ROBUSTNESS EVALUATION OF UNIFORM POWER ALLOCATION WITH ANTENNA SELECTION FOR SPATIAL TOMLINSON-HARASHIMA PRECODING

Miquel Payaró, Ana Pérez-Neira, and Miguel Ángel Lagunas

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
c/ Gran Capità 2-4, Rooms 202-203, 08034 Barcelona, Spain
e-mail: {miquel.payaro,ana.isabel.perez,m.a.lagunas}@cttc.es

ABSTRACT

The performance of any communications precoding scheme relies on the quality of the channel state information that is made available at the transmitter side. In this paper, we study the robustness of Spatial Tomlinson-Harashima Precoding (STHP) in terms of rate loss when errors are considered in the feedback link. We will evaluate the performance degradation of the uniform power allocation with antenna selection strategy for STHP (which is quasi-optimal when there are no feedback errors) and compare it with a robust power allocation designed to maximize the achievable rates for the worst-case feedback noise. It will be shown that the performance difference between both approaches is very small. Consequently, uniform power allocation with antenna selection arises as a good solution for STHP even in the presence of feedback errors.

1. INTRODUCTION

More than thirty years ago, Tomlinson [1] and Harashima [2] independently proposed a non-linear precoding scheme for temporal intersymbol interference mitigation. Recently, Fischer *et al.* [3] extended the precoding structure to equalize spatial interference for multiple-input multiple-output (MIMO) systems. Next, in [4], the scheme proposed in [3] was modified by allowing variations of the transmitted power per antenna. Moreover, the authors stated the problem of finding the maximum achievable rates for this modified Spatial Tomlinson-Harashima Precoding (STHP) scheme, and solved it, finding that uniform power allocation with antenna selection (UPA-AS) is a quasi-optimal transmission scheme.

The main idea behind precoding schemes is to pre-subtract, at the transmitter side, the interference that will be introduced by the channel, so that the received signal is interference free. This implies that channel state information (CSI) has to be made available at the transmitter, since the precoder has to be matched to the channel. When the CSI at the transmitter is imperfect the system suffers from performance degradation. The effects of imperfect CSI for STHP, in terms of signal-to-noise ratio (SNR) loss, were studied in [5].

In this paper, a different approach will be taken. We will address the problem of evaluating the *capacity* robustness of STHP when feedback errors are present. First, we will present two precoding structures and will find analytic expressions for the received signal for each scheme. Next, their performance in the presence of feedback errors will be evaluated. From that point, we will focus our attention on the evaluation of different power allocation strategies for the best precoder. We will compare the rates achieved by the quasi-optimal UPA-AS scheme with those achieved by a robust design. It will be shown that, despite the optimal power allocation obtained by the robust design differs substantially from the UPA-AS approach, the capacity loss of the latter is small compared to that of the former. Thus, UPA-AS arises as a good solution in terms of simplicity-performance tradeoff among power allocation strategies for STHP if feedback errors are considered.

The remainder of the paper is organized as follows: next section presents the system and signal models and, two different precoding schemes. In section 3, the effect of feedback errors is calculated and compared for both precoding structures. A robust power allocation, in the maximin sense, is obtained in section 4. Finally, in section 5, simulations results and conclusions are presented.

2. SYSTEM MODEL

The model of the received signal vector, $\mathbf{y}$, for a flat fading MIMO channel with n_T transmit and n_R receive antennas is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n},$$

where $\mathbf{x} \in \mathbb{C}^{n_T}$ is the transmitted vector, $\mathbf{n} \in \mathbb{C}^{n_R}$ is the noise vector, and $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ is the channel matrix, where $[\mathbf{H}]_{kl}$ represents the base band complex attenuation from l-th transmitter to k-th receiver. Throughout this paper $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_R})$ and $n_R \geq n_T$ will be assumed.

For STHP (see Fig. 1), the transmission vector, $\mathbf{x}$, is obtained from data symbols, $\mathbf{s} \in \mathbb{C}^{n_T}$, in a similar way to the recursive spatial relation described in [3]:

$$x_k = M_{t_k} \left[s_k - \sum_{l=1}^{k-1} b_{kl} x_l \right], \quad \text{for } k = 1, \ldots, n_T,$$

where b_{kl} are the elements of the strictly lower triangular matrix $\mathbf{B} \in \mathbb{C}^{n_T \times n_T}$ and $M_{t_k}[x]$ is a complex modulo reduction of x into the interval $[-t_k, t_k) \times [-j\, t_k, j\, t_k)$, see [3] for further details. An alternative and more compact notation can be used for transmitted vector $\mathbf{x}$,

$$\mathbf{x_B} = (\mathbf{B} + \mathbf{I}_{n_T})^{-1} (\mathbf{s} + \mathbf{a}), \tag{1}$$

This work was partially supported by the Spanish Ministry of Science and Technology (MCYT) under projects PU-2002-56, TIC2002-04594-C02 (jointly financed by FEDER), HU2002-0032, FIT-070000-2003-257, the European Commission IST-2001-38835 ANWIRE, and the Generalitat de Catalunya 2003FI 00195.

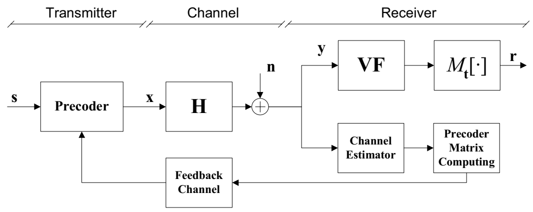

Fig. 1. MIMO communications system with STHP.

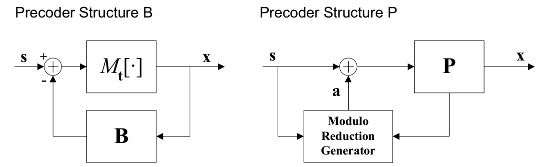

Fig. 2. Two possible precoding structures.

where the elements of vector $\mathbf{a}$ must be of the form $a_k = 2t_k p_k + j2t_k q_k$ for some $p_k, q_k \in \mathbb{Z}$, which must be chosen such that the resulting elements of the transmitted vector $\mathbf{x}$ lie inside the modulo region defined by the elements of vector $\mathbf{t}$, *i.e.* $x_k \in [-t_k, t_k) \times [-j\,t_k, j\,t_k)$. If $\mathbf{P} = (\mathbf{B} + \mathbf{I})^{-1}$ is defined, the transmitted vector can be equivalently expressed as

$$\mathbf{x_P} = \mathbf{P}(\mathbf{s} + \mathbf{a}). \tag{2}$$

Both alternative expressions for the transmitted vector $\mathbf{x}$ in (1) and (2) allow to build two precoding structures depending on which is the precoding matrix that is fedback to the transmitter, $\mathbf{B}$ or $\mathbf{P}$ (see Fig. 2).

The two precoders, $\mathbf{B}$ and $\mathbf{P}$, and the receiver, $\mathbf{VF}$, design for a zero-forcing criterion is [3],

$$\begin{aligned} \mathbf{H}^H\mathbf{H} = \mathbf{S}^H\mathbf{S}, \quad \mathbf{V} = \mathbf{diag}(s_{ii}^{-1}), \\ \mathbf{F} = \mathbf{S}^{-H}\mathbf{H}^H, \quad \mathbf{C} = \mathbf{VS}, \\ \mathbf{B} = \mathbf{C} - \mathbf{I}_{n_T}, \quad \mathbf{P} = \mathbf{C}^{-1}, \end{aligned} \tag{3}$$

where $\mathbf{S}$ is a lower triangular matrix obtained by means of a modified Cholesky factorization and where $\mathbf{F}$ is unitary. Since $\mathbf{B}$ is a strictly lower triangular matrix, using (3) it can be easily seen that $\mathbf{P}$ is a lower triangular matrix with ones in the main diagonal. With these definitions, the received signal vector, $\mathbf{r} \in \mathbb{C}^{n_T}$, can be expressed as

$$\mathbf{r} = M_\mathbf{t}\left[\mathbf{Cx} + \mathbf{V\tilde{n}}\right], \tag{4}$$

where $\mathbf{\tilde{n}} = \mathbf{Fn}$ has the same statistical behavior as $\mathbf{n}$, because $\mathbf{F}$ is a unitary transformation matrix. Using (3) in (1) or in (2) the transmitted signal vector can be expressed as $\mathbf{x} = \mathbf{C}^{-1}(\mathbf{s} + \mathbf{a})$ and thus, the received vector becomes

$$\mathbf{r} = M_\mathbf{t}\left[\mathbf{s} + \mathbf{a} + \mathbf{V\tilde{n}}\right] = M_\mathbf{t}\left[\mathbf{s} + \mathbf{V\tilde{n}}\right],$$

as the $\mathbf{a}$ term can be eliminated since it is inside the modulo operation. Notice that $M_\mathbf{t}[\mathbf{d}]$ performs a modulo-t_k operation for each element d_k.

3. FEEDBACK ERROR IN THE PRECODING MATRICES

If no feedback errors are considered, both precoding structures, (1) and (2), are obviously equivalent in the sense that given a data vector, $\mathbf{s}$, they produce the same output, $\mathbf{x_B} = \mathbf{x_P}$. Consequently, it would be of no importance for the communications system performance to feedback either the lower diagonal elements of matrix $\mathbf{B}$ or the lower diagonal elements of matrix $\mathbf{P}$ (there is no need to feedback the diagonal elements of $\mathbf{P}$ as they are all equal to 1). However, if some kind of transmission error is considered in the feedback link, $\widehat{\mathbf{B}} = \mathbf{B} + \boldsymbol{\Delta}$ and $\widehat{\mathbf{P}} = \mathbf{P} + \boldsymbol{\Delta}$, with $\boldsymbol{\Delta}$ strictly lower triangular, the two precoding structures become,

$$\begin{aligned} \mathbf{x}_{\widehat{\mathbf{B}}} &= (\widehat{\mathbf{B}} + \mathbf{I}_{n_T})^{-1}(\mathbf{s} + \mathbf{a}_{\widehat{\mathbf{B}}}) = \\ &= (\mathbf{C} + \boldsymbol{\Delta})^{-1}(\mathbf{s} + \mathbf{a}_{\widehat{\mathbf{B}}}), \end{aligned} \tag{5}$$

$$\mathbf{x}_{\widehat{\mathbf{P}}} = \widehat{\mathbf{P}}(\mathbf{s} + \mathbf{a}_{\widehat{\mathbf{P}}}) = (\mathbf{C}^{-1} + \boldsymbol{\Delta})(\mathbf{s} + \mathbf{a}_{\widehat{\mathbf{P}}}), \tag{6}$$

which are obviously different, $\mathbf{x}_{\widehat{\mathbf{B}}} \neq \mathbf{x}_{\widehat{\mathbf{P}}}$. Notice that $\mathbf{a}_{\widehat{\mathbf{B}}}$ in (5) and $\mathbf{a}_{\widehat{\mathbf{P}}}$ in (6) must now be chosen so that $\mathbf{x}_{\widehat{\mathbf{B}}}$ and $\mathbf{x}_{\widehat{\mathbf{P}}}$ lie inside the modulo region defined by vector $\mathbf{t}$. In addition, it is important to state that, in general, $\mathbf{a}_{\widehat{\mathbf{B}}} \neq \mathbf{a}$, $\mathbf{a}_{\widehat{\mathbf{P}}} \neq \mathbf{a}$, and $\mathbf{a}_{\widehat{\mathbf{B}}} \neq \mathbf{a}_{\widehat{\mathbf{P}}}$. However, if the variance of the elements of the error matrix $\boldsymbol{\Delta}$ is kept small it can be considered that

$$\mathbf{a} = \mathbf{a}_{\widehat{\mathbf{B}}} = \mathbf{a}_{\widehat{\mathbf{P}}}. \tag{7}$$

Assuming that this simplification is true (see section 5 for a validity evaluation), the received signal for both precoding structures (named Structure $\mathbf{B}$ and Structure $\mathbf{P}$ according to which matrix is actually fedback) will be found in the following two subsections.

3.1. Precoder Structure $\mathbf{B}$

Substituting (5) in (4) and using the matrix inversion lemma [6], $(\mathbf{C}+\boldsymbol{\Delta})^{-1} = \mathbf{C}^{-1} - \mathbf{C}^{-1}(\boldsymbol{\Delta}\mathbf{C}^{-1} + \mathbf{I}_{n_T})^{-1}\boldsymbol{\Delta}\mathbf{C}^{-1}$, the received signal with precoder $\mathbf{B}$ can be expressed as

$$\mathbf{r}_{\widehat{\mathbf{B}}} = M_\mathbf{t}[\mathbf{s} + (\boldsymbol{\Delta}\mathbf{C}^{-1} + \mathbf{I}_{n_T})^{-1}\boldsymbol{\Delta}\mathbf{C}^{-1}(\mathbf{s} + \mathbf{a}) + \mathbf{V\tilde{n}}].$$

Notice that $\mathbf{C}^{-1}(\mathbf{s} + \mathbf{a}) = \mathbf{x}$, *i.e.* we have reduced $(\mathbf{s} + \mathbf{a})$ to $\mathbf{x}$, which would be the transmitted signal if no errors were present in $\widehat{\mathbf{B}}$ matrix, and whose components are bounded in $[-t_k, t_k) \times [-j\,t_k, j\,t_k)$. Defining $\mathbf{L} = (\boldsymbol{\Delta}\mathbf{C}^{-1} + \mathbf{I}_{n_T})^{-1}\boldsymbol{\Delta}$ the received signal finally reads as

$$\mathbf{r}_{\widehat{\mathbf{B}}} = M_\mathbf{t}[\mathbf{s} + \mathbf{Lx} + \mathbf{V\tilde{n}}], \tag{8}$$

where $\mathbf{L}$ is a strictly lower triangular random matrix. Attention must be paid to the fact that if the elements of $\boldsymbol{\Delta}$ are sufficiently low compared with $\mathbf{C}$, the first order approximation $\mathbf{L} = \boldsymbol{\Delta} + o(\boldsymbol{\Delta})$ becomes valid.

3.2. Precoder Structure $\mathbf{P}$

Substituting (6) in (4) and using some basic algebraic manipulations the received signal with precoder $\mathbf{P}$ is

$$\mathbf{r}_{\widehat{\mathbf{P}}} = M_\mathbf{t}[\mathbf{s} + \mathbf{C}\boldsymbol{\Delta}(\mathbf{s} + \mathbf{a}) + \mathbf{V\tilde{n}}]. \tag{9}$$

Notice that in this case, the reduction of $(\mathbf{s} + \mathbf{a})$ into $\mathbf{x}$ can not be done.

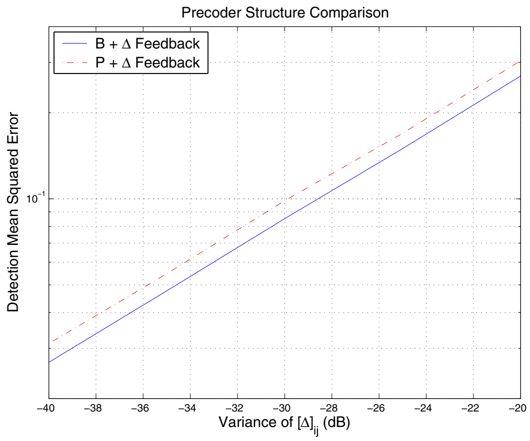

Fig. 3. Comparison of the mean squared error of the received signal for both precoding schemes for $n_T = 3$.

3.3. Precoder comparison

Gathering both expressions for the received signal, (8) and (9), it can be seen that the difference between them lies in the terms $\mathbf{Lx}$ in (8) and $\mathbf{C\Delta}(\mathbf{s}+\mathbf{a})$ in (9). Since $\mathbf{L}$ and $\mathbf{C}$ depend non-trivially on the channel realization $\mathbf{H}$, which, in addition, is random, it is very difficult to decide by formulae inspection which precoder is better, so we compared them by simulation. In Fig. 3 there is a plot of the mean squared error between the received signal vector and the transmitted data, $\mathbb{E}\{\|\mathbf{r}-\mathbf{s}\|^2\}$, versus the variance of the elements of $\mathbf{\Delta}$, $\delta_{ij} \sim \mathcal{CN}(0, \sigma_\delta^2)$, for both precoding structures. The simulation parameters have been set up to $n_T = n_R = 3$ and $\sigma^2 = 0$, the entries of $\mathbf{H}$ have been assumed to be zero mean unit variance i.i.d. circularly symmetric complex gaussian random variables and the entries of $\mathbf{s}$ follow i.i.d. uniform distributions in the interval $[-1, 1) \times [-j, j)$. It can be seen that the precoder $\mathbf{B}$ performs better than precoder $\mathbf{P}$. The authors conjecture that this can be due to the fact that in the received signal for precoder $\mathbf{P}$ no reduction of $(\mathbf{s}+\mathbf{a})$ into $\mathbf{x}$ can be done and, while $\mathbf{x}$ is bounded, $(\mathbf{s}+\mathbf{a})$ is not. Similar results have been obtained for different values of n_T.

4. ROBUST POWER ALLOCATION

In last section, it was found that the precoder structure $\mathbf{B}$ is slightly more robust to feedback errors than precoder structure $\mathbf{P}$. In this section, a robust power allocation that maximizes the worst-case achievable rates for the precoder structure $\mathbf{B}$ will be found.

As it was described in [4], the power transmitted through k-th antenna, P_k, is controlled by t_k by the relation $P_k = 2t_k^2/3$. In [4], it was found that, if no feedback errors are present, the Spatial Tomlinson-Harashima Precoder gets very close to achieve its capacity when each element of data vector $\mathbf{s}$ is uniformly distributed in the interval defined by vector $\mathbf{t}$ and the transmission power, P_T, is equally distributed among a set of active antennas. From now on, the set of active antennas will be numbered from 1 to N, *i.e.* $P_1, \ldots, P_N > 0$, $P_{N+1}, \ldots, P_{n_T} = 0$. In the following subsections, the power distribution will be adapted to maximize the worst-case achievable rates when feedback errors are present.

4.1. Problem Statement

The first order approximation in $\mathbf{\Delta}$ of the received signal vector for precoding structure $\mathbf{B}$ is

$$\mathbf{r}_{\hat{\mathbf{B}}} \approx M_{\mathbf{t}}[\mathbf{s} + \mathbf{\Delta x} + \mathbf{V}\tilde{\mathbf{n}}]. \tag{10}$$

If $\mathbf{\Delta x}$ is treated as an unknown interference the mutual information between k-th element of data vector, s_k, and the corresponding element of received signal, r_k, is

$$I(s_k, r_k) = \log(6P_k) - h\left(M_{t_k}\left[v_{kk}\tilde{n}_k + \sum_{j<k}\delta_{kj}x_j\right]\right), \tag{11}$$

where $\delta_{kj} = [\mathbf{\Delta}]_{kj}$ and $h(\cdot)$ denotes differential entropy (see [7]). Notice that, as it was stated in [4], x_k is uniformly distributed in $[-t_k, t_k) \times [-jt_k, jt_k)$ and thus its variance equals P_k.

Let us define the random variable $z_k = v_{kk}\tilde{n}_k + \sum_{j<k}\delta_{kj}x_j$ with power $\mathbb{E}\{|z_k|^2\} = v_{kk}^2\sigma^2 + \sum_{j<k}|\delta_{kj}|^2 P_j$. It can be easily verified that z_k can be approximately modeled as a complex gaussian random variable as long as

$$\max_j |\delta_{kj}|^2 P_j \lesssim v_{kk}^2\sigma^2/3. \tag{12}$$

For the cases where equation above is true, it was found in [4] that the mutual information expression (11) can be very well approximated by

$$I(s_k, r_k) \approx \log^+\left(\frac{6P_k}{\pi e(v_{kk}^2\sigma^2 + \sum_{j<k}|\delta_{kj}|^2 P_j)}\right), \tag{13}$$

where $\log^+(x) = \max(\log(x), 0)$. The achievable rates for the STHP structure will then be the sum of the mutual information of each active stream, $C = \sum_{k=1}^{N} I(s_k, r_k)$.

In order to describe the noise worst-case scenario we will consider that the squared moduli of the components of the error matrix $\mathbf{\Delta}$ are bounded, *i.e.* $|\delta_{kj}|^2 < \alpha_{kj}$. In addition, for the sake of simplicity we will assume that $\alpha_{kj} = \alpha, \forall k, j$. Notice that, since we are interested in the worst case, no generality is lost with last restriction, because, in case that the values of α_{kj} were different for some k, j, it would suffice to let $\alpha = \max_{k,j}\alpha_{kj}$.

From all the considerations above, the power distribution that will maximize the worst-case achievable rates when feedback errors are present will be the solution of the following maximin problem:

$$\begin{aligned} C^{rob} &= \max_{\{P_i\}} \min_{\{\delta_{ij}\}} \sum_{k=1}^{N} I(s_k, r_k), \\ &\quad \text{s.t.} \sum_{k=1}^{N} P_k = P_T, \quad |\delta_{ij}|^2 \le \alpha, \ \forall i, j. \end{aligned}$$

4.2. Problem Solution

The solution of the minimization part is trivial, since each term $I(s_k, r_k)$ is a decreasing function of $|\delta_{ij}|^2$ and each $|\delta_{ij}|^2$ is upper bounded independently of the others. Thus, the minimum will be attained when $|\delta_{ij}|^2 = \alpha, \forall i, j$. The resulting maximization problem obtained is a standard constrained optimization problem, and

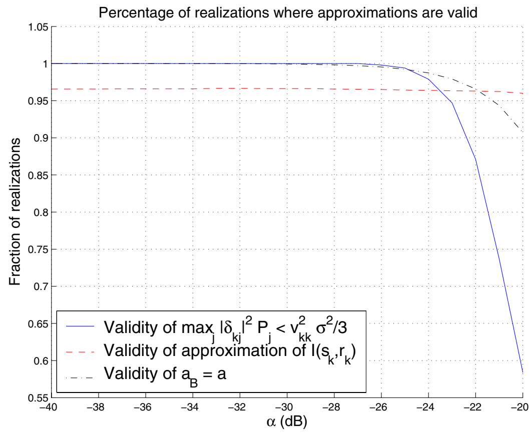

Fig. 4. Validation of the approximations done in the analysis.

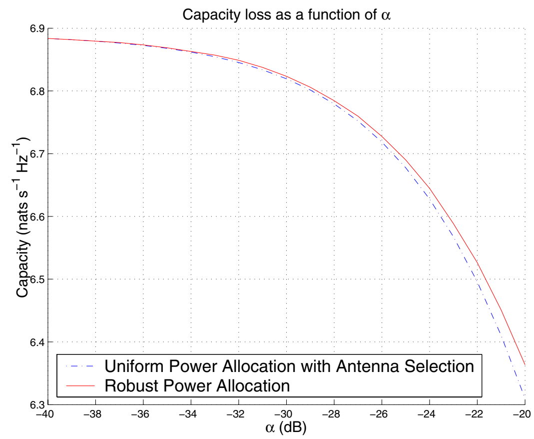

Fig. 5. Capacity for different power allocation strategies as a function of α.

can be solved with the use of the Lagrange method. The Lagrange equation is, up to a constant,

$$\mathcal{L} = \sum_{k=1}^{N} \log \left(\frac{P_k}{(v_{kk}^2 \sigma^2 + \alpha \sum_{j<k} P_j)} \right) + \lambda \left(\sum_{k=1}^{N} P_k - P_T \right). \tag{14}$$

The optimal power allocation should satisfy

$$\frac{\partial \mathcal{L}}{\partial P_k} = 0 \text{ for } k = 1, \dots, N \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \lambda} = 0, \tag{15}$$

with the additional constraint that $\{P_k\}$ is non-negative $\forall k$. With some basic manipulations from (14) and (15) a recursive relation of the type $P_{N-k} = f(P_{N-k+1}, \dots, P_N)$ for $k = 1, \dots, N-1$ between the assigned power to each antenna can be found as

$$P_{N-k} = P_{N-k+1} \frac{n_{N-k+1} + \alpha \left(P_T - \sum_{j>N-k} P_j \right)}{n_{N-k+1} + \alpha \left(P_T - \sum_{j>N-k+1} P_j \right)}, \tag{16}$$

for $k = 1, \dots, N-1$. Notice that the second factor in last equation is always lower than unity, which implies necessarily that $P_1 \leq P_2 \leq \dots \leq P_{N-1} \leq P_N$, which is reasonable since power interference is progressive in the sense that P_k interferes with streams from $k+1$ to N, see (13). The set of equations in (16) together with $\sum_{k=1}^{N} P_k = P_T$ can be solved numerically obtaining the robust power allocation.

5. SIMULATIONS AND CONCLUSIONS

For simulation purposes we have considered $n_T = n_R = 3$, $P_T/\sigma^2 = 15$ dB, and the entries of $\mathbf{H}$ have been assumed to be zero mean unit variance i.i.d. circularly symmetric complex gaussian random variables. The simulations were conducted using 10^5 channel realizations. As the robust capacity analysis has been done using numerous approximations (7), (10), (12), and (13), before presenting the simulations results, the validity of the approximations is shown in Fig. 4 by plotting the fraction of channel realizations in which the approximations are valid for different values of α. It can be seen that, for the particular values of the simulation parameters taken in this section, the capacity analysis is valid for values of α up to -22 dB. Notice that the validity of (10) has not been plotted as the differences between $\mathbf{L}$ and $\mathbf{\Delta}$ have been negligible for all channel realizations.

In Fig. 5, we have plotted the mean of the maximum achievable rates for the UPA-AS scheme (dotted) and for the robust power allocation strategy (solid). The capacity loss due to the presence of feedback errors between $\alpha = -40$ dB and $\alpha = -22$ dB is around 6%. In contrast, the maximum loss of UPA-AS approach with respect to the robust power allocation scheme is just 0.3%. Although both power allocation strategies yield almost the same capacity, it must be added that the robust power allocation differs substantially from the uniform power allocation, *e.g.* for $\alpha = -22$ dB, $P_1^{rob} \approx 0.28 P_T$, $P_2^{rob} \approx 0.32 P_T$, and $P_3^{rob} \approx 0.4 P_T$.

As a conclusion, it must be stated that UPA-AS strategy with precoding structure $\mathbf{B}$ is a very robust and simple structure which behaves almost optimally in terms of achievable rates, even in the presence of errors in the feedback channel.

6. REFERENCES

[1] M. Tomlinson, "New automatic equalizer employing modulo arithmetic," *Electronic Letters*, vol. 7, no. 2, pp. 138–139, Mar. 1971.

[2] H. Miyakawa and H. Harashima, "Information transmission rate in matched transmission systems with peak transmitting power limitation," *Nat. Conf. Rec., Inst. Electron., Inform., Commun. Eng. of Japan*, vol. 7, no. 2, pp. 138–139, Aug. 1972.

[3] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "Space-time transmission using Tomlinson-Harashima precoding," in *Proc. Proceedings of 4th ITG Conference on Source and Channel Coding*, Jan. 2002.

[4] M. Payaró, A. Pérez-Neira, and M.A. Lagunas, "Achievable rates for generalized spatial Tomlinson-Harashima precoding in MIMO systems," in *Proc. IEEE Vehic. Tech. Conf. Fall*, Sep. 2004.

[5] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "Tomlinson-Harashima precoding in space-time transmission for low-rate backward channel," in *Proc. IEEE International Zurich Seminar on Broadband Communications (IZS '02)*, Feb. 2002.

[6] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[7] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, New York: John Wiley & Sons, Inc., 1991.