Fifth IEEE Workshop on Signal Processing Advances in Wireless Communications, Lisboa, Portugal, July 11-14, 2004

Diversity and Multiplexing Tradeoff of Beamforming for MIMO channels

Alba Pagès-Zamora[1]
Signal Processing and Communications Group
Technical University of Catalonia (UPC)
Jordi Girona 1-3, Campus Nord, Edifici D5
08034 Barcelona (SPAIN)
e-mail: alba@gps.tsc.upc.es

Javier R. Fonollosa[1]
Signal Processing and Communications Group
Technical University of Catalonia (UPC)
Jordi Girona 1-3, Campus Nord, Edifici D5
08034 Barcelona (SPAIN)
e-mail: fono@gps.tsc.upc.es

Luis García-Ordóñez[1]
Signal Processing and Communications Group
Technical University of Catalonia (UPC)
Jordi Girona 1-3, Campus Nord, Edifici D5
08034 Barcelona (SPAIN)
e-mail: luisg@gps.tsc.upc.es

Abstract — **This paper derives a tight upper bound on the diversity and spatial multiplexing gain for MIMO systems when the transmitter has perfect channel knowledge, generalizing the pioneering work of Zheng and Tse developed in the context of no channel state information at the transmit side. This result is obtained looking at the channel performance limits when both the SNR and the transmission rate tend to infinity, such that the relation between the rate and the capacity is constant. This result is then particularized for the optimal single beamforming scheme illustrating the performance loss incurred by any coding scheme based on this structure in the diversity versus rate tradeoff.**

I. Introduction

The increasing demand of higher data rates, specially in wireless systems, has motivated interest in the analysis of Multiple Input Multiple Output (MIMO) channels. In [1] and [2] it was shown that a MIMO channel formed by M antennas at the transmitter and N antennas at the receiver provides up to $\min\{M,N\}$ times the capacity of a single-input single-output (SISO) channel, without any increase of the required bandwidth or transmitted power. In practice, this promising capacity is difficult to achieve mainly due to the poor scattering conditions of the channel or the amount and quality of the Channel State Information (CSI) available at the receiver (CSI-R) and/or transmitter (CSI-T). This fact motivated a prolific research activity aimed at developing communication schemes specially conceived for MIMO channels. Traditionally, the design of these MIMO systems has been done under two different perspectives: the maximization of the diversity (to increase the transmission reliability) or the maximization of the transmission rate. Another common design criterion of MIMO systems is to assume that a certain degree of channel knowledge is present only at the receiver or at both, the receiver and the transmitter.

A new framework for comparison among MIMO systems has been recently proposed in [3]. Specifically, the optimal tradeoff curve between the diversity gain, denoted by d, and the spatial multiplexing gain, denoted by r, for a MIMO system without CSI-T is presented. A system has a diversity gain of d if the error probability decays like $1/SNR^d$, and a spatial multiplexing gain of r if the rate of the scheme is $r\log SNR$. The fundamental tradeoff $d(r)$ presented in [3] establishes that the diversity gain can not be increased without penalizing the spatial multiplexing gain. It is important to remark that, in this new framework, not only the SNR tends to infinity but also the transmission rate does, which is set to $R = r\log SNR$. The focus is then on the performance of any generic coding scheme formed by a family of codes, one code for each SNR level, such that the coding rate scales with $\log SNR$. Therefore the spatial multiplexing gain r represents the fraction of the ergodic capacity the MIMO system is able to sustain, and this fraction is evaluated asymptotically at high SNR, as illustrated by the following equation,

$$r = \min\{M,N\} \lim_{SNR\to\infty} \frac{R}{C}.$$

where C is the capacity of the multiple antenna channel. If we choose a family of MIMO coding schemes (with specific codes suited to every SNR value) with spatial multiplexing gain r close to $\min\{M,N\}$ we are forcing a data rate R close to the capacity limit C and thus the diversity gain will be reduced. Alternatively if the coding schemes are chosen such that their normalized spatial multiplexing gain r is much smaller than $\min\{M,N\}$ then they might benefit from higher diversity gains.

Note that the conventional derivation of the diversity gain of a given scheme at high SNR for a fixed rate code, as for instance the one obtained in [4], corresponds to the curve $d(r)$ for $r = 0$.

In this paper, we obtain the curve $d(r)$ for the MIMO single beamforming scheme. Since this MIMO system requires CSI available at the transmitter, Section 2 is devoted to discuss the optimal diversity-multiplexing tradeoff curve for a MIMO system when perfect CSI is assumed at the transmitter. The diversity-multiplexing curve for the single beamforming design in a MIMO system is addressed in Section 3. Finally, the last section summarizes the main results of the paper.

II. Outage and error probability with CSI-T

In this section we will first derive the outage probability of a MIMO Rayleigh channel when both the transmitter and the receiver have perfect CSI. Afterwards, we will discuss the relation between this outage probability and the error probability for any coding scheme transmitted through the channel.

Outage probability with perfect CSI-T

We are interested in the outage probability of a nonergodic fading MIMO channel in which the matrix channel $\mathbf{H}$ is chosen randomly but is fixed during the transmission. A total of M and N antennas are used in the transmit and receive side respectively. In this case, the received signal at time t, $\mathbf{y}_t \in \mathcal{C}^N$, is

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{n}_t, \quad \text{for } t = 1, 2, \cdots, \infty \tag{1}$$

[1]This work was financed by the Spanish/Catalan Science and Technology Commissions: TIC2003-05482, and 2001SGR-00268.

with $\mathbf{x}_t \in \mathcal{C}^M$ the transmitted signal at time t, $\mathbf{H} \in \mathcal{C}^{N\times M}$ is the channel matrix whose coefficients $[h_{ij}]$ (assumed to be known at both the transmit and receive sides) are i.i.d. complex Gaussian distributed with unit variance and $\mathbf{n}_t \sim \mathcal{C}N(\mathbf{0}, \mathbf{I}_N)$ is the additive noise with zero mean i.i.d. components.

The outage probability is the probability that the mutual information between the input and the output of the channel, $I(\mathbf{x}_t; \mathbf{y}_t)$ lays below a given rate R. When minimizing this probability, one can restrict its attention to the case in which $\mathbf{x}_t$ is stationary circularly symmetric and complex Gaussian with zero mean and covariance matrix equal to $\mathbf{Q_x} = E[\mathbf{x}_t\mathbf{x}_t^\dagger]$ since this distribution maximizes the mutual information for any given channel realization [1]. The mutual information $I(\mathbf{x}_t; \mathbf{y}_t)$ depends on the input distribution and the channel realization as follows

$$I(\mathbf{x}; \mathbf{y}|\mathbf{H} = H) = \log\det(\mathbf{I}_N + H\mathbf{Q_x}H^\dagger). \quad (2)$$

The distribution of $\mathbf{x}_t$ should satisfy trace$(\mathbf{Q_x}) \leq P_T$, with P_T the total transmitted power, and also equal to the average SNR at each receive antenna since the noise variance is normalized to 1. The outage probability for a given $\mathbf{Q_x}$ is obtained averaging over the random channel matrix $\mathbf{H}$ as shown in (3).

$$P_{out}(R, \mathbf{Q_x}) = P\left[\log\det(\mathbf{I}_N + \mathbf{H}\mathbf{Q_x}\mathbf{H}^\dagger) < R\right]. \quad (3)$$

In general, the outage probability is conditioned on the strategy adopted by the transmitter to select the input covariance matrix $\mathbf{Q_x}$. With perfect CSI at the transmitter, it is well known that the mutual information is maximized for

$$\mathbf{Q_x} = \mathbf{U}^\dagger\widetilde{\mathbf{Q}}_\mathbf{x}\mathbf{U} \quad (4)$$

[1], with $\mathbf{U}^\dagger \in \mathcal{C}^{M\times m}$ a unitary matrix whose columns are the eigenvectors of $\mathbf{H}^\dagger\mathbf{H}$, $\widetilde{\mathbf{Q}}_\mathbf{x} \in \mathcal{R}^{m\times m}$ a diagonal matrix and $m =$ rank$(\mathbf{H}^\dagger\mathbf{H})$. The optimal diagonal entries of $\widetilde{\mathbf{Q}}_\mathbf{x}$ are found via de "water-filling" approach and are equal to

$$\widetilde{q}_{ii} = \left(\mu - \lambda_i^{-1}\right)^+, \quad \text{for } i = 1, \cdots, m; \quad (5)$$

where $(x)^+$denotes $\max\{0, x\}$, $\{\lambda_i; i = 1, \cdots, m\}$ are the nonzero eigenvalues of $(\mathbf{H}^\dagger\mathbf{H})$ and $\mu \in \mathcal{R}$ is a constant value selected to satisfy the power constraint, that is, trace$(\mathbf{Q_x}) =$ trace$(\widetilde{\mathbf{Q}}_\mathbf{x}) = P_T = SNR$. Moreover, the "water-filling" solution minimizes the probability (3) with respect to the input distribution $\mathbf{Q_x}$, that is,

$$\begin{aligned} &P_{out}\left(R, \mathbf{Q_x}\right) \geq P_{out}\left(R, \mathbf{U}^\dagger\widetilde{\mathbf{Q}}_\mathbf{x}\mathbf{U}\right) \\ &= P\left[\log\prod_{i=1}^m\left(1 + \widetilde{q}_{ii}\lambda_i\right) < R\right]. \end{aligned} \quad (6)$$

We are particularly interested in the outage probability of the "water-filling" solution, that is, $P_{out}(R, \mathbf{U}^\dagger\widetilde{\mathbf{Q}}_\mathbf{x}\mathbf{U})$, at high SNR. Before going on, let us define the symbol $\doteq$ to denote *exponential equality*, i.e., we will write $P(SNR) \doteq SNR^b$ to denote

$$\lim_{SNR\to\infty}\frac{\log P(SNR)}{\log SNR} = b. \quad (7)$$

The symbols $\dot{\geq}$, $\dot{\leq}$ are similarly defined. At high SNR, since μ increases with the SNR, all the $\widetilde{q}_{ii}$ values of (5) tend to μ. Moreover, since $\widetilde{\mathbf{Q}}_\mathbf{x}$ has to satisfy the power constraint, at high SNR the optimal diagonal entries are all equal to

$$\lim_{SNR\to\infty}\widetilde{q}_{ii} = \lim_{SNR\to\infty}\left(\mu - \lambda_i^{-1}\right)^+ = \lim_{SNR\to\infty}\mu = \frac{SNR}{m}. \quad (8)$$

Substituting (8) into (6), the outage probability of the "water-filling" solution at high SNR follows:

$$P_{out}(R, \mathbf{U}^\dagger\widetilde{\mathbf{Q}}_\mathbf{x}\mathbf{U}) \doteq P\left[\log\left(\prod_{i=1}^m(1 + SNR\lambda_i)\right) < R\right]. \quad (9)$$

Moreover, if we consider a family of codes in which the rate increases as $R = r\log SNR$ the outage probability can be expressed as

$$\begin{aligned} &P_{out}(R, \mathbf{U}^\dagger\widetilde{\mathbf{Q}}_\mathbf{x}\mathbf{U}) \doteq \\ &\doteq P\left[\log\left(\prod_{i=1}^m(1 + SNR\lambda_i)\right) < \log SNR^r\right] = \\ &= P\left[\left(\prod_{i=1}^m(1 + SNR\lambda_i)\right) < SNR^r\right] \end{aligned} \quad (10)$$

As shown in [3], the SNR exponent of (10) can be explicitly computed and is equal to

$$P_{out}(R, \mathbf{U}^\dagger\widetilde{\mathbf{Q}}_\mathbf{x}\mathbf{U}) \doteq SNR^{-d_{out}(r)}, \quad (11)$$

where $d_{out}(r)$ is the piecewise-linear function with connecting points $\{k, (M-k)(N-k)\}$ for $k = 0, .., \min\{M, N\}$ shown in dashed line in Figure (1). This SNR exponent is therefore exactly the same as the one of the outage probability of a MIMO system without CSI-T. That is, the diversity and normalized coding rate tradeoff expressed in terms of the outage probability for a MIMO channel is not conditioned on the potential knowledge the transmitter could have about the CSI.

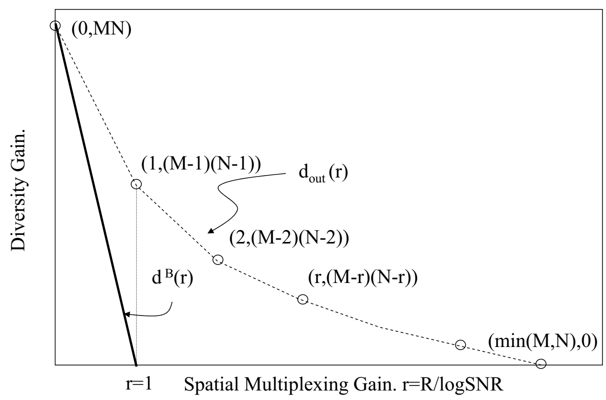

Figure 1: (dashed) SNR exponent of the P_{out} for a MIMO channel with and without CSI-T. (solid) Diversity-multiplexing tradeoff curve for the single beamforming scheme.

Error probability with perfect CSI-T

In the previous section, the diversity and rate tradeoff in terms of the outage probability offered by a MIMO channel with perfect CSI-T and CSI-R has been evaluated. Now we explore its relation with the error probability associated to any coding scheme transmitted through the channel. In order to do so we generalize the channel model of (1) to a block of l

symbol periods in which the channel matrix $\mathbf{H}$ is assumed to remain constant, i.e., we consider that the coding block length l is much smaller than the coherence time of the channel:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N}, \tag{12}$$

where $\mathbf{Y} \in \mathcal{C}^{N\times l}$, is formed by the vectors $\mathbf{y}_t$ for $t = 1, \cdots, l$ and $\mathbf{X} \in \mathcal{C}^{M\times l}$ and $\mathbf{N} \in \mathcal{C}^{N\times l}$ are similarly defined. Using Fano's inequality, it can be shown that the outage probability provides a lower bound on the error probability, that is,

$$P_e(SNR) \dot{\geq} SNR^{-d_{out}(r)}, \tag{13}$$

for any coding scheme transmitted through the MIMO channel and (12) optimized with the "water-filling" solution of (5).

Proof: Let us assume we use a codebook $\mathcal{C}$ of size 2^{Rl} and let $\mathbf{X}$ be the input to the channel, which is uniformly drawn from the codebook $\mathcal{C}$. Since CSI is available at the transmitter the distribution of $\mathbf{X}$ depends on the channel realization $\mathbf{H}$. Conditioned on a specific channel realization $\mathbf{H} = H$, and using Fano's inequality we have:

$$1 + P(\text{error}|\mathbf{H} = H)Rl \geq \mathcal{H}(\mathbf{X}|\mathbf{Y}, \mathbf{H} = H), \tag{14}$$

for any coding scheme and where $\mathcal{H}(\mathbf{X})$ indicates the entropy of $\mathbf{X}$. By direct substitution of the definition of the mutual information, (14) is equivalent to

$$1 + P(\text{error}|\mathbf{H} = H)Rl \geq \mathcal{H}(\mathbf{X}|\mathbf{H} = H) - I(\mathbf{X}; \mathbf{Y}|\mathbf{H} = H). \tag{15}$$

Considering $\mathbf{X} = [\mathbf{x}_t]$ and $\mathbf{Y} = [\mathbf{y}_t]$ for $t = 1, \cdots, l$, the right hand side of (15) can be bounded as follows. First, the conditional entropy of $\mathbf{X}$ is lower bounded by the conditional entropy of any of the $\mathbf{x}_t$

$$\mathcal{H}(\mathbf{X}|\mathbf{H} = H) \geq \mathcal{H}(\mathbf{x}_t|\mathbf{H} = H); \quad \forall t = 1, \cdots, l, \tag{16}$$

whereas, on the other hand, the conditional entropy of $\mathbf{Y}$ can be upper bounded by the sum of the conditional entropies of $\mathbf{y}_t$ $t = 1, \cdots, l$. Also, making use of the chain rule for the entropy,

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}|\mathbf{H} = H) &= \mathcal{H}(\mathbf{Y}|\mathbf{H} = H) - \mathcal{H}(\mathbf{Y}|\mathbf{X}, \mathbf{H} = H) \leq \\ &\leq \sum_{t=1}^{l} \mathcal{H}(\mathbf{y}_t|\mathbf{H} = H) - \sum_{t=1}^{l} \mathcal{H}(\mathbf{y}_t|\mathbf{y}_{t-1}, \cdots, \mathbf{y}_1, \mathbf{X}, \mathbf{H} = H). \end{aligned}$$

Now, since the MIMO channel in (12) is memoryless $\mathbf{y}_t$ depends only on $\mathbf{x}_t$ and is conditionally independent of everything else,

$$\begin{aligned} & I(\mathbf{X}; \mathbf{Y}|\mathbf{H} = H) \leq \\ &\leq \sum_{t=1}^{l} \mathcal{H}(\mathbf{y}_t|\mathbf{H} = H) - \sum_{t=1}^{l} \mathcal{H}(\mathbf{y}_t|\mathbf{x}_t, \mathbf{H} = H) = \\ &= \sum_{t=1}^{l} I(\mathbf{x}_t; \mathbf{y}_t|\mathbf{H} = H). \end{aligned} \tag{17}$$

Substituting (16) and (17) back into (15), we obtain

$$\begin{aligned} & 1 + P(\text{error}|\mathbf{H} = H)Rl \geq \\ & \geq \mathcal{H}(\mathbf{x}_t|\mathbf{H} = H) - \sum_{t=1}^{l} I(\mathbf{x}_t; \mathbf{y}_t|\mathbf{H} = H). \end{aligned} \tag{18}$$

The conditional entropy of the input distribution can be evaluated explicitly since this follows the "water-filling" solution, that is, $\mathbf{Q_x} = \mathbf{U}^\dagger \widetilde{\mathbf{Q}}_\mathbf{x} \mathbf{U}$, with $\mathbf{U}^\dagger \in \mathcal{C}^{\mathbf{M}\times\mathbf{m}}$ a unitary matrix whose columns are the eigenvectors of $\mathbf{H}^\dagger\mathbf{H}$ and $\widetilde{\mathbf{Q}}_\mathbf{x} \in \mathcal{R}^{m\times m}$ a diagonal matrix with diagonal entries of $\widetilde{\mathbf{q}}_{ii}$ defined in (5). In this case, the entropy of $\mathbf{x}_t$ for a given channel realization is equal to

$$\begin{aligned} \mathcal{H}(\mathbf{x}_t|\mathbf{H} = H) &= \log\det(\pi e \mathbf{Q_x}) = \log\det\left(\pi \mathbf{e} \widetilde{\mathbf{Q}}_\mathbf{x}\right) = \\ &= \log\det(\pi e)^m + \sum_{i=1}^{m} \log \widetilde{\mathbf{q}}_{ii}. \end{aligned} \tag{19}$$

Considering $\widetilde{\mathbf{q}}_{ii}$ tends to $\frac{SNR}{m}$ at high SNR as shown in (8), (19) becomes exponentially equivalent to

$$\mathcal{H}(\mathbf{x}_t|\mathbf{H} = H) \doteq m \log SNR. \tag{20}$$

Then, assuming a rate $R = r \log SNR$ and the fact that the $\mathbf{x}_t$ are equally distributed for $\forall t = 1, \cdots, l$, i.e., $\sum_{i=1}^{l} I(\mathbf{x}_i; \mathbf{y}_i|\mathbf{H} = H) = lI(\mathbf{x}_t; \mathbf{y}_t|\mathbf{H} = H)$ the inequality of (18) can be expressed as

$$P(\text{error}|\mathbf{H} = H) \dot{\geq} \frac{m}{l\,r} - \frac{I(\mathbf{x}_t; \mathbf{y}_t|\mathbf{H} = H)}{r \log SNR} - \frac{1}{l\,r \log SNR}. \tag{21}$$

From now on, we can follow a similar procedure as the one shown in [3] to prove the outage probability provides a lower bound on the error probability associated to any coding scheme used in the channel. First of all, the last term goes to 0 as $SNR \to \infty$. Then, in order to get the average error probability, we average over $\mathbf{H}$.

$$P_e(SNR) = \varepsilon_H \left[P(\text{error}|\mathbf{H} = H)\right] \tag{22}$$

Now for any $\delta > 0$, for any H in the set

$$\mathcal{D}_\delta \triangleq \{H : I(\mathbf{x}_t; \mathbf{y}_t|\mathbf{H} = H) < (r - \delta) \log SNR\} \tag{23}$$

the probability of error is lower-bounded by

$$P_e(SNR) \dot{\geq} \left(\frac{m}{l\,r} - \frac{r - \delta}{r} + o\,(1)\right) P\left(\mathcal{D}_\delta\right). \tag{24}$$

Since the "water-filling" solution maximizes the mutual information $I\left(\mathbf{x}_t; \mathbf{y}_t|\mathbf{H} = H\right)$, $P(\mathcal{D}_\delta)$ is therefore minimized and becomes equivalent to the outage probability (11). Therefore,

$$\begin{aligned} P_e(SNR) &\dot{\geq} \left(\frac{m}{l\,r} - \frac{r - \delta}{r} + o(1)\right) SNR^{-d_{out}(r-\delta)} \\ &\doteq SNR^{-d_{out}(r-\delta)} \end{aligned}$$

Taking $\delta \to 0$, by the continuity of $d_{out}(r)$, we have

$$P_e(SNR) \dot{\geq} SNR^{-d_{out}(r)} \qquad \square$$

The previous proof shows that the outage probability provides a lower bound on the probability of error of any coding scheme based on the "water-filling" solution. Moreover, it can also be shown that, even without CSI-T, this lower bound of the error probability is achievable, that is the bound is tight, since a code can be constructed for which $P_e(SNR) \doteq SNR^{-d_{out}(r)}$, whenever the block coding length is $l \geq M + N - 1$ [3]. In the case $l < M+N-1$ however, we can only conclude that the P_e of a MIMO scheme with perfect CSI-T is lower bounded by the SNR exponent of the outage probability as stated in (13).

III. Tradeoff curve of beamforming

This section is devoted to derive the diversity-multiplexing tradeoff curve associated to the single beamforming scheme in MIMO channels, and to compare this result to the general tradeoff offered by the channel as derived in Section 2.

Optimal MIMO beamforming scheme

We consider the communication system of Figure (2) with a beamvector $\mathbf{b} \in \mathcal{C}^M$ at the transmitter and a linear equalizer $\mathbf{a}^\dagger \in \mathcal{C}^N$ at the receiver. Each symbol $x_t \in \mathcal{C}$ (we consider zero mean unit-variance time-uncorrelated symbols, i.e. $\sigma_x^2 = E\left[|x_t|^2\right] = 1$ is multiplied by the beamvector $\mathbf{b}$ before being transmitted with $\mathbf{s}_t = \mathbf{b}x_t$. For simplicity, we will drop the time index notation in x and $\mathbf{s}$. The signal $\mathbf{s} \in \mathcal{C}^M$ is sent through the MIMO channel and the received signal $\mathbf{y} \in \mathcal{C}^N$is equal to

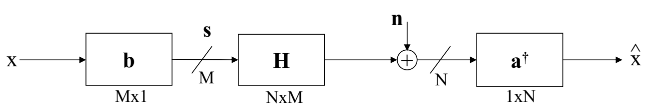

Figure 2: Single beamforming scheme for a MIMO channel.

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} \qquad (25)$$

where $\mathbf{n} \in \mathcal{C}^N$ is zero mean additive white Gaussian noise, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$. The received signal is then linearly equalized with $\mathbf{a}^\dagger$ and $\widehat{x} \in \mathcal{C}$ is the linear estimation of x.

$$\widehat{x} = \mathbf{a}^\dagger \mathbf{y} \qquad (26)$$

As shown in e.g. [6], the classical single beamforming design arises from the minimization of the Mean Square Error (MSE) of the estimated symbol, $\widehat{x}$. The optimal receiver is $\mathbf{a} = \left(\mathbf{H}\mathbf{b}\mathbf{b}^\dagger\mathbf{H}^\dagger + \mathbf{I_N}\right)^{-1}\mathbf{H}\mathbf{b}$ and the resulting MSE is $(1 + \mathbf{b}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{b})^{-1}$. The problem reduces to the design of $\mathbf{b}$ solving the following optimization problem.

$$\min_{\mathbf{b}} (1 + \mathbf{b}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{b})^{-1} \quad \text{subject to } \mathbf{b}^\dagger\mathbf{b} \leq P_T. \qquad (27)$$

The optimal solution to this optimization problem is $\mathbf{b} = \sqrt{P_T}\mathbf{u}_{\max}$, where $\mathbf{u}_{\max}$ is the eigenvector of $(\mathbf{H}^\dagger\mathbf{H})$ corresponding to the largest eigenvalue, denoted by $\lambda_{\max}$. The resulting optimal beamforming scheme is equivalent to the following SISO system

$$\widehat{x} = \widehat{h}x + \widehat{a}w, \qquad (28)$$

where w is a normalized equivalent unit-variance white Gaussian noise and with $\widehat{h} \in \mathcal{R}$ and $\widehat{a} \in \mathcal{R}$ defined as follows

$$\widehat{h} = \frac{P_T\lambda_{\max}}{1 + P_T\lambda_{\max}} \quad ; \ \widehat{a} = \frac{\sqrt{P_T}\lambda_{\max}^{1/2}}{1 + P_T\lambda_{\max}}. \qquad (29)$$

Outage probability and error probability of the optimal beamforming scheme

We are first interested in obtaining the outage probability when transmitting using the optimal beamforming scheme. For this purpose, we deal with the equivalent scalar channel provided by (28) and (29). Thus, given a channel realization $\mathbf{H} = H$ the mutual information for this system is equal to

$$\begin{aligned} I\left(x; \widehat{x}|\mathbf{H} = H\right) &= \mathcal{H}\left(\widehat{x}|\mathbf{H} = H\right) - \mathcal{H}\left(\widehat{x}|x, \mathbf{H} = H\right) \\ &= \mathcal{H}\left(\widehat{x}|\mathbf{H} = H\right) - \mathcal{H}\left(\widehat{a}w\right) \end{aligned} \qquad (30)$$

Since $\widehat{x}$ is zero mean complex Gaussian, the mutual information becomes

$$I(x; \widehat{x}|\mathbf{H} = H) = \log(\sigma_{\widehat{x}}^2) - \log(\sigma_{\widehat{a}w}^2), \qquad (31)$$

where $\sigma_{\widehat{a}w}^2$ is the variance of the output noise and $\sigma_{\widehat{x}}^2$ is the variance of the output signal. Recalling $\sigma_x^2 = \sigma_w^2 = 1$, it can easily be seen that

$$\sigma_{\widehat{x}}^2 = \widehat{h}^2 + \sigma_{\widehat{a}w}^2 = \widehat{h}^2 + \widehat{a}^2. \qquad (32)$$

Substituting (29) and (32) in (31), the mutual information between the input and the output can be expressed as follows

$$I\left(x; \widehat{x}|\mathbf{H} = H\right) = \log\left(1 + P_T\lambda_{\max}\right). \qquad (33)$$

We will consider a target rate $R = r\log SNR$ and recall $P_T \triangleq SNR$. The diversity and rate tradeoff curve for the outage probability of the optimal beamforming scheme is

$$\begin{aligned} P_{out}^B(R) &= P\left[\log\left(1 + SNR\lambda_{\max}\right) < r\log SNR\right] \\ &\doteq P\left[\lambda_{\max} < SNR^{r-1}\right] \doteq SNR^{-d_{out}^B(r)}. \end{aligned} \qquad (34)$$

From (13) we know that the error probability of any channel coding scheme used in combination with single beamforming is lower bounded by $d_{out}^B(r)$. We will now look for an upper bound of this probability to proof that this bound is tight. For this purpose, we will assume a QAM modulation for the transmitted symbols. In order to obtain a transmission rate of $R = rlogSNR$ the constellation size is chosen to be SNR^r such that the minimum distance between symbols tends to $SNR^{-r/2}$ as the SNR increases. Let x_0 and x_1 be two arbitrary symbols of the QAM constellation, the Pairwise Error Probability (PEP) of the beamforming system is given by

$$\begin{aligned} P(x_0 \to x_1) = P(|\widehat{x} - x_1|^2 < |\widehat{x} - x_0|^2) = \\ = P(\left|\frac{1}{2}SNR^{1/2}\lambda_{\max}^{1/2}(x_0 - x_1)\right|^2 < |w|^2) \\ \dot{\leq} P(\lambda_{\max} < 4\,|w|^2\,SNR^{r-1}). \end{aligned} \qquad (35)$$

Then, since when using QAM modulation there are at most four nearest neighbors to x_0, the overall error event is the union of these four pairwise error events which means that the probability of error is upper bounded by four times the PEP. Therefore, using (34), (13) and (35) we can conclude that the following inequalities hold

$$P\left[\lambda_{\max} < SNR^{r-1}\right] \dot{\leq} P_e^B \dot{\leq} P(\lambda_{\max} < 4\,|w|^2\,SNR^{r-1}).$$

Since both the upper and the lower bounds have the same SNR exponent, the tradeoff curve of the error probability of the single beamforming scheme, denoted by $d^B(r)$, can be computed as

$$P_e^B \doteq P\left[\lambda_{\max} < SNR^{r-1}\right] \doteq SNR^{-d^B(r)}. \qquad (36)$$

Tradeoff curve for the optimal beamforming

In this section, we derive the diversity-multiplexing tradeoff curve $d^B(r)$ at high SNR for the optimal single beamforming scheme, that is, we evaluate (36). The evaluation of (36) requires the distribution function of $\lambda_{\max}$ of $\mathbf{H}^\dagger\mathbf{H}$ [1], which, as shown in Appendix A, is given by

$$p(\lambda_{\max} < \lambda) = \frac{\left|\mathbf{\Sigma}^{-1}\lambda\right|^N}{K_{M,N}} \int_{\mathbf{0}<\mathbf{X}<\mathbf{I}} e^{-\lambda tr \mathbf{\Sigma}^{-1}\mathbf{X}} \left|\mathbf{X}\right|^{N-M} (d\mathbf{X}). \qquad (37)$$

where $K_{M,N}$ is a normalizing factor, $(d\mathbf{X})$ is defined in the appendix in equation (42) and $\mathbf{\Sigma} \in \mathcal{C}^{M\times M}$ is the covariance matrix of the rows of $\mathbf{H}$. Since we consider the channel matrix $\mathbf{H}$ with coefficients h_{ij} i.i.d. Rayleigh distributed, it can easily be seen that $\mathbf{\Sigma} = \mathbf{I}_M$. Substituting λ by SNR^{r-1} in (37) and solving the limit, we have

$$\lim_{SNR\to\infty} \frac{\log\left[\frac{\left|SNR^{r-1}\mathbf{I}_M\right|^N \int_{\mathbf{0}<\mathbf{X}<\mathbf{I}} e^{-SNR^{r-1} tr\mathbf{X}} |\mathbf{X}|^{N-M} (d\mathbf{X})}{K_{M,N}}\right]}{\log SNR}. \qquad (38)$$

The integral term has no effect in this calculation as

$$\lim_{SNR\to\infty} \frac{\log\left[\frac{1}{K_{M,N}} \int_{0<\mathbf{X}<I} e^{-SNR^{r-1} tr\mathbf{X}} \left|\mathbf{X}\right|^{N-M} (d\mathbf{X})\right]}{\log SNR} = 0, \qquad (39)$$

because the term $e^{-SNR^{r-1} tr\mathbf{X}}$ tends to a constant value as the SNR increases for $0 \le r \le 1$ and for $\mathbf{0} < \mathbf{X} < \mathbf{I}$.

Using (39), (38) leads to

$$\lim_{SNR\to\infty} \frac{\log\left[\left|SNR^{r-1}\mathbf{I}_M\right|^N\right]}{\log SNR} = MN(r-1), \qquad (40)$$

and the resulting diversity-multiplexing curve for the optimal single beamforming scheme is (see also Figure (1))

$$d^B(r) = MN(1-r). \qquad (41)$$

IV. Conclusions

In this paper, we obtain a tight bound for the relation $d(r)$ between the diversity and the spatial multiplexing gain that any coding scheme combined with single beamforming can achieve in a MIMO channel. We then compare it to the optimal relation offered by the channel assuming perfect CSI-T. This relation is obtained looking at the ability of the system to accommodate higher data rates as the SNR increases. The optimal beamforming scheme attains the maximum diversity given by the channel but only for $r = 0$, that is, for a fixed transmission rate R. Conversely, if r is larger than 0 (for instance, when an adaptive constellation is employed such that R linearly increases with $\log SNR$), the diversity attained by the optimal beamforming scheme is far from the one that the channel can provide.

[1] We consider that $\mathbf{H}^\dagger\mathbf{H}$ is $W_M(N;\mathbf{I})$, with $W_M(N;\mathbf{\Sigma})$ the Wishart distribution defined as the distribution of a $M\times M$ random matrix $\mathbf{A} = \mathbf{Z}^\dagger\mathbf{Z}$ where $\mathbf{Z} \in \mathbf{C}^{N\times M}$ is normal distributed $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{\Sigma})$ with $\mathbf{\Sigma} = \mathbf{I}_M$.

Appendix A

In this appendix we want to derive the distribution function of the largest eigenvalue of $\mathbf{S}$, denoted by $\lambda_{\max}$, where $\mathbf{S} = \mathbf{H}^\dagger\mathbf{H}$ with $\mathbf{H}$ is a $N \times M$ matrix with complex variates distributed in the complex multivariate normal distribution with $E[\mathbf{H}] = 0$. Then, $\mathbf{S} = \mathbf{H}^\dagger\mathbf{H}$ is a positive definite Hermitian matrix distributed as follows

$$p(\mathbf{S}) = \frac{1}{K_{M,N}\left|\mathbf{\Sigma}\right|^N} e^{-tr\mathbf{\Sigma}^{-1}\mathbf{S}} \left|\mathbf{S}\right|^{N-M} (d\mathbf{S}),$$

where $|\cdot|$ is the determinant operator, $K_{M,N}$ is a normalising factor and $(d\mathbf{S})$ is equal to

$$(d\mathbf{S}) = \prod_{i=1}^{M} ds_{ii} \prod_{1\le i<j\le M}^{M} d(\mathrm{Re}(s_{ij}))d(\mathrm{Im}(s_{ij})) \qquad (42)$$

We are first interested in the probability that $\mathbf{\Omega}-\mathbf{S}$ is positive definite, that is

$$p(\mathbf{S} < \mathbf{\Omega}) = \frac{1}{K_{M,N}|\mathbf{\Sigma}|^N} \int_{0<\mathbf{S}<\mathbf{\Omega}} e^{-tr\mathbf{\Sigma}^{-1}\mathbf{S}} \left|\mathbf{S}\right|^{N-M} (d\mathbf{S}).$$

Letting $\mathbf{S} = \mathbf{\Omega}\mathbf{X}$ so that $(d\mathbf{S}) = |\mathbf{\Omega}|^M (d\mathbf{X})$ (see [5] p. 57) and $|\mathbf{S}| = |\mathbf{\Omega}|\,|\mathbf{X}|$, this becomes

$$p(\mathbf{S} < \mathbf{\Omega}) = \frac{\left|\mathbf{\Sigma}^{-1}\mathbf{\Omega}\right|^N}{K_{M,N}} \int_{0<\mathbf{X}<\mathbf{I}} e^{-tr\mathbf{\Sigma}^{-1}\mathbf{\Omega}\mathbf{X}} \left|\mathbf{X}\right|^{N-M} (d\mathbf{X}) \qquad (43)$$

In fact, we are interested in probability that $\lambda_{\max} < \lambda$, that is, the distribution function of the largest eigenvalue of $\mathbf{S}$. Note that the inequality $\lambda_{\max} < \lambda$ is equivalent to $\mathbf{S} < \lambda\mathbf{I}_M$ (i.e. $\lambda\mathbf{I}_M - \mathbf{S}$ is positive definite). The result, provided in (37), follows from substituting $\mathbf{\Omega} = \lambda\mathbf{I}_M$ in (43).

References

[1] I. E. Telatar, "Capacity of multi–antenna gaussian channel", European Trans. on Telecommunications, vol. 10, pp. 585-95, Nov/Dec 1999.

[2] G. Foschini, M.Gans, "On limits of wireless communications in a fading environment when using multiple antennas", Wireless Personal Communications, vol. 6, pp. 311-335, 1998.

[3] L. Zheng, D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels", IEEE Trans. on Information Theory, Vol. 49, No. 5, pp. 1073-1096, May 2003.

[4] V. Tarokh, N. Seshadri, A. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction", IEEE Trans. on Information Theory, Vol. 44, pp. 744-765, Mar. 1998.

[5] R. J. Muirhead, "Aspects of Multivariate Statistical Theory". John Wiley and Sons, 1982.

[6] D. P. Palomar, "A Unified Framework for Communications through MIMO Channels", Ph.D. dissertation, Technical University of Catalonia (UPC), Barcelona, Spain, May 2003.