MIMO Transceiver Optimization With Linear Constraints on Transmitted Signal Covariance Components

Ching-Chih Weng and P. P. Vaidyanathan

Abstract—This correspondence revisits the joint transceiver optimization problem for multiple-input multiple-output (MIMO) channels. The linear transceiver as well as the transceiver with linear precoding and decision feedback equalization are considered. For both types of transceivers, in addition to the usual total power constraint, an individual power constraint on each antenna element is also imposed. A number of objective functions including the average bit error rate, are considered for both of the above systems under the generalized power constraint. It is shown that for both types of systems the optimization problem can be solved by first solving a class of MMSE problems (AM-MMSE or GM-MMSE depending on the type of transceiver), and then using majorization theory. The first step, under the generalized power constraint, can be formulated as a semidefinite program (SDP) for both types of transceivers, and can be solved efficiently by convex optimization tools. The second step is addressed by using results from majorization theory. The framework developed here is general enough to add any finite number of linear constraints to the covariance matrix of the input.

Index Terms—BER optimization, MIMO transceivers, per-antenna constraints, Schur convexity, semidefinite programming.

I. INTRODUCTION

In this correspondence we revisit the optimization of multiple-input multiple-output (MIMO) communication systems. In the MIMO system, the transmitter has M antennas sending independent information to the receiver equipped with N antennas. The signal vector consisting of M substreams 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 at both ends of the communication link, in [10] the authors considered the linear transceiver optimization problem under the total power constraint. By introducing majorization theory, several objective functions, which can be categorized as Schur-convex or Schur-concave [8] 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 [4], [12].

Instead of only the total power constraint, in this correspondence we also consider the more realistic per-antenna power constraints on the transmitter [6], [18]. This is because in practice each antenna is limited individually by its equipped power amplifier. In [6], the MMSE problem under individual power constraints is solved suboptimally using a numerical approach. In [18], the multiuser down-link transceiver design problem is considered. The total power constraint might still be needed since the antennas might rely on a common power supply. Under these constraints, we consider the linear transceiver case and also the simple nonlinear case, i.e., linear precoding with DFE at the equalizer.[1] For each case, we will see that the optimization process consists of two steps.

For the linear transceiver case, we first consider the minimum arithmetic mean of mean-square errors (AM-MSE) design. We show that it can be reformulated as a semidefinite program (SDP), which can be solved numerically by convex optimization tools. Then, among the family of minimum AM-MSE linear transceivers, we develop a method to find the one that minimizes the average bit error rate as well as many other objective functions. This second step is achieved by appealing to majorization theory. Similarly, for the transceivers with linear precoding and DFE, we first consider the minimum geometric mean of mean-square errors (GM-MSE) design. We show that it can also be reformulated as an SDP, and solved efficiently. Then, among the family of minimum GM-MSE designs, we develop a method to find the one that minimizes the average bit error rate as well as many other objective functions.

Based on majorization theory [8], we will argue that the minimal average BER transceiver design method developed in this correspondence can also be applied to a wider class of objective functions. Also, we will show that under the framework developed in the correspondence, any additional linear constraints on the covariance matrix of the transmitted signals can be further added, and the problem is solved both in theory and practice with no difficulty.

A. Relation to Previous Work

The problem of optimizing linear transceivers subject to individual power constraints was addressed in [6], [9], [10], and [18], in different contexts, but transceivers with DFE were not considered. In [6], the authors considered the MMSE problem, and solved it suboptimally using numerical methods. In [10], the authors considered only Schur-concave objective functions subject to the individual power constraints. However, the problem of optimizing the transceiver for other important objective functions (e.g., Schur-convex functions, including average BER) was not addressed. Direct use of the results of [10] to address this case is nontrivial. In [9] the author considered shaping constraints on the transmitted signal covariance matrix. However, as acknowledged by the author in [9], the paper introduced a stronger artificial constraint which leads to a suboptimal solution. This will be elaborated in Section V-B. We will also argue that under our framework, the problem of optimizing transceivers subject to other meaningful constraints on the transmitting covariance matrix can be solved. However, there are some shortcomings and restrictions of our framework. These will be elaborated in Section V-B as well.

Luo *et al.* [7] formulated the MMSE transceiver design for the multiuser case as an SDP problem. For linear transceivers, we adopt a similar mathematical derivation to further consider the average BER minimization as well as other Schur-convex problems for MIMO communication systems. For the system with DFE and linear precoding, we formulate the minimization of GM-MSE to be a log determinant maximization problem. The determinant maximization problem for a positive definite matrix is well studied in the literature [15]. It turns out that this GM-MSE minimization problem has exactly the same form as the maximization of the mutual information of communication systems over Gaussian channels [2], [15], [17].

B. Outline of the Correspondence

This correspondence is structured as follows. In Section II, we will discuss the communication models used in the correspondence, and

[1]Throughout the correspondence, the phrase "simple nonlinear transceiver" refers to a transceiver with a linear precoder and a DFE equalizer.
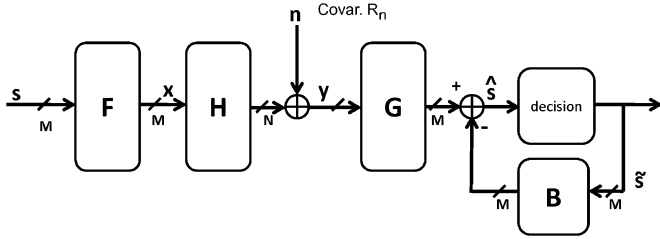
Fig. 1.   System with linear precoding and DFE.

also give explicit problem formulations. The review of the optimal receiver structures will also be given. In Sections III and IV, we formulate the optimal transceiver design problems for the linear and DFE cases, respectively. We also derive the optimal designs which achieve minimum BER. In Section V, we will give further discussions on the framework developed in this correspondence. The relationship between this correspondence and [9] is also addressed. Section VI presents the numerical simulation results related to the topics discussed in the correspondence. The final conclusions are summarized in Section VII.[2]

II. SYSTEM MODELS AND PROBLEM FORMULATION

The transceiver model can be represented as in Fig. 1. $\mathbf{R}_n$ is the covariance matrix of the additive Gaussian noise; $\mathbf{H}$ is the channel matrix; $\mathbf{F}$ is the precoder; $\mathbf{G}$ is the receiving filter; $\mathbf{B}$ is $\mathbf{0}$ for linear transceiver case, and strictly lower triangular [16] for the system with linear precoding and DFE. The per-antenna power constraints can be formulated as

$$\left(E[\mathbf{F}\mathbf{s}\mathbf{s}^\dagger\mathbf{F}^\dagger]\right)_{ii} = (\mathbf{F}\mathbf{F}^\dagger)_{ii} \leq P_i, \ \forall i = 1, 2, \cdots, M. \quad (1)$$

The total power constraint can be written as

$$\mathrm{Tr}(\mathbf{F}\mathbf{F}^\dagger) \leq P_{\mathrm{total}}. \quad (2)$$

In this work we want to minimize the uncoded bit error rate subject to the individual and total power constraints. This problem formulation is slightly different from those considered by previous papers [4], [10], [12], and [6], which considered only either the total power constraint or the individual power constraints. It will be elaborated later that because of the introduction of the constraints (1), the nice quasi-analytic form of the optimal solutions obtained in [10] is gone. However, since the convexity of the problem still holds true, the optimal solution can be obtained by numerical search as we show in the following sections. In this correspondence, we will establish the link between the optimal transceiver design problems and the SDPs. Also, the majorization theory is used to further simplify the considered problems.

It can be observed that the optimal receiver structures do not change due to the individual power constraints when $\mathbf{F}$ is given in both cases. In the following we briefly review the optimal receiver structures [10]–[12]. In the linear transceiver case, it turns out that with no loss of optimality [11] we can use the Wiener filter to be the

[2]The following notations are used in the correspondence. Boldface upper-case letters denote matrices, boldface lower-case letters denote the column vectors, and italics denote scalars. The superscript $(\cdot)^\dagger$ denotes transpose conjugation, and $(\cdot)^T$ denotes the transpose operation. $\mathbf{A}_{ij}$ denotes the (i,j)th element of the matrix $\mathbf{A}$. By $\mathbf{A} \succeq \mathbf{B}$, we mean $\mathbf{A} - \mathbf{B}$ is positive semidefinite. For vector $\mathbf{x}$, the notation $\mathrm{diag}(\mathbf{x})$ denotes the diagonal matrix with diagonal terms equal to the elements in the vector $\mathbf{x}$. For matrix $\mathbf{X}$, the notation $\mathrm{diag}(\mathbf{X})$ denotes the column vector whose elements are the diagonal terms of the matrix $\mathbf{X}$. The notation $\mathbf{a} \succ_+ \mathbf{b}$ means that the vector $\mathbf{a}$ majorizes $\mathbf{b}$ additively [8], [10]. Similarly $\mathbf{a} \succ_\times \mathbf{b}$ means that the vector $\mathbf{a}$ majorizes $\mathbf{b}$ multiplicatively [4], [8].

receiver, i.e., $\mathbf{G}_{\mathrm{opt}} = \mathbf{F}^\dagger\mathbf{H}^\dagger(\mathbf{H}\mathbf{F}\mathbf{F}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}$. If the Wiener filter $\mathbf{G}_{\mathrm{opt}}$ is used, the resulting error covariance matrix (MSE matrix) can be written as $\mathbf{E} := E[\mathbf{e}\mathbf{e}^\dagger] = \mathbf{I} - \mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{W}\mathbf{H}\mathbf{F}$, where $\mathbf{W} = (\mathbf{H}\mathbf{F}\mathbf{F}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}$. This can be rewritten in the following form by using matrix inversion lemma [10]: $\mathbf{E} = (\mathbf{I} + \mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F})^{-1}$.

Note that $\mathbf{G}_{\mathrm{opt}}$ is both optimal in the sense of maximizing SINR in each substream as well as minimizing the mean square error. In this case, the SINR can be related to the MSE as [10]

$$\mathrm{SINR}_i = \frac{1}{\mathrm{MSE}_i} - 1. \quad (3)$$

The optimum decision feedback equalization with successive decoding for MIMO channels is considered in [16]. First, the feedforward filter is $\mathbf{G}_{\mathrm{opt}} = \mathbf{C}\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{R}_y^{-1}$, and the resulting MSE matrix can be written as $\mathbf{E} := E[\mathbf{e}\mathbf{e}^\dagger] = \mathbf{C}(\mathbf{I} + \mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F})^{-1}\mathbf{C}^\dagger = \mathbf{C}\mathbf{M}\mathbf{C}^\dagger$, where $\mathbf{M}$ is defined as: $\mathbf{M} := (\mathbf{I} + \mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F})^{-1}$, and $\mathbf{C} := \mathbf{I} + \mathbf{B}$. It can also be shown that the optimal $\mathbf{C}$ can be chosen as [12] $\mathbf{C} = \mathrm{diag}([\mathbf{L}_{11}, \cdots, \mathbf{L}_{MM}]^T)\mathbf{L}^{-1}$, where $\mathbf{L}$ is the lower triangular Cholesky factor of $\mathbf{M}$, i.e., $\mathbf{M} = \mathbf{L}\mathbf{L}^\dagger$. The resulting MSE matrix will be $\mathbf{E} = \mathrm{diag}([\mathbf{L}_{11}^2, \cdots, \mathbf{L}_{MM}^2]^T)$. Under these choices, the SINR and MSE in each substream also have a nice relation as in (3) as shown in [4] and [16].

III. OPTIMAL LINEAR TRANSCEIVER SYSTEMS

In this section we will focus on solving the problem of minimizing BER subject to individual and total power constraints for the linear transceiver case. We use the two-step approach. In the first step we will minimize the AM-MSE (Arithmetic mean of mean square error) of the system. This is done by reformulating the problem as an SDP as we shall see. In the second step, we will argue that there is a set of systems in the minimum AM-MSE family which minimize the average error probability among all linear transceivers $\{\mathbf{F}, \mathbf{G}\}$. That is, the average BER is the smallest possible. An approach to find one of such optimal transceivers will also be given.

The minimum AM-MSE problem with per-antenna and total power constraints can be cast as the minimization of $\mathrm{Tr}(\mathbf{E})$, where $\mathbf{E}$ is the MSE matrix as discussed in Section II.

In the following we will adopt the trick in [7] to formulate the current problem to be an SDP. By similar derivation as in [7] we have $\mathrm{Tr}(\mathbf{E}) = M - N + \mathrm{Tr}(\mathbf{W}\mathbf{R}_n)$. Since M and N are constants and $\mathbf{R}_n$ is known, the AM-MSE depends only on $\mathbf{W}$, which is a function of $\mathbf{F}$. Furthermore, if we define $\mathbf{U} := \mathbf{F}\mathbf{F}^\dagger$, we can write $\mathbf{W} = (\mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}$.

This equation can be replaced with $\mathbf{W}_0 \succeq (\mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n)^{-1}$ (as discussed in [7]). Also it holds true if and only if the following linear matrix inequality holds [3, p. 472] $\begin{pmatrix} \mathbf{H}\mathbf{U}^\dagger\mathbf{H}^\dagger + \mathbf{R}_n & \mathbf{I} \\ \mathbf{I} & \mathbf{W}_0 \end{pmatrix} \succeq \mathbf{0}$.

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

$$\min_{\mathbf{U}, \mathbf{W}_0} \quad \mathrm{Tr}(\mathbf{W}_0\mathbf{R}_n)$$

s.t. (a) $(\mathbf{U})_{ii} \leq P_i, \quad \forall i = 1, 2, \cdots, M$

(b) $\mathrm{Tr}(\mathbf{U}) \leq P_{\mathrm{total}}$

(c) $\mathbf{U} \succeq \mathbf{0}$

(d) $\begin{pmatrix} \mathbf{H}\mathbf{U}\mathbf{H}^\dagger + \mathbf{R}_n & \mathbf{I} \\ \mathbf{I} & \mathbf{W}_0 \end{pmatrix} \succeq \mathbf{0}. \quad (4)$

In (4), the objective function is linear, and the constraints are either linear or positive semidefinite. Therefore, the problem (4) is an SDP problem [14]. This ensures that the global minimum of (4) can be found in polynomial time, when the precision of the solution is specified.

Now consider any given precoder $\mathbf{F}$, where a unitary matrix $\boldsymbol{\Psi}$ is further inserted in front of $\mathbf{F}$. We notice that this substitution does not change the individual power in each antenna nor the AM-MSE [1]. In

the high SNR region, the average BER is an increasing Schur-convex function [10] in the vector $\mathrm{diag}(\mathbf{E})$. Therefore, we have

$$P_{br} = \frac{1}{M} \sum_{i=1}^{M} \alpha Q \left(\sqrt{ \beta \cdot \left(\frac{1}{\mathbf{E}_{ii}} - 1 \right) } \right)$$

$$\geq \alpha Q \left(\sqrt{ \beta \cdot \left(\frac{1}{\frac{1}{M} \mathrm{Tr}(\mathbf{E})} - 1 \right) } \right) \qquad (5)$$

where α and β are constants depending on the QAM constellation. It is now clear that the lower bound is minimized by minimum AM-MSE. The equality is achieved by choosing the matrix $\mathbf{\Psi}$ to equalize the MSE in each substream.

Now we provide an approach to obtain one of the optimal minimum BER solutions. Taking any solution of $\mathbf{U}_{\mathrm{AM}}$ to the problem (4), the optimal minimum AM-MSE solution $\mathbf{F}_{\mathrm{AM}}$ can be taken as any Cholesky factor of $\mathbf{U}_{\mathrm{AM}}$. Let $\mathbf{V}$ denote the unitary matrix which diagonalizes $\mathbf{F}_{\mathrm{AM}}^{\dagger} \mathbf{H}^{\dagger} \mathbf{R}_n^{-1} \mathbf{H} \mathbf{F}_{\mathrm{AM}}$: $\mathbf{F}_{\mathrm{AM}}^{\dagger} \mathbf{H}^{\dagger} \mathbf{R}_n^{-1} \mathbf{H} \mathbf{F}_{\mathrm{AM}} = \mathbf{V} \mathbf{\Sigma} \mathbf{V}^{\dagger}$. The optimal precoder can be taken as $\mathbf{F}_{\mathrm{opt}} = \mathbf{F}_{\mathrm{AM}} \mathbf{V} \mathbf{\Phi}$, where $\mathbf{\Phi}$ denotes the unitary matrix such that the MSE matrix has the identical diagonal elements. The existence of such unitary matrix $\mathbf{\Phi}$ is given by [8]. $\mathbf{\Phi}$ can be taken as a matrix with constant magnitude in each of its entries. Examples of such $\mathbf{\Phi}$ are the Hadamard matrix and the discrete Fourier transform (DFT) matrix [10]. Note that such $\mathbf{\Phi}$ are not unique, which means the minimum BER system is not unique.

IV. OPTIMAL LINEAR PRECODING WITH DFE SYSTEMS

In this section we will focus on solving the problem of minimizing the BER subject to individual and total power constraints for the system with DFE and linear precoding. We will take the two-step approach. In the first step we will minimize the GM-MSE (Geometric mean of mean square error) of the system. This is done by formulating the problem as an SDP as we shall see. In the second step, we will argue there is a set of systems in the minimum GM-MSE family, which yield the minimum average BER among all transceivers $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$. A method to find one such optimal transceiver is also discussed in this section.

Since $\mathbf{E} = \mathbf{C} \mathbf{M} \mathbf{C}^{\dagger}$, and $\mathbf{C}$ is a lower triangular matrix with diagonal terms equal to the identity, we have the relation $\det(\mathbf{E}) = \det(\mathbf{C} \mathbf{M} \mathbf{C}^{\dagger}) = \det(\mathbf{M}) = \prod_{i=1}^{M} \mathbf{L}_{ii}^2$, which is the product of the MSE in each substream. Therefore, the minimization of the geometric mean of the MSEs is equivalent to the minimization of the determinant of $\mathbf{M}$.

Since $\mathbf{M}$ has the form as in Section II, we have $\det(\mathbf{M}) = \det(\mathbf{I}_M + \mathbf{F}^{\dagger} \mathbf{H}^{\dagger} \mathbf{R}_n^{-1} \mathbf{H} \mathbf{F})$. Note that for any $m \times n$ matrix $\mathbf{X}$, we have the equality $\det(\mathbf{I}_m + \mathbf{X} \mathbf{X}^{\dagger}) = \det(\mathbf{I}_n + \mathbf{X}^{\dagger} \mathbf{X})$. Therefore, we have $\det(\mathbf{I}_M + \mathbf{F}^{\dagger} \mathbf{H}^{\dagger} \mathbf{R}_n^{-1} \mathbf{H} \mathbf{F}) = \det(\mathbf{I}_N + \mathbf{R}_n^{-(1/2)} \mathbf{H} \mathbf{F} \mathbf{F}^{\dagger} \mathbf{H}^{\dagger} \mathbf{R}_n^{-(1/2)})$, where $\mathbf{R}_n^{-(1/2)}$ is the Cholesky factor of the noise covariance matrix $\mathbf{R}_n$. By setting $\mathbf{U} := \mathbf{F} \mathbf{F}^{\dagger}$, we can rewrite the problem of minimizing the GM-MSE in the following form:

$$\begin{aligned} \max_{\mathbf{U}} \quad & \log \det \left(\mathbf{I}_N + \mathbf{R}_n^{-\frac{1}{2}} \mathbf{H} \mathbf{U} \mathbf{H}^{\dagger} \mathbf{R}_n^{-\frac{1}{2}} \right) \\ \mathrm{s.t.} \quad & \mathrm{(a)} \quad (\mathbf{U})_{ii} \leq P_i, \quad \forall i = 1, 2, \cdots, M \\ & \mathrm{(b)} \quad \mathrm{Tr}(\mathbf{U}) \leq P_{\mathrm{total}} \\ & \mathrm{(c)} \quad \mathbf{U} \succeq \mathbf{0}. \end{aligned} \qquad (6)$$

This reformulation holds true because the $\log(\cdot)$ function is a monotone function when the argument is positive. Problem (6) has been considered by several authors [14], [15]. It is an SDP-representable problem, and can be solved numerically by the interior point method efficiently [15]. See [15] and the references therein for more detailed discussions about the determinant maximization problem. To summarize, the minimum GM-MSE problem can be solved numerically efficiently, to a specified precision by the typical SDP solver.

First we observe that substituting any $\mathbf{F}$ with $\mathbf{F}_{new} := \mathbf{F} \mathbf{\Psi}$ for some unitary matrix $\mathbf{\Psi}$ does not change the GM-MSE nor the individual power in each antenna. In high SNR region P_{br} is a Schur-convex increasing function in the vector $\mathbf{g} = [\log(\mathbf{E}_{11}) \log(\mathbf{E}_{22}) \cdots \log(\mathbf{E}_{MM})]$ [12]. Based on those observations, we have

$$P_{br} = \frac{1}{M} \sum_{i=1}^{M} \alpha Q \left(\sqrt{ \beta \cdot \left(\frac{1}{\mathbf{E}_{ii}} - 1 \right) } \right)$$

$$\geq \alpha Q \left(\sqrt{ \beta \cdot \left(\frac{1}{\sqrt[M]{\prod_{i=1}^{M} \mathbf{E}_{ii}}} - 1 \right) } \right)$$

$$= \alpha Q \left(\sqrt{ \beta \cdot \left(\frac{1}{\sqrt[M]{\det(\mathbf{E})}} - 1 \right) } \right). \qquad (7)$$

It is now clear that the lower bound is minimized by minimum GM-MSE. The equality is achieved by choosing the matrix $\mathbf{\Psi}$ to equalize the MSE in each substream.

Now we provide a way to compute one solution for the optimal precoders. Suppose we already found the solution $\mathbf{U}_{\mathrm{GM}}$ to the problem (6). The minimum GM-MSE precoder $\mathbf{F}_{\mathrm{GM}}$ can be taken as any Cholesky factor of $\mathbf{U}_{\mathrm{GM}}$. Suppose $\mathbf{V}$ is the unitary matrix diagonalizing $\mathbf{F}_{\mathrm{GM}}^{\dagger} \mathbf{H}^{\dagger} \mathbf{R}_n^{-1} \mathbf{H} \mathbf{F}_{\mathrm{GM}}$:

$$\mathbf{F}_{\mathrm{GM}}^{\dagger} \mathbf{H}^{\dagger} \mathbf{R}_n^{-1} \mathbf{H} \mathbf{F}_{\mathrm{GM}} = \mathbf{V} \mathbf{\Sigma} \mathbf{V}^{\dagger}.$$

Recall that the MSE matrix will be as in Section II. In [5], the authors showed that there exist unitary matrices $\mathbf{Q}$ and $\mathbf{\Phi}$ such that $(\mathbf{I} + \mathbf{\Sigma})^{-(1/2)} = \mathbf{Q} \mathbf{R} \mathbf{\Phi}^{\dagger}$, where $\mathbf{R}$ is an upper triangular matrix with diagonal terms all equal to the geometric mean of the diagonal terms of $(\mathbf{I} + \mathbf{\Sigma})^{-(1/2)}$. The algorithm for computing these unitary matrices is given in [5] and [19]. The optimal $\mathbf{F}_{\mathrm{opt}}$ can be taken as $\mathbf{F}_{\mathrm{opt}} = \mathbf{F}_{\mathrm{GM}} \mathbf{V} \mathbf{\Phi}^{\dagger}$, where $\mathbf{\Phi}$ is the unitary matrix obtained by the triangular decomposition discussed above. Note that such $\mathbf{\Phi}$ are not unique [4], which means the minimum BER system is not unique.

V. FURTHER DISCUSSION

A. Schur-Convex Objective Functions and Additional Linear Constraints

It can be observed that the discussion 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 optimization problems are provided in [10] for the linear transceiver case, and in [4] and [12] for the DFE case. For all such objective functions, the systems discussed in Sections III and IV are optimal.

It can also be seen that in the framework developed in this correspondence, any finite number of linear constraints on the covariance matrix of the transmitted signals can be further added with no difficulty. This is because when the problem (4) or the problem (6) has one more constraint added: $f(\mathbf{F} \mathbf{F}^{\dagger}) \geq 0$, where $f(\mathbf{F} \mathbf{F}^{\dagger})$ is a linear function in the elements of covariance matrix $\mathbf{F} \mathbf{F}^{\dagger}$ of the transmitted signal, it still remains an SDP.

Several examples of such linear constraints were addressed in [9], such as spectral masks in cable systems to control the crosstalk among DSL users, and limiting the power transmitted along some directions in wireless systems. Here we elaborate further about spatial masks constraints in the wireless systems. Suppose $\mathbf{a}$ is some spatial steering vector of interest, then the power along the direction is proportional to $\mathbf{a}^{\dagger} \mathbf{F} \mathbf{F}^{\dagger} \mathbf{a}$. Suppose we want to limit the power transmitted along this

direction of interest, the constraint on the transmitted signal covariance becomes $\mathbf{a}^\dagger \mathbf{F}\mathbf{F}^\dagger \mathbf{a} \le \alpha$, for some constant α. This equation can be rewritten as $\mathbf{a}^\dagger \mathbf{F}\mathbf{F}^\dagger \mathbf{a} = \mathrm{Tr}(\mathbf{F}\mathbf{F}^\dagger \mathbf{a}\mathbf{a}^\dagger) \le \alpha$, which is a linear constraint in the transmitted covariance matrix $\mathbf{F}\mathbf{F}^\dagger$. Therefore, the framework proposed in this correspondence can be easily modified to include this kind of constraints, both for the linear transceiver and the DFE with linear precoding.

B. Further Remarks

An idea similar to this correspondence is proposed in [9] where the author considered shaping constraints on the linear transceivers. In [9], the constraint on the covariance matrix of the transmitted signals is $\mathbf{F}\mathbf{F}^\dagger \preceq \mathbf{S}$, which means the matrix $\mathbf{S} - \mathbf{F}\mathbf{F}^\dagger$ is positive semidefinite. However, there is a difference between our work and [9], i.e., our constraint is componentwise while the constraint in [9] is the positive definiteness constraint. This is why our approach needs to be more involved (reformulating the problem to be an SDP for solving the minimum-AM MSE). Our approach has some advantages over that in [9].

1) Our work as well as some work in the literature, for example [18] and [6], precisely capture the individual power constraints. As acknowledged by the author in [9], the individual power constraints (1) are replaced with the tighter constraint (as in [9, eq. (7)]): $\mathbf{F}\mathbf{F}^\dagger \preceq \mathrm{diag}([P_1, P_2, \cdots, P_M])$. This artificial replacement yields a solution in which the non-diagonal elements of $\mathbf{F}\mathbf{F}^\dagger$ are zero. We will see that the optimized solutions to the individual power constraint problem need not have zero non-diagonal elements for $\mathbf{F}\mathbf{F}^\dagger$. Therefore, the solution obtained in [9] is a suboptimal solution to the individual power constraint problem. This fact will be amply shown in the numerical simulations.

2) As we argued in Section V-A, the power constraint along a direction should be like what is discussed in V-A. It is shown in this correspondence that this problem can be optimally solved under our framework. In [9] the author needs to find a shaping upper bound for the covariance matrix. However, the procedure of finding the tight upper bound was not trivial [9].

3) Actually, our framework can incorporate the problem discussed in [9]. This can be seen from the fact that the constraint in [9]: $\mathbf{U} \preceq \mathbf{S}$ can be added in our framework, and the problem remains SDP. In our work, we also consider the optimization of the transceivers with DFE and linear precoder, while in [9] only linear transceivers are considered.

However, there are some disadvantages of our formulations compared to [9] when dealing with the shaping constraint problems.

1) Our framework can only deal with square precoding matrix.

2) Given the SDP formulation, the optimal signaling direction cannot be characterized, whereas [9] gives a nice interpretation.

3) The computational complexity of our approach is much higher than [9] because of the need to solve the SDP. Solving an SDP requires $O(M^6)$ flop counts for each iteration.

C. Rectangular Precoder

If the precoder matrix $\mathbf{F}$ is not square, for example, when the channel matrix $\mathbf{H}$ is $N \times P$ and $P > M$, the rank of $\mathbf{U}$ should be no greater than M. This rank constraint should be further added into the problem formulation (4) and (6), which will destroy the convexity of the problem. Generally speaking, the rank constrained problem is difficult to solve optimally. Therefore, we propose a heuristic way to take care of this issue.

When $P > M$, suppose we first relax the rank constraint, then the rank-relaxed covariance matrix $\mathbf{U}$ is solved by the SDP solver as in Sections III and IV. Now we denote $\mathbf{U} = \sum_{i=1}^{P} \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger$, where λ_i are the eigenvalues of $\mathbf{U}$ with non-increasing order and $\mathbf{u}_i$ are the corresponding eigenvectors with dimension $P \times 1$. Then we can take
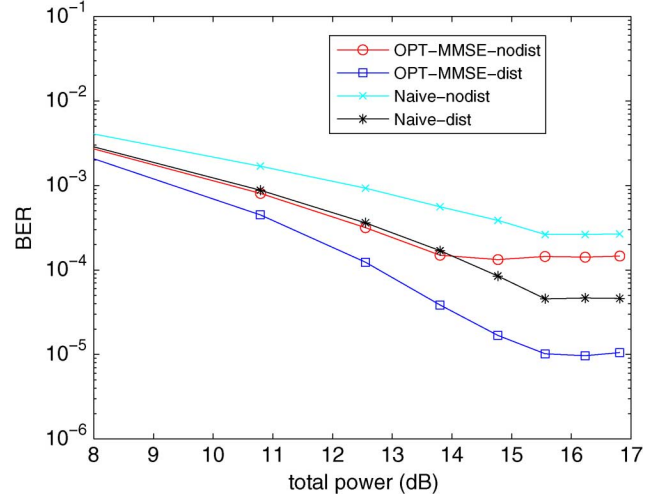

Fig. 2. Comparing four transceivers for 100 channel realizations, with each antenna power ≤ 9. The x axis represents the total power constraint.

$\mathbf{F} = [\sqrt{\lambda_1}\mathbf{u}_1 \cdots \sqrt{\lambda_M}\mathbf{u}_M]\mathbf{\Phi}$, where $\mathbf{\Phi}$ is a $M \times M$ unitary matrix. This will result in $\mathbf{F}\mathbf{F}^\dagger = \sum_{i=1}^{M} \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger$, which is a rank-$M$ approximation of $\mathbf{U}$. Matrix $\mathbf{\Phi}$ can be obtained later as in Sections III and IV, to distribute the MSE equally in each substream. It can be easily shown that with this approximation, the individual and total power constraint are still satisfied provided the original $\mathbf{U}$ is also in the feasible set of the problem formulation (4) and (6). The remaining eigenvalues may be scaled at this point to improve the system performance while maintaining the power constraints.

VI. NUMERICAL RESULTS

In this section we present our simulation results relevant to the discussions carried out in previous sections. For linear transceiver, after formulating the problem as in (4), we use the convex optimization tool "SeDuMi" [13] to obtain the optimal solution.

We choose $M = 4$, $N = 5$, and per-antenna power constraints to be $[P_1, P_2, P_3, P_4] = [9, 9, 9, 9]$. The total power is varied in the simulation. The constellations are all QPSK. The noise is additive white Gaussian, with covariance matrix $\mathbf{R}_n = \mathbf{I}$. "OPT-MMSE-nodist" denotes the optimal MMSE design but without distributing the MSE in each substream. "OPT-MMSE-dist" denotes the optimal MMSE design with the MSE in each substream identical, which is the method proposed in this correspondence. "Naive-nodist" denotes the case where the power constraints are satisfied by using the simple choice $\mathbf{F}_{\mathrm{naive}} = \mathrm{diag}([P_1, P_2, \cdots, P_M])$. if the total power constraint is not violated. If the choice violates the total power constraint, then we take $\mathbf{F}_{\mathrm{naive}} = (P_{\mathrm{total}} / \sum_{i=1}^{M} P_i) \times \mathrm{diag}([P_1, P_2, \cdots, P_M])$ "Naive-dist" corresponds to the case where the precoder matrix is $\mathbf{F} = \mathbf{F}_{\mathrm{naive}}\mathbf{V}\mathbf{\Phi}$, where $\mathbf{\Phi}$ is to force the MSE matrix to have identical diagonal elements. Note that this method is exactly the one that was proposed in [9]. In Fig. 2 we provide the simulation of the linear transceivers for BER averaged over 100 channel realizations. The channel entries are drawn from an i.i.d. Gaussian distribution. It can be seen that the typical performance of the proposed method is significant. When the total power constraint is more than $\sum_{i=1}^{M} P_i = 36$ (or 15.56 dB), the total power constraint is actually inactivated. Therefore, we can see the performance saturate after this point.

VII. CONCLUDING REMARKS

We have presented a convex formulation for optimal transceiver design with per-antenna and total power constraints for both linear and DFE cases. The problems are reformulated as SDPs, which can

be solved efficiently by convex optimization tools. The optimization for average BER as well as other Schur-convex functions were also discussed. The general framework provided in this correspondence is capable of dealing with any finite number of linear constraints added on the covariance matrix of the transmitted signals. Several examples of such problems are discussed. The simulations presented here demonstrate remarkable improvement in performance for jointly optimal transceivers.

REFERENCES

[1] S. S. Chan, T. N. Davidson, and K. M. Wong, "Asymptotically minimum BER linear block precoders for MMSE equalisation," *IEE Proc. Commun.*, pp. 297–304, Aug. 2004.
[2] J. M. Cioffi and G. D. Forney, "Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise," in *Communications, Computation, Control and Signal Processing: A Tribute to Thomas Kailath, A. Paulraj, V. Roychowdhury, C. D. Shaper.* Norwell, MA: Kluwer, 1997, Eds..
[3] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, U.K.: Cambridge Univ. Press, 1985.
[4] Y. Jiang, D. P. Palomar, and M. K. Varanasi, "Precoder optimization for nonlinear MIMO transceiver based on arbitrary cost function," in *Proc. Conf. Information Sciences Systems*, Mar. 2007, pp. 119–124.
[5] Y. Jiang, W. W. Hager, and J. Li, "Generalized triangular decomposition," *Math. Comput.*, vol. 77, no. 262, pp. 1037–1056, Apr. 2008.
[6] K. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. 24, no. 12, pp. 1283–1290, Dec. 1976.
[7] Z. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1037–1052, Apr. 2004.
[8] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications.* New York: Academic, 1979.
[9] D. P. Palomar, "Unified framework for linear MIMO transceivers with shaping constraints," *IEEE Commun. Lett.*, vol. 8, no. 12, pp. 697–699, Dec. 2004.
[10] 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. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.
[11] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Found. Trends Commun. Inf. Theory*, vol. 3, no. 4, pp. 331–551, Nov. 2006.
[12] M. B. Shenouda and T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson–Harashima precoding," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 401–411, Feb. 2008.
[13] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11–12, pp. 625–653, 1999.
[14] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, Mar. 1996.
[15] L. Vandenberghe, S. Boyd, and S. P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 19, no. 2, pp. 499–533, Apr. 1998.
[16] M. K. Varanasi and T. Guess, "An information-theoretic framework for deriving canonical decision-feedback receivers in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 173–187, Jan. 2005.
[17] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 21, no. 12, pp. 3221–3231, Dec. 1994.
[18] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
[19] J. K. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 154–172, Jan. 2005.

Asymptotic Spectral Efficiency of MIMO Ad hoc Networks

Zohreh Motamedi and M. Reza Soleymani

Abstract—In this correspondence, we present an asymptotic analysis of spectral efficiency of MIMO ad hoc networks. We consider a MIMO ad hoc network where L links, each consisting of a transmitter/receiver pair, communicate independent information over the network, simultaneously. Each transmitter node has N_t antennas and each receiver node is equipped with N_r antennas. It is assumed that for each user, perfect channel state information (CSI) is available at its transmitter and receiver. The exact asymptotic spectral efficiency of large MIMO ad hoc networks has been considered to be an unsolved problem until now. A lower bound has been found by Chen and Gans based on using beamforming as a sub-optimal transmission scheme. In this work, we show that the previously presented lower bound is also an upper bound when L goes to infinity, all links use full power and single-user decoding is used; hence, it represents the exact asymptotic spectral efficiency (ASE) of the network. Therefore, the network ASE is equal to the mean value of the largest eigenvalue of the channel covariance matrix which is upper bounded by $(\sqrt{N_t} + \sqrt{N_r})^2$ for large values of N_t and N_r.

Index Terms—Interfering MIMO links, MIMO ad hoc networks.

I. INTRODUCTION

It is well known that in a rich-scattering environment the use of multiple antennas at both transmit and the receive sides significantly improves the capacity of wireless links. The channel capacity and optimum signaling for a single-user MIMO channel, where the only impairment is thermal noise, have been studied in pioneering works of Foschini and Telatar [1], [2].

In multiuser MIMO systems, given the fact that often the co-channel interference is the dominant impairment, the problem of finding the optimal signaling for each user becomes more difficult. Recently, some researchers have studied the use of multiple antennas in ad hoc networks, where MIMO users mutually interfere. In general, the capacity of the interference channel, even for two-user channel, is still an open problem. Therefore, most of the works done so far, have investigated the problem, considering the constraint of employing single-user detection at the receivers [3]–[6].

To find the optimum transmission strategy for a set of interfering MIMO links, an iterative water-filling algorithm is proposed in [3]. At each iteration, assuming that perfect knowledge of CSI and interference is available at each transmitter, each link attempts to selfishly optimize its own channel capacity.

The scenario of a group of MIMO interfering links is also studied in [4] and [5]. Blum [4] investigates the capacity of a system consisting of L interfering transmit/receive pairs when the receivers use single user detection and the transmitters have no knowledge of the CSI. Assuming that all users employ the usual form of the optimum signal in MIMO problems [2], [7], i.e., Gaussian codebooks, Blum shows that in this case, depending on the strength of the interference, the transmitter should either put equal power into each antenna (weak interference)