ERGODIC CAPACITY ACHIEVING TRANSMIT STRATEGY IN MIMO SYSTEMS WITH STATISTICAL AND SHORT-TERM NORM CSI

Eduard A. Jorswieck, David Hammarwall, and Björn Ottersten

Signal Processing, School of Electrical Engineering, Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden
firstname.lastname@ee.kth.se

ABSTRACT

The type and quality of the channel state information at the transmitter of a fading multiple-input multiple-output system greatly affects the ergodic capacity of the wireless link. In order to compare and unify the different proposals of transmit strategies for different scenarios, recently classes of MIMO channels are introduced that share a common optimal transmit strategy. In this work, we derive the ergodic capacity achieving transmit strategy for the class of unitary invariant norm feedback which complements statistical channel information at the transmitter. The impact of the short-term feedback quality is illustrated by the beamforming optimality range. The higher the feedback norm is the more likely is single stream beamforming to be optimal.

***Index Terms*—** MIMO system, ergodic capacity, short-term and long-term CSI, unitary invariant norm feedback.

1. INTRODUCTION

The increasing need for fast and reliable wireless communication links opens the discussion about systems with multiple antennas both located at the transmitter and the receiver, so called multiple-input multiple-output (MIMO) systems [1]. Recently, [2] shows that MIMO systems have the ability to reach higher transmission rates than one-sided array links.

Many results regarding the capacity of multiple-input single-output (MISO) and MIMO systems under different levels of CSI and the corresponding transmission strategies are recently published [3]. For example, it is shown that even partial CSI at the transmitter can increase the capacity of a MISO system and transmission schemes for optimizing capacity in MISO mean- and covariance-feedback systems were analyzed [4]. The complete characterization of the impact of correlation on the ergodic capacity in MISO systems is provided in [5].

Recently, the question about the ergodic capacity achieving transmit strategy for the most general scenario with covariance and mean feedback is discussed in [6, 7, 8]. Further on, the optimal transmit strategy for a certain class of MIMO channels is described in [9]. In contrast to a certain class of channels, in this work we consider a certain class of short term feedback strategies that complement the long-term statistics known at the transmitter. The transmitter is assumed to know either the channel covariance or the channel mean. The short-term feedback is assumed to belong to the class of *unitary invariant norms* of the channel matrix. This class of short-term feedback comprehends many common feedback norms, e.g. ℓ-2 norm feedback or trace feedback. The optimal transmit directions for this type of CSI at the transmitter are shown to correspond to the eigenvectors of the channel covariance matrix or the channel mean gram matrix irrespective of the short-term feedback of this type. However, the optimal power allocation is affected by the short-term feedback. The impact of the channel norm is characterized in terms of the beamforming optimality range, i.e. the SNR range in which single-stream beamforming achieves the ergodic capacity.

With statistical CSI at the transmitter, the following relation between channel quality and optimal number of multiplexed data streams was observed: The better the channel, i.e. the higher the SNR, the more channels are supported. Interestingly, with additional short-term norm feedback, the relation between channel quality and optimal number of multiplexed data streams is the other way round: The larger the channel norm (for fixed SNR), the less channels are supported.

2. CHANNEL MODEL

Consider the quasi-static block flat-fading channel model. The received vector is given by $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$. The channel matrix $\boldsymbol{H}$ has dimension $n_R \times n_T$. Denote the minimum $n = \min(n_T, n_R)$ and the maximum $m = \max(n_T, n_R)$. The singular values of the channel matrix $\boldsymbol{H}$ are ordered in descending order and given by $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$. Denote the vector of singular values as $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_n]$. The AWGN vector $\boldsymbol{n}$ is complex independent and identically Gaussian distributed (iid) with zero mean and variance σ_n^2, i.e. $\boldsymbol{n} \sim \mathcal{CN}(0, \sigma_n^2)$. Denote the inverse noise power by $\Gamma = \frac{1}{\sigma_n^2}$. The SNR is then $\text{SNR} = \frac{P}{\sigma_n^2} = \Gamma P$.

In this work, we study either a correlated Rayleigh-fading MIMO channel or an uncorrelated Ricean-fading MIMO channel. The non line-of-sight (NLOS) channel matrix $\boldsymbol{H}$ for the case in which we have correlated transmit and correlated receive antennas is modelled as $\boldsymbol{H}_{NLOS} = \boldsymbol{R}_R^{\frac{1}{2}} \cdot \boldsymbol{W} \cdot \boldsymbol{R}_T^{\frac{1}{2}}$ with transmit correlation matrix $\boldsymbol{R}_T = \boldsymbol{U}_T \boldsymbol{D}_T \boldsymbol{U}_T^H$ and receive correlation matrix $\boldsymbol{R}_R = \boldsymbol{U}_R \boldsymbol{D}_R \boldsymbol{U}_R^H$. The random matrix $\boldsymbol{W}$ has zero-mean independent complex Gaussian identically distributed entries, i.e. $\boldsymbol{W} \sim \mathcal{CN}(0, \boldsymbol{I})$. $\boldsymbol{U}_T$ and $\boldsymbol{U}_R$ are the matrices with the eigenvectors of $\boldsymbol{R}_T$ and $\boldsymbol{R}_R$ respectively, and $\boldsymbol{D}_T$, $\boldsymbol{D}_R$ are diagonal matrices with the eigenvalues of the matrix $\boldsymbol{R}_T$ and $\boldsymbol{R}_R$, respectively, i.e. $\boldsymbol{D}_T = \text{diag}[\lambda_1^T, ..., \lambda_{n_T}^T]$ and $\boldsymbol{D}_R = \text{diag}[\lambda_1^R, ..., \lambda_{n_R}^R]$. Denote the vector of eigenvalues with $\boldsymbol{\lambda}_T$ and $\boldsymbol{\lambda}_R$ respectively. Without loss of generality, we assume that all eigenvalues are ordered with decreasing order, i.e. $\lambda_1^T \geq \lambda_2^T \geq ... \geq \lambda_{n_T}^T$.

Further on, we model the Ricean-fading channel matrix $\boldsymbol{H}$ by an additional deterministic matrix $\boldsymbol{H}_0$ with $\text{tr}(\boldsymbol{H}_0 \boldsymbol{H}_0^H) = n_T n_R$ as $\boldsymbol{H} = \sqrt{\frac{1}{K+1}} \boldsymbol{H}_{NLOS} + \sqrt{\frac{K}{K+1}} \boldsymbol{H}_0$ such that $\mathbb{E}[\boldsymbol{H}] = \sqrt{\frac{K}{K+1}} \boldsymbol{H}_0$. Let the singular value decomposition of $\boldsymbol{H}_0$ be given by $\boldsymbol{H}_0 = \boldsymbol{U}_0 \boldsymbol{\Lambda}_0 \boldsymbol{V}_0^H$ and define $\zeta = \sqrt{\frac{K}{K+1}}$ and $\zeta' = \sqrt{\frac{1}{K+1}}$. The Ricean

K-factor describes the ratio of the power of the LOS and NLOS component. The LOS component depends on the geometry of the transmission scenario and can have rank equal to one up to rank equal to $\min(n_T, n_R)$.

3. FEEDBACK MODEL, PERFORMANCE MEASURE, AND NORM-FEEDBACK

The receiver is assumed to have gained perfect CSI due to e.g. pilot assisted channel estimation. The transmitter has gathered a certain knowledge about the channel statistics. It is assumed that the transmitter knows either the channel correlation or the channel mean but not both. Additionally, the receiver feeds back some kind of short-term CSI that is a function of the instantaneous channel realization $f(\boldsymbol{H})$.

In the block fading model, the channel is constant for the coherence time T. It is assumed that the coherence time T is large enough to code over many blocks in order to achieve almost the mutual information. Then the mutual information maximized over the input distribution has its usual meaning as the instantaneous capacity [10]. Denote the instantaneous mutual information for a certain channel state $\boldsymbol{H}$ by

$$C(\boldsymbol{H}, \Gamma, \boldsymbol{Q}) = \log\det\left(\boldsymbol{I} + \Gamma \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\right). \tag{1}$$

The average with respect to $C(\boldsymbol{H}, \Gamma, \boldsymbol{Q})$[1] describes the overall performance and should be maximized. Since the transmitter does not know $\boldsymbol{H}$ exactly, it relies on the long-term CSI and the short-term CSI $f(\boldsymbol{H})$ in choosing $\boldsymbol{Q}$. Therefore, the average capacity is given by

$$C(\Gamma) = \mathbb{E}_\rho\left[\max_{\boldsymbol{Q}} \mathbb{E}_{\boldsymbol{H}}\left(C(\boldsymbol{H}, \Gamma, \boldsymbol{Q})\right)|f(\boldsymbol{H}) = \rho\right]. \tag{2}$$

For maximization of (2), the optimization problem for each ρ has to be solved under short-term power constraints.

Define a function f that acts on the instantaneous channel matrix $\boldsymbol{H}$, i.e. it maps from the set of $\mathbb{C}^{n_R \times n_T}$ matrices to the positive real numbers. The receiver feeds back the value of this function evaluated for the current channel realization. Then the optimal transmit strategy $\boldsymbol{Q}$ under long-term and short-term CSI is

$$\max_{\boldsymbol{Q}} \mathbb{E}\left[\log\det\left(\boldsymbol{I} + \Gamma \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\right)|f(\boldsymbol{H}) = \rho\right] \quad \text{s.t.} \quad \boldsymbol{Q} \succeq \boldsymbol{0}, \mathrm{tr}\boldsymbol{Q} \leq 1 \tag{3}$$

There are various ways how to define the short-term CSI function $f(\boldsymbol{H})$. In [11], the spatially correlated MISO Rayleigh fading channel with feedback of the channel norm was studied. This corresponds to $f(\boldsymbol{h}) = \lambda_1^2(\boldsymbol{h})$, i.e. the largest eigenvalue. Later in [12], the model was extended to include Ricean MISO channels as well.

Most common and often used short-term CSI functions f belong to the class of matrix-norms $f(\boldsymbol{H}) = ||\boldsymbol{H}||$. An important subclass of matrix-norms that we will use in the following are the *unitary invariant norms* $f(\boldsymbol{H}) = |||\boldsymbol{H}|||$. These norms have the property that $|||\boldsymbol{U}\boldsymbol{H}\boldsymbol{V}||| = |||\boldsymbol{H}|||$ for all unitary $\boldsymbol{V}$ and $\boldsymbol{U}$. Further on, the unitary invariant norms are related to symmetric gauge functions Φ on $\mathbb{R}^n$ by [13, Theorem IV.2.1]

$$|||\boldsymbol{H}|||_\Phi = \Phi(\boldsymbol{\lambda}(\boldsymbol{H})). \tag{4}$$

[1]Note that the function C is used with different sets of parameters.

The vector $\boldsymbol{\lambda}(\boldsymbol{H})$ denotes the vector with singular values of $\boldsymbol{H}$. The symmetric gauge function Φ is a norm, permutation invariant $\Phi(\boldsymbol{P}\boldsymbol{x}) = \Phi(\boldsymbol{x})$, invariant against plus minus one weighting, i.e. $\Phi(\pm x_1, ..., \pm x_n) = \Phi(x_1, ..., x_n)$, and normalized $\Phi(1, 0, ..., 0) = 1$. For each unitary invariant norm there is a corresponding symmetric gauge function and vice versa.

Consider the trace feedback of the outer product of the instantaneous channel matrix, i.e. $f(\boldsymbol{H}) = \mathrm{tr}\left(\boldsymbol{H}^H\boldsymbol{H}\right)$. This feedback corresponds to $\Phi(\boldsymbol{x}) = \sum_{k=1}^n x_k^2$ which is obviously a symmetric gauge function on $\mathbb{R}^n$. Therefore this feedback belongs to the class of unitary invariant norm feedback. Note that the feedback of ℓ-2 norm in MISO channels mentioned above is a special case of trace feedback. The ℓ-2 norm is the only unitary invariant vector norm [14].

In the following we will focus on the unitary invariant norm feedback. For fixed norm ρ, the average mutual information is given by

$$C(\rho, \Gamma, \boldsymbol{Q}) = \mathbb{E}\left[\log\det\left(\boldsymbol{I} + \Gamma\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\right) \mid |||\boldsymbol{H}||| = \rho\right]. \tag{5}$$

4. OPTIMAL TRANSMIT DIRECTIONS

The problem (3) is usually solved in two steps. The transmit covariance matrix is eigenvalue decomposed into the eigenvectors (or beamforming vectors) and eigenvalues (corresponding to power allocation). Then the optimal beamforming vectors are derived and based on this result the remaining power allocation is performed. For several important cases, the optimal beamforming vectors correspond to the eigenvectors of the known component of the channel, i.e. the correlation matrix or the mean matrix. As Myth 1 in [15] shows this does not hold in general but should be carefully proven. As pointed out in [12], the norm information provides in general also spatial information. However, the proof of the next theorem follows closely the lines of [16] and [17] and is omitted.

Theorem 1 *Assume that the short-term CSI is a unitary invariant norm of $\boldsymbol{H}$, i.e. $f(\boldsymbol{H}) = |||\boldsymbol{H}|||$. Then for the correlated Rayleigh-fading case, the optimal eigenvectors of $\boldsymbol{Q}$ that solve (3) are given by $\boldsymbol{U} = \boldsymbol{U}_T$. For the uncorrelated Ricean-fading case, the optimal eigenvectors are $\boldsymbol{U} = \boldsymbol{V}_0$. In both cases, the achievable transmission rate does not depend on $\boldsymbol{U}_T$ or $\boldsymbol{V}_0$.*

The expectation in the objective function is with respect to the conditional pdf $p(\boldsymbol{H}|\ |||\boldsymbol{H}|||)$. This pdf can be evaluated by Bayes law as

$$p(\boldsymbol{H}|\ |||\boldsymbol{H}|||) = \frac{p(|||\boldsymbol{H}|||\ |\boldsymbol{H})p(\boldsymbol{H})}{p(|||\boldsymbol{H}|||)}. \tag{6}$$

Note that $p(|||\boldsymbol{H}|||\ |\boldsymbol{H})$[2] only depends on the eigenvalues of $\boldsymbol{H}$. $p(|||\boldsymbol{H}|||)$ also depends only on the eigenvalues of $\boldsymbol{H}$.

5. OPTIMAL POWER ALLOCATION

Let us define the transformed channel matrix as $\tilde{\boldsymbol{H}} = \boldsymbol{U}\boldsymbol{H}\boldsymbol{U}^H$ with the optimal beamforming matrix $\boldsymbol{U}$ from Theorem 1. Further,

[2]This pdf is basically an indicator function, i.e. it equals either zero or one.

$\tilde{\boldsymbol{h}}_k$ is the k-th column of $\tilde{\boldsymbol{H}}$. For short-term unitary invariant norm-feedback, the remaining power allocation problem reads

$$\max \mathbb{E}\left[\log\det\left(\boldsymbol{I}+\Gamma\sum_{l=1}^{n_T} p_l \tilde{\boldsymbol{h}}_l \tilde{\boldsymbol{h}}_l^H\right)\Big| |||\boldsymbol{H}||| = \rho\right]$$
$$\text{s.t.} \quad \sum_{l=1}^{n_T} p_l \leq 1, p_l \geq 0. \tag{7}$$

The problem is convex and therefore the Karush-Kuhn-Tucker optimality conditions [18] apply. They are given by

$$\mathbb{E}\left[\tilde{\boldsymbol{h}}_l^H \left(\boldsymbol{I}+\Gamma\sum_{l=1}^{n_T} p_l \tilde{\boldsymbol{h}}_l \tilde{\boldsymbol{h}}_l^H\right)^{-1} \tilde{\boldsymbol{h}}_l \Big| |||\boldsymbol{H}||| = \rho\right] = \mu - \nu_l$$
$$p_k \nu_k = 0, p_k \geq 0, \nu_k \geq 0,$$
$$\sum_{l=1}^{n_T} p_l - 1 \leq 0, \mu\left(\sum_{l=1}^{n_T} p_l - 1\right) = 0, \mu \geq 0. \tag{8}$$

For all active transmit directions $p_l > 0$, the expectation in (8) is equal to the waterlevel μ. For all inactive transmit directions $p_l = 0$, the expectation in (8) is smaller than or equal to the waterlevel. Therefore, the beamforming optimality condition [19] is described in the next result.

Theorem 2 *The optimal power allocation is single-stream beamforming $p_1 = 1, p_2 = p_3 = ... = p_{n_T} = 0$ if and only if the following condition is satisfied*

$$\mathbb{E}\left[\tilde{\boldsymbol{h}}_1^H \left(\boldsymbol{I}+\Gamma \tilde{\boldsymbol{h}}_1 \tilde{\boldsymbol{h}}_1^H\right)^{-1} \tilde{\boldsymbol{h}}_1 \Big| |||\boldsymbol{H}||| = \rho\right] \geq$$
$$\mathbb{E}\left[\tilde{\boldsymbol{h}}_2^H \left(\boldsymbol{I}+\Gamma \tilde{\boldsymbol{h}}_1 \tilde{\boldsymbol{h}}_1^H\right)^{-1} \tilde{\boldsymbol{h}}_2 \Big| |||\boldsymbol{H}||| = \rho\right]. \tag{9}$$

Iterative algorithms for solving the power allocation problem are proposed e.g. in [16] and can be extended to the case in (7) including the conditional expectation. The characterization of the optimal power allocation in previous work showed that the higher the SNR the more streams are supported, i.e. $p_k > 0$.

We give the following characterization of the beamforming optimality range. Define the difference in Theorem 2 as a function $\Delta(\Gamma, \rho, \boldsymbol{\lambda}_T, \boldsymbol{\lambda}_R) = LHS - RHS$ of (9).

Lemma 3 *For fixed channel norm ρ and spatial correlation $\boldsymbol{\lambda}_T$, $\boldsymbol{\lambda}_R$, there always exists a SNR point Γ^* with*

$$\Delta(\Gamma, \rho, \boldsymbol{\lambda}_T, \boldsymbol{\lambda}_R) \begin{cases} \geq 0 & \text{for } \Gamma \leq \Gamma^* \\ < 0 & \text{for } \Gamma > \Gamma^* \end{cases}. \tag{10}$$

The Lemma means that for all channel norms ρ and spatial correlations there is a certain low SNR range in which single-stream beamforming achieves capacity. Define $\Delta(0) = \Delta(0, \rho, \boldsymbol{\lambda}_T, \boldsymbol{\lambda}_R)$.

Proof: The proof is based on the analytical properties of the function Δ with respect to Γ. Define the outer product of the k-th column of $\boldsymbol{H}$ as $\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{h}}_k \tilde{\boldsymbol{h}}_k^H$. Decompose the difference into the two parts

$$f_1(\Gamma) = \mathbb{E}\left[\text{tr}\left(\tilde{\boldsymbol{H}}_1 \left[\boldsymbol{I}+\Gamma\tilde{\boldsymbol{H}}_1\right]^{-1}\right) \Big| |||\boldsymbol{H}||| = \rho\right]$$

and $f_2(\Gamma)$ analogue. Note that $|||\boldsymbol{H}||| = |||\tilde{\boldsymbol{H}}|||$.

1. The functions $f_1(\Gamma)$ and $f_2(\Gamma)$ are monotonic decreasing with Γ. Further on, they are convex with respect to Γ.
2. It holds $\Delta(0) > 0$ and $\lim_{\Gamma\to\infty} \Delta(\Gamma) < 0$.

From these properties it follows, that there is exactly one intersection point between $f_1(\Gamma)$ and $f_2(\Gamma)$, and (10).

The first property of $f_1(\Gamma)$ and $f_2(\Gamma)$ follows easily from the first and second derivative of Δ with respect to Γ, i.e.

$$\begin{aligned} f_l'(\Gamma) &= -\mathbb{E}\left[\text{tr}\left(\tilde{\boldsymbol{H}}_l \left[\boldsymbol{I}+\Gamma\tilde{\boldsymbol{H}}_1\right]^{-2} \tilde{\boldsymbol{H}}_l\right) \Big| |||\boldsymbol{H}||| = \rho\right] \\ f_l''(\Gamma) &= +2\mathbb{E}\left[\tilde{\boldsymbol{h}}_l^H \tilde{\boldsymbol{H}}_l \left[\boldsymbol{I}+\Gamma\tilde{\boldsymbol{H}}_1\right]^{-3} \tilde{\boldsymbol{H}}_l \tilde{\boldsymbol{h}}_l \Big| |||\boldsymbol{H}||| = \rho\right]. \end{aligned}$$

The second property follows from

$$\Delta(0) = \mathbb{E}\left[||\tilde{\boldsymbol{h}}_1||^2 - ||\tilde{\boldsymbol{h}}_2||^2 \Big| |||\boldsymbol{H}||| = \rho\right] > 0$$

and the representation of $\Delta(\Gamma)$ after applying the matrix inversion lemma [14] to f_1 and f_2 given by

$$\begin{aligned} \Delta(\Gamma) &= \mathbb{E}\Big[\frac{||\tilde{\boldsymbol{h}}_1||^2 - ||\tilde{\boldsymbol{h}}_2||^2_\Gamma ||\tilde{\boldsymbol{h}}_2^H \tilde{\boldsymbol{h}}_1||^2 - \Gamma||\tilde{\boldsymbol{h}}_1||^2||\tilde{\boldsymbol{h}}_2||^2}{1+\Gamma||\tilde{\boldsymbol{h}}_1||^2} \\ &\qquad \Big| |||\boldsymbol{H}||| = \rho\Big]. \end{aligned} \tag{11}$$

Taking the limit $\Gamma \to \infty$ in (11) yields

$$\lim_{\Gamma\to\infty} \Delta(\Gamma) = \mathbb{E}\left[\frac{||\tilde{\boldsymbol{h}}_2^H \tilde{\boldsymbol{h}}_1||^2 - ||\tilde{\boldsymbol{h}}_1||^2||\tilde{\boldsymbol{h}}_2||^2}{||\tilde{\boldsymbol{h}}_1||^2} \Big| |||\boldsymbol{H}||| = \rho\right] < 0.$$

6. ILLUSTRATION

In Fig. (1), the optimal power allocation as well as the optimality condition from Theorem 2 is shown for fixed SNR Γ over the channel norm ρ. The zero of the optimality condition corresponds to the point from which on beamforming is optimal. This fits well to the optimal power allocation. A higher channel norm ρ corresponds to better channel conditions. In the scenario without short-term but long-term CSI available at the transmitter, the beamforming optimality range is for small SNR values. If short-term CSI is available, the beamforming optimality range is in the range of high ρ.

An explanation of the high norm behavior can be given by the observation in [11, Section 3.2]. The probability that the channel has rank one increases for increasing channel norm.

In Fig. (2), the gain by having channel norm information is shown. The norm helps in scenarios where either the best user can be chosen from a set of users [12] or if power allocation over norm realizations can be done, e.g. spectral power allocation. The ergodic capacity with covariance knowledge and with or without channel norm information at the transmitter is shown over the SNR. Without norm information, a random user is picked from a set of $K = 10$ users. If norm information is available, the user with highest norm is chosen. We plotted also the ergodic capacity with perfect CSI and user selection based on the largest eigenvalue of the channel matrix as an upper bound.

7. CONCLUSION

In this paper, the ergodic capacity achieving transmit strategy for a single-user MIMO system with perfect CSI at the receiver and statistical CSI as well as short-term unitary invariant norm feedback is

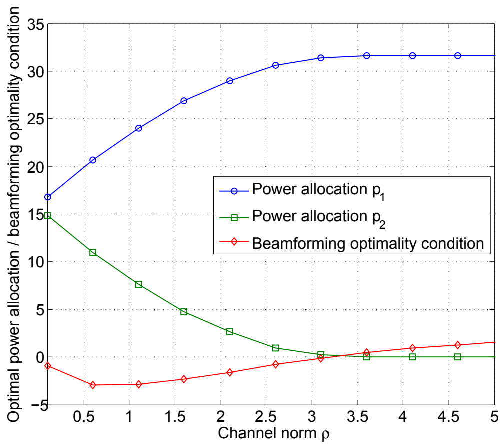

Fig. 1. Optimal power allocation (p_1 for stream 1 and p_2 for stream 2) and beamforming optimality function over channel norm ρ.

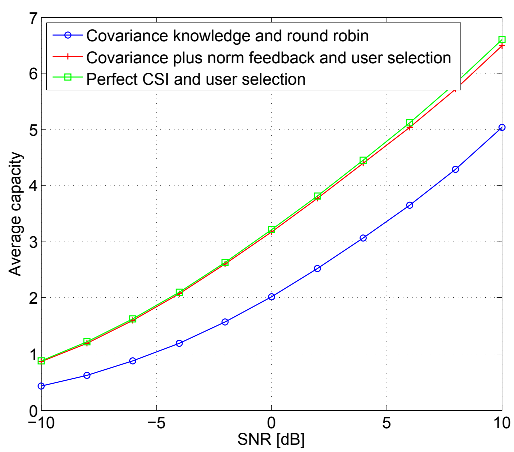

Fig. 2. Gain by channel norm feedback 10 users, 2×2 systems with correlation eigenvalues $\lambda_1 = 1.8, \lambda_2 = 0.2$.

characterized. For this class of short-term norm feedback the optimal beamforming vectors correspond to the eigenvectors of the long-term CSI component. The optimal power allocation is described in terms of the beamforming optimality range. In contrast to the common case of only statistical CSI, it turns out that the higher the channel norm, the fewer streams are optimally multiplexed.

8. REFERENCES

[1] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications 6: 311-335*, 1998.

[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov/Dec 1999.

[3] A. J. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Jourunal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[4] G. Barriac and U. Madhow, "Space time communication for OFDM with implicit channel feedback," *IEEE Trans. on Information Theory*, vol. 50, no. 12, pp. 3111– 3129, December 2004.

[5] E. A. Jorswieck and H. Boche, "Optimal transmission strategies and impact of correlation in multi-antenna systems with different types of channel state information," *IEEE Trans. on Signal Processing*, vol. 52, no. 12, pp. 3440–3453, Dec. 2004.

[6] M. Vu and A. Paulraj, "Optimal linear precoder for mimo wireless correlated channels with non-zero mean in space-time coded systems," *IEEE Trans. Signal Processing, to appear*, 2006.

[7] L. W. Hanlen and A. J. Grant, "Optimal transmit covariance for ergodic MIMO channels," *submitted to IEEE Trans. on Information Theory*, 2005.

[8] X. Zhang, D. P. Palomar, and B. Ottersten, "Robust design of linear MIMO transceivers under channel uncertainty," *Proc. of IEEE Intern. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2006.

[9] W. Rhee and G. Taricco, "On the ergodic capacity-achieving covariance matrix of certain classes of MIMO channels," *Proc. of IEEE Intern. Symp. on Information Theory*, pp. 1219–1223, 2006.

[10] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. on Information Theory*, vol. 44, no. 6, pp. 2619–2692, October 1998.

[11] D. Samuelsson (Hammarwall), M. Bengtsson, and B. Ottersten, "Improved multiuser diversity using smart antennas with limited feedback," in *Proc. European Signal Processing Conference*, 2005.

[12] D. Hammarwall and B. Ottersten, "Exploiting the spatial information provided by channel statistics and snr feedback," in *Proc. IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, 2006.

[13] R. Bhatia, *Matrix Analysis*, Springer - Verlag, 1997.

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

[15] A. Lozano, A. Tulino, and S. Verdu, *Space-Time Wireless Systems: From Array Processing to MIMO Communications*, chapter Multiantenna capacity: Myths and realities, Cambridge University Press, 2006.

[16] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. on Wireless Communications*, vol. 5, no. 3, pp. 662–671, Mar. 2006.

[17] D. Hösli and A. Lapidoth, "The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean," *Proc. 5th International ITG Conference on Source and Channel Coding*, January 2004.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[19] S. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *Trans. on Wireless Communication*, vol. 3, pp. 1165–1175, July 2004.