SUCCESSIVE OPTIMIZATION TOMLINSON-HARASHIMA PRECODING (SO THP) FOR MULTI-USER MIMO SYSTEMS

Veljko Stankovic and Martin Haardt

Ilmenau University of Technology, Communications Research Laboratory
P.O. Box 10 05 65, 98684 Ilmenau, Germany
{veljko.stankovic, martin.haardt}@tu-ilmenau.de

ABSTRACT

Multi-user multiple-input multiple-output (MU MIMO) systems have the advantage of combining the high capacity achievable with MIMO processing and the benefits of space division multiple access. Previously proposed techniques that use the channel state information at the transmitter to improve the performance of the downlink suffer from a capacity loss due to a zero multi-user interference (MUI) constraint or they have to allow some MUI to improve the system performance. In this paper we propose a combination of one linear pre-coding technique named successive optimization (SO) and a non-linear precoding technique, Tomlin-son-Harashima precoding (THP). It uses all of the subspaces available in a MU MIMO system and can completely eliminate multi-user interference. We show that SO THP provides a very high capacity and a better BER performance than similar minimum mean-square-error (MMSE) THP precoding techniques at low SNRs especially for the users equipped with multiple antennas. SO THP is also less sensitive to channel estimation errors than MMSE THP precoding. Thereby, we minimize the capacity loss due to the MUI cancellation and we reduce the complexity of the receiver since there is no MUI in the system.

1. INTRODUCTION

In recent years, there has been a considerable interest in wireless multiple-input, multiple output (MIMO) communications systems because of their promising improvement in terms of performance and bandwidth efficiency. An important research topic is the study of multi-user (MU) MIMO system [1], [2], [3], [4], [5]. Such systems have the potential to combine the high capacity achievable with MIMO processing with the benefits of space division multiple access. In the downlink scenario, a base station (BS) or an access point (AP) is equipped with multiple antennas and it simultaneously transmits to a group of users. Each of these users is also equipped with multiple antennas. Motivated by the need for cheap mobiles with low power consumption, we focus on systems where the computationally demanding signal processing is performed at the BS/AP. The BS/AP will use the channel state information (CSI) available at the transmitter to allow these users to share the same channel and mitigate or ideally completely eliminate multi-user interference (MUI) at the transmitter by intelligent beamforming or by the use of "dirty-paper" codes.

This work has been partially performed in the framework of the IST project IST-2003-507581 WINNER, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues.

Block diagonalization (BD) is a linear pre-coding technique for the downlink of MU MIMO systems [6]. It decomposes a MU MIMO downlink channel into multiple parallel independent single-user MIMO downlink channels. The signal of each user is pre-processed at the transmitter using a modulation matrix that lies in the null space of other users' channel matrices. Thereby, the MUI in the system is efficiently set to zero. BD can be used with any other previously defined single-user MIMO techniques [7], as the different users do not interfere with each other. BD is attractive if the users are equipped with more than one antenna. However, the zero MUI constraint can lead to a significant capacity loss when users' subspaces significantly overlap. Another technique also proposed in [6], named successive optimization (SO), addresses the problem of minimizing the total transmit power while achieving a predefined Quality-of-Service (QoS) level for each user in the network and the near-far problem. It can yield better results in some situations but its performance depends on the power allocation and the order in which the users' signals are pre-processed. The zero MUI constraint is relaxed and a certain amount of interference is allowed. Tomlinson-Harashima precoding (THP) is a non-linear pre-coding technique developed for single-input, single-output (SISO) multipath channels. Recently it has been also proposed for the equalization of MUI in MIMO systems [8], where it performs spatial pre-equalization instead of temporal pre-equalization for ISI channels.

In this paper we propose a novel technique that combines SO and THP in order to reduce the capacity loss due to the overlapping of different users' subspaces and to eliminate the MUI. After the precoding, the resulting equivalent combined channel matrix of all users is again block diagonal. This also facilitates the definition of a new ordering algorithm. Unlike in [9], this technique allows more than one antenna at the mobile terminals and has no performance loss due to the cancellation of interference between the signals transmitted to two closely spaced antennas at the same terminal.

The paper is organized as follows. In Section 2 we describe the MU MIMO downlink channel and the model for the channel estimation errors. In Section 3 we describe SO THP and its performance is presented in Section 4. In Section 5 we give a short summary.

2. SYSTEM MODEL

We consider a MU MIMO downlink channel, where M_T transmit antennas are located at the base station, and M_{R_i} receive antennas are located at the i-th mobile station (MS), $i = 1, 2, \ldots, K$. There are K users (or MSs) in the system. The total number of receive antennas is $M_R = \sum_{i=0}^{K} M_{R_i}$. We will use the notation $\{M_{R_1}, \cdots, M_{R_K}\} \times M_T$ to describe the antenna configuration of the system. In this paper, we assume a flat fading channel. The MIMO channel to user i is denoted as $\boldsymbol{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$. Moreover, the combined channel matrix of all users is given by

$$\boldsymbol{H} = \left[\begin{array}{cccc} \boldsymbol{H}_1^T & \boldsymbol{H}_2^T & \cdots & \boldsymbol{H}_K^T \end{array} \right]^T.$$

In order to take into account channel estimation errors we use a "nominal-plus-perturbation" model. The estimated combined channel matrix can be represented as

$$\widehat{\boldsymbol{H}} = \boldsymbol{H} + \boldsymbol{E}$$

where $\boldsymbol{H}$ denotes the flat fading combined channel matrix of all users, and $\boldsymbol{E}$ is a complex random Gaussian matrix distributed according to $\mathcal{CN}\left(\boldsymbol{0}_{M_R \times M_T}, M_R \sigma_n^2 \boldsymbol{I}_{M_T}\right)$. Let $\rho = \|\boldsymbol{H}\|^2 / \|\boldsymbol{E}\|^2$ be the SNR of the channel state information at the transmitter.

3. MULTI-USER PRECODING

In this section we will briefly describe BD, SO, and the proposed technique that combines SO and THP.

3.1. Block diagonalization (BD)

Block diagonalization was first proposed in [6]. It can be applied to solve either the problem of maximizing the total system throughput under a transmit power constraint or to minimize the total transmit power for a predefined QoS level. It is restricted to channels where the number of transmit antennas M_T is not smaller than the total number of receive antennas in the network M_R.

Let us define the precoder matrices as

$$\boldsymbol{F} = \left[\begin{array}{cccc} \boldsymbol{F}_1 & \boldsymbol{F}_2 & \cdots & \boldsymbol{F}_K \end{array} \right] \in \mathbb{C}^{M_T \times r}$$

where $\boldsymbol{F}_i \in \mathbb{C}^{M_T \times r_i}$ is the i-th user's precoder matrix. Moreover, $r \leq M_R$ is the total number of the transmitted data stream sequences, while $r_i \leq M_{R_i}$ is the number of data stream sequences transmitted to the i-th user. We can find the optimal precoding matrix $\boldsymbol{F}$ such that all MUI is zero by choosing a precoding matrix $\boldsymbol{F}_i$ that lies in the null space of the other users' channel matrices. Thereby, a MU MIMO downlink channel is decomposed into multiple parallel independent SU MIMO channels [10], [7].

If we define $\widetilde{\boldsymbol{H}}_i$ as

$$\widetilde{\boldsymbol{H}}_i = \left[\begin{array}{cccccc} \boldsymbol{H}_1^T & \cdots & \boldsymbol{H}_{i-1}^T & \boldsymbol{H}_{i+1}^T & \cdots & \boldsymbol{H}_K^T \end{array} \right]^T \quad (1)$$

the zero MUI constraint forces $\boldsymbol{F}_i$ to lie in the null space of $\widetilde{\boldsymbol{H}}_i$. From the singular value decomposition (SVD) of $\widetilde{\boldsymbol{H}}_i$ whose rank is $\widetilde{L}_i$

$$\widetilde{\boldsymbol{H}}_i = \widetilde{\boldsymbol{U}}_i \widetilde{\boldsymbol{\Sigma}}_i \left[\begin{array}{cc} \widetilde{\boldsymbol{V}}_i^{(1)} & \widetilde{\boldsymbol{V}}_i^{(0)} \end{array} \right]^H \quad (2)$$

we choose the last right $M_T - \widetilde{L}_i$ singular vectors $\widetilde{\boldsymbol{V}}_i^{(0)} \in \mathbb{C}^{M_T \times M_T - \widetilde{L}_i}$ which form an orthogonal basis for the null space of $\widetilde{\boldsymbol{H}}_i$. The equivalent channel of user i after eliminating the MUI is identified as $\boldsymbol{H}_i \widetilde{\boldsymbol{V}}_i^{(0)}$, whose dimension is $M_{R_i} \times \left(M_T - \widetilde{L}_i\right)$ and is equivalent to a system with $M_T - \widetilde{L}_i$ transmit antennas and M_{R_i} receive antennas. Each of these equivalent SU MIMO channels has the same properties as a conventional SU MIMO channel. Define the SVD

$$\boldsymbol{H}_i \widetilde{\boldsymbol{V}}_i^{(0)} = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \left[\begin{array}{cc} \boldsymbol{V}_i^{(1)} & \boldsymbol{V}_i^{(0)} \end{array} \right]^H \quad (3)$$

and let the rank of the i-th user's equivalent channel matrix be L_i. The product of the first L_i singular vectors $\boldsymbol{V}_i^{(1)}$ and $\widetilde{\boldsymbol{V}}_i^{(0)}$ produces an orthogonal basis of dimension L_i and represents the transmission vectors that maximize the information rate for user i subject to the zero MUI constraint. The demodulation matrix of the i-th user is chosen as $\boldsymbol{D}_i = \boldsymbol{U}_i^H$.

3.2. Successive optimization (SO)

As mentioned before, by applying BD on the combined channel matrix of all users the MU MIMO channel can be transformed into a set of parallel single-user MIMO channels. However, there is a capacity loss due to the cancellation of overlapping subspaces of different users. In [6], the authors propose a successive precoding algorithm in order to define a simplified solution of the power control problem. By allowing a certain amount of interference, this algorithm reduces the capacity loss due to the subspace cancellation.

First, we have to assume or determine a certain optimum ordering of the users, similar to VBLAST [11] or MMSE THP [12]. Using SO, the modulation matrix for each user is designed in such a way that it lies only in the null space of the channel matrices of previous users. As a consequence, only they will generate the interference to this user. Let us define the previous $i-1$ users' combined channel matrix as

$$\widehat{\boldsymbol{H}}_i = \left[\begin{array}{cccc} \boldsymbol{H}_1^T & \boldsymbol{H}_2^T & \cdots & \boldsymbol{H}_{i-1}^T \end{array} \right]^T$$

and its corresponding SVD as

$$\widehat{\boldsymbol{H}}_i = \widehat{\boldsymbol{U}}_i \widehat{\boldsymbol{\Sigma}}_i \left[\begin{array}{cc} \widehat{\boldsymbol{V}}_i^{(1)} & \widehat{\boldsymbol{V}}_i^{(0)} \end{array} \right]^H. \quad (4)$$

If the rank of $\widehat{\boldsymbol{H}}_i$ is $\hat{L}_i$, then $\widehat{\boldsymbol{V}}_i^{(0)}$ contains $M_T - \hat{L}_i$ right singular vectors. As in the BD solution, we force the modulation matrix $\boldsymbol{F}_i$ to lie in the null space of $\widehat{\boldsymbol{H}}_i$ by setting $\boldsymbol{F}_i = \widehat{\boldsymbol{V}}_i^{(0)} \boldsymbol{F}_i'$ for some choice of $\boldsymbol{F}_i'$. Thereby, the i-th user does not see any interference from any subsequent user $(i+1, \ldots, K)$.

3.3. Combination of SO and THP

In this section we will describe how to combine SO and THP in order to improve the use of the available subspace of different users and eliminate any residual MUI. The resulting equivalent channel matrix is also block diagonal which facilitates the definition of an ordering algorithm of the users.

The combination of SO and THP (SO THP) is performed by successively calculating the BD, the reordering of users, and in the end precoding with THP. Instead of examining all $K!$ possibilities for ordering to minimize the total capacity loss in the system, we propose a heuristic simplification to minimize the capacity loss of each user in the presence of the other co-channel users separately. The whole SO THP algorithm is summarized in Table 1.

In Table 1, we use the following notation: BD () is BD as explained in Section 3.1, $\boldsymbol{P}_k$ is an auxiliary matrix where we store the precoding matrices generated using BD, $\mathcal{S}$ is a set of indices of the users to be processed, $\boldsymbol{D}_i$ is the i-th user demodulation matrix obtained by using the BD algorithm and $\boldsymbol{B}$ is the THP feedback matrix. In short, we first calculate the capacity that an individual user can achieve assuming there are no other users in the system. Then, we look for the user for whom the difference between its capacity when there are no other users and its BD capacity is minimum and generate the precoding matrix of this user such that it lies in the null space of the remaining users' channel matrices. In for $i = 1 : K$

$$\boldsymbol{H}_i = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \begin{bmatrix} \boldsymbol{V}_i^{(1)} & \boldsymbol{V}_i^{(0)} \end{bmatrix}^H ;$$
$$\boldsymbol{F}_{\max,i} = \boldsymbol{V}_i^{(1)} ;$$
$$C_{\max,i} = \log_2 \det \left(\boldsymbol{I} + \boldsymbol{R}_{n,i}^{-1} \boldsymbol{H}_i \boldsymbol{F}_{\max,i} \boldsymbol{F}_{\max,i}^H \boldsymbol{H}_i^H \right) ;$$

end;

$S = \{1, \ldots, K\}$;

$\boldsymbol{G} = \boldsymbol{H}$;

for $i = K : 1$

$$\begin{bmatrix} \boldsymbol{P}_1, & \ldots & \boldsymbol{P}_i, & \boldsymbol{U}_1, & \ldots & \boldsymbol{U}_i \end{bmatrix} = \mathrm{BD}(\boldsymbol{G}) ;$$

for $k = 1 : i$

$$C_k = \log_2 \det \left(\boldsymbol{I} + \boldsymbol{R}_{n,k}^{-1} \boldsymbol{H}_k \boldsymbol{P}_k \boldsymbol{P}_k^H \boldsymbol{H}_k^H \right) ;$$

end;

$$k_i = \arg\min_{k \in S} \left(C_{\max,k} - C_k \right) ;$$
$$\boldsymbol{F}_i = \boldsymbol{P}_{k_i} ;$$
$$\boldsymbol{D}_i = \boldsymbol{U}_{k_i}^H ;$$
$$S = S \backslash \{k_i\} ;$$
$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{H}_1^T & \ldots & \boldsymbol{H}_{k_i-1}^T & \boldsymbol{H}_{k_i+1}^T & \ldots & \boldsymbol{H}_K^T \end{bmatrix}^T ;$$

end;

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1 & \ldots & \boldsymbol{F}_K \end{bmatrix} ;$$
$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{D}_1 & & \\ & \ddots & \\ & & \boldsymbol{D}_K \end{bmatrix} ;$$
$$\boldsymbol{B} = \text{lower triangular} \left(\boldsymbol{DHF} \cdot \mathrm{diag} \left([\boldsymbol{DHF}]_{ii}^{-1} \right) \right) ;$$

Table 1. SO THP algorithm.

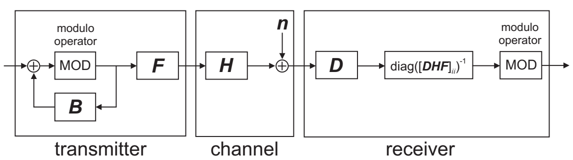

Fig. 1. Block diagram of the SO THP system.

each step we find the user with the minimum capacity loss and place it as the last one. Afterwards, we form the new combined channel matrix $\boldsymbol{G}$ without this user's channel matrix $\boldsymbol{H}_{k_i}$. We repeat these steps until the combined channel matrix is empty.

The order of the users is the reverse of the order in which their precoding matrices are generated. With this reordering of the users we achieve that the equivalent combined channel matrix after precoding and demodulation is lower triangular with the singular values on the main diagonal. The lower triangular feedback matrix $\boldsymbol{B}$, used in THP precoding [12], is generated from this equivalent combined channel matrix after the elements in each row are divided by the elements on the main diagonal, i.e., the corresponding singular values, as it can be seen from the last equation in Table 1.

In Figure 1 we show the block diagram of the SO THP system. The individual users' channel matrices and demodulation matrices are grouped in matrices $\boldsymbol{H}$ and $\boldsymbol{D}$. The feedback matrix $\boldsymbol{B}$, generated in the last step of the SO THP algorithm is now used to precode the users' data streams starting with the data stream of the first user whose precoding matrix $\boldsymbol{F}_1$ was generated as the last one.

By using THP at the transmit side we significantly increase the transmit power. That is why we have to introduce the modulo operator at the transmitter and the receiver in order to reduce the constellation size into certain boundaries. Before applying the modulo operator at the receiver we have to divide each data stream by the corresponding singular value so that the constellation boundaries at the receiver are the same as at the transmitter, [12].

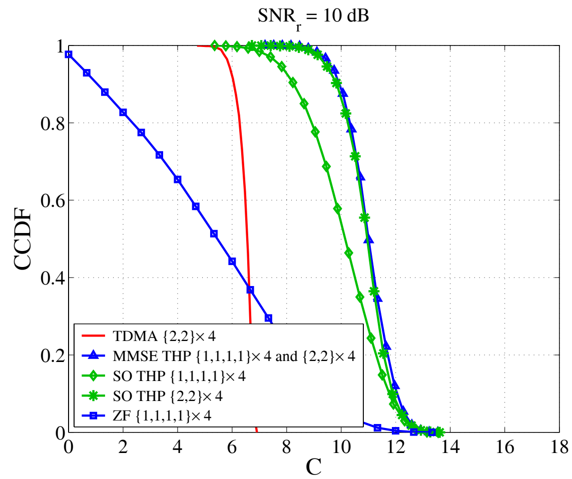

Fig. 2. Complementary cumulative distribution function of the MU downlink system capacity in bps/Hz.

4. SIMULATION RESULTS

In this section we will compare the performance of SO THP and minimum mean-square-error (MMSE) THP transmit filtering, proposed in [12]. The channel $\boldsymbol{H}$ is assumed to be spatially white and flat fading. First, we use the complementary cumulative distribution function (CCDF) and a 10% outage capacity to compare the system with configuration $\{1,1,1,1\} \times 4$ employing MMSE THP pre-filtering and a system employing SO THP with the configurations $\{1,1,1,1\} \times 4$ and $\{2,2\} \times 4$. The antenna configuration $\{1,1,1,1\} \times 4$ for MMSE THP is equivalent to $\{2,2\} \times 4$ because each data stream is processed separately. We also present capacity results for a TDMA system as a comparison. The capacity of a TDMA system is calculated as the average capacity of the system when the users transmit one at the time. We employ the following transmit and receive SNR definitions, respectively:

$$\mathrm{SNR_t} = 10 \log_{10} \left(\frac{P_T}{M_R \sigma_n^2} \right) \text{ and } \mathrm{SNR_r} = 10 \log_{10} \left(\frac{P_T}{\sigma_n^2} \right) .$$

In Fig. 2, we compare the CCDF performance of MMSE THP and SO THP. As a comparison we show also the CCDF function of a TDMA system and of a system employing the zero forcing solution (ZF), where the precoding matrix $\boldsymbol{F}$ is defined as the pseudo-inverse of the combined channel matrix $\boldsymbol{H}$. From this figure we see that MMSE THP provides a higher capacity than SO THP for the users equipped with one antenna each. However, for multiple-antenna users, SO THP provides approximately the same capacity as MMSE THP but with one difference, there is no MUI which enables the use of a simpler receiver. The capacity is calculated using the results on the capacity of MIMO broadcast channels in [3].

In Figures 3 and 4 we compare these techniques using 10% outage capacity as a function of receive SNR and the number of transmit antennas M_T. In Fig. 3 we show that for high SNR ratios SO THP can provide a higher capacity for both single- and multiple-antenna users. However, at low SNR ratios MMSE THP has an advantage over SO THP when the users are equipped with only one antenna.

The BER performance of SO THP and MMSE THP is shown in Fig. 5. In order to keep the same data rate in both systems, we

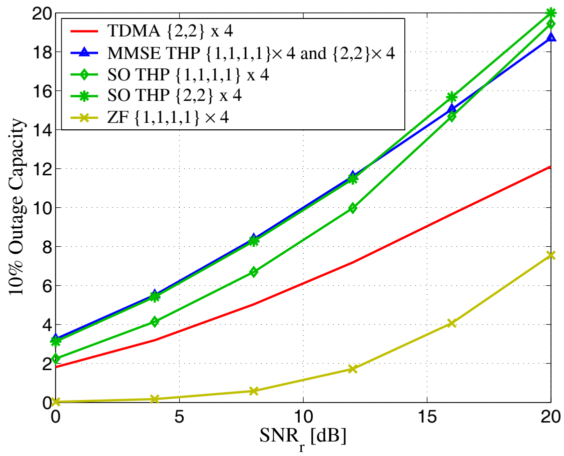

Fig. 3. 10 % outage capacity in bps/Hz as a function of the receive SNR$_r$.

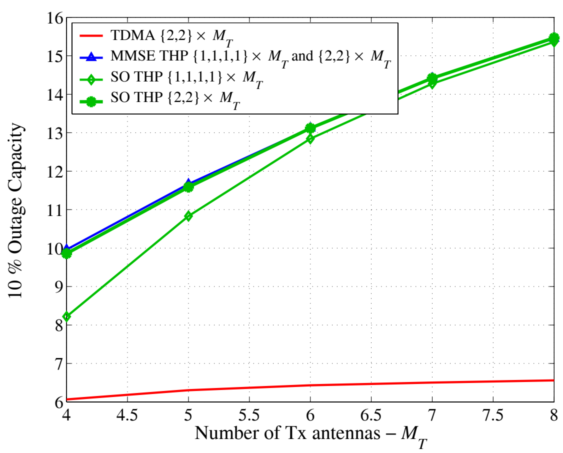

Fig. 4. 10 % outage capacity in bps/Hz as a function of the number of transmit antennas M_T. SNR$_r$ = 10 dB

use SO THP dominant eigenmode transmission and QAM modulation, while for MMSE THP we use BPSK. Note that SO THP provides a better performance at low SNR ratios and it is less sensitive to the channel estimation errors. In this case, the BER performance of a SO THP system can even be better by on order of magnitude than a system employing MMSE THP.

5. CONCLUSION

In this paper we propose the combination of a linear pre-coding technique called successive optimization (SO) and a non-linear technique, Tomlinson-Harashima precoding (THP). By combining these two techniques we are able to completely eliminate the MUI when there is perfect channel state information available at the transmitter and use all of the available subspace of different users in a mobile MU MIMO communication system. The equivalent channel matrix is block diagonal after precoding as in the case of precoding using BD. SO THP is especially attractive in cases when the users and the base station/access point are equipped with multiple antennas. In these cases SO THP provides the same capacity as for example MMSE THP but without any MUI. SO THP provides a higher capacity than MMSE THP for high SNR ratios

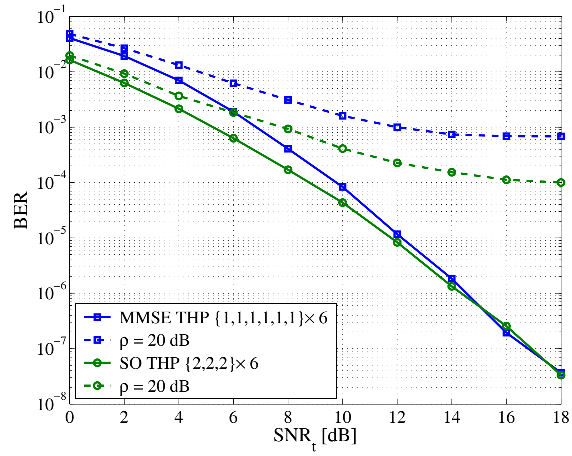

Fig. 5. BER as a function of SNR$_t$.

regardless of the number of receive antennas. When the users are equipped with multiple antennas SO THP provides a better BER performance than MMSE THP by transmitting only on the dominant eigenmodes of each user. This advantage is especially important when we do not have perfect channel state information available at the transmitter. Moreover, SO THP is less sensitive to channel estimation errors and can give results that are better by an order of magnitude.

6. 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] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, “An introduction to the multi-user MIMO downlink,” *IEEE Communications Magazine*, vol. 42, no. 10, pp. 60–67, October 2004.

[3] S. Vishwanath, N. Jindal, and A. J. Goldsmith, “On the capacity of multiple input multiple output broadcast channels,” in *Proc. of the IEEE International Conference on Communications (ICC), New York, NY*, April 2002.

[4] Z. Pan, K. K. Pan, and T. Ng, “MIMO antenna system for multi-user multi-stream orthogonal space time division multiplexing,” in *In Proc. of the IEEE International Conference on Communicaions, Anchorage, Alaska*, May 2003.

[5] K. K. Wong, “Adaptive space-division-multiplexing and bit-and-power allocation in multiuser MIMO flat fading broadcast channel,” in *Proc. of the IEEE 58th Vehicular Technology Conference, Orlando, FL*, October 2003.

[6] Q. Spencer and M. Haardt, “Capacity and downlink transmission algorithms for a multi-user MIMO channel,” in *Proc. 36th Asilomar Conf. on Signals, Systems, and Computers, Pacific Grove, CA, IEEE Computer Society Press*, November 2002.

[7] L. U. Choi and R. D. Murch, “A transmit preprocessing technique for multi-user MIMO systems using a decomposition approach,” *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 20–24, January 2004.

[8] G. Cinis and J. Cioffi, “A multi-user precoding scheme achieving crosstalk cancellation with application to DSL systems,” in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, November 2000, vol. 2, pp. 1627–1637.

[9] G. Caire and S. Shamai, “On the achievable throughput of a multiantenna gaussian broadcast channel,” *IEEE Trans. on Inf. Theory.*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[10] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in Multiuser MIMO channels,” *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, February 2004.

[11] P.W. Wolniansky, G.J Foschini, G.D. Golden, and R.A. Valenzuela, “V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel,” in *Proc. ISSSE 98*, September 1998.

[12] M. Joham, J. Brehmer, and W. Utschick, “MMSE approaches to multiuser spatio-temporal Tomlinson-Harashima precoding,” in *Proc. 5th International ITG Conference on Source and Channel Coding (ITG SCC’04)*, January 2004, pp. 387–394.