Per-Antenna Power Constrained MIMO Transceivers Optimized for BER

Ching-Chih Weng and P. P. Vaidyanathan
Dept. of Electrical Engineering, MC 136-93
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: cweng@caltech.edu, ppvnath@systems.caltech.edu

***Abstract*— This paper considers the linear transceiver optimization problem for multi-carrier multiple-input multiple-output (MIMO) channels with per-antenna power constraints. Because in practical implementations each antenna is limited individually by its equipped power amplifier, this paper adopts the more realistic per-antenna power constraints, in contrast to the conventional sum-power constraint on the transmitter antennas. Assuming perfect channel knowledge both at the transmitter and the receiver, the optimization problem can be transformed into a semi-definite program (SDP), which can be solved by convex optimization tools. Furthermore, several objective functions of the MIMO system, including average bit error rate, can also be optimized by the introduction of the majorization theory.** [1]

***Index Terms* — Per-antenna Constraints, BER Optimization, Schur Convexity, Semi-definite Programming, MIMO Transceivers.**

I. Introduction

In this paper we consider the optimization of multi-carrier multiple-input multiple-output (MIMO) communication systems with per-antenna power constraints. In the MIMO system, the transmitter has M antennas sending independent information to the receiver equipped with J antennas. The signal vector consisting of M sub-streams is assumed to be linearly transformed by the channel matrix $\mathbf{H}$, and corrupted by the additive Gaussian noise.

Based on the assumption of having perfect channel state information, the transmitter can use appropriate precoding, and jointly with the equalization scheme at the receiver side, better performance can be achieved. It is known that the optimal equalization technique is the maximum likelihood receiver. However, due to the heavy computation load, usually the linear precoding and equalization scheme, or other suboptimal techniques, such as linear precoder with decision feedback equalizers, are utilized. Under those schemes, several authors considered the optimization of the MIMO communication systems. In [12], the authors considered the linear transceiver optimization problem under the total power constraint. By introducing the majorization theory, several objective functions, which can be categorized as Schur-convex or Schur-concave [10] functions of the mean squared errors, were optimized. Recently, the optimal solution for decision feedback equalizer along with precoder, under the total power constraint, was also found [13], [6].

Instead of total power constraint, in this paper we consider the more realistic per-antenna power constraints on the transmitter [7], [11], since in practice each antenna is limited individually by its equipped power amplifier. The optimization process consists of two steps. We first formulate the MMSE (Minimum Mean Square Error) linear transceiver design with per-antenna power constraints to be a semi-definite program (SDP), which can be efficiently solved by convex optimization tools. Then, among the family of MMSE transceivers, we develop a method to find the one that minimizes the average bit error rate as well as many other Schur-convex objective functions. The proof of the optimality is also given in the paper.

This paper is structured as follows. In Section II, after a brief introduction, the optimal receiver matrix as well as the resulting error covariance matrix will be given. The formulation of MMSE transceiver optimization problem in the form of a SDP will also be described. In Section III, the optimization of average BER problem will be discussed. Section IV presents the numerical simulation results related to the topics discussed in the paper. The final conclusions of the paper are summarized in section V.

II. MMSE Transceiver Design with Per-Antenna Power Constraints

In this section we describe the MMSE transceiver optimization with per-antenna power constraints. This serves as the first step toward minimizing the average BER. We consider a communication system with M transmit and J receive antennas. This gives rise to a MIMO channel that can be represented by a channel matrix. Consider the system model in Fig. 1, where $\mathbf{s}$ is the $M \times 1$ transmitted symbol vector, $\mathbf{H}$ is the channel matrix with dimension $J \times M$, which is assumed to be fixed during the transmission of a block, and known at both sides of the communication links. We assume zero-mean unit-energy uncorrelated (white) symbols, i.e., $E[\mathbf{s}\mathbf{s}^\dagger] = \mathbf{I}_M$. The received signal prior to equalization is

$$\mathbf{y} = \mathbf{HFs} + \mathbf{n}. \quad (1)$$

where $\mathbf{F}$ is the precoding matrix, and $\mathbf{n}$ is the additive Gaussian noise with covariance matrix $\mathbf{R}_n$, i.e., $\mathbf{n} \sim CN(\mathbf{0}, \mathbf{R}_n)$.

[1] This work is supported in parts by the ONR grant N00014-06-1-0011, and the TMS scholarship 94-2-A-018 of the National Science Council of Republic of China, Taiwan.

The per-antenna power constraints can be formulated as

$$(E[\mathbf{F}\mathbf{s}\mathbf{s}^\dagger\mathbf{F}^\dagger])_{ii} = (\mathbf{F}\mathbf{F}^\dagger)_{ii} \le P_i, \ \forall i = 1, 2, \cdots, M \tag{2}$$

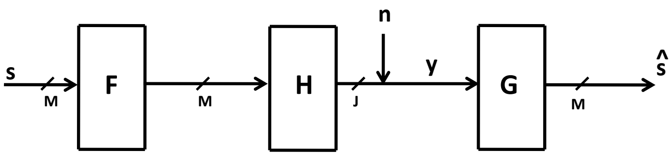

Fig. 1. Communication model.

Usually the performance of the communication system depends on the mean square error (MSE). To design the system, we first derive the optimum receiver matrix $\mathbf{G}$ which minimizes the MSE, when $\mathbf{H}$, $\mathbf{F}$, and $\mathbf{R}_n$ are given. The error vector in Fig. 1 of the transmitted symbols before the hard decision device is

$$\mathbf{e} = \mathbf{G}\mathbf{y} - \mathbf{s} = (\mathbf{G}\mathbf{H}\mathbf{F} - \mathbf{I})\mathbf{s} + \mathbf{G}\mathbf{n}. \tag{3}$$

Under the assumption that the transmitted symbols have no correlation with the additive noise, the covariance matrix of the error vector is

$$E[\mathbf{e}\mathbf{e}^\dagger] = (\mathbf{G}\mathbf{H}\mathbf{F} - \mathbf{I})(\mathbf{G}\mathbf{H}\mathbf{F} - \mathbf{I})^\dagger + \mathbf{G}\mathbf{R}_n\mathbf{G}^\dagger \tag{4}$$

It is well known that if $\mathbf{F}$ is given [12], [9], the optimum receiver matrix would be the linear minimum MSE (LMMSE) receiver or so called Wiener filter:

$$\mathbf{G}_{opt} = \mathbf{F}^\dagger\mathbf{H}^\dagger(\mathbf{H}\mathbf{F}\mathbf{F}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}, \tag{5}$$

and the resulting error covariance matrix (MSE matrix) can be written as

$$\mathbf{E} \equiv E[\mathbf{e}\mathbf{e}^\dagger] = \mathbf{I} - \mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{W}\mathbf{H}\mathbf{F} \tag{6}$$

where

$$\mathbf{W} = (\mathbf{H}\mathbf{F}\mathbf{F}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n)^{-1} \tag{7}$$

This can be rewritten using matrix inversion lemma [4]

$$\mathbf{E} \equiv E[\mathbf{e}\mathbf{e}^\dagger] = (\mathbf{I} + \mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F})^{-1} \tag{8}$$

In the following we will describe how to solve the MMSE problem with the per-antenna power constraints. This is done by formulating the problem as a semi-definite program (SDP) as we shall see.

The MMSE problem with per-antenna power constraints can be formulated as

$$\begin{aligned} \min_{\mathbf{F}} \quad & Tr(\mathbf{E}) \\ \text{s.t.} \quad & (\mathbf{F}\mathbf{F}^\dagger)_{ii} \le P_i, \quad \forall i = 1, 2, \cdots, M \end{aligned} \tag{9}$$

where $Tr(\mathbf{E})$ denotes the trace of the matrix $\mathbf{E}$. From (6)

$$\begin{aligned} Tr(\mathbf{E}) = & \quad Tr(\mathbf{I} - \mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{W}\mathbf{H}\mathbf{F}) \\ = & \quad M - Tr(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{W}\mathbf{H}\mathbf{F}) \\ = & \quad M - Tr(\mathbf{W}\mathbf{H}\mathbf{F}\mathbf{F}^\dagger\mathbf{H}^\dagger) \\ = & \quad M - Tr(\mathbf{W}(\mathbf{W}^{-1} - \mathbf{R}_n)) \\ = & \quad M - J + Tr(\mathbf{W}\mathbf{R}_n) \end{aligned}$$

Since M and J are constants, $\mathbf{R}_n$ is known, and the MMSE depends only on $\mathbf{W}$, which is a function of $\mathbf{F}$ as in (7). Furthermore, if we define

$$\mathbf{U} \equiv \mathbf{F}\mathbf{F}^\dagger \tag{10}$$

then according to (7) we can write

$$\mathbf{W} = (\mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n)^{-1} \tag{11}$$

By the above formulation, we can re-cast the problem (9) as follows:

$$\begin{aligned} \min_{\mathbf{U}} \quad & Tr(\mathbf{W}\mathbf{R}_n) \\ \text{s.t.} \quad & (a)\ (\mathbf{U})_{ii} \le P_i, \quad \forall i = 1, 2, \cdots, M \\ & (b)\ \mathbf{U} \succeq \mathbf{0} \\ & (c)\ \mathbf{W} = (\mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n)^{-1} \end{aligned} \tag{12}$$

where $\mathbf{U} \succeq \mathbf{0}$ means that the matrix $\mathbf{U}$ is positive semi-definite. The problem formulation in (12) is not convex though. Similar to [9], by introducing an auxiliary matrix variable $\mathbf{W}_0$, we can cast the problem in (12) as the following equivalent problem:

$$\begin{aligned} \min_{\mathbf{U},\mathbf{W}_0} \quad & Tr(\mathbf{W}_0\mathbf{R}_n) \\ \text{s.t.} \quad & (a)\ (\mathbf{U})_{ii} \le P_i, \quad \forall i = 1, 2, \cdots, M \\ & (b)\ \mathbf{U} \succeq \mathbf{0} \\ & (c)\ \mathbf{W}_0 \succeq (\mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n)^{-1} \end{aligned} \tag{13}$$

The equivalence of problem formulation (12) and (13) can be easily explained as follows: In (13) we are minimizing the trace of matrix $\mathbf{W}_0\mathbf{R}_n$, where $\mathbf{R}_n$ is positive semi-definite. Therefore we have [9]

$$Tr(\mathbf{W}_0\mathbf{R}_n) \ge Tr((\mathbf{H}\mathbf{U}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}\mathbf{R}_n)$$

since $\mathbf{W}_0 \succeq (\mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}$. The optimal solution of problem (13) ensures the equality

$$\mathbf{W}_0 = (\mathbf{H}\mathbf{U}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}$$

holds, where $\mathbf{U}$ is also the optimal solution for (12). Note that the constraint

$$\mathbf{W}_0 \succeq (\mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}$$

holds if only if the following linear matrix inequality holds (p. 472 in [4])

$$\begin{pmatrix} \mathbf{H}\mathbf{U}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n & \mathbf{I} \\ \mathbf{I} & \mathbf{W}_0 \end{pmatrix} \succeq \mathbf{0} \tag{14}$$

Therefore, our final form of problem formulation can be written as

$$\begin{aligned} \min_{\mathbf{U},\mathbf{W}_0} \quad & Tr(\mathbf{W}_0\mathbf{R}_n) \\ \text{s.t.} \quad & (a)\ (\mathbf{U})_{ii} \le P_i, \quad \forall i = 1, 2, \cdots, M \\ & (b)\ \mathbf{U} \succeq \mathbf{0} \\ & (c)\ \begin{pmatrix} \mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n & \mathbf{I} \\ \mathbf{I} & \mathbf{W}_0 \end{pmatrix} \succeq \mathbf{0} \end{aligned} \tag{15}$$

In (15) the objective function is linear, and the constraints are either linear or positive semi-definite, which implies the problem (15) is a SDP problem [15]. This ensures that the global minimum of (15) can be found in polynomial time. The algorithmic complexity of solving SDP using interior points methods is $O(n^{6.5}log(1/\varepsilon))$, where ε is the solution accuracy, and n is the problem size [15], [1]. Once we solve the problem (15), the precoder matrix $\mathbf{F}$ is the Cholesky factorization of the solution $\mathbf{U}$ in (15).

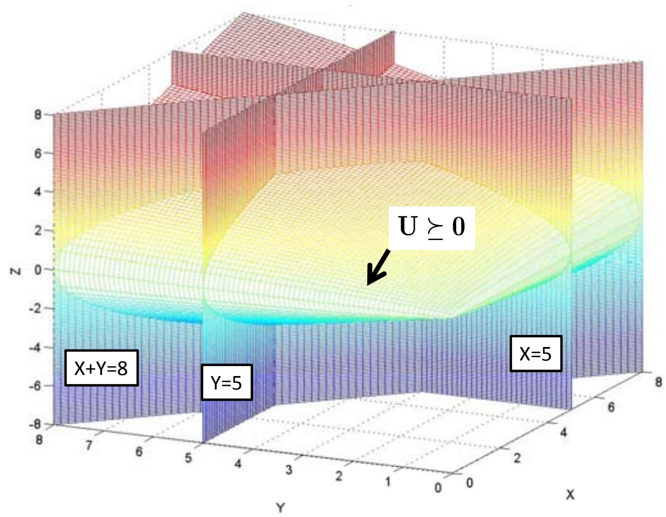

Fig. 2. Visualization of the feasible region of the solution of (15) for the two antenna case.

Fig. 2 shows the visualization of the feasible region of the solution $\mathbf{U}$ in problem (15) for the two antenna case. For simplicity, we assume

$$\mathbf{U} = \begin{pmatrix} X & Z \\ Z & Y \end{pmatrix}$$

where Z is a real number. Therefore the constraint $U \succeq \mathbf{0}$ forms a positive semi-definite cone. The constraints $\mathbf{U}_{kk} \leq 5$ define the halfspaces on one side of the hyperplanes $\mathbf{U}_{kk} = 5$. The total power constraint $\text{Tr}(\mathbf{U}) \leq 8$, is also a halfspace formed by the hyperplane cutting through the semi-definite cone. The convex problem is solved by running the interior point method [2] in the feasible region, which is formed by the intersection of the semi-definite cone and the halfspaces.

III. BER Optimization of Transceiver with Per-Antenna Power Constraints

In this section we derive the optimal transceiver for the average bit error rate (BER), with per-antenna power constraints. This method is based on distributing the MSE identically in each substream.

Suppose we already found the solution $\mathbf{U}_{mmse}$ to the problem (15). The optimal MMSE solution $\mathbf{F}_{mmse}$ can be taken as any Cholesky factor of $\mathbf{U}_{mmse}$, so that

$$\mathbf{F}_{mmse}\mathbf{F}_{mmse}^{\dagger} = \mathbf{U}_{mmse}. \tag{16}$$

Note that for any unitary matrix $\mathbf{\Psi}$, the product $\mathbf{F}_{mmse}\mathbf{\Psi}$ will still be an MMSE solution. This is the freedom we have, to optimize the average bit error rate. Suppose matrix $\mathbf{V}$ is the unitary matrix which diagonalizes $\mathbf{F}_{mmse}^{\dagger}\mathbf{H}^{\dagger}\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F}_{mmse}$:

$$\mathbf{F}_{mmse}^{\dagger}\mathbf{H}^{\dagger}\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F}_{mmse} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^{\dagger}. \tag{17}$$

Then the MSE matrix in (8) is

$$\begin{aligned} E[\mathbf{e}\mathbf{e}^{\dagger}] = & \quad (\mathbf{I} + \mathbf{F}_{mmse}^{\dagger}\mathbf{H}^{\dagger}\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F}_{mmse})^{-1} \\ = & \quad \mathbf{V}(\mathbf{I} + \mathbf{\Sigma})^{-1}\mathbf{V}^{\dagger} \end{aligned} \tag{18}$$

In view of Theorem 1, which will be proved later, the optimal $\mathbf{F}$ has the form

$$\mathbf{F} = \mathbf{F}_{mmse}\mathbf{V}\mathbf{\Phi} \tag{19}$$

where $\mathbf{\Phi}$ denotes the unitary matrix such that the MSE matrix as in (8) has the identical diagonal elements. Examples of such $\mathbf{\Phi}$ are the Hadamard matrix and the discrete Fourier transform (DFT) matrix [8], [3].

The performance of a digital communication system ultimately is given by the bit error rate (BER). Assuming a Gray encoding is used to map the bits into the constellation points, the BER can be approximately obtained from the symbol error rate (SER), i.e.,

$$BER \cong SER/k,$$

where k is the number of bits per-symbol. Under the Gaussian assumption the SER can be approximately expressed as a function of SINR:

$$Pe(SINR) = \alpha Q(\sqrt{\beta SINR}), \tag{20}$$

where Q is Q-function defined as in [12] and α and β are constants depending on the constellation. The following lemma is useful in developing the BER optimality of $\mathbf{F}$ in (19).

Lemma 1. In the high SINR region, the bit error rate is a convex increasing function of the MSE. That is, $BER(E_{ii})$ is a convex increasing function of E_{ii}, where E_{ii} is defined as the ith diagonal term of the MSE matrix in (8).

Proof: See Appendix H in [12]. ■

The high SINR regions represent the low BER regions, which are often the focus of the analysis in communication systems. The following theorem shows that the system with $\mathbf{F}$ in (19), and with the corresponding $\mathbf{G}$ in (5), has the minimum average BER.

Theorem 1. Among all systems in Fig. 1 which satisfy (2), the $\mathbf{F}$ in (19) and the corresponding $\mathbf{G}$ in (5) are the optimal solutions that minimize the average BER, which is defined as

$$\overline{BER} = \frac{1}{M}\sum_{i=1}^{M} BER(E_{ii}) \tag{21}$$

.

Proof: For the system $\{\mathbf{F}_{mmse}, \mathbf{G}_{mmse}\}$, suppose the diagonal terms of the MSE matrix are $\underline{E} = \{E_{11}, E_{22}, \cdots, E_{MM}\}$, where $\mathbf{F}_{mmse}$ is a Cholesky factor of $\mathbf{U}_{mmse}$, and $\mathbf{G}_{mmse}$ is the corresponding LMMSE filter.

Suppose there is another system, $\{\mathbf{F}^{'}, \mathbf{G}^{'}\}$, which has the diagonal terms of the MSE matrix $\underline{E}^{'} = \{E_{11}^{'}, E_{22}^{'}, \cdots, E_{MM}^{'}\}$. Trivially we have

$$\underline{E}^{'} \succ \overline{\underline{E}}'$$

where $\overline{\underline{E}}'$ is defined as the vector with all elements equal to the arithmetic mean of $\underline{E}'$, and '$\succ$' denotes the majorization as in [10]. According to **Lemma 1**, $BER(\cdot)$ is a convex function of E_{ii}. Therefore the average BER is a Schur-convex function of $\underline{E}^{'}$ (p. 70 in [10]). Then we have

$$\begin{aligned} \overline{BER}^{'} &= \tfrac{1}{M}\textstyle\sum_{i=1}^{M} BER(E_{ii}^{'}) && (22)\\ &\geq BER(\tfrac{1}{M}\textstyle\sum_{i=1}^{M} E_{ii}^{'}) && (23)\\ &\geq BER(\tfrac{1}{M}\textstyle\sum_{i=1}^{M} E_{ii}) && (24)\\ &= \overline{BER}_{(19)} && (25) \end{aligned}$$

where $\overline{BER}_{(19)}$ is defined as the BER obtained by using $\mathbf{F}$ in (19). The inequality in (23) follows from the fact that Schur-convex function preserves majorization. The inequality in (24) is due to the fact that $\underline{E}$ is the corresponding MSE of the solution to problem (15), and the fact that $BER(\cdot)$ is an increasing function. The equality in (25) is because $\mathbf{F}$ in (19) is the one that makes all diagonal terms of the MSE matrix identical.

Therefore, any other system satisfying (2) will have larger average BER than the system with $\mathbf{F}$ as in (19) and the corresponding LMMSE receiver $\mathbf{G}$ as in (5). This concludes our proof. ■

Note that the proof given above relies only on the fact that the average BER is a Schur-convex and increasing function of MSE's. Therefore the same concept can also be applied to other objective functions that have these two properties. Many examples of such objective functions are provided in [12]. For all such objective functions, the system with $\mathbf{F}$ as in (19) and the corresponding LMMSE receiver $\mathbf{G}$ as in (5) is always the optimal one!

IV. Numerical Results

In this section we present our simulation results relevant to the discussions carried out in previous sections. After formulating the problem as in (15), we use the convex optimization tool "SeDuMi" [14] to obtain the optimal solution.

In both Fig. 3 and Fig. 4 we choose $M = 4$, $J = 5$, and per-antenna power constraints to be $[P_1, P_2, P_3, P_4] = \alpha[5, 4, 3, 2]$, where α controls the total power. The constellation are all QPSK. The noise is additive white Gaussian for each case, with covariance matrix $\mathbf{R}_n = 4\mathbf{I}$ for Fig. 3, and $\mathbf{R}_n = 6.25\mathbf{I}$ for Fig. 4. The channels used in the simulations are $\mathbf{H}_1$ and $\mathbf{H}_2$, for Fig. 3 and Fig. 4 respectively, given as follows:

$$\mathbf{H}_1 = \begin{pmatrix} -1.59-0.62i & -0.94-0.85i & -0.44-0.43i & -0.70-0.50i \\ 1.27+0.88i & 0.44-1.66i & 0.75+0.34i & -0.58+1.17i \\ -0.40-0.05i & 0.17-0.70i & -2.48+0.51i & 0.31-0.18i \\ 0.22-1.16i & -0.00-0.80i & 1.07-1.23i & 0.80+0.79i \\ -0.80-0.99i & -2.09+0.13i & 0.88-0.22i & -1.86+0.18i \end{pmatrix}$$

and

$$\mathbf{H}_2 = \begin{pmatrix} 0.01+0.47i & -0.19+0.43i & 0.84+0.86i & 2.27-1.37i \\ 1.81-0.23i & -1.14-0.62i & -1.43+2.48i & 0.63-0.83i \\ 2.47+1.34i & 3.66-2.19i & 1.78+0.27i & 1.71-1.41i \\ -0.82-0.45i & -1.19-1.51i & 0.96-0.74i & -0.08-1.67i \\ 1.38-0.15i & 0.82-0.17i & 0.26-0.97i & -1.23+1.64i \end{pmatrix}$$

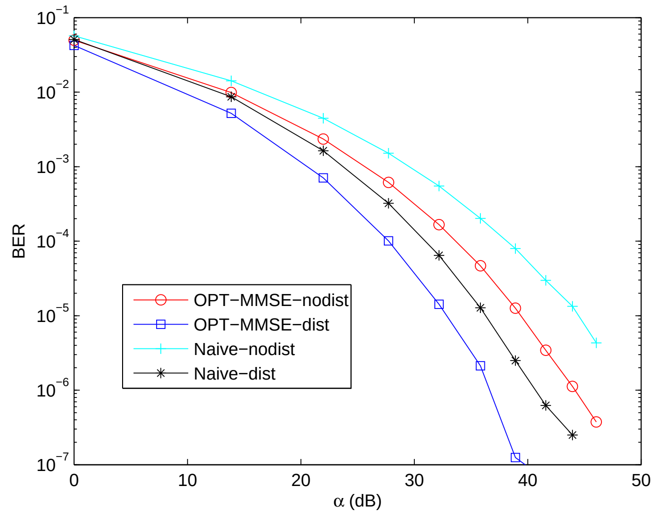

Fig. 3. Comparing four different transceivers for channel $\mathbf{H}_1$, with per-antenna power constraints $\alpha[5, 4, 3, 2]$.

The plots in Fig. 3 and Fig. 4 are with respect to the variable α, where α controls the total power. "OPT-MMSE-nodist" denotes the optimal MMSE design but without distributing the MSE in each sub-stream, i.e., the precoder is taken to be a Cholesky factor $\mathbf{F}_{mmse}$ as in (16). "OPT-MMSE-dist" denotes the optimal MMSE design with the MSE in each sub-stream identical, which corresponds to using $\mathbf{F}$ as in (19). "Naive-nodist" denotes the case where the power constraints are satisfied by using the simple choice $\mathbf{F}_{naive} = diag(\alpha[P_1, P_2, \cdots, P_M])$. "Naive-dist" corresponds to the case where the precoder matrix is $\mathbf{F} = \mathbf{F}_{naive}\mathbf{V}\mathbf{\Phi}$ as in (19), so that (8) has identical diagonal elements.

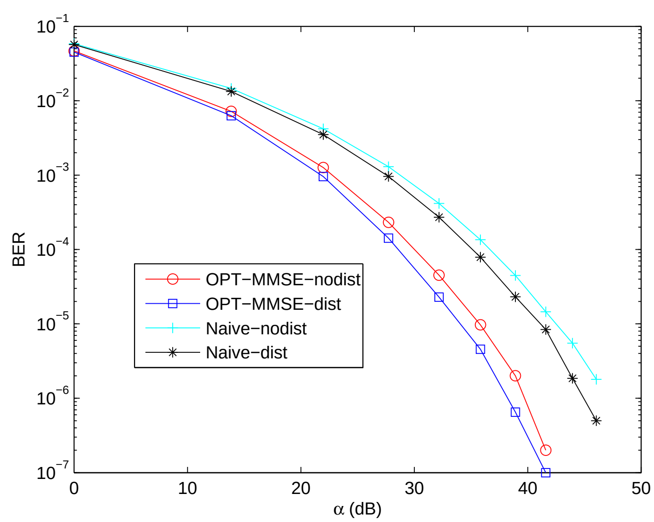

Fig. 4. Comparing four different transceivers for channel $\mathbf{H}_2$, with per-antenna power constraints $\alpha[5, 4, 3, 2]$.

From the simulations, we can see that the proposed optimization method is superior to naive designs. Also, both the MSE-minimization step and MSE-distribution step are beneficial for minimizing the average BER.

V. Concluding Remarks

We have presented a convex formulation for optimal transceiver design with per-antenna power constraints. The problem is re-formulated as a SDP, which can be solved efficiently by convex optimization tools. The optimization for average BER as well as other Schur-convex functions are also discussed. Several simulation results were presented to demonstrate the advantages of the optimized designs.

References

[1] A. Antoniou, and W. S. Lu, "Practical optimization: algorithms and engineering applications," Springer. 2007.

[2] S. Boyd, and L. Vandenberghe, *Convex optimization,* Cambridge. 2004.

[3] Y. Ding, T. N. Davidson, Z. Luo, and K. M. Wong, "Minimum BER precoders for zero-forcing equalization," *IEEE Trans. Sig. Proc.* pp. 2410 - 2423, Sep. 2003.

[4] R. A. Horn, C. R. Johnson, "Matrix analysis," Cambridge Univ. Press 1985.

[5] Y. Jiang, J. Li, and W. W. Hager "Joint transceiver design for MIMO communications using geometric mean decomposition, " *IEEE Trans. Sig. Proc.,* pp.3791-3803, Oct. 2005.

[6] Y. Jiang, D. P. Palomar, and M. K. Varanasi, "Precoder Optimization for Nonlinear MIMO Transceiver Based on Arbitrary Cost Function," *Proc. Conference on Information Sciences and Systems.* March, 2007.

[7] K. H. Lee, and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Comm.* pp. 1283 - 1290, Dec. 1976.

[8] Y. P. Lin, and S. M. Phoong, "BER minimized OFDM systems with channel independent precoders," *IEEE Trans. Sig. Proc.* pp. 2369 - 2380, Sep. 2003.

[9] Z. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Sig. Proc.* pp. 1037 - 1052, Apr. 2004.

[10] A. W. Marshall, I. Olkin, "Inequalities: Theory of majorization and its applications,"Academic Press. 1979.

[11] H. S. Malvar, and D. H. Staelin, "Optimal pre- and postfilters for multichannel signal processing," *IEEE Trans. Acoustics, Speech, and Signal Processing* pp. 287 - 289, Feb. 1988.

[12] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Sig. Proc.* pp. 2381 - 2401, Sept. 2003.

[13] M. B. Shenouda, T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding," *Proc. of the ICASSP,* pp. 209-212, April. 2007.

[14] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones" *Optim. Meth. Software.* vol. 11-12 pp. 625 - 653, 1999.

[15] L. Vandenberghe, and S. Boyd, "Semidefinite programming," *SIAM Rev.* pp. 49 - 95, 2001.

[16] F. Xu, T. N. Davidson, J. K. Zhang, S. S. Chan, and K. M. Wong, "Design of block transcivers with MMSE decision feedback detection," *Proc. of the ICASSP,* pp. 1109 - 1112, March. 2004.