Rate Maximization in Multi-Antenna Broadcast Channels with Linear Preprocessing *

MIHAILO STOJNIC, HARIS VIKALO AND BABAK HASSIBI

Department of Electrical Engineering

California Institute of Technology, Pasadena, CA 91125

e-mail: mihailo,hvikalo,hassibi@systems.caltech.edu

Abstract

The sum rate capacity of the multi-antenna broadcast channel has recently been computed. However, the search for efficient practical schemes that achieve it is still ongoing. In this paper, we focus on schemes with linear preprocessing of the transmitted data. We propose two criteria for the precoding matrix design: one maximizing the sum rate and the other maximizing the minimum rate among all users. The latter problem is shown to be quasiconvex and is solved exactly via a bisection method. In addition to precoding, we employ a signal scaling scheme that minimizes the average bit-error-rate (BER). The signal scaling scheme is posed as a convex optimization problem, and thus can be solved exactly via efficient interior-point methods. In terms of the achievable sum rate, the proposed technique significantly outperforms traditional channel inversion methods, while having comparable (in fact, often superior) BER performance.

Index Terms: Multi-antenna broadcast channel, convex problem, quasiconvex problem, bisection method, interior-point method

1 Introduction

Recently, the achievable limits of performance of multi-antenna broadcast channels have been intensively studied (see, e.g., [1], [2], and the references therein). In [3], [4], non-linear techniques that attempt to approach those limits have been considered. However, these schemes are often computationally prohibitive when the number of transmit antennas is large. In this paper, we limit ourselves to *linear* data preprocessing at the transmitter and, under such a constraint, find

*This work was supported in part by the National Science Foundation under grant no. CCR-0133818, by the Office of Naval Research under grant no. N00014-02-1-0578, and by Caltech's Lee Center for Advanced Networking.

a precoding scheme maximizing the sum-rate of the broadcast channel. We also consider linear precoding schemes that maximize the minimum rate among the users. The latter problem is shown to be quasiconvex and is solved exactly using efficient interior point methods. In addition to the precoding, we minimize the average BER among the users by performing an appropriate signal scaling. The best performance is obtained when the optimal preprocessing and signal scaling are combined.

We assume a standard system model for the broadcast channel with M transmit antennas and M users, described by

$$\mathbf{r} = H\mathbf{s} + \mathbf{w}, \tag{1}$$

where H is an $M \times M$ fading channel matrix whose entries are i.i.d. zero-mean, unit variance, complex Gaussian random variables, and $\mathbf{w}$ is an $M \times 1$ vector whose entries are also i.i.d. zero-mean, variance σ^2 complex Gaussian random variables which represent additive noise at each receiver. Furthermore, $\mathbf{s}$ is an $M \times 1$ vector of signals sent from the transmit antennas, and $\mathbf{r}$ is an $M \times 1$ vector whose components are the received signals at each user. The transmitted vector $\mathbf{s}$ is assumed to be obtained by linear preprocessing of the information vector $\mathbf{u}$, i.e. $\mathbf{s} = kG\mathbf{u}$, where $\mathbf{u} = \begin{bmatrix} u_1, u_2, ..., u_M \end{bmatrix}^T$, u_i is the symbol intended for the i-th user, $1 \leq i \leq M$, and where k is a scaling coefficient which ensures that the power constraint is satisfied.

We organize the paper in the following way; first in Section 2 we propose two possible schemes for designing the preprocessing matrix G. In Section 3, we propose a possible scheme for determining the optimal value of the scaling coefficient k under the constraint of linear preprocessing at the transmitter. In Section 4, we describe how to combine the schemes from Sections 2 and 3. Finally, in Sections 5 and 6 we give simulation results, a brief discussion and several conclusions.

Preliminary version of this paper will appear in [10].

2 Finding optimal preprocessing matrix G

In this section, we find the optimal preprocessing matrix G assuming an average transmit power constraint, $E\|\mathbf{s}\|^2 = 1$. Without loss of generality, we will assume that $E\mathbf{u}\mathbf{u}^* = I$. Then

$E\|G\mathbf{u}\|^2 = E\text{tr}(G^*G\mathbf{u}\mathbf{u}^*) = \text{tr}(G^*G)$ and thus $k = 1/\sqrt{\text{tr}(G^*G)}$. Hence, from (1) we obtain

$$\mathbf{r} = \frac{HG\mathbf{u}}{\sqrt{\text{tr}(G^*G)}} + \mathbf{w}. \tag{2}$$

The matrix G in (2) should be designed to optimize the performance of the overall system in terms of both the rate as well as the bit error rate. Often encountered in the literature is the solution employing a regularized pseudo-inverse of the channel matrix H, i.e., $G = H^*(\beta I + HH^*)^{-1}$, where the coefficient β is typically chosen to maximize the signal-to-interference and noise ratio (SINR) (see, e.g., [4]). However, optimizing for SINR does not necessarily imply that the total sum rate will be maximized. This justifies the search for a better choice for the matrix G.

We consider two optimization criteria for the design of the preprocessing matrix G. First, we maximize the total sum rate over the space of all $M \times M$ complex matrices G. As we shall see, this optimization results in a strategy where at each channel use, a subset of users is chosen and data transmitted only to those users. Second, we consider the problem of optimal preprocessing that maximizes the minimum rate among all the users. Extensive simulations imply that the best BER performance of the system is achieved when the two strategies are combined, i.e., when a subset of users is selected and then the minimum rate among the users in that subset is maximized.

2.1 Maximizing the sum rate over G

We assume that each user treats the interference as noise. Therefore the sum rate of the broadcast channel (2) is given by $R = \sum_{m=1}^{M} \log\left(1 + \frac{\left|\sum_p H_{mp}G_{pm}\right|^2}{\sigma^2\text{tr}(G^*G) + \sum_{n\neq m}\left|\sum_p H_{mp}G_{pn}\right|^2}\right)$. The optimal choice for the matrix G is the solution to the optimization problem

$$\max_G R \tag{3}$$

A closed-form analytic solution to (3) does not appear easy to find and thus we solve it iteratively. Before proceeding any further, we will find it useful to define $\text{num}_m = \left|\sum_{p=1}^{M} H_{mp}G_{pm}\right|^2$, and $\text{den}_m = \sigma^2\text{tr}(G^*G) + \sum_{n=1,n\neq m}^{M}\left|\sum_{p=1}^{M} H_{mp}G_{pn}\right|^2$. The following lemma gives a necessary condition for the optimal G.

Lemma 1. *Denote* $\Delta = diag(\frac{(HG)_{11}}{den_1}, .., \frac{(HG)_{ll}}{den_l}, ..., \frac{(HG)_{MM}}{den_M})$ *and* $D = diag(\frac{num_1}{den_1(den_1+num_1)}, ..., \frac{num_l}{den_l(den_l+num_l)}, ..., \frac{num_M}{den_M(den_M+num_M)})$*, Then any* G *which is solution of (3) is of the form* $G = ((\sigma^2 trD)I + H^*DH)^{-1}H^*\Delta$.

Proof. It is sufficient to show that $\frac{\partial R}{\partial G_{kl}} = 0 \Rightarrow G = ((\sigma^2 \mathrm{tr} D)I + H^*DH)^{-1}H^*\Delta$. It is straightforward to show that $\frac{\partial R}{\partial G_{kl}} = \frac{H_{lk}(HG)^*_{ll}}{\mathrm{den}_l} - \sum_{m=1}^{M} \frac{\mathrm{num}_m H_{mk}(HG)^*_{ml}}{\mathrm{den}_m(\mathrm{num}_m+\mathrm{den}_m)} - \sum_{m=1}^{M} \frac{\sigma^2 G^*_{kl}\mathrm{num}_m}{\mathrm{den}_m(\mathrm{den}_m+\mathrm{num}_m)}$. Setting each of these derivatives to zero, we obtain $H^*\Delta - H^*DHG - (\sigma^2\mathrm{tr}D)G = 0$, or equivalently $G = ((\sigma^2\mathrm{tr}D)I + H^*DH)^{-1}H^*\Delta$. This implies that $\frac{\partial R}{\partial G_{kl}} = 0 \Rightarrow G = ((\sigma^2\mathrm{tr}D)I + H^*DH)^{-1}H^*\Delta$, which concludes the proof. □

Using Lemma 1, we state the following iterative algorithm for solving (3).

$D_0 = I, \Delta_0 = I, i = 0$

Repeat

1. $G_i = ((\sigma^2 trD_i)I + H^*D_iH)^{-1}H^*\Delta_i$
2. $num_m = \left|\sum_{p=1}^{M}(HG_i)_{mm}\right|^2$, $den_m = \sigma^2 tr(G_i^*G_i) + \sum_{n=1,n\neq m}^{M}\left|\sum_{p=1}^{M}(HG_i)_{mn}\right|^2$
3. $D_{i+1} = diag(\frac{num_1}{den_1(den_1+num_1)}, ..., \frac{num_l}{den_l(den_l+num_l)}, ..., \frac{num_M}{den_M(den_M+num_M)})$,
4. $\Delta_{i+1} = diag(\frac{(HG_i)_{11}}{den_1}, .., \frac{(HG_i)_{ll}}{den_l}, ..., \frac{(HG_i)_{MM}}{den_M})$, *i=i+1*

end

We refer to using the matrix G obtained from the previous iterative procedure as Method 2.1. Since $H^*((\sigma^2\mathrm{tr}D)I + HH^*)^{-1} = ((\sigma^2\mathrm{tr}D)I + H^*H)^{-1}H^*$, the initial value G_0 coincides with the one obtained by the regularized pseudo-inverse (see, e.g., [4]). Simulation results presented in following sections imply that such a choice of initial value leads to an iterative process that converges after a fairly small number of iterations (roughly 15 on average).

2.2 Maximizing the minimum rate over G

Instead of maximizing the sum rate, one may demand that the worst (active) user gets as large rate as possible. This criterion leads to the following optimization problem

$$\max_G \min_i \log\left(1 + \frac{|(HG)_{ii}|^2}{\sigma^2\mathrm{tr}(G^*G) + \sum_{j,j\neq i}|(HG)_{ij}|^2}\right). \quad (4)$$

The previous problem (or problems similar to it) have been studied and various algorithms for solving it have been suggested throughout the literature (see, e.g. [5, 6, 7, 8]). Here we suggest another way of solving it based on interior point methods. Define $B = HG$. Then (4) can be written as

$$\max_B \min_i \frac{|B_{ii}|^2}{\sigma^2 \mathrm{tr}(B^* H^{-*} H^{-1} B) + \sum_{j,j\neq i} |B_{ij}|^2}. \tag{5}$$

Without loss of generality, we can assume that the optimal B_{ii} are real and positive. Let $\mathbf{vec}(B)$ denote a vector comprised of columns of matrix B. Then we can write $\sigma^2\mathrm{tr}(B^* H^{-*} H^{-1} B) = \sigma^2 \mathbf{vec}(B)^* (I \otimes H^{-*}H^{-1}) \mathbf{vec}(B)$. Denoting $F = I \otimes H^{-*}H^{-1}$, $\mathbf{x} = \begin{bmatrix} \Re(\mathbf{vec}(B)) \\ \Im(\mathbf{vec}(B)) \end{bmatrix}$ and $T = \begin{bmatrix} \Re(F) & -\Im(F) \\ \Im(F) & \Re(F) \end{bmatrix}$, we have $\sigma^2\mathrm{tr}(B^* H^{-*} H^{-1} B) = \sigma^2 \mathbf{x}^* T \mathbf{x}$. Define $2M^2 \times 2M^2$ matrix $K^{(ij)}$ with $K^{(ij)}_{(j-1)M+i,(j-1)M+i} = K^{(ij)}_{M^2+(j-1)M+i,M^2+(j-1)M+i} = 1$ and zeros otherwise. Combining all of the above, (5) can be rewritten as

$$\begin{aligned} \min_{\mathbf{x}} \max_i \quad & \frac{\mathbf{x}^* W_i \mathbf{x}}{x^2_{(i-1)M+i}} \\ \text{subject to} \quad & x_{(i-1)M+i} > 0, \qquad 1 \leq i \leq M \\ & x_{M^2+(i-1)M+i} = 0, \qquad 1 \leq i \leq M, \end{aligned} \tag{6}$$

where $W_i = \sigma^2 T + \sum_{j=1,j\neq i}^{M} K^{(ij)}$. Note that W_i is positive semidefinite because matrices T and $K^{(ij)}$ are positive semidefinite. To solve (6), we first prove the following lemma.

Lemma 2. *The optimization problem (6) is quasiconvex.*

Proof. We first need to prove that function $f_i(\mathbf{x}) = \frac{\mathbf{x}^* W_i \mathbf{x}}{x^2_{(i-1)M+i}}$ is quasiconvex. We can write $f_i(\mathbf{x}) = \frac{g_i(\mathbf{x})}{x_{(i-1)M+i}}$, where $g_i(\mathbf{x}) = \frac{\mathbf{x}^* W_i \mathbf{x}}{x_{(i-1)M+i}}$. Let us show that the function $g_i(\mathbf{x})$ is convex for $x_{(i-1)M+i} > 0$. To do so, we need to show that $g_i(\theta\mathbf{x} + \gamma\mathbf{y}) \leq \theta g_i(\mathbf{x}) + \gamma g_i(\mathbf{y})$, where $\theta + \gamma = 1, 0 \leq \theta, \gamma \leq 1$. This is equivalent to showing that $\frac{y_{(i-1)M+i}}{x_{(i-1)M+i}} \mathbf{x}^* W_i \mathbf{x} - 2\mathbf{x}^* W_i \mathbf{y} + \frac{x_{(i-1)M+i}}{y_{(i-1)M+i}} \mathbf{y}^* W_i \mathbf{y} \geq 0$. Since W_i is symmetric and positive semidefinite, it can be written as $W_i = R_i^* R_i$. From Cauchy-Schwartz inequality it follows that $\mathbf{x}^* W_i \mathbf{y} = \mathbf{x}^* R_i^* R_i \mathbf{y} \leq ||R_i \mathbf{x}||_2 ||R_i \mathbf{y}||_2 = \sqrt{\mathbf{x}^* W_i \mathbf{x} \mathbf{y}^* W_i \mathbf{y}}$, from which it follows that

$$\frac{y_{(i-1)M+i}}{x_{(i-1)M+i}}\mathbf{x}^*W_i\mathbf{x}-2\mathbf{x}^*W_i\mathbf{y}+\frac{x_{(i-1)M+i}}{y_{(i-1)M+i}}\mathbf{y}^*W_i\mathbf{y} \geq \left(\sqrt{\frac{y_{(i-1)M+i}}{x_{(i-1)M+i}}\mathbf{x}^*W_i\mathbf{x}} - \sqrt{\frac{x_{(i-1)M+i}}{y_{(i-1)M+i}}\mathbf{y}^*W_i\mathbf{y}}\right)^2 \geq 0$$

therefore, function $g_i(\mathbf{x})$ is convex for $x_{(i-1)M+i} > 0$. Since the ratio of convex and linear function is qausiconvex and since pointwise maximum of quasiconvex functions is quasiconvex (see, e.g., [9]) we conclude that the objective function in (6) is quasiconvex. □

Remark: When preparing the final version of this paper, we became aware of related work [8], where the authors deal with a similar problem. There they present another proof of quasiconvexity of (6), using different approach.

We use the bisection method combined with the interior-point method to solve (6). Once we find the optimal $\mathbf{x}$ in (6), we determine B such that $\mathbf{x} = \begin{bmatrix} \Re(\mathbf{vec}(B)) \\ \Im(\mathbf{vec}(B)) \end{bmatrix}$. Then we calculate G as $G = H^{-1}B$. We refer to using the matrix G found by the aforementioned procedure as Method 2.2.

The technique described in Subsection 2.1 maximizes the sum rate of the multi-antenna broadcast system under the linear data processing constraint. The individual rates resulting from the maximization (3), however, may differ significantly. This disparity is inherent to the optimization (3) since (3) essentially denotes the maximization of $\|\mathbf{v}\|_1$ (i.e., norm-1 of the vector $\mathbf{v}$). It is well known that in the process of maximizing the norm-1 of a vector, a few components of the vector are suppressed while the remaining ones are boosted up. Thus in Subsection 2.1 the sum rate is maximized at the expense of the weakest few users which are ignored. [Note: Transmitting data over many channel uses provides fairness.] The symbols intended for the remaining strong users may be modulated with higher modulation schemes, thus overcompensating for the sum rate seemingly lost by transmitting only to a subset of users.

On the other hand, as a result of the disparity among the individual rates (and hence among the SINRs and BERs of individual users), the average BER of the system may suffer. To compensate for the loss in average BER, we employ Method 2.2 on the subset of strong users selected for transmission by Method 2.1. We formalize this combination of Method 2.1 and Method 2.2 in the following way

1. *obtain G using Method 2.1*

2. *denote the set of indices which correspond to zero-columns of G by $\mathcal{I}_0$*

3. *denote a submatrix of H comprised of rows $1 \leq i \leq N$, $i \notin \mathcal{I}_0$ by H_{sub}*

4. *Apply Method 2.2 on H_{sub} to obtain B; set $G = H^*_{sub}(H_{sub}H^*_{sub})^{-1}B$.*

As it turns out, maximizing the minimum individual rate among the selected strong users results in fairly equal (and high) SINRs. We refer to the previous combination of Method 2.1 and Method 2.2 as Method 2.

3 Finding the optimal scaling coefficient k

Recall our basic model (1). Let us assume that the preprocessing matrix G is obtained by simple inversion of the channel matrix H, i.e., $G = H^{-1}$. In this section, we propose a way of scaling the magnitudes of the information signal $\mathbf{u}$ so as to minimize the average BER. [Note that in Section 4 we will show how to employ this signal scaling scheme to the optimal G obtained in Section 2.]

To minimize the average BER, one needs to maximize the minimum SINR at receivers. To this end, in [4] authors suggest perturbation of information signals by appropriately translating original M-QAM signal constellation in complex space. In this section we suggest a similar idea but focus on perturbations (in fact, radial scaling) of M-PSK constellation. An advantage of constraining ourselves to PSK constellations is in the simplicity of decoding. Since the signal points are perturbed only radially, rather than vertically or horizontally as in QAM, the angular information has not changed. Therefore, no side information about the signaling scheme (i.e., the nature of the perturbation) is needed at the receiver. In other words, each user's decoder makes simple angular decisions. The decoder is no longer necessarily ML but it is efficient and practical since it requires no additional information from the transmitter. [Our simulation results indicate that the performance of this sub-optimal ML decoder is almost identical to the optimal one.]

By fixing $G = H^{-1}$ and representing $\mathbf{u}$ via its phases and magnitudes, we can rewrite (1) as $\mathbf{r} = kHH^{-1}\Phi\mathbf{u}_m + \mathbf{w}$, where $\mathbf{u} = \Phi\mathbf{u}_m$ and Φ is the diagonal matrix of phases of $\mathbf{u}$, and where $\mathbf{u}_m$ is the vector of magnitudes of $\mathbf{u}$. Note that due to the use of a PSK modulation scheme, the information to be transmitted is contained in Φ. We are concerned with designing optimal magnitudes of the signals, i.e., designing the $\mathbf{u}_m$. The relevant power constraint now becomes the one on instantaneous, rather than average, transmission power. This means that the corresponding form to (2) can be written as

$$\mathbf{r} = \frac{HH^{-1}\Phi\mathbf{u}_m}{\sqrt{\text{tr}(\mathbf{u}_m^*\Phi^*H^{-*}H^{-1}\Phi\mathbf{u}_m)}} + \mathbf{w}. \tag{7}$$

Now we want to optimize scaling coefficient while keeping magnitudes of $\mathbf{u}$ greater or equal to 1. This will result in magnitudes of the components of the received vector $\mathbf{r}$ that are at least as large as if there were no signal scaling at all. This requires solving the following optimization problem

$$\begin{aligned} \min \quad & \mathbf{u}_m^*\Phi^*H^{-*}H^{-1}\Phi\mathbf{u}_m \\ \text{subject to} \quad & u_{m_i} \geq 1, \qquad 1 \leq i \leq M \end{aligned} \tag{8}$$

This problem is convex and can easily be solved exactly by a host of numerical methods (see, e.g., [9] and the references therein). More importantly, we can show that the solution of this problem is equal to the solution to

$$\begin{aligned} \max_{u_{m_1},u_{m_2},\ldots,u_{m_M}} \min_i \quad & \frac{u_{m_i}^2}{\mathbf{u}_m^*\Phi^*H^{-*}H^{-1}\Phi\mathbf{u}_m} \\ \text{subject to} \quad & u_{m_i} \geq 1, \qquad 1 \leq i \leq M, \end{aligned} \tag{9}$$

which is the problem of maximizing the minimum SINR in system (7). Denoting by $\widehat{\mathbf{u}_m}$ the solution to (9), we see that the transmitted signal should have the form of $\mathbf{s} = \frac{H^{-1}\Phi\widehat{\mathbf{u}_m}}{\sqrt{\widehat{\mathbf{u}_m}^*\Phi^*H^{-*}H^{-1}\Phi\widehat{\mathbf{u}_m}}}$. We refer to this signal scaling policy as Method 3. As said earlier, although the magnitudes of optimal $\mathbf{u}$ will generally be different from 1, the receivers will still be able to decode the received signals by considering their angle, since $\mathbf{s}$ has the same phase matrix Φ as $\mathbf{u}$.

4 Combined method

In Section 3, we employed the signal scaling scheme to optimize the BER in a system that uses $G = H^{-1}$ for data preprocessing. In this section, we combine the signal scaling with the optimal preprocessing matrices G found in Section 2. This is done in stages. In particular, assume that Method 2.1 is used to find G which maximizes the sum rate of the channel. Then to minimize the average BER of the users, we employ signal scaling for such G. Instead of solving (8) (which assumed $G = H^{-1}$), we now need to solve optimization

$$\begin{aligned} \min \quad & \mathbf{u}_m^* \Phi^* \widehat{G}^* \widehat{G} \Phi \mathbf{u}_m \\ \text{subject to} \quad & u_{m_i} \geq 1, \qquad 1 \leq i \leq M \end{aligned} \tag{10}$$

where $\widehat{G}$ is G found by Method 2.1. The above problem is convex and thus can be solved exactly via efficient convex optimization techniques. If we denote solution of (10) by $\widehat{\mathbf{u}_m}$, the optimal transmitted signal $\mathbf{s}$ is given by $\mathbf{s} = \frac{\widehat{G}\Phi\widehat{\mathbf{u}_m}}{\sqrt{\widehat{\mathbf{u}_m}^* \Phi^* \widehat{G}^* \widehat{G} \Phi \widehat{\mathbf{u}_m}}}$. We refer to the above algorithm as Method 4.

5 Simulation results

In this section we briefly discuss simulation results of the suggested methods for linear preprocessing. Figures 1 and 2 show that Method 2.1 performs at least as good as the regularized pseudo inverse in terms of BER while, due to the use of a higher modulation scheme, provides significantly higher sum rate. Figure 1 also shows that Method 2, due to the additional minmax optimization of SINRs, performs even better than Method 2.1 in terms of BER. Figure 3 shows that the simple scaling strategy gives a better BER performance than the pseudo inverse. Finally, Figures 4 and 5 show that both, Method 2 and Method 4, outperform the pseudo inverse in terms of both, the BER and the sum rate. All plots were done using uncoded sequences of information bits at the transmitter modulated with symbols from standard PSK constellations as denoted below the figures.

6 Conclusion

In this paper, we have proposed two criteria for the design of the precoding matrix in a multi-antenna broadcast system. First, we maximized the sum rate, and then we showed how to maximize the minimum rate among all users. The latter problem is shown to be quasiconvex and solved exactly. The precoding techniques are constrained to linear preprocessing at the transmitter. In addition to precoding, we have employed a signal scaling scheme that minimizes the average BER of the users. The signal scaling scheme is posed as a convex optimization problem, and solved exactly via interior-point methods. Finally, we have combined the precoding with signal scaling. The combined scheme can be efficiently applied in practice. In terms of the achievable sum rate, the proposed technique significantly outperforms traditional channel inversion methods, while having comparable (in fact, often superior) BER performance.

References

[1] G.Caire and S.Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Trans. on Inf. Theory*, July 2003.

[2] S.Wishwanath, N.Jindal, and A.Goldsmith, "Achievable rates, and sum-rate capacity of gaussian mimo broadcast channels," *IEEE Trans. on Inf. Theory*, Oct. 2003.

[3] R.Zamir, S.Shamai, and U.Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. on Inf. Theory*, June 2002.

[4] C.Peel, B.Hochwald, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multi-user," *submitted to IEEE Trans. of Communications*, June 2003.

[5] J.-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter receiver diversity for efficient space division multiaccess," *IEEE Trans. on Wireless Communications*, Jan. 2002.

[6] J. Bertrand and P. Forster, "Optimal weights computation of an emitting antenna array-the obele algorithm," *IEEE Trans. on Signal Processing*, July 2003.

[7] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," *IEEE Vehicular Technology Conference*, May 1999.

[8] A.Wiesel, Y.C.Eldar, and S.Shamai, "Multiusers precoders for fixed receivers," *International Zurich Seminar on Communications*, 2004.

[9] S.Boyd and L.Vandenberghe, *Convex Optimization*, Cambridge University Press, 2003.

[10] M. Stojnic, H. Vikalo, and B. Hassibi, "Rate maximization in multi-antenna broadcast channels with linear preprocessing," *Proc. IEEE Globecom*, Nov. 2004.

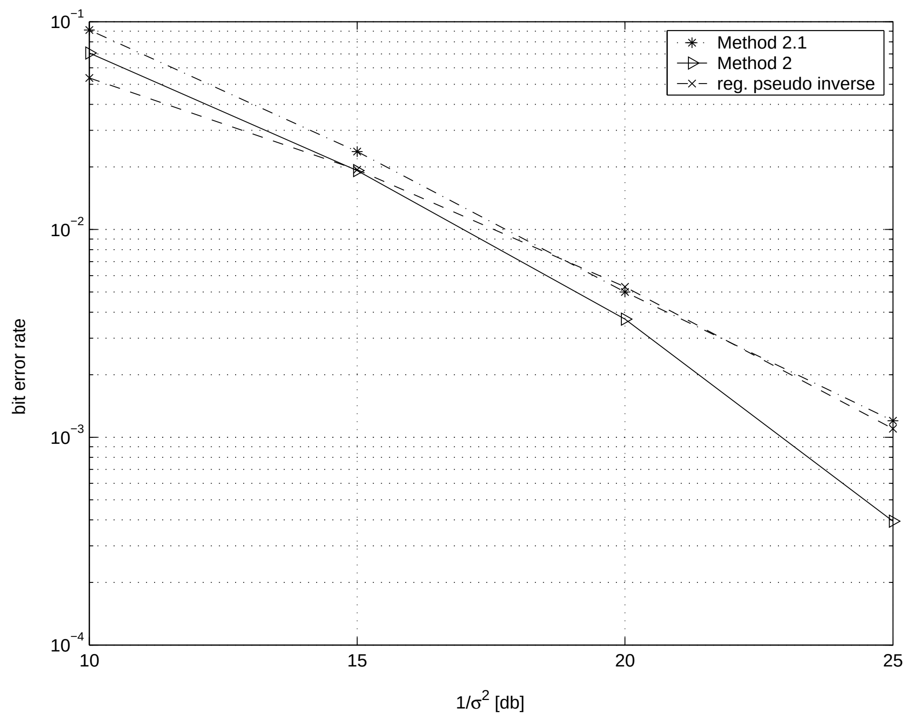

Figure 1: Comparison of BER, M=6 antennas/users, 8PSK-Method 2, 8PSK-Method 2.1, 4PSK-regularized pseudo inverse

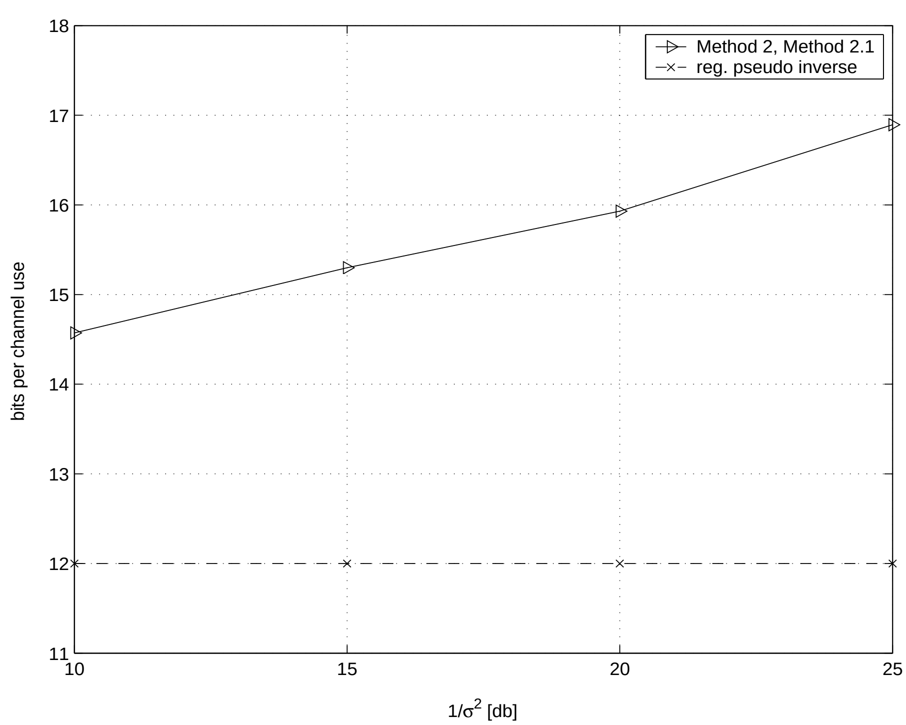

Figure 2: Comparison of rates, M=6 antennas/users, 8PSK-Method 2, 8PSK-Method 2.1, 4PSK-regularized pseudo inverse

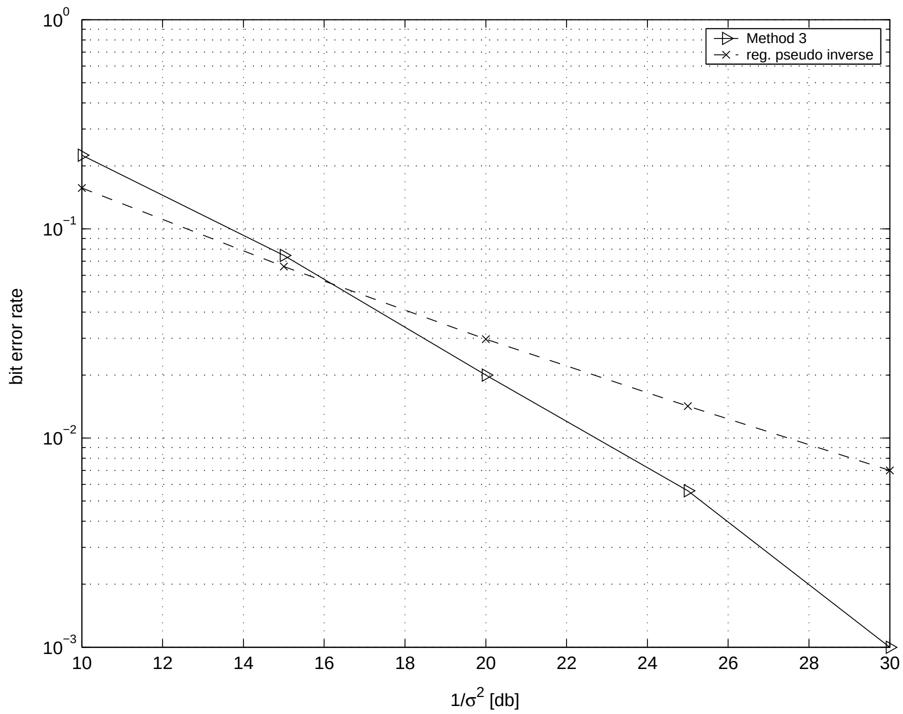

Figure 3: Comparison of BER, M=20 antennas/users,8PSK-Method 3, 8PSK-regularized pseudo inverse

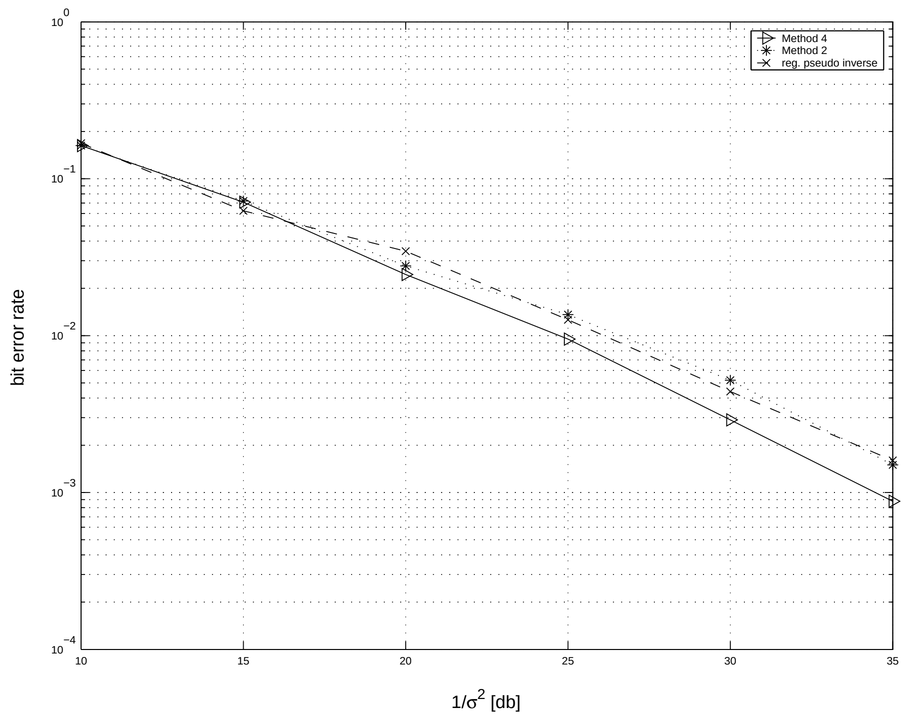

Figure 4: Comparison of BER, M=6 antennas/users, 16PSK-Method 4, 16PSK-Method 2,8PSK-regularized pseudo inverse

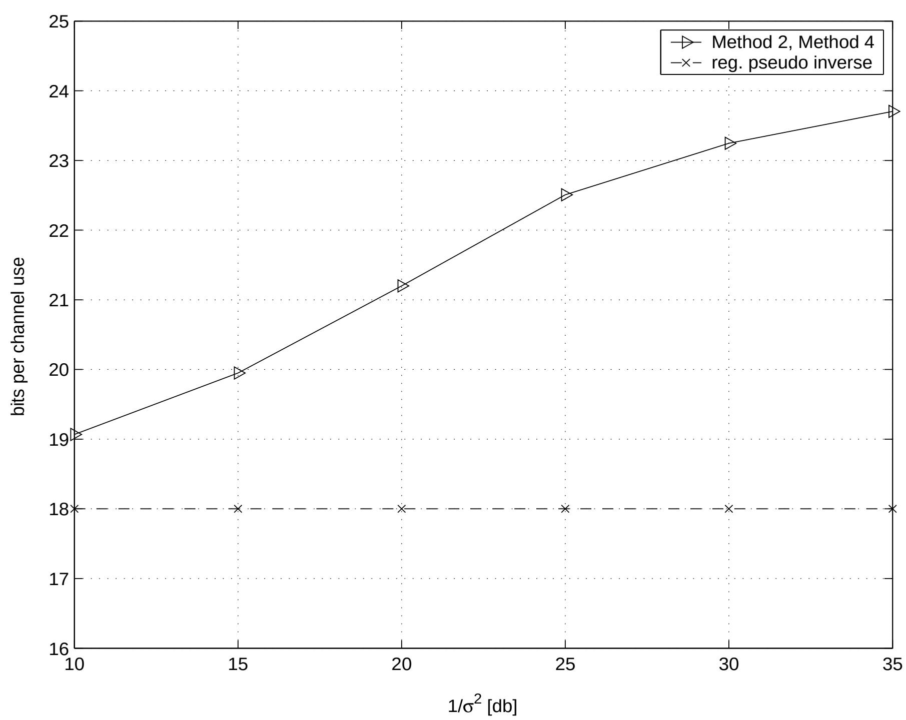

Figure 5: Comparison of rates,M=6 antennas/users, 16PSK-Method 4, 16PSK-Method 2,8PSK-regularized pseudo inverse