Bit Loading for MIMO with Statistical Channel Information at the Transmitter and MMSE Receivers

Li Hong*, Ana Garcia Armada* and Angel Lozano†
*Universidad Carlos III de Madrid, 28911 Leganes, Madrid, Spain
†Universitat Pompeu Fabra, 08018 Barcelona, Spain

***Abstract*—Bit and power loading schemes for a single-user multiple-input multiple-output channel using a minimum mean-square error receiver are proposed. Such schemes are derived without the benefit of instantaneous channel state information (CSI) at the transmitter, on the basis of only statistical CSI. They exploit the relationships between the average bit error probability at the receiver and the power allocation at the transmitter, in the context of spatially correlated Rayleigh fading. The performance of the proposed schemes is then benchmarked with respect to both non-adaptive transmission and to an adaptive counterpart with a zero-forcing receiver.**

***Index Terms*—bit loading, minimum mean-square error (MMSE), bit error rate (BER), multiple-input multiple-output (MIMO), correlated channel.**

I. Introduction

In multiple-input multiple-output (MIMO) communication, the adaptation of modulation and coding at the transmitter side, based on channel characteristics, is an important feature to reduce the transmission power and/or enhance the bit rates. It is reasonable to assume that the receiver has perfect instantaneous channel state information (CSIR). For the case that such instantaneous CSI is also available at the transmitter, the design of linear minimum mean-square error (MMSE) transmitters and receivers is addressed in [1]. Other contributions in this context include [2] and [3].

Often, the transmitter does not have such accurate channel knowledge, but only knowledge of the CSI distribution (CDIT). Under CDIT, the spatial covariance of the optimal Gaussian signal that achieves capacity is characterized in [4]. From the viewpoint of implementation, the emphasis shifts from the optimal Gaussian signals onto families of discrete constellations, with bit loading [5] the standard way of adapting the transmission to the channel conditions. In [6], the authors propose some bit and power loading schemes for single-user MIMO with CDIT and a linear zero-forcing (ZF) receiver. In this paper, those schemes are modified to incorporate a superior linear MMSE receiver. Altogether, this work enables addressing the following questions:

- What bit rates can be achieved, with discrete constellations, at a certain average bit error probability (BER)?
- What is the optimum power loading?
- How much can be gained, in terms of signal-to-noise ratio (SNR), with respect to a nonoptimized transmission or to an optimized transmission with a ZF receiver?

This work was supported by the Spanish Ministry of Science and Innovation (Ref. CONSOLIDER-INGENIO CSD2008-00010 ”COMONSENS”).

In order to answer these questions, it is essential to establish relationships between the average BER after the MMSE receiver and the powers allocated at the transmitter under the premise of CDIT. For that purpose, our derivations leverage two earlier results for MIMO MMSE receivers over a correlated Rayleigh-fading channel. First, the characterization of the signal-to-interference-plus-noise-ratio (SINR) at the receiver output, furnished in [7], and second, the bit error rate (BER) approximations for M-QAM signals, given in [8].

II. Channel Model

Let n_{T} and n_{R} denote the number of transmit and receive antennas, respectively, with $n_{\mathrm{T}} \leq n_{\mathrm{R}}$. The $n_{\mathrm{T}} \times 1$ transmit vector is denoted by $\mathbf{x} = \mathbf{V}\mathbf{P}^{\mathbf{1/2}}\mathbf{s}$, where $\mathbf{s}$ is a vector of independent unit-variance M-QAM symbols. The unitary matrix $\mathbf{V}$ represents a linear precoder whose columns define the signalling directions. In turn, $\mathbf{P} = \mathrm{diag}\left[p_1\, p_2\, \ldots\, p_{n_{\mathrm{T}}}\right]$ with $p_j \geq 0$ is the power allocated to the jth signaling vector, normalized such that $\mathrm{E}\left[\mathrm{Tr}\left\{\mathbf{x}^H\mathbf{x}\right\}\right] = \mathrm{Tr}\left\{\mathbf{P}\right\} \leq n_{\mathrm{T}}$. Then, the $n_{\mathrm{R}} \times 1$ received vector is

$$\mathbf{y} = \sqrt{\frac{\gamma}{n_{\mathrm{T}}}}\mathbf{H}_c\mathbf{V}\mathbf{P}^{1/2}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{n}$ is a vector with independent zero-mean unit-variance complex Gaussian noise samples, γ is the average signal-to-noise ratio (SNR) per receive antenna, and $\mathbf{H}_c$ is the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ matrix representing the MIMO wireless channel. Allowing for arbitrary correlations among the transmit antennas,

$$\mathbf{H}_c = \mathbf{H}_w\mathbf{R}^{1/2} \tag{2}$$

where the (i, j)th entry of $\mathbf{R}$ equals the correlation between the ith and jth transmit antennas, normalized such that $\mathrm{Tr}\{\mathbf{R}\} = n_{\mathrm{T}}$, while $\mathbf{H}_w$ contains independent zero-mean unit-variance complex Gaussian entries.

Defining an effective channel matrix

$$\mathbf{H} = \sqrt{\frac{\gamma}{n_{\mathrm{T}}}}\mathbf{H}_c\mathbf{V}\mathbf{P}^{1/2}, \tag{3}$$

$$= \sqrt{\frac{\gamma}{n_{\mathrm{T}}}}\mathbf{H}_w\mathbf{R}^{1/2}\mathbf{V}\mathbf{P}^{1/2}. \tag{4}$$

we can simply write $\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}$. The output of a linear MMSE receiver is $\left(\mathbf{I} + \mathbf{H}^{\dagger}\mathbf{H}\right)^{-1}\mathbf{H}^{\dagger}\mathbf{y}$.

For a rather general class of channel models, the precoder $\mathbf{V}$ is capacity-achieving when its columns coincide with the eigenvectors of $\mathbf{R}$ [4]. This same precoder is also shown in [9] to be optimal, from an SINR standpoint, with a linear MMSE receiver. Thus, we shall use this precoder matrix for our derivations.

III. Average Bit Error Probability

It is shown in [7] that, with Rayleigh fading, the SINR distribution corresponding to the jth signaling direction at the output of a linear MMSE receiver is given by the sum of two terms, $S_j + T_j$, where S_j is the SINR at the output of a ZF receiver while T_j is an independent term. Furthermore:

- The term S_j follows the Gamma distribution

$$f_{S_j}(s) = \frac{s^{\alpha-1}e^{-s/\theta}}{\Gamma(\alpha)\theta^{\alpha}} \tag{5}$$

  with

$$\alpha = n_{\mathrm{R}} - n_{\mathrm{T}} + 1 \tag{6}$$

$$\theta_j = \frac{\gamma/n_{\mathrm{T}}}{[\mathbf{P}^{1/2}\mathbf{V}^{\dagger}\mathbf{R}\mathbf{V}\mathbf{P}^{1/2}]^{-1}_{j,j}} \tag{7}$$

  where $[\cdot]^{-1}_{j,j}$ denotes the jth diagonal entry of the inverse of a matrix. Since our choice for the precoder $\mathbf{V}$ diagonalizes $\mathbf{R}$,

$$\theta_j = \frac{p_j\lambda_j\gamma}{n_{\mathrm{T}}} \tag{8}$$

  where λ_j is the jth eigenvalue of $\mathbf{R}$.
- The term T_j is well approximated by a Generalized Gamma distribution. The Generalized Gamma is defined by three parameters, herein labeled α_j, β_j and ξ_j, which are conveniently expressed via the first three moments of T_j. Such moments, denoted respectively by μ_j, σ_j^2 and η_j, can be approximated by numerically solving [7]

$$\mu_j = \frac{1}{n_{\mathrm{T}}-1}\sum_{i=1}^{n_{\mathrm{T}}-1}\frac{1}{1+\gamma\tau_i\left(1-\upsilon+\upsilon\mu\right)} \tag{9}$$

$$\sigma_j^2 = \frac{\dfrac{1}{n_{\mathrm{T}}-1}\displaystyle\sum_{i=1}^{n_{\mathrm{T}}-1}\frac{1+\gamma\tau_i\upsilon\mu}{\left(1+\gamma\tau_i(1-\upsilon+\upsilon\mu)\right)^2}}{1+\dfrac{1}{n_{\mathrm{T}}-1}\displaystyle\sum_{i=1}^{n_{\mathrm{T}}-1}\frac{\gamma\tau_i\upsilon}{\left(1+\gamma\tau_i\left(1-\upsilon+\upsilon\mu\right)\right)^2}} \tag{10}$$

$$\eta_j = \frac{\dfrac{2}{n_{\mathrm{T}}-1}\displaystyle\sum_{i=1}^{n_{\mathrm{T}}-1}\frac{\gamma\tau_i\upsilon\sigma^2}{\left(\gamma\tau_i\left(1-\upsilon+\upsilon\mu\right)+1\right)^2}}{\left(1+\dfrac{1}{n_{\mathrm{T}}-1}\displaystyle\sum_{i=1}^{n_{\mathrm{T}}-1}\frac{\gamma\tau_i\upsilon}{\left(\gamma\tau_i\left(1-\upsilon+\upsilon\mu\right)+1\right)^2}\right)} + \frac{\dfrac{2}{n_{\mathrm{T}}-1}\displaystyle\sum_{i=1}^{n_{\mathrm{T}}-1}\frac{\left(\gamma\tau_i\upsilon\mu-\gamma\tau_i\upsilon\sigma^2+1\right)}{\left(\gamma\tau_i\left(1-\upsilon+\upsilon\mu\right)+1\right)^3}}{\left(1+\dfrac{1}{n_{\mathrm{T}}-1}\displaystyle\sum_{i=1}^{n_{\mathrm{T}}-1}\frac{\gamma\tau_i\upsilon}{\left(\gamma\tau_i\left(1-\upsilon+\upsilon\mu\right)+1\right)^2}\right)} \tag{11}$$

  where $\upsilon = (n_{\mathrm{T}}-1)/n_{\mathrm{R}}$ while τ_i is the ith eigenvalue of the matrix $\mathbf{P}^{1/2}\mathbf{V}^{\dagger}\mathbf{R}\mathbf{V}\mathbf{P}^{1/2}$ with the jth row and the jth column removed. For our choice of $\mathbf{V}$, then, $\{\tau_i\}_{i=1}^{n_{\mathrm{T}}-1} = \{p_i\lambda_i\}_{i\neq j}$, which determines how the moments depend on j.

From μ_j, σ_j^2 and η_j, the parameters of Generalized Gamma distribution for T_j can then be written as

$$\alpha_j = \frac{\left(n_{\mathrm{R}} - n_{\mathrm{T}} + 1 + \left(n_{\mathrm{T}}-1\right)\mu_j\right)^2}{n_{\mathrm{R}} - n_{\mathrm{T}} + 1 + \left(n_{\mathrm{T}}-1\right)\sigma_j^2} \tag{12}$$

$$\beta_j = \theta_j\frac{n_{\mathrm{T}}}{n_{\mathrm{R}}}\frac{n_{\mathrm{R}} - n_{\mathrm{T}} + 1 + \left(n_{\mathrm{T}}-1\right)\sigma_j^2}{n_{\mathrm{R}} - n_{\mathrm{T}} + 1 + \left(n_{\mathrm{T}}-1\right)\mu_j} \tag{13}$$

where θ_j in (13) is as defined in (8). See (14) in the next page for ξ_j.

By integrating well-known instantaneous uncoded BER expressions over the SINR distribution and using its moment generating function, the exact uncoded average BER can be obtained [8], [10]. Handy approximations to such uncoded average BER are given in [8] and reflected in Table I (see the next page), parameterized by the type of constellation and by whether ξ_j is above or below unity.

Furthermore, from [6], [11], for M-QAM constellations in conjunction with a convolutional code of rate $r_c = k_c/n_c$ and minimum distance d_f, the BER at the output of a hard decision decoder can be approximated, at high SNR, by

$$P_c \approx \frac{1}{k_c}A_{d_f}2^{d_f}P_e^{d_f/2} \tag{15}$$

where P_e is the corresponding uncoded error probability while A_{d_f} is the number of paths in the trellis with distance d_f from the all-zero path that merge with this all-zero path for the first time. Using

$$D = \frac{1}{k_c}A_{d_f}2^{d_f/2}\pi^{d_f/4} \tag{16}$$

and the uncoded approximations in [8], the average BER corresponding to the j signaling vector with coded M-QAM $P_{c,j}$ can be approximated, at high SNR, as shown in Table I.

IV. Optimization Problems

Letting M_j be the cardinality of the modulation applied to the jth signaling vector in $\mathbf{V}$, two dual optimization problems can be posed.

A. Bit rate maximization. In this case, the goal is to find the combination $\{M_j\}_{j=1}^{n_{\mathrm{T}}}$ that solves

$$\xi_j = \frac{2\left(1-\frac{n_{\mathrm{T}}-1}{n_{\mathrm{R}}}+\frac{n_{\mathrm{T}}-1}{n_{\mathrm{R}}}\mu_j\right)\left(1-\frac{n_{\mathrm{T}}-1}{n_{\mathrm{R}}}+\frac{n_{\mathrm{T}}-1}{n_{\mathrm{R}}}\eta_j\right)}{\left(1-\frac{n_{\mathrm{T}}-1}{n_{\mathrm{R}}}+\frac{n_{\mathrm{T}}-1}{n_{\mathrm{R}}}\sigma_j^2\right)^2} - 1 \tag{14}$$

Table I
AVERAGE BER HIGH-SNR APPROXIMATIONS FOR CODED MODULATIONS. (FOR UNCODED, SET $D = 1/(2\sqrt{\pi})$ AND $d_f = 1$.)

$\xi_j \geq 1$
$P_{c,j}^{2-\mathsf{PAM}} \approx D \exp\left\{\frac{\alpha_j}{\xi_j - 1}\left(1-\left(1+d_f\beta_j\xi_j\right)^{1-1/\xi_j}\right)\right\}$
$P_{c,j}^{4-\mathsf{QAM}} \approx D \exp\left\{\frac{\alpha_j}{\xi_j - 1}\left(1-\left(1+\frac{d_f\beta_j\xi_j}{2}\right)^{1-1/\xi_j}\right\}\right)$
$P_{c,j}^{M-\mathsf{QAM}} \approx \frac{4D}{\log_2 M} \exp\left\{\frac{\alpha_j}{\xi_j - 1}\left(1-\left(1+\frac{3}{2\,(M-1)}d_f\beta_j\xi_j\right)^{1-1/\xi_j}\right)\right\}$
$\xi_j \leq 1$
$P_{c,j}^{2-\mathsf{PAM}} \approx D \exp\left\{\frac{\alpha_j}{1-\xi_j}\left(\left(1+d_f\beta_j\xi_j\right)^{1-1/\xi_j}-1\right)\right\}$
$P_{c,j}^{4-\mathsf{QAM}} \approx D \exp\left\{\frac{\alpha_j}{1-\xi_j}\left(\left(1+\frac{d_f\beta_j\xi_j}{2}\right)^{1-1/\xi_j}-1\right)\right\}$
$P_{c,j}^{M-\mathsf{QAM}} \approx \frac{4D}{\log_2 M} \exp\left\{\frac{\alpha_j}{1-\xi_j}\left(\left(1+\frac{3d_f\beta_j\xi_j}{2\,(M-1)}\right)^{1-1/\xi_j}-1\right)\right\}$

$$\begin{aligned} \max \quad & R = \sum_{j=1}^{n_{\mathrm{T}}} \log_2 M_j \\ \text{s.t.}\ & P_{c,j} \leq \overline{\mathrm{BER}},\ 1 \leq j \leq n_{\mathrm{T}} \\ & \mathrm{Tr}\{\mathbf{P}\} = n_{\mathrm{T}},\ p_j \geq 0 \end{aligned} \tag{17}$$

where $\overline{\mathrm{BER}}$ is the target average BER.

B. Power minimization. In this case, the goal is to find the combination $\{M_j\}_{j=1}^{n_{\mathrm{T}}}$ that solves

$$\begin{aligned} \min \quad & \mathrm{Tr}\{\mathbf{P}\} \\ \text{s.t.}\ & P_{c,j} \leq \overline{\mathrm{BER}},\ 1 \leq j \leq n_{\mathrm{T}} \\ & \sum_{j=1}^{n_{\mathrm{T}}} \log_2 M_j = R_{\mathsf{tot}},\ p_j \geq 0 \end{aligned} \tag{18}$$

where R_{tot} is the target bit rate.

The proposed solution to the optimization problems is an exhaustive search for the best combinations of bit (i.e., constellation cardinalities and coding rates when applicable) and power loads. For each possible bit allocation, and using the approximations in Table I that relate the BER with the power matrix $\mathbf{P} = \mathrm{diag}\,[p_1\, p_2\, \ldots\, p_{n_{\mathrm{T}}}]$, the appropriate constraints can be enforced. Clearly, for a given set of constraints, the total number of bits with an MMSE receiver cannot be less than with a ZF receiver. Therefore, the starting point for the search is always the solution derived in [6] for ZF receivers, and the search progresses from that point towards higher loads until the constraints set in. With this choice of the starting point, the process is usually quick and, even if truncated before termination, it always returns a performance point that is superior to that achieved with a ZF receiver.

V. EXAMPLES

For the performance evaluations that follow, a transmit covariance matrix corresponding to suburban/rural outdoor environments with small angular spreads at the transmitter is used [12]. This matrix, given by

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho^4 & \cdots & \rho^{(n_{\mathrm{T}}-1)^2} \\ \rho & 1 & \rho & \cdots & \rho^{(n_{\mathrm{T}}-2)^2} \\ \rho^4 & \rho & 1 & \cdots & \rho^{(n_{\mathrm{T}}-3)^2} \\ \vdots & & & \ddots & \vdots \\ \rho^{(n_{\mathrm{T}}-1)^2} & \cdots & \rho^4 & \rho & 1 \end{bmatrix}. \tag{19}$$

is both representative and convenient, as it is a function of a single parameter ρ.

The following scenarios are considered in terms of modulation and coding schemes available at the transmitter:

- 6 uncoded bit rates, with $M_j \in \{0, 2, 4, 8, 16, 32, 64\}$. This is a realistic uncoded scenario.
- 12 bit rates with convolutional coding, where $M_j \in \{0, 2, 4, 8, 16, 32, 64\}$. Here, the transmission is either uncoded or else coded with a rate-$1/2$ convolutional code of generator polynomials $(133, 171)$ and $d_f = 10$.

A. Bit Rate Maximization

Figs. 1 and 2 show the achievable bit rate R for $n_{\mathrm{T}} = 3$, $n_{\mathrm{R}} = 4$, and $\overline{\mathrm{BER}} = 10^{-2}$, with two distinct values for the correlation parameter ρ. Fig. 1 corresponds to 6 uncoded bit rates while Fig. 2 corresponds to 12 bit rates with convolutional coding. The following observations can be made:

- Transmit correlation is seen to be beneficial at low SNR, but detrimental at high SNR. This agrees with the impact of transmit correlation on channel capacity [13].

- Predictably, the performance of MMSE and ZF receivers become progressively similar as the SNR grows.

Fig. 3 illustrates the applicability of the proposed optimization to large-dimensional settings, and specifically to the case $n_{\mathrm{T}} = 8$, $n_{\mathrm{R}} = 10$.

B. Power Minimization

We define as *adaptation power gain* the difference between the power required without and with bit load optimization for given $\overline{\mathrm{BER}}$ and R_{tot}. Hence, the adaptation power gain equals $n_{\mathrm{T}} - \mathrm{Tr}\{\mathbf{P}\}$.

Figs. 4 and 5 portray the adaptation power gain as function of the correlation parameter ρ for $n_{\mathrm{T}} = 3$ and $n_{\mathrm{R}} = 4$, respectively with $\overline{\mathrm{BER}} = 10^{-2}$ and $\overline{\mathrm{BER}} = 10^{-4}$. The results are parameterized by R_{tot}. The adaptation power gain increases for higher ρ since the optimization allows adapting to the fact that the output SINRs become very different in that case. We can see a minor dependence on the values of $\overline{\mathrm{BER}}$.

Fig. 6 depicts two extreme cases with $\overline{\mathrm{BER}} = 10^{-2}$, when the required bit rate is either $R_{\mathsf{tot}} = n_{\mathrm{T}}$ or $R_{\mathsf{tot}} = 6n_{\mathrm{T}}$. These bit rates push the channel into low- and high-SNR behaviors, respectively, revealing that for a given correlation the adaptation power gain at low SNR is larger than at high SNR. At high SNR, the fact that the constellation cardinalities are capped restricts the ability to optimize the bit and power loads.

Finally, we define as *MMSE power gain* the difference between the powers required with ZF and MMSE receivers, respectively. Leveraging the ZF results in [6], Figs. 7 and 8 present MMSE power gains as function of ρ, parameterized by R_{tot}, for two distinct values of $\overline{\mathrm{BER}}$. We can see substantial gains. Here the dependence on ρ is residual because both schemes leverage the difference among output SINRs.

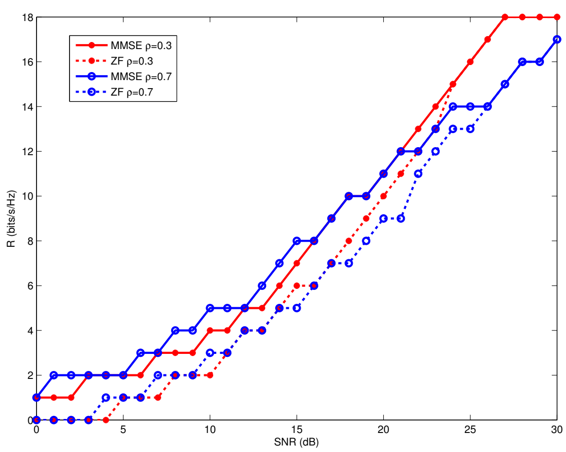

Figure 1. Bit rate vs. SNR; 6 uncoded bit rates with $n_{\mathrm{T}} = 3$, $n_{\mathrm{R}} = 4$ and $\overline{\mathrm{BER}} = 10^{-2}$.

VI. Summary

For either uncoded or convolutionally encoded transmission in single-user MIMO with linear MMSE reception, a method

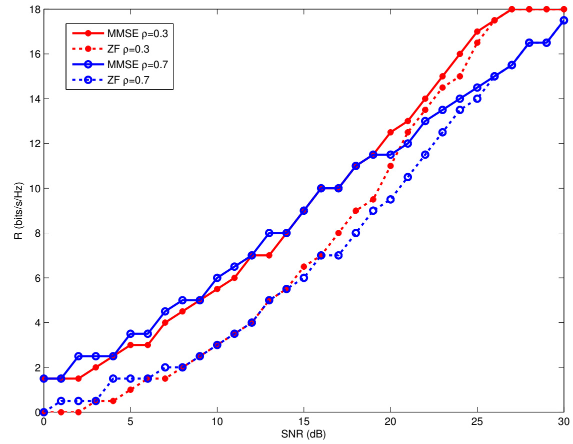

Figure 2. Bit rate vs. SNR; 12 bit rates with convolutional coding, $n_{\mathrm{T}} = 3$, $n_{\mathrm{R}} = 4$ and $\overline{\mathrm{BER}} = 10^{-2}$.

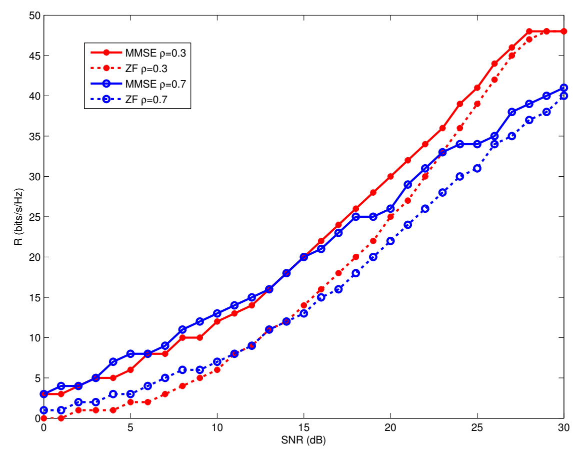

Figure 3. Bit rate vs. SNR; 6 uncoded bit rates with $n_{\mathrm{T}} = 8$, $n_{\mathrm{R}} = 10$ and $\overline{\mathrm{BER}} = 10^{-2}$.

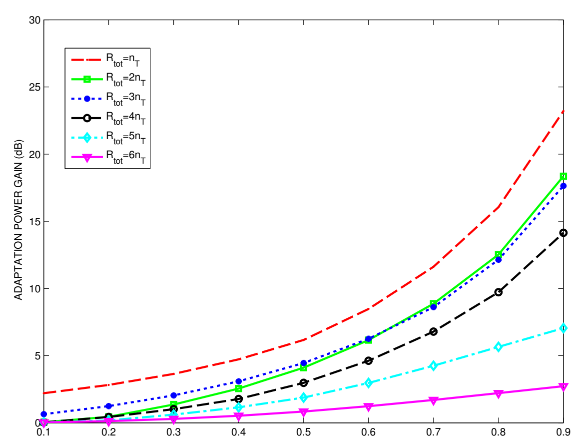

Figure 4. Adaptation power gain vs. ρ; 6 uncoded bit rates with $n_{\mathrm{T}} = 3$, $n_{\mathrm{R}} = 4$ and $\overline{\mathrm{BER}} = 10^{-2}$.

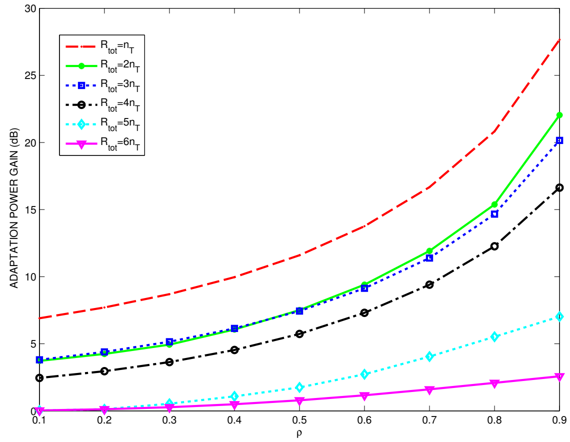

Figure 5. Adaptation power gain vs. ρ; 6 uncoded bit rates with $n_{\mathrm{T}} = 3$, $n_{\mathrm{R}} = 4$ and $\overline{\mathrm{BER}} = 10^{-4}$.

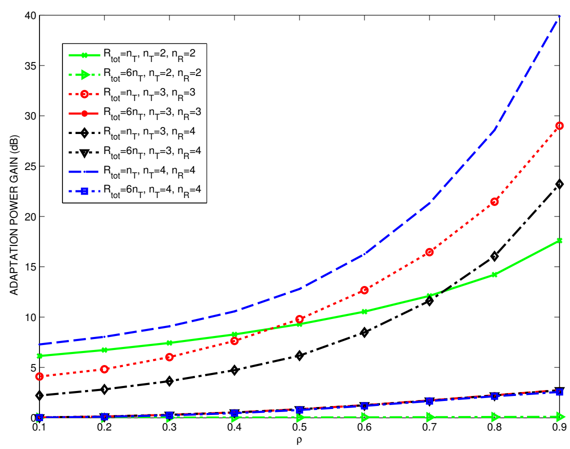

Figure 6. Adaptation power gain vs. ρ; 6 uncoded bit rates with $\overline{\mathrm{BER}} = 10^{-2}$.

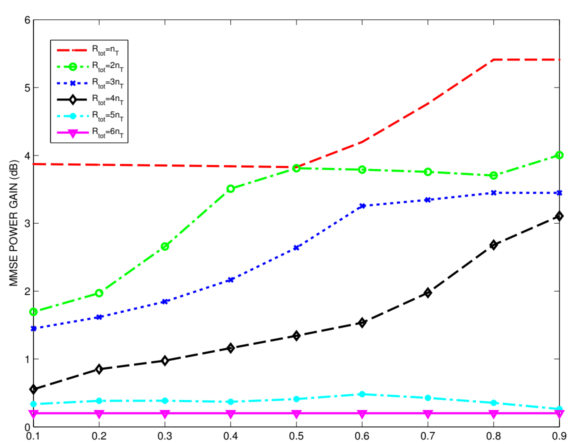

Figure 7. MMSE power gain vs. ρ; 6 uncoded bit rates with $n_{\mathrm{T}} = 3$, $n_{\mathrm{R}} = 4$ and $\overline{\mathrm{BER}} = 10^{-2}$.

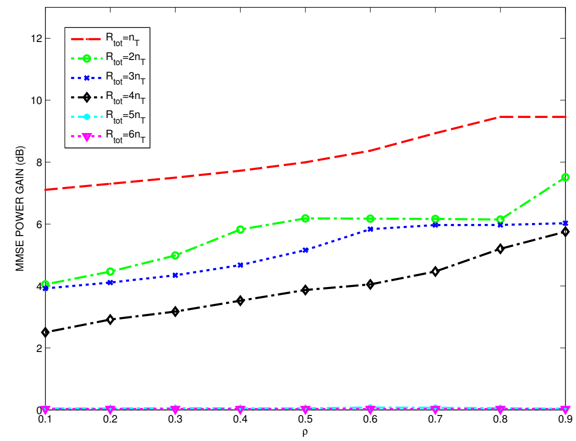

Figure 8. MMSE power gain vs. ρ; 6 uncoded bit rates with $n_{\mathrm{T}} = 3$, $n_{\mathrm{R}} = 4$ and $\overline{\mathrm{BER}} = 10^{-4}$.

has been proposed to solve the two dual optimization problems of bit rate maximization and power minimization. As shown through simulations, higher bit rates and/or lower power consumption can be achieved. The relative improvement relative to ZF reception has also been illustrated.

References

[1] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal Designs for Space-Time Linear Precoders and Decoders," IEEE Trans. on Signal Processing, vol. 50, no. 5, pp. 1051-1064, May 2002.

[2] C.-C. Li, Y.-P. Lin, S.-H. Tsai, and P. P. Vaidyanathan, "Optimization of Transceivers with Bit Allocation to Maximize Bit Rate for MIMO Transmission," IEEE Trans. on Communications, vol. 57, no. 12, pp. 3556- 3560, Dec. 2009.

[3] C. Xiao, Y. R. Zheng, "Transmit Precoding for MIMO Systems with Partial CSI and Discrete-Constellation Inputs," Proc. Int'l Conf. on Communications (ICC'09), Dresden, Germany, pp. 1-5, Jun. 2009.

[4] 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.

[5] J. Campello, "Practical bit loading for DMT," Proc. Int'l Conf. on Communications (ICC'99), Vancouver, Canada, Vol. 2, pp. 801-805, Jun. 1999.

[6] A. García.Armada, L. Hong and A. Lozano, "Bit Loading for MIMO with Statistical Channel Information at the Transmitter and ZF Receiver," Proc. Int'l Conf. on Communications (ICC'09), Dresden, Germany, pp. 1-5, Jun. 2009.

[7] P. Li, D. Paul, R. Narasimhan, "On the Distribution of SINR for the MMSE MIMO Receiver and Performance Analysis," IEEE. Trans. on Information Theory, vol. 52, no. 1, Jan. 2006.

[8] L. Hong, A. García-Armada, "Bit Error Rate Performance of MIMO MMSE Receivers in Correlated Rayleigh Flat-fading Channels," IEEE Trans. on Vehicular Technology, vol. 60, no. 1, pp. 313-317, Jan. 2010.

[9] A. Scaglione, G. Giannakis, and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers. I. Unification and Optimal Designs," IEEE Trans. on Signal Processing, vol. 47, no. 7, pp. 1988-2006, Jul. 1999.

[10] M. K. Simon, M.-S. Alouini, Digital Communication over Fading Channels. Wiley Series in Telecommunications and Signal Processing, 2005.

[11] J. Proakis, Digital Communications, 4th ed. McGraw Hill, 2000.

[12] T.-S. Chu and L. J. Greenstein, "A semiempirical representation of antenna diversity gain at cellular and PCS base stations," IEEE Trans. on Communications, vol. 45-46, pp. 644-646, Jun. 1997.

[13] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of Antenna Correlation the Capacity of Multiantenna Channels," IEEE Trans. on Information Theory, vol. 51, no. 7, pp. 2491-2509, Jul. 2005.