetilde{\boldsymbol{\Sigma}}_i^{(l)\,T}\widetilde{\boldsymbol{\Sigma}}_i^{(l)} + \frac{\sigma_n^2}{P_t}\boldsymbol{I}_{M_B}\right)^{-1/2}\widetilde{\boldsymbol{V}}_i^{(l)\,T} \tag{20}$$

where

$$\widetilde{\boldsymbol{H}}_i^{(l)} = \widetilde{\boldsymbol{U}}_i^{(l)}\widetilde{\boldsymbol{\Sigma}}_i^{(l)}\widetilde{\boldsymbol{V}}_i^{(l)\,H}. \tag{21}$$

Note that the computational complexity of one iteration of the iRBD algorithm reduces after the first iteration since the dimension of the UT matrices reduce from $(M_{U_i} \times M_B)$ to $(r_i \times M_B)$. The objective in (16) is a concave function of the receive matrices, and the optimization over each $\boldsymbol{D}_{a_i}^{(l)}$ is separable. In such situation, it is generally sufficient to optimize with respect to one variable while holding all other variables constant, then optimize to the next variable while holding all other variables constant. This is referred to as the block-coordinate ascent algorithm and convergence can be shown under relatively general conditions. At each step of the algorithm, we optimize the receive matrix of only one user, user i, we update its corresponding matrix $\boldsymbol{H}_i^{(l)}$ in $\widetilde{\boldsymbol{H}}_{j\neq i}^{(l)}$ for all other co-channel users, and then move on to the next user.

Successive minimum mean-sqare error (SMMSE) filtering

If we use the "regular" MMSE optimization criterion that is given in equation (8) to design the receive matrices, by treating the signals transmitted from the collocated antennas at the same terminal as interference, system performance is degraded. Since each user can coordinate the processing over all of its antennas, these signals can be combined in order to extract higher spatial processing gains.

A straightforward way of reducing this problem is to design receive matrix row by row, by neglecting the contribution of the signals from the collocated antennas at the same terminal to this UT's MSE. The interference of other co-channel users to the signal from the i-th user's j-th antenna signal is suppressed independently from the other antennas at the same terminal. This is done for each antenna at the same UT successively. Therefore, the j-th row of the i-th user's receive matrix $\boldsymbol{D}_{a_i}$, corresponding to the i-th user's j-th antenna, is equal to the first row of the matrix $\overline{\boldsymbol{D}}_{a_{i,j}}$ which is obtained from the following optimization

(22)
$$\overline{\boldsymbol{D}}_{a_{i,j}} = \min_{\overline{\boldsymbol{D}}_{a_{i,j}}} \mathrm{E}\left\{\left\|\overline{\boldsymbol{D}}_{a_{i,j}}\left(\overline{\boldsymbol{H}}_i^{(j)\,T}\overline{\boldsymbol{z}}_i^{(j)} + \boldsymbol{n}\right) - \overline{\boldsymbol{z}}_i^{(j)}\right\|^2\right\}$$

Matrix $\overline{\boldsymbol{H}}_i^{(j)}$ and the vector $\overline{\boldsymbol{z}}_i^{(j)}$ corresponding to the i-th user's, $i = 1, \ldots, K$, j-th transmit antenna, $j = 1, \ldots, M_{U_i}$, are defined as

$$\overline{\boldsymbol{H}}_i^{(j)} = \begin{bmatrix} \boldsymbol{h}_{i,j} & \widetilde{\boldsymbol{H}}_i^T \end{bmatrix}^T \in \mathbb{C}^{(M_U - M_{U_i}+1)\times M_B}, \tag{23}$$

and

$$\overline{\boldsymbol{z}}_i^{(j)} = \begin{bmatrix} z_{i,j} & \tilde{\boldsymbol{z}}_i^T \end{bmatrix}^T \in \mathbb{C}^{(M_U - M_{U_i}+1)\times 1}, \tag{24}$$

where $\boldsymbol{h}_{i,j}^T$ is the j-th row of the i-th user's channel matrix $\boldsymbol{H}_i$ and $z_{i,j}$ is the j-th element of the i-th user's vector $\boldsymbol{z}_i$. The elements of the vector $\boldsymbol{z}_i$ are i.i.d. complex uniform random variables with zero mean and variance P_t. Matrix $\widetilde{\boldsymbol{H}}_i$ and the vector $\tilde{z}_i$ are defined as in equations (10) and (11), respectively.

The rows of the receive matrix $\boldsymbol{D}_{a_i}$, each corresponding to one of the UTs' antennas, are calculated successively. The j-th row of the i-th UT's receive matrix $\boldsymbol{D}_{a_i}$ is equal to:

$$\boldsymbol{d}_{a_{i,j}} = \boldsymbol{h}_{i,j}^H\left(\overline{\boldsymbol{H}}_i^{(j)\,T}\overline{\boldsymbol{H}}_i^{(j)\,*} + \alpha\boldsymbol{I}_{M_B}\right)^{-1} \tag{25}$$

The parameter α is equal to $\alpha = \sigma_n^2/P_t$ as in (12). The combination of SMMSE and SIC was introduced in [4].

Iterative SMMSE (iSMMSE) receive filtering

Iterative SMMSE (iSMMSE) receive filter is derived from the following optimization:

(26)
$$\overline{\boldsymbol{D}}_{a_{i,j}}^{(l)} = \min_{\overline{\boldsymbol{D}}_{a_{i,j}}^{(l)}} \mathrm{E}\left\{\left\|\overline{\boldsymbol{D}}_{a_{i,j}}^{(l)}\left(\overline{\boldsymbol{H}}_i^{(j,l)\,T}\overline{\boldsymbol{z}}_i^{(j)} + \boldsymbol{n}\right) - \overline{\boldsymbol{z}}_i^{(j)}\right\|^2\right\}$$

where the matrix $\overline{\boldsymbol{H}}_i^{(j,l)}$ is defined as

$$\overline{\boldsymbol{H}}_i^{(j,l)} = \begin{bmatrix} \boldsymbol{h}_{i,j}^{(l)} & \widetilde{\boldsymbol{H}}_i^{(l)\,T} \end{bmatrix}^T \in \mathbb{C}^{(r-r_i+1)\times M_B}. \tag{27}$$

Vector $\boldsymbol{h}_{i,j}^{(l)\,T}$ is the j-th row of the i-th user's channel matrix $\boldsymbol{H}_i^{(l)}$. The elements of the vector $\boldsymbol{z}_i$ are i.i.d. complex uniform random variables with zero mean and variance P_t. Matrices $\widetilde{\boldsymbol{H}}_i^{(l)}$, $\boldsymbol{H}_i^{(l)}$ and the vector $\tilde{z}_i$ are defined as in equations (17), (18) and (11), respectively. Same as for previous algorithms, when we generate matrices $\boldsymbol{D}_{a_i}^{(l)}$, in each iteration we assume that the matrices $\boldsymbol{D}_{b_i}^{(l)}$ and $\boldsymbol{Q}_i^{(l)}$ are unitary, which corresponds to the assumption that each UT is transmitting over all of its available channel subspace. In the first iteration the number of elements of $\tilde{\boldsymbol{z}}_i$ will be M_{U_i} and in every other iteration r_i. The j-th row of the i-th UT's receive matrix $\boldsymbol{D}_{a_i}^{(l)}$ is obtained as:

$$\boldsymbol{d}_{a_{i,j}}^{(l)} = \boldsymbol{h}_{i,j}^{(l)\ H} \left(\overline{\boldsymbol{H}}_i^{(j,l)\ T} \overline{\boldsymbol{H}}_i^{(j,l)\ *} + \alpha \boldsymbol{I}_{M_B} \right)^{-1} \tag{28}$$

The parameter α is equal to $\alpha = \sigma_n^2 / P_t$ as in (12).

Numerical results

In this section we compare the performance of MU MIMO algorithms introduced in this paper and VBLAST. To this end we simulate a purely stochastic spatially white frequency flat channel $\boldsymbol{H}_w$. The elements of the channel matrices are zero mean, unit variance complex Gaussian variables. Data is modulated using 4QAM. In case the UTs do not use STCs on the uplink a part of the system throughput is used for the feedback of the UTs' uplink precoding matrices $\boldsymbol{Q}_i$ or there are additional pilots used for estimation of the matrices $\boldsymbol{Q}_i$ at the BS. In the results presented in this paper we have assumed that the matrices $\boldsymbol{Q}_i$ are generated at the BS and then sent to the UTs.

In Figure 1 we compare the performance of algorithms for MU MIMO detection in an overloaded system with the antenna configuration $\{2,2,2\}\times 4$. In this case only the iterative MU MIMO detection algorithms with $\boldsymbol{Q}_i$ feedback do not have error floor and still extract the maximum diversity order inherent in the system. Although iRBD and iSMMSE algorithms with the feedback of the uplink UT precoding matrices $\boldsymbol{Q}_i$ extract full diversity also in the overloaded system, they require part of the throughput to be reserved for the feedback of matrices $\boldsymbol{Q}_i$ from the BS to the UTs. In the following figure we investigate the influence of the feedback overhead size on the performance of these algorithms. Consider a system where

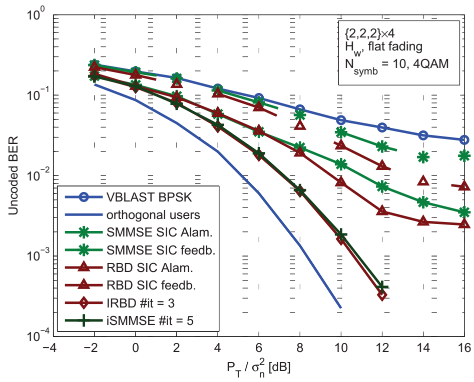

Fig. 1. Uncoded BER performance of MU MIMO detection algorithms in a system with the antenna configuration $\{2,2,2\}\times 4$.

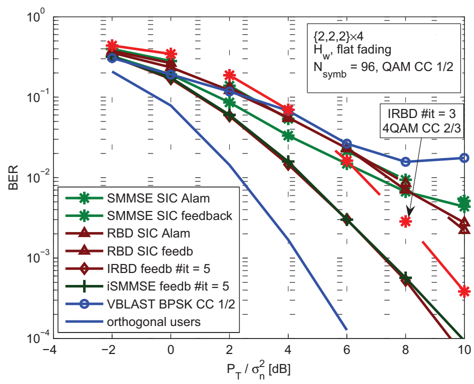

Fig. 2. Influence of feedback overhead on coded BER performance of MU MIMO detection algorithms in a $\{2,2,2\}\times 4$ system.

the data is transmitted in packets of size 96 symbols. In case UTs transmit data using the Alamouti code the data is coded using a convolutional code rate $1/2\ (561,753)_{\text{oct}}$. In case the uplink UT precoding matrices $\boldsymbol{Q}_i$ are generated at the BS and then fedback to the UTs, we assume that the data is coded using a convolutional code rate $2/3\ (561,753)_{\text{oct}}$ with the puncturing pattern $\begin{bmatrix} 1 & 0 & 1 & 1 \end{bmatrix}$. Higher data rate is introduced in order to account for the throughput used for feedback overhead and to keep the comparison between the algorithms using $\boldsymbol{Q}_i$ feedback and the algorithms that combine MU MIMO detection techniques with the STC fair.

The results shown in Figure 2 for iRBD, taking into account the feedback overhead, show that the iterative MU MIMO detection algorithms in combination with the feedback of the UT uplink precoding matrices still provide full spatial diversity gain. The BER performance of iRBD that takes into account the feedback overhead is represented using a red dashed curve. Due to the feedback there is a 2 dB SNR loss. However, even with this SNR loss, the iterative MU MIMO detection algorithms in combination with $\boldsymbol{Q}_i$ feedback still outperform algorithms that do not require CSI at the transmitter at the SNR range of interest.

Conclusions

In this paper we have used the modified MMSE criterion to design MU MIMO uplink detection matrices in a generalized way which would allow easy adaptation to any optimization criterion or the combination of different optimization criteria for different users. We have shown that the iterative MU MIMO detection techniques in combination with the feedback of the UT uplink precoding matrices are able to extract full spatial diversity gain inherent in the system and without the limitations of other algorithms previously presented.

REFERENCES

[1] R. W. Heath, M. Airy, and A. J. Paulraj, "Multiuser diversity for MIMO wireless systems with linear receivers," in *Proc. 35th Asilomar Conf. on Signals, Systems, and Computers, Pacific Grove, CA, IEEE Computer Society Press*, November 2001.

[2] A.F. Naguib, N. Seshadri, and A.R. Calderbank, "Applications of space-time block codes and interference suppression for high capacity and high data rate wireless systems," in *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers, Pacific Grove, CA*, November 1998.

[3] L. Dai, S. Sfar, and K. Ben Letaief, "An efficient detector for combined space time coding and layered processing," *IEEE Trans. on Comm.*, vol. 53, no. 9, pp. 1438–1442, September 2005.

[4] V. Stankovic and M. Haardt, "Improved diversity on the uplink of multi-user MIMO systems," in *Proc. European Conf. on Wir. Tech.(ECWT 2005), Paris, France*, October 2005.

[5] A. Wolfgang, S. Chen, and L. Hanzo, "Parallel interference cancellation based turbo space-time equalization in the SDMA uplink," *IEEE Trans. on Wireless Comm.*, vol. 6, no. 2, pp. 609–616, February 2007.

[6] V. Stankovic, *Multi-user MIMO wireless communications*, Ph.D. thesis, Ilmenau University of Technology, Germany, March 2007.

[7] A. Scaglione at. el., "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. on Sig. Proc.*, vol. 50, no. 5, pp. 1051–1064, May 2002.

Authors: Ph.D.P. Spalević, Ph.D. V. Stanković, State University of Novi Pazar, Vuka Karadzića bb, 36300 Novi Pazar, Serbia, email: petarspalevic@yahoo.com, Prof. M. Stefanović, Ph.D. S. Panić, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, email: stefanpnc@yahoo.com, Ph.D. Ana Savić, High school of electrical engineering and computing vocational study, Belgrade, Serbia, email:ana.savic@viser.edu.rs