REVERSE CHANNEL TRAINING FOR RECIPROCAL MIMO SYSTEMS WITH SPATIAL MULTIPLEXING

B. N. Bharath and Chandra R. Murthy

Electrical Communication Engineering Department Indian Institute of Science, Bangalore-560 012, India. email:{bharath, cmurthy}@ece.iisc.ernet.in

ABSTRACT

This paper investigates the problem of designing reverse channel training sequences for a TDD-MIMO spatial-multiplexing system. Assuming perfect channel state information at the receiver and spatial multiplexing at the transmitter with equal power allocation to the m dominant modes of the estimated channel, the pilot is designed to ensure an estimate of the channel which improves the forward link capacity. Using perturbation techniques, a lower bound on the forward link capacity is derived with respect to which the training sequence is optimized. Thus, the reverse channel training sequence makes use of the channel knowledge at the receiver. The performance of orthogonal training sequence with MMSE estimation at the transmitter and the proposed training sequence are compared. Simulation results show a significant improvement in performance.

Keywords: Channel estimation, training sequence design, reciprocity.

1. INTRODUCTION

Over the past decade or so, the use of multiple antennas at both the transmitter and receiver of a wireless communication system has received tremendous attention, as it offers a linear increase in capacity with the minimum of the number of transmit and receive antennas [1],[2]. However, much of the gain in capacity is at the cost of perfect Channel State Information (CSI) at Receiver (CSIR). CSIR can be acquired by sending a known training sequence (pilot) from the transmitter to the receiver. The accuracy of the estimate depends on the structure and duration of training sequence. It has been shown in [3] that orthogonal training sequences with training duration equal to the number of transmit antennas is optimal in an MMSE sense, when there is no prior knowledge of channel at the transmitter.

It is also well-known that a further gain in capacity is possible when there is perfect CSI at the transmitter (CSIT) [4] in addition to CSIR. Channel knowledge at the transmitter can be obtained in two ways: i) feedback of *quantized CSI* from the receiver to the transmitter, and ii) training in the reverse link, which we term *Reverse channel training (RCT)*. Much of the earlier work on CSI feedback has focussed on the feedback of quantized CSI to the transmitter, and in particular, the design of a FDD-MIMO spatial multiplexing system with equal power allocation and quantized CSI feedback was considered in [5]. Training in the reverse link, on the other hand, is possible only when the channel is reciprocal, i.e., for Time Division Duplex (TDD) systems. In this paper, therefore, we consider a TDD-MIMO system with perfectly reciprocal channels, and allow the receiver to exploit its channel knowledge in designing the training sequences. However, reciprocity requires well-calibrated transmit and receive RF chain components, which we assume here. Furthermore, in order to isolate the effect of channel estimation quality, we assume that the receiver has perfect CSI, which is a standard assumption in quantized CSI feedback related research as well. We also assume that the transmitter employs the m dominant modes of the estimated channel with pure spatial multiplexing, i.e., it sends m independent data streams using the estimated eigenvectors as beamforming vectors [5]. In this scenario, qualitatively speaking, one would expect that the receiver must allocate the available training power to the mdominant channel modes based on the gain of each mode, to guarantee the best possible estimate of the channel at the transmitter in terms of maximizing the forward link capacity. One of the key features of the proposed training sequence is that the entire knowledge of the sequence at the transmitter is not required in order to estimate the channel and hence can dynamically adapt to the varying channel.

The main contributions of this paper are as follows. Using a fairly general structure on the training sequence, we derive a lower bound on the forward link capacity with the estimated channel matrix at the transmitter. We then lean on tools from matrix perturbation theory to formulate the training sequence design problem as an optimization problem, using which the training parameters that maximize the lower bound on the channel capacity are determined. Simulation results show that significant performance improvement can be obtained compared to conventional orthogonal pilot based training schemes.

The rest of the paper is organized as follows. In Section 2, we present the system model. The design of training sequence is outlined in Section 3. One basic result in perturbation analysis is recapitulated in Section 4. The bound on down-link capacity is derived in Section 5. Optimization of training sequence is given in Section 6. Numerical results are presented in Section 7. Section 8 concludes the paper.

We use the following notation in this paper. $\mathbb{E}[.]$ denotes expected value of [.]. Capital letter will be used for matrix and small letter will be used for vectors. (.)^{*H*} denotes the transpose conjugate of a matrix, and |.| denotes the determinant of a matrix or the absolute value depending on the context. I_m is the $m \times m$ identity matrix.

2. SYSTEM MODEL

Consider a TDD-MIMO system in which users *A* and *B* wish to communicate with each other. The wireless link between *A* and *B* is assumed to be a quasi-static Rayleigh flat-fading channel. Let n_A and n_B be the number of antennas at *A* and *B* respectively. $H_{AB} \in \mathbb{C}^{n_B \times n_A}$ and $H_{BA} \in \mathbb{C}^{n_A \times n_B}$ represent the channel from *A* to *B* and *B* to *A* whose entries are independent and identically distributed (i.i.d.) zero mean, circularly symmetric complex

This work was supported by a FAST track fellowship from the Department of Science and Technology, Govt. of India.

Gaussian $\mathbb{CN}(0, 1)$. Due to reciprocity, we assume that $H_{AB} = H_{BA}^H$. Then $H_{AB} = U\Sigma V^H$ represent the SVD of the channel matrix where $U \in \mathbb{C}^{n_B \times n_B}$ and $V \in \mathbb{C}^{n_A \times n_A}$ are eigen-vectors of $H_{AB}H_{AB}^H$ and $H_{AB}^HH_{AB}$ respectively. Let *n* denote the rank of H_{AB} . $\Sigma \in \mathbb{R}^{n_B \times n_A}$ contains distinct singular values $\sigma_1 > \sigma_2 >, ..., > \sigma_n$ of H_{AB} . We assume that the channel is full rank i.e., $n = \min(n_A, n_B)$, which is true with probability one for a Rayleigh fading channel. The transmission involves training phase and data phase. During the training phase, user B sends a pilot sequence to enable A to obtain (partial) CSI, and during the data phase, user A transmits data to user B using the estimated channel to construct a precoding matrix. On either of the users, the received signal $Y_i \in \mathbb{C}^{n_i \times T}$, is

$$Y_i = H_{ji}X_j + W_i, \quad i, j \in \{A, B\}, \quad i \neq j$$
(1)

where $X_j \in \mathbb{C}^{n_j \times T}, j \in \{A, B\}$ is the signal matrix transmitted from the corresponding user, and T represents the number of channel uses. The average power constraint at either A or B during transmission is $\mathbb{E}[\operatorname{Tr}(X_j X_j^H)] \leq P$, where P is the total power. $W_i \in \mathbb{C}^{n_j \times T}$ represents thermal noise matrix of user $i \in \{A, B\}$ whose entries are distributed as $\mathbb{CN}(0, \sigma_i^2)$.

Consider the situation where user *B* has a perfect knowledge of the channel and *A* transmits data to *B*. User *B* transmits a training sequence $S_{\tau} \in \mathbb{C}^{n_B \times T_{\tau}}$ of duration T_{τ} to enable user *A* to estimate the channel. The received signal at *A*, $Y_A \in \mathbb{C}^{n_B \times T_{\tau}}$, is given by

$$Y_A = H^H_{AB} S_\tau + W_A. \tag{2}$$

Here, the total power constraint on the training sequence is given by $\operatorname{Tr}(S_{\tau}S_{\tau}^{H}) \leq P_{\tau}$. User A will make use of Y_A in (2) to estimate the channel for spatial multiplexing of data.

3. TRAINING SEQUENCE DESIGN

In this section, we propose a training sequence that attempts to maximize a lower bound on the capacity of the link from A to B. Since the spatial multiplexing of data is assumed, user A should know matrix V whose columns are the eigen-vectors of $H_{AB}^H H_{AB}$. The equation governing the *data* transmission from A to B is given by,

$$Y_B = H_{AB}\hat{V}_s X_d + W_B,\tag{3}$$

where $X_d \in \mathbb{C}^{m \times T_d}$ is the matrix of data symbols of duration T_d with the total power constraint given by $\mathbb{E}[X_d X_d^H] \leq (P_d T_d/m) I_m$. Here $\hat{V}_s \in \mathbb{C}^{n_A \times m}$ is a matrix containing the first m columns of V corresponding to the dominant eigen-vectors of H_{AB} . Its reasonable to assume that m is at most equal to the rank of H_{AB} , since we do spatial multiplexing of data. The receiver B processes the received *data* by pre-multiplying Y_B by U^H ,

$$\tilde{Y}_B = U^H Y_B = \Sigma V^H \hat{V}_s X_d + U^H W_B \tag{4}$$

The above equation can be viewed as a MIMO system with the effective channel matrix $\Sigma V^H \hat{V}_s$ as seen by the receiver *B*. At each symbol time, (4) reduces to

$$\tilde{y}_B = \Sigma V^H \hat{V}_s x_d + U^H w_B \tag{5}$$

where $\tilde{y}_B \in \mathbb{C}^{n_B \times 1}$, $x_d \in \mathbb{C}^{m \times 1}$ and $w_B \in \mathbb{C}^{n_B \times 1}$ are received data, transmitted data and noise vectors respectively. Since the user A is assumed to employ pure spatial multiplexing, the power constraint on x_d is given by $\mathbb{E}[x_d x_d^H] = \frac{P_d}{m} I_m$. Also, since the user B has perfect knowledge of the channel, we consider the following general structure on the training sequence,

$$S_{\tau} = \sqrt{P_{\tau} T_{\tau}} U D V^H \tag{6}$$

where, we restrict $D \in \mathbb{C}^{n_B \times n_A}$ to be a diagonal matrix with nonnegative entries. $||D||_F = 1$ ensures that the training power constraint is satisfied. Note that the most general pilot structure would allow D to be arbitrary matrix satisfying $||D||_F = 1$, but it will be shown that the structure proposed in this paper is simple, analytically tractable and offers a significant performance improvement compared to existing training schemes. The proposed training sequence has training length $T_{\tau} = n_A$. Since the user B has perfect CSI, the choice of D is made to depend on the current instantiation of the channel. Therefore, the goal is to set D that maximizes a lower bound on the down-link capacity. Substituting (6) in (2) and dividing by $\sqrt{P_{\tau}n_A}$, we get

$$\widetilde{Y}_A = V\Sigma DV^H + \widetilde{W}_A + \beta I_{n_A} \tag{7}$$

where $\tilde{Y}_A = Y_A/\sqrt{P_{\tau}n_A}$ and $\beta > 0$ is a regularization term which ensures that \tilde{Y}_A is a full rank matrix, and $\tilde{W}_A = W_A/\sqrt{P_{\tau}n_A}$. The full rank condition is necessary for the analysis to follow (cf. Sec. 4), in practical computation, it is in fact superfluous. Also, note that one can improve the accuracy of estimation by using a Hermitiansymmetrized version of \tilde{Y}_A . The symmetrizing operation has the effect of reducing the noise variance by 3dB which is not taken into account in the analysis. The 3dB reduction in noise power comes from eliminating the non-Hermitian symmetric part of \tilde{Y}_A . User A will make an estimate of V by eigen-value decomposition of (7). This is shown schematically below,

$$S_{\tau} \to \tilde{Y}_A \to \hat{V}\hat{\Sigma}\hat{V}^H$$
 (8)

We can consider \tilde{Y}_A as the perturbation of $V(\Sigma D + \beta I_{n_A})V^H$ which is Hermitian symmetric. Notice that due to diagonal the diagonal structure on D, knowledge of the training sequence is not required at A in order to estimate the eigen vectors. The next section will recapitulate the basic results of perturbation analysis which we use in this paper.

4. PERTURBATION ANALYSIS

Consider a first-order perturbation of an $n \times n$ full-rank Hermitian symmetric matrix G by an error matrix ΔG to get \hat{G} , i.e., $\hat{G} = G + \Delta G$. If the eigen-values are distinct, then the eigen-vector \hat{s}_k of \hat{G} can be approximated in terms of eigen-vector s_k of G as [6]

$$\hat{s}_k \approx s_k + \sum_{r=1, r \neq k}^n \frac{s_r^H \Delta G s_k}{\lambda_k - \lambda_r} s_r \tag{9}$$

where *n* is the rank of *G*, λ_k is the k^{th} largest eigen-value of *G*. For k = 1, 2, ..., n, we have $\hat{s}_i = S\tilde{d}_i$ where $S_n = [s_1, ..., s_n]$ is the $n \times n$ matrix of eigenvectors of *G* and

$$\tilde{d}_i = [\Delta d_{i1}, \dots, 1, \dots, \Delta d_{in}]^T,$$
(10)

where the '1' is at the i^{th} position, and Δd_{ij} is defined as,

$$\Delta d_{ij} \approx \frac{s_i^H \Delta G s_j}{\lambda_j - \lambda_i}, i \neq j.$$
⁽¹¹⁾

Since \hat{s}_k are orthonormal vectors, normalizing each of \tilde{d}_i 's and doing a similar approximation as in [7][8] we get,

$$S = S(I_n + E) \tag{12}$$

where,

$$E = \begin{pmatrix} \Delta d_1 & -\Delta d_{12} & \dots & -\Delta d_{1n} \\ \Delta d_{21} & \Delta d_2 & \dots & -\Delta d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta d_{n1} & \Delta d_{n2} & \dots & \Delta d_n \end{pmatrix}$$
(13)

and
$$\Delta d_k \approx -\frac{1}{2} \sum_{\substack{i=1, \\ i \neq k}}^n |\Delta d_{ki}|^2$$
 (14)

5. BOUNDS ON CAPACITY

Consider the following equation, which relates \hat{V} and V

$$\hat{V}_s = V\left(\left[\begin{array}{c}I_m\\0\end{array}\right]_{n_A \times m} + E_p\right) \tag{15}$$

where the $n_A \times m$ matrix E_p is defined as,

$$E_{p} \triangleq \begin{pmatrix} \Delta d_{1} & -\Delta d_{12} & \dots & -\Delta d_{1m} \\ \Delta d_{21} & \Delta d_{2} & \dots & -\Delta d_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta d_{n_{A}1} & \Delta d_{n_{A}2} & \dots & \Delta d_{n_{A}m} \end{pmatrix}.$$
(16)

and let v_i be the i^{th} column of V and with a slight abuse of notation, we let

$$\Delta d_{ij} \triangleq \begin{cases} \frac{v_i^H \widetilde{W}_A v_j}{\sigma_j d_j - \sigma_i d_i} & i \neq j, 1 \le i \le n, 1 \le j \le m \\ \frac{v_i^H \widetilde{W}_A v_j}{\sigma_j d_j} & n+1 \le i \le n_A, 1 \le j \le m \end{cases}$$
(17)

The above equation is obtained by observing that \tilde{Y}_A in (7) is the perturbation of $V\Sigma DV^H + \beta I_{n_A}$ and applying the analysis outlined in the previous section. Note that the off-diagonal terms are complex Gaussian. Then, from Lemma (2) in the Appendix, the $(i, j)^{th}$ non-diagonal entries have variance equal to $\sigma_A^2/(P_\tau n_A(\sigma_i d_i - \sigma_j d_j)^2)$, if $i \neq j, 1 \leq i \leq n, 1 \leq j \leq m$ and $\sigma_A^2/(p_\tau n_A \sigma_j^2 d_j^2)$, if $n + 1 \leq i \leq n_A, 1 \leq j \leq m$. Using (15) in (5) we get

$$\tilde{y}_B \approx \Sigma \left(\begin{bmatrix} I_m \\ 0 \end{bmatrix}_{n_A \times m} + E_p \right) x_d + \tilde{w}_B$$
(18)

where $\tilde{w}_B \triangleq U^H w_B \in \mathbb{C}^{n_A \times 1}$ is the noise vector. Equation (18) can be written in the following form

$$\tilde{y}_B \approx \Sigma_s x_d + e \tag{19}$$

where $e \triangleq \Sigma E_p x_d + \tilde{w}_B$ and $\Sigma_s \in \mathbb{R}^{n_B \times m}$ contains first m columns of Σ . The effective noise term e is uncorrelated with the data vector x_d i.e., $\mathbb{E}[ex_d^H] = 0_{n_B \times m}$, where the expectation is taken with respect to the training noise \widetilde{W}_A , the data noise w_B and the data symbol x_d . A capacity lower bound can be obtained by considering a suboptimal receiver that treats the effective noise as Gaussian, which is the essence of the following theorem.

Theorem 1 : For the system described by (19), a lower bound on the ergodic capacity is given by

$$C_L \triangleq \mathbb{E}_{\Sigma_s} \left(\log_2 \left| I_m + (P_d/m) \frac{\Sigma_s^H \Sigma_s}{\sigma_{eff}^2} \right| \right), \qquad (20)$$

where,

$$\sigma_{eff}^2 \approx \frac{P_d \sigma_A^2}{P_\tau n_A n_B m} \left[\sum_{i=1}^m \sum_{j=1, j \neq i}^{n_A} \frac{\sigma_i^2}{(\sigma_i d_i - \sigma_j d_j)^2} \right] + \sigma_B^2 \quad (21)$$

and $d_i = D_{ii}$, is the *i*th diagonal component of D and *i* = 1,..., min(n_A, n_B).

Proof: Since $\mathbb{E}[ex_d] = 0_{n_B \times m}$, the capacity can be lower bounded by [3][9],

$$C_{\Delta} = \mathbb{E}_{\Sigma_s} \left(\log_2 \left| I_m + \frac{P_d}{m} \Sigma_s^H (\mathbb{E}_e[ee^H])^{-1} \Sigma_s \right| \right).$$
(22)

Let $A \triangleq H_{AB}V \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}_{n_A \times n_A} V^H H_{AB}^H$. Then, $\log_2 |I + cA| = \log_2 |I + c\Sigma_s \Sigma_s^H|$ for any c > 0. Since H_{AB} is

 $\log_2 |I + cA| = \log_2 |I + c\Sigma_s \Sigma_s^H|$ for any c > 0. Since H_{AB} is rotationally invariant, it follows that A is also rotationally invariant, i.e, $\Pi(\Theta A) = \Pi(A)$ and $\Pi(A\Theta) = \Pi(A)$, where Π denotes the probability distribution of A and Θ is any unitary matrix. Therefore it follows from [3] that (22) can be further lower bounded to get (20), where $\sigma_{\text{eff}}^2 = \frac{1}{n_B} \mathbb{E}_e [\text{Tr}(ee^H)]$. Substituting for e, we get,

 $\sigma_{\text{eff}}^2 = \frac{P_d \sigma_A^{2^{(D)}}}{P_\tau n_A n_B m} \mathbb{E}_{w_A} \left[\text{Tr}(\Sigma E_p E_p^H \Sigma^H) \right] + \sigma_B^2.$ Further simplification is shown in the Appendix.

6. REVERSE TRAINING SEQUENCE OPTIMIZATION

We now wish to choose D that maximizes C_L , the lower bound on the capacity of the link from A to B. The maximization problem can be stated as

$$\max_{\substack{\{d_i:d_i\geq 0, ||D||_F^2=1\}}} \mathbb{E}_{\Sigma_s}\left(\log_2\left|I + (P_d/m)\frac{\Sigma_s^H \Sigma_s}{\sigma_{eff}^2}\right|\right).$$

which is equivalent to minimizing (21) :

$$\min_{\substack{d_i:d_i \ge 0, \\ ||D||_F = 1}} \sum_{i=1}^m \left[\sum_{j=1, j \ne i}^n \frac{\sigma_i^2}{(\sigma_i d_i - \sigma_j d_j)^2} + \frac{n_A - n}{d_i^2} \right].$$
(23)

Intuitively, when m = 1 (maximum ratio transmission), we would expect that user *B* would put all its training power solely on the dominant mode of the channel, which is established in the following lemma:

Lemma 1 : For a $2 \times n_A$ MIMO system with m = 1, the optimization problem given by (23) has solution $d_1 = 1, d_i = 0, \forall i \ge 2$.

Proof: Follows upon substituting m = 1, n = 2 in (23).

Now consider a $2 \times n_A$ MIMO system with m = 2. The minimization problem becomes

$$\min_{\substack{l_1, d_2 \ge 0, \\ l_1^2 + d_2^2 = 1}} \frac{\sigma_1^2 + \sigma_2^2}{(\sigma_1 d_1 - \sigma_2 d_2)^2} + \frac{(n_A - 2)}{d_1^2} + \frac{(n_A - 2)}{d_2^2}$$
(24)

The solution space is limited to $d_i \in [0, 1]$, but (24) does not have a closed form solution. However, if, for computational purposes, we let $\sigma_1 \gg \sigma_2$, we get

$$\min_{\{d_1:d_1\in[0,1]\}} \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2 d_1^2} + \frac{(n_A - 2)}{d_1^2 (1 - d_1^2)}$$
(25)

The solution to (25) is given by

$$d_1 = \sqrt{1 + \rho(n_A - 2)} \sqrt{1 - \sqrt{\frac{\rho(n_A - 2)}{1 + \rho(n_A - 2)}}},$$
 (26)

where $\rho \triangleq \sigma_1^2/(\sigma_1^2 + \sigma_2^2)$. We can use (26) for choosing d_1 and $d_2 = \sqrt{1 - d_1^2}$ as a simple, sub-optimal solution for all values of σ_1 and σ_2 . In our simulations, we have found that the loss in performance due to using the sub-optimal solution above relative to using a zero-finding algorithm that computes the optimal d_1 is very small.

7. SIMULATION RESULTS

The simulation consists of a MIMO-TDD Rayleigh flat fading channel with spatial multiplexing of data. The noise variances are assumed to be unity, i.e., $\sigma_A^2 = \sigma_B^2 = 1$. The mutual information for each noise and channel instantiation is evaluated using [5]

$$C = \log_2 \left| I + \frac{P_d}{m} H_{AB} \hat{V}_s \hat{V}_s^{\ H} H_{AB}^{\ H} \right| \tag{27}$$

The average mutual information is obtained by averaging (27) over all training noise and channel instantiations. The optimal values of d_1 and d_2 are evaluated using (26). Fig. 1 portrays the capacity performance of a 2 × 4 system and a 2 × 3 system. For comparison, we also plot the performance of orthogonal training with MMSE channel estimate and a system with perfect CSI. At all training powers, the proposed sequence outperforms the orthogonal training sequence. For example, $P_{\tau} = 5dB$, the proposed training sequence offer a training power gain of approximately 5dB. Fig. 2 shows the percent capacity loss with respect to the number of transmit antennas n_A for a given P_d and P_{τ} . For any number of transmit antennas the proposed RCT sequence outperforms the orthogonal training.



Fig. 1. Capacity vs P_{τ} in dB for a 2 × 3 and a 2 × 4 system with $P_d = 10$ dB and m = 2.

8. CONCLUSIONS

We have investigated the problem of reverse channel training in a TDD-MIMO system whose channel is reciprocal. We have assumed that the receiver (user *B*) is aware of the channel and hence can design a training sequence based on the number of modes used during down-link data transmission and the channel eigenvalues and eigenvectors. We proposed a specific training sequence which allocates different training powers to different modes. We used tools from perturbation analysis to derive a lower bound on the down-link capacity. This lower bound was maximized with respect to the parameters of the proposed training sequence and the training power allocation to different modes that maximize the lower bound were derived. The proposed RCT was compared with orthogonal training sequence with the same training power and duration. Simulation results showed a significant improvement in down-link capacity.

9. APPENDIX

Lemma 2 Let $v_i \in \mathbb{C}^{1 \times n}$, i = 1, ..., n be orthonormal vectors. Let the entries of $N \in \mathbb{C}^{n \times n}$ be i.i.d. circularly symmetric complex gaussian $\mathbb{CN}(0, 1)$. Then, $E[|a_{ij}|^2] = 1$, where $a_{ij} = v_i N v_j^H$. *Proof:* Follows directly from the orthonormality of v_i .

Proof of (21): Using Tr[AB] = Tr[BA] we can write, $Tr[\Sigma \mathbb{E}_{w_A}(E_p E_p^H) \Sigma^H] = Tr[\Sigma^H \Sigma \mathbb{E}_{w_A}(E_p E_p^H)]$. From Lemma 2 and (17), for $1 \leq j \leq m$, the off-diagonal elements of E_p are $\mathbb{CN}(0, \sigma_A^2/(P_\tau n_A(\sigma_i d_i - \sigma_j d_j)^2))$ for $1 \leq i \leq n$ and $\mathbb{CN}(0, \sigma_A^2/(P_\tau n_A \sigma_j^2 d_j^2))$ for $n + 1 \leq i \leq n_A$. The result follows by substituting for E_p and neglecting the diagonal terms $|\Delta d_1|^2, |\Delta d_2|^2, \dots$ which are higher order terms relative to $|\Delta d_{ij}|^2$.



Fig. 2. Percent capacity loss vs number of transmit antennas n_A for a system with $P_d = 1dB$, $P_{\tau} = 2$ dB, $n_B = 2$ and m = 2

10. REFERENCES

- I. Teletar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecomm*, vol. 10, no. 4, pp. 585 595, Dec. 1999.
- [2] 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, no. 4, pp. 311–335, Mar. 1998.
- [3] B. Hassibi and B. Hochwald, "How much training is needed in multi-antenna wireless links?" Wireless Personal Commun, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [4] T. Marzetta and B. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Trans on signal processing*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.
- [5] J. C. Roh and B. D. Rao, "Design and analysis of mimo spatial multiplexing systems with quantized feedback," *IEEE Trans. on Signal Processing*, vol. 54, no. 8, pp. 2874–2886, Aug. 2008.
- [6] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, 1st ed. Oxford U.K: Oxford Univ. Press, 1965.
- [7] C. R. Murthy, A. K. Jagannatham, and B. D. Rao, "Trainingonly and semi-blind channel estimation for maximum ratio transmission based mimo systems," *IEEE Trans on Signal processing*, vol. 54, no. 7, pp. 2546–2558, May 2000.
- [8] M. Kaveh and A. J. Barabell, "The statistical performance of the music and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust., Speech, Signal Process*, vol. 34, no. 2, pp. 331–341, Feb. 1986.
- [9] M. Medard, "The effect upon channel capacity in wireless communications with perfect and imperfect knowledge of the channel," *IEEE Trans on Inform Theory*, vol. 46, no. 3, pp. 933–946, May 2000.